SLAC-R-661

A STUDY OF NEUTRAL B MESON TIME EVOLUTION USING EXCLUSIVELY RECONSTRUCTED SEMILEPTONIC DECAYS*

Timothy Isaac Meyer

Stanford Linear Accelerator Center
Stanford University
Stanford, CA 94309

SLAC-Report-661
August 2002

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 94309.

A STUDY OF NEUTRAL B MESON TIME EVOLUTION USING EXCLUSIVELY RECONSTRUCTED SEMILEPTONIC DECAYS

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

Timothy Isaac Meyer
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 Rose 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.

Scott Thomas

Approved for the University Committee on Graduate Studies:

Abstract

The Standard Model of particle physics describes the fundamental building blocks of the Universe and their basic interactions. The model naturally describes the time evolution of the basic particles, of which lifetime and mixing are two examples. The neutral B meson, consisting of a bottom quark and an oppositely charged down quark, enjoys a lifetime of about 1.5 ps and the special property of mixing with its antiparticle partner, the $\overline{B}^0$. That is, due to second order weak interactions, the B^0 meson can change into a $\overline{B}^0$ meson and back again as it evolves through time. The details of this behavior offer an opportunity to closely examine the Standard Model.

In this dissertation, I report on a measurement of the lifetime and mixing frequency of the neutral B meson. Using the semileptonic decay channel $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$, we select more than 68,000 signal and background candidates from about 23 million $B\overline{B}$ pairs collected in 1999-2000 with the *BABAR* detector located at the Stanford Linear Accelerator Center. The other B in the event is reconstructed inclusively. By constructing a master probability density function that describes the distribution of decay time differences in the sample, we use a maximum likelihood technique to simultaneously extract the B^0 lifetime and mixing parameters with precision comparable to the year 2000 world average. The results are $\tau_{B^0} = (1.523^{+0.024}_{-0.023} \pm 0.022)$ ps and $\Delta m_d = (0.492 \pm 0.018 \pm 0.013)$ ps^{-1}. The statistical correlation coefficient between τ_{B^0} and Δm_d is -0.22.

I describe in detail several cutting-edge strategies this analysis uses to study these phenomena, laying important groundwork for the future. I also discuss several extensions of this work to include possible measurements of higher order parameters such as $\Delta\Gamma_d$.

Foreword

Spokesperson Stewart Smith recently introduced the *BABAR* experiment before the announcement of a new result with the words, "Our experiment has two products: new science, and highly trained people." This is a fundamental statement. I believe the field of physics has two purposes: discoveries about our Universe and the relationships that govern it, and the rigorous training of individuals in the skills of critical thinking, careful observation, and creative vision. This duality is not divisive; rather it is inclusive – it asserts that there is more to the discipline of physics than merely a derived understanding of the Universe. There is a practice, a paradigm, a set of skills that can be developed in physics that has a greater value. It is this dual character that first inspired me to join the field.

This dissertation is about my training in those areas, and it has been an education filled with joy, challenge, and spirit.

FORGET THE WORLD; THE MUSIC IS NOW. - *DJ Zenith*

Acknowledgements

I am grateful for the many gifts in my life. Earning a Ph.D. in Physics at Stanford is no exception, and I will take a moment here to pay my respects before I take the next fork in the road. I am proud of my connection to these people, and shout out my sincere thanks and praise with great enthusiasm.

Pat – As a mentor and advisor she has been a constant source of support and inspiration, embodying the skills of creative vision and mindful observation. As a friend, she's been a model of balance, openness, and sincerity.

David K – Everyone needs a guide when embarking on something unknown, and David has been that older brother, wiser colleague, and the slightly more sober drinking partner. Compassionate, genuine, and dedicated!

Chris and Chih-Hsiang – My drinking buddies, especially when there wasn't even any alcohol around. Brilliant, talented, and good-natured.

Stanford crew – Fil & Nicci, Travis & Taska, John & Danna, Tali & Barbara, Phil & Veronica, Aaron & Kelsey, Gary & Rosemary, Tarek, Aaron & Holly, Pablo & Cathryn, Jason & Heather (& Zoe), and others; all have been the key ingredients in my Bay Area community of family, intimacy, and joy.

The (Extended) Gomez Family – Axis and Allies, "That was a Murali-style hand of cards," and "Arun, you called it on what?!!" I tried to get away by moving to California, but still they found me. And thank goodness they did, for my love and respect for them all has only increased! (Although I'd swear I can still beat them in euchre!)

The Crew at Emilia's – Truly the greatest part of my experience here has been in meeting these friends, and sharing in the quest for understanding what it means to be human and under authority in this life. Especially Tim the Elder, who has been

both father and friend, and who gave me the courage to seek my true intentions.

Tom – Clearly thinking, and thinking clearly. No one could be more honest, understanding, or insightful. A source of unwavering support and confidence, and especially late-night advice. (And thanks to Janis for keeping him around!)

Todd and Jeff – Compadres from a bygone era, these two have always given me the boost I needed to reach the cookie jar on the top shelf.

The Discoverers of Volleyball – Exhilaration, focus, and smacking something really hard. All necessary parts of the happy graduate student existence.

Mom M, Mom & Dad H – A statement on origins: I couldn't have made it here without them. I treasure their encouragement, excitement, and their excellent examples of being; and I thank them for sharing Allison and Emily with me!

BABAR, SLAC – "Crazy" Italians and all, I am indebted to all these wonderful persons' insight, good humor and advice. (Even while on shift!)

Stanford Physics Department – A tremendous environment that nurtured my personal and professional development, and which has always had its heart in the right place.

Emily - My sweetpea, who knew it all was possible. Her deep love for heart, mind, soul, and strength are my inspiration. She's brought unconditional grace, love, and beauty into my life, and her presence has been one of my greatest blessings.

and lastly, Mr. J – he taught me all about Source and Direction.

Finally, if there are some whom I've forgotten or neglected, do write me; the check will be in the mail.

Contents

List of Tables

List of Figures

Chapter 1

Executive Overview

This study examines the time evolution of neutral B mesons produced at a dedicated collider-detector facility at the Stanford Linear Accelerator Center (SLAC). The collider produces copious amount of e^+e^- annilihation events at the $\Upsilon(4S)$ resonance. These bound states in turn decay predominantly to $B\overline{B}$. Due to the difference in the beam energies in the lab frame, the B mesons are boosted in the "forward" direction, and therefore travel about $250\,\mu$m before decaying. We use the BABAR detector to record the life cycle of the B meson.

Our strategy is to reconstruct candidates in the decay mode $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$, and assign the remaining particle tracks and energy in the event to the other B meson decay. By measuring the vertex of each B, we estimate the decay distance difference that we convert to the decay time difference Δt using our knowledge of the boost between the $\Upsilon(4S)$ frame and the lab. We also use charge correlations between the B meson decay daughters and the parent candidate to determine the flavor of the constituent b quark at the time of decay.

The quality of the reconstructed $D^*\ell$ candidate is estimated through a series of fits to the $m(D^*) - m(D^0)$ spectrum (after grouping the events into similarly-behaved sub-groups) and by the use of relative efficiencies in similarly selected background control samples. (For instance, a sample of data recorded off the $\Upsilon(4S)$ peak is enriched in non-$B\overline{B}$ background contributions.) A probability density function for the distribution of all classes of events (signal and backgrounds) is constructed using the control subsamples and taking into account the cross-contaminations, flavor misidentification rates, and decay time difference resolution functions. The 72-parameter model is fit to about $68,000$ events from the 1999-2000 BABAR dataset that match our selection criteria. The results include the physics parameters τ_{B^0} and Δm_d. Consistency checks are performed to validate the technique, and variations of the process are used to estimate the degree of systematic uncertainty.

The phenomenon of particle-antiparticle oscillations 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], and more recently in the neutrino sector. [2] The time evolution of B^0 mesons is governed by both the overall decay rate $1/\tau_{B^0}$ and the B^0-$\overline{B}^0$ oscillation frequency Δm_d. In the Standard Model of particle physics, this mixing is the result of second-order charged weak interactions involving box diagrams containing virtual quarks with charge 2/3. In B 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]. In the neutral K meson system, mixing also has contributions from real intermediate states accessible to both a K^0 meson and a $\overline{K}^0$. These contributions are expected to be small for B mixing and are neglected in most of this analysis.

We present a measurement of the B^0 lifetime τ_{B^0} and the oscillation frequency Δm_d based on a sample of approximately 14,000 exclusively reconstructed $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ decays selected from a sample of 23 million $B\overline{B}$ events recorded during 1999-2000 at the $\Upsilon(4S)$ resonance with the *BABAR* detector at the Stanford Linear Accelerator Center. In the experiment, 9 GeV electrons and 3.1 GeV positrons annihilate to produce $B\overline{B}$ pairs moving along the electron beam direction (z axis) with a known Lorentz boost of $\beta\gamma = 0.55$, which allows a measurement of the time between the two B decays, Δt.

The proper decay-time difference Δt between two neutral B mesons produced in a coherent P-wave state in an $\Upsilon(4S)$ event is governed by the following probabilities to observe an unmixed event,

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

or a mixed event,

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

Therefore, if we measure Δt and identify the b-quark flavor of each B meson when it decays, we can extract the B^0 lifetime τ_{B^0} and the mixing frequency Δm_d. In this

analysis, one B is reconstructed in the mode $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$, which has a measured branching fraction of $(4.60 \pm 0.21)\%$ [5]. Although the neutrino cannot be detected, the requirement of a reconstructed $D^{*-} \to \overline{D}^0\pi^-$ decay and a high-momentum lepton satisfying kinematic constraints consistent with the decay $B \to D^*\ell\overline{\nu}_\ell$ 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 muon. The charges of the final-state particles are used to identify the meson as a B^0 or a $\overline{B}^0$. The remaining charged particles in the event, which originate from the other B (referred to as $B_{\rm tag}$), are used to identify, or "tag", its flavor as a B^0 or a $\overline{B}^0$. The time difference $\Delta t = t_{D^*\ell} - t_{\rm tag} \approx \Delta z/\beta\gamma c$ is determined from the separation Δz along the boost direction of the decay vertices for the $D^*\ell$ candidate and the tagging B. The average separation is $250\,\mu$m.

The oscillation 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 the measured Δt distributions of events that have been classified as mixed and unmixed. This is in contrast to previous measurements in which only Δm_d is measured with τ_{B^0} fixed to the world average, or in which τ_{B^0} alone is measured. There are several reasons to measure the lifetime and oscillation frequency simultaneously. Since statistical precision for both τ_{B^0} and Δm_d is comparable to the uncertainty on the world average, it is appropriate to measure both quantities rather than fixing the lifetime to the world average. Secondly, since mixed and unmixed events have different Δt distributions, the mixing information for each event improves sensitivity to the Δt resolution function, and a smaller statistical uncertainty on τ_{B^0} results. Also, since $B^0\overline{B}^0$ and B^+B^- 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 B^+B^- background events in the joint lifetime and mixing measurement.

There are three main experimental complications that affect the Δt distributions given in Eqs. 1.1 and 1.2. First, the tagging algorithm, which classifies events into categories c depending on the source of the available tagging information, incorrectly identifies the flavor of $B_{\rm tag}$ with a probability w_c with a consequent reduction of the observed amplitude for the mixing oscillation by a factor $(1 - 2w_c)$. Second, the resolution for Δt is comparable to the lifetime and must be well understood. The

probability density functions (PDF's) for the unmixed (+) and mixed (−) signal events can be expressed as the convolution of the underlying $\Delta t_{\rm true}$ distribution for tagging category c,

$$\frac{e^{-|\Delta t_{\rm true}|/\tau_{B^0}}}{4\tau_{B^0}}\,[1 \pm (1-2w_c)\cos\Delta m_d \Delta t_{\rm true}],$$

with a resolution function $\mathcal{R}(\Delta t_{\rm meas} - \Delta t_{\rm true}; \vec{q}_c)$ that depends on parameters $\vec{q}_c$. A final complication is that the sample of selected $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ candidates is not pure signal.

This study is described in this document with the following organization. The theoretical framework for B^0 time evolution is presented in Chapter 3, while Chapter 4 discusses other techniques for these measurements and summarizes the current state-of-the-art. An overview of the experimental apparatus is described in Chapter 5. The measurement of Δz and the determination of Δt and $\sigma_{\Delta t}$ for each event is discussed in Chapter 6. The b-quark tagging algorithm is described in Chapter 7. In addition to signal events, control samples of events enhanced in each type of background are selected as described in Chapter 8. Chapter 9 details the general strategy involving fits to the $m(D^*) - m(D^0)$ spectrum and the Δt distributions. The method of determining the signal and background probabilities for each event in the signal and background control samples is discussed in Chapter 10. The physics models and Δt resolution functions used to describe the measured distribution for signal and backgrounds are given in Chapter 11. The likelihood is maximized in a simultaneous fit to the signal and background control samples to extract the B^0 lifetime τ_{B^0}, the mixing frequency Δm_d, the mistag probabilities w_c, the signal Δt resolution parameters $\vec{q}_c$, the background Δt model parameters, and the fraction of $B^\pm \to D^{*0}\ell^\pm\nu_\ell X$ decays in the signal sample. The results of the fit are given in Chapter 12. Cross-checks are described in Chapter 13 and systematic uncertainties are summarized in Chapter 14. Finally, some future experimental challenges are discussed in Chapter 15.

Chapter 2

Introduction and Motivation

Contemporary understanding of particle physics is encapsulated in the Standard Model, a description of fundamental particles (leptons and quarks) and the relationships between them (mediated by force-carrying bosons). One of these fundamental particles is the b quark, the second most massive of the six quarks. Although only discovered in 1974 [6], we know a lot about this quark by analogy to its partners, the s and d quarks and because it has been the subject of intense scrutiny. The Standard Model treats matter and antimatter almost identically (the difference is parameterized by CP violation); so the existence of antimatter is almost unavoidable in this framework. The antiparticle of the b quark is the $\overline{b}$, which we expect to behave similarly to the b. Quarks do not manifest themselves freely in Nature; they are bound into composite particles. The two-quark composite is called a meson, and here we study the $\overline{B}$ meson, which contains a b quark and a complementary $\overline{u}$ or $\overline{d}$ quark to make a B^- or $\overline{B}^0$ meson. The motivation for the study described in this dissertation is to precisely measure the life-cycle of the B^0 meson in order to confirm (or reject) the description provided by the Standard Model.

One prediction of the Standard Model is that B^0 and $\overline{B}^0$ mesons can "mix" or transform into each other and back again according to strict guidelines. Two parameters describing the time evolution (i.e., life-cycle) of the B^0 are its lifetime τ_{B^0} (how long it lives) and its mixing frequency Δm_d (the average rate at which B^0 transforms in to $\overline{B}^0$ and vice versa). This study uses a novel technique to measure both of these parameters in the same data sample, which consists of a collection of events wherein which we try to identify the "birth" and "death" of each B meson, produced at a special facility called PEP-II at the Stanford Linear Accelerator Center. The BABAR experiment is used to record and analyze the B meson data, which involves the work of more than 500 scientists. In this study, we seek to identify and characterize a subset of the B mesons in the sample by identifying those which decay in a particular way: the semileptonic decay $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$.

2.1 The Standard Model of Particle Physics

The Standard Model

Current understanding of particle physics is described within the framework of the Standard Model. This picture enumerates the essential sets of particles and describe the forces that act between them. It is under this (somewhat) short list of rules that most of the Universe operates today. However, the Standard Model isn't complete; or rather, it is not yet satisfying as a scientific theory. * There are phenomena in Nature that are not described in the Model (such as gravity), and furthermore, we suspect that there are connections between phenomena that we have not yet elucidated. For instance, one needs the physics of Standard Model and 18 parameters [7] in order to cook up the Universe as we presently see it. The 18 independent parameters are divided up between particle masses and strengths of their interactions (couplings); most of these have been determined experimentally (*i.e.,* measured in the lab for our Universe), but not all. Depending on what we learn about the neutrino portion of the Standard Model, we could even see the number of input parameters grow by as much as 9.

The Particles

The Standard Model describes fundamental particles. In general, particles are distinguished from each other by quantum numbers, which are (by construction) those properties of a particle which are conserved under certain interactions. The most familiar quantum numbers include (electric) charge, color, spin, and flavor. Quantum numbers are the key to "telling things apart," and they are used to distinguish individual particles and multi-particle systems (states). Noether's Theorem in quantum field theory [8] guarantees us a conserved quantity for each symmetry of a given system, which can be labelled by the quantum number. Electric charge, for instance, is just the quantum number for a system which corresponds to the conserved quantity associated with a local gauge symmetry. [9].

*This is an aesthetic impression, perhaps anthropomorphic, but it is widespread.

The basic particles are divided into two groups depending on their spin: *fermions* are those particles who spin is half-integer, and *bosons* are particles with integer spin. The fermion or boson nature of a particle determines what "statistics" it obeys, that is, the rules which must be followed when combining them into multi-particle systems. For instance, fermions obey Fermi statistics, which results in the Pauli Exclusion Principle [10], a statement that no two identical fermions can occupy the same state at the same time. That is, no two identical fermions can have the same quantum numbers. Bosons on the other hand, can occupy the same state, and in fact, multiply occupied states are favored. Fig. 2.1 [11] shows the set of fermions and bosons that are included in the Standard Model.

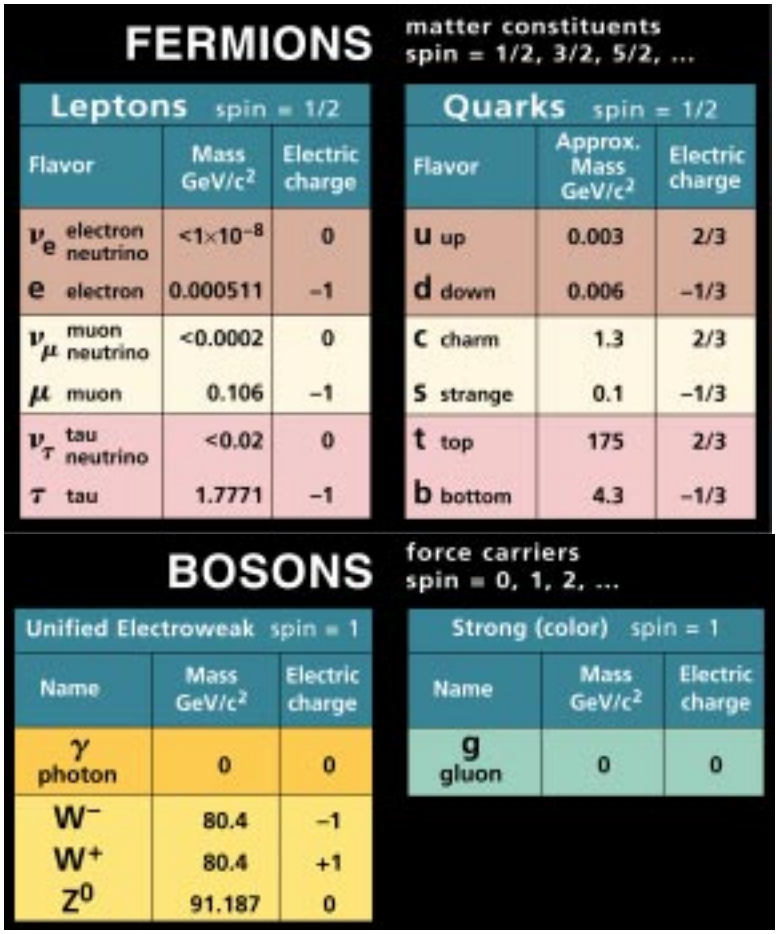
Matter is made up of the fermions, divided into leptons and quarks, in turn grouped into three families or generations, while all of the force-carrying particles are bosons. Each generation contains one "up-type" quark, one "down-type" quark, a lepton, and a corresponding neutrino. "Up-type" quarks have charge $+2/3$ while "down-type" quarks have charge $-1/3$. Each particle has an antimatter partner that has opposite electric charge and flavor. Quarks can be combined into particles called hadrons, identified as mesons ($q_1\overline{q}_2$) and bayrons ($q_1q_2q_3$).

The meson we study here is the B^0 meson, a $\overline{b}$ quark paired with a d quark. This particle is electrically neutral, and has a corresponding neutral antiparticle, $\overline{B}^0$, composed of a b quark and a $\overline{d}$ quark. Another B meson considered in this analysis is the B^+ meson, which is the charged version, containing instead a u quark partnered with the $\overline{b}$.

The Interactions

The Standard Model describes four fundamental forces in the Universe: gravity, electromagnetism, the weak force, and the strong force (in increasing order of relative strength). Each force is mediated by a set of particles, or rather, represented in particle interactions, by force-carrying bosons. The four forces are depicted in Fig. 2.2 [11].

A key aspect of the Standard Model is the notion of renormalization; this sophisticated notion helps describe how to use the theory once we "place a cut to cancel



FERMIONS matter constituents spin = 1/2, 3/2, 5/2, ...

Leptons spin = 1/2

Flavor	Mass GeV/c^2	Electric charge
ν_e electron neutrino	$<1\times10^{-8}$	0
e electron	0.000511	−1
ν_μ muon neutrino	<0.0002	0
μ muon	0.106	−1
ν_τ tau neutrino	<0.02	0
τ tau	1.7771	−1

Quarks spin = 1/2

Flavor	Approx. Mass GeV/c^2	Electric charge
u up	0.003	2/3
d down	0.006	−1/3
c charm	1.3	2/3
s strange	0.1	−1/3
t top	175	2/3
b bottom	4.3	−1/3

BOSONS force carriers spin = 0, 1, 2, ...

Unified Electroweak spin = 1

Name	Mass GeV/c^2	Electric charge
γ photon	0	0
W^-	80.4	−1
W^+	80.4	+1
Z^0	91.187	0

Strong (color) spin = 1

Name	Mass GeV/c^2	Electric charge
g gluon	0	0

Figure 2.1: Fundamental bosons and fermions defined in the Standard Model of particle physics.

out the infinities" that arise in field theory calculations. But it is actually much more about the observation that the forces can behave differently at different energy scales because we generally work with an *effective* theory. In fact, the Standard Model suggests that unification of the forces is possible at a very high energy scale: although the four forces manifest themselves in an apparently unrelated fashion under ordinary conditions , they can be viewed as different aspects of the same fundamental interaction (at a sufficiently high energy), and so become fundamentally connected at a deep level. One might liken this situation to an ideal gas: at high temperature in the gaseous phase, all of the particles behave identically. As one lowers the temperature,

however (*i.e.,* the average energy scale), different substances will begin to liquefy and even solidify before others, taking on very different forms and behaviors. Although this analogy is quite limited, it illustrates one point: in the Standard Model, the four distinct forces (well, except for gravity) we observe now are actually the "frozen-out" components of some higher-energy Grand Unified Force. For instance, we often refer to the combination of the weak and electromagnetic forces above a certain energy threshold (known as the electroweak symmetry breaking scale) as the "electroweak" force.

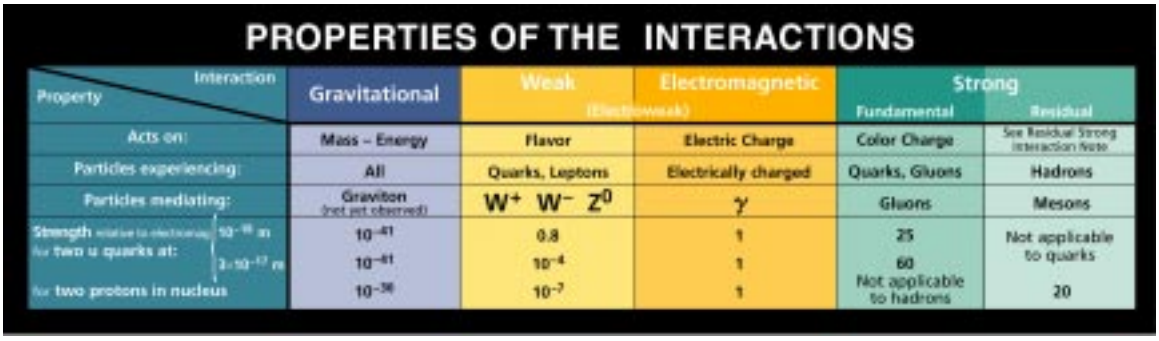


Figure 2.2: Fundamental interactions defined in the Standard Model of particle physics.

The interactions (forces) described by the Standard Model are elegantly summarized in the Standard Model Lagrangian. A Lagrangian is a function that describes the state of a dynamic system in terms of generalized coordinates and their time derivatives. In this study, we are primarily interested in the electroweak portion of the Lagrangrian, because it is responsible for the phenomenon of B^0-$\overline{B}^0$ mixing.

2.1.1 Symmetries

We've already seen several examples of the important role that symmetries play in physics. A particular interaction is often observed to obey many different conservations laws and so the mathematical description of the phenomena has to fulfill stringent invariance requirements. (Conserved quantum numbers are associated with operators that commute with the Hamiltonian.) There are several so-called *discrete*

symmetries that, if true in the Universe, restrict the class of possible models. These symmetries are C (charge conjugation), P (parity), and T (time reversal).[†] Let us define these symmetries here, as we may revisit some of them in Chapter 15.

Quantum field theory, by construction, seeks to be relativistic, such that the Minkowski interval $t^2 - x^2$ is preserved under transformations both discrete and continuous. The set of continuous transformations that preserve this norm are the Lorentz transformations, including translation, rotations, and boosts. It can be shown that with the addition of three discrete symmetry transformations to the Lorentz group of operators and gauge transformations, one forms a complete basis of the Minkowski interval preserving transformations [13]. These separate transformation can be performed in sequence as well, and one could find that, for instance, while the electroweak interaction violates *CP* symmetry, it preserves the more general *CPT* symmetry. *CPT* has a rather special role in quantum field theory, as it's conservation is guaranteed solely from the reasonable assumptions of Lorentz invariance and locality. [8]

Parity

The *parity* transformation reverses the signs of the 3 spatial components of a four-vector: $(t, x) \rightarrow (t, -x)$ and $(E, \vec{p}) \rightarrow (E, -\vec{p})$. That is, the effect is to reverse the handedness of a particle, often likened to viewing the particle in a mirror. The effect is to reverse the sign of a particle's momentum while preserving it's spin. The parity operator is unitary, suggested by the property that a second application returns the original state.

Parity was shown to be violated in weak decays in 1957 by Wu *et al.* [14] when examining β decays of ^{60}Co after aligning the cobalt nuclei with an external magnetic field and analyzing the angular distribution of the emitted electron spectrum. This is the classic example of parity violation: the absence of right-handed neutrinos, as

[†]One of the primary goals of the *BABAR* experiment is the study of *CP*-violation in the B^0 meson system. That study is parallel but separate from this analysis, so we choose to avoid that interesting but distracting discussion. See Ref. [12] for more discussion.

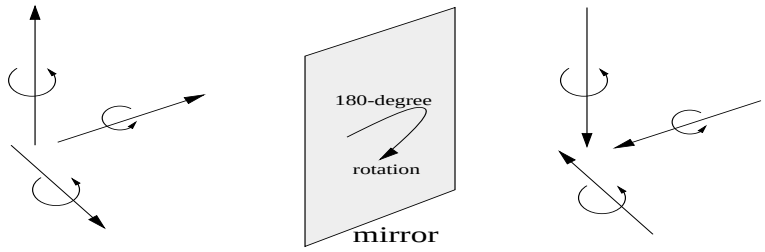


Figure 2.3: Pictorial representation of the parity transformation.

in this decay rate and it's parity-conjugated decay.:

$$\Gamma(\pi^+ \to \mu^+ \nu_L) \neq \Gamma(\pi^+ \to \mu^+ \nu_R) = 0.$$

Charge Conjugation

The charge conjugation transformation is defined to be the transformation of a particle into its antiparticle without changing its momentum or spin (*i.e.,* the particle-antiparticle symmetry operation). Electrodynamics explicitly conserves charge conjugation: although the fields and potentials change their sign under the C symmetry transformation, the Lorentz force incorporates the sign of the charge, and the final physics is invariant. The only eigenstates of the C operator are neutral particles that are their own antiparticle. Charge conjugation symmetry is violated in weak decays, as indicated in this pair of decay rates (there are no left-handed anti-neutrinos):

$$\Gamma(\pi^+ \to \mu^+ \nu_L) \neq \Gamma(\pi^- \to \mu^- \overline{\nu}_L) = 0.$$

Time Reversal

The time reversal transformation reverses the momentum and spin of a particle and reverses the sign of the time component (*i.e.,* interchanging the forward and backward light-cones). This is often likened to watching a film of the process in reverse. If the *CPT* theorem is true for our Universe, we expect to see time reversal violation in weak decays, since we've already seen that *CP* is not conserved. Because the experiments are challenging, evidence for time reversal violation has only recently

observed by CPLEAR. [‡] [15]

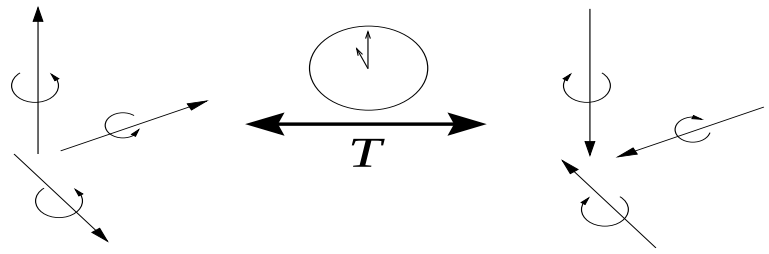


Figure 2.4: Pictorial representation of the time-reversal transformation.

	C	P	T	CP	CPT
Scalar	$+1$	$+1$	$+1$	$+1$	$+1$
Pseudoscalar	$+1$	-1	-1	-1	$+1$
Vector	-1	$\begin{pmatrix} +1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} +1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} -1 \\ +1 \\ +1 \\ +1 \end{pmatrix}$	-1
Pseudovector	$+1$	$\begin{pmatrix} -1 \\ +1 \\ +1 \\ +1 \end{pmatrix}$	$\begin{pmatrix} +1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} -1 \\ +1 \\ +1 \\ +1 \end{pmatrix}$	-1
Tensor	-1	$\begin{pmatrix} +1 & -1 & -1 & -1 \\ -1 & +1 & +1 & +1 \\ -1 & +1 & +1 & +1 \\ -1 & +1 & +1 & +1 \end{pmatrix}$	$\begin{pmatrix} -1 & +1 & +1 & +1 \\ +1 & -1 & -1 & -1 \\ +1 & -1 & -1 & -1 \\ +1 & -1 & -1 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & +1 & +1 & +1 \\ +1 & -1 & -1 & -1 \\ +1 & -1 & -1 & -1 \\ +1 & -1 & -1 & -1 \end{pmatrix}$	$+1$
Derivative Operator	$+1$	$\begin{pmatrix} +1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} -1 \\ +1 \\ +1 \\ +1 \end{pmatrix}$	$\begin{pmatrix} +1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$	-1

Table 2.1: Table of how the discrete symmetry operators, C, P, and T affect the basic Lorentz objects: eigenvalues and transformed objects are shown as appropriate.

[‡]Indeed, one prospect for *BABAR* is the observation of time reversal violation in b quark system.

2.1.2 Comments on the Electroweak Interaction

The electroweak Lagrangian is based on the SU(2)×U(1) symmetry group containing four bosons: a massless isovector triplet $\mathbf{W}_\mu$ for the SU(2) of weak isospin, and a massless isosinglet B_μ for the the U(1) of weak hypercharge. [16] Above the spontaneous symmetry-breaking threshold, all four bosons ($W^\pm, W^0, B^0$) are massless and the SU(2)×U(1) symmetry is unbroken. Below the symmetry-breaking threshold, the charged bosons and one linear combination of the neutral bosons acquire mass ($W^\pm, Z^0$) while the orthogonal combination of neutral bosons remains massless (the photon). The SU(2)×U(1) symmetry is thus not completely broken. There remains a U(1) symmetry which gives rise to electromagnetism and the photon.

Following the discussion in Ref. [17], we can write out the interaction term in explicit form:

$$\mathcal{L}_{\mathrm{EW}} = -i\left[g\mathbf{J}^\mu \cdot \mathbf{W}_\mu + \frac{g'}{2}(J^Y)^\mu B_\mu\right]. \tag{2.1}$$

This is elegant and compact, but requires some explanation. Loosely speaking, we define the components as

- $\mathbf{W}_\mu$ is the isotriplet of massless Yang-Mills gauge fields (bosons) introduced to preserve the SU(2) symmetry of the Lagrangian,
- B_μ is the isosinglet gauge field associated with the U(1) symmetry for weak hypercharge Y,
- $\mathbf{J}^\mu$ is the weak isopin current, *i.e.,* a combination of fermions fields[§] that is conserved due to the SU(2) symmetry of weak isospin [¶] (in fact, this symmetry is only applied to left-handed fermion fields),
- J^Y is the weak hypercharge current, *i.e.,* a combination of fermion fields that is conserved due to the U(1) symmetry (operating on both left- and right-handed fermion fields),

[§]In fact, if we impose Lorentz invariance upon our Lagrangian, we are restricted to a finite set of ways in which we can combine fermion fields to form these currents!

[¶]There are *three* conserved charges, corresponding to the three isospin components of $\mathbf{J}^\mu$.

- g is the coupling of the isotriplet of vector fields $\mathbf{W}_\mu$ to the weak isospin current $\mathbf{J}^\mu$,

- $\frac{g\prime}{2}$ is the coupling of the vector field B_μ to the weak hypercharge current, J^Y

This model assumes all the particles (both the fermions and bosons) are massless. To observe the unification of electromagnetism and the weak interaction, we note that the fields W_μ^3 and B_μ are neutral fields, and can be linearly combined (*mixed*) to form the physical fields A_μ and Z_μ:

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \\ A_\mu &= B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W \\ Z_\mu &= -B_\mu \sin\theta_W + W_\mu^3 \cos\theta_W. \end{aligned} \tag{2.2}$$

Identifying $g/g\prime = \tan\theta_W$ and substituting back into the Lagrangian term with $J_\mu^\pm \equiv J_\mu^1 \pm J_\mu^2$, we arrive at:

$$\mathcal{L} = \frac{g}{\sqrt{2}}(J_\mu^- W_\mu^+ + J_\mu^+ W_\mu^-) + \frac{g}{\cos\theta_W}(J_\mu^3 - \sin^2\theta_W J_\mu^{e.m.})Z_\mu + g\sin\theta_W J_\mu^{e.m.} A_\mu, \tag{2.3}$$

where the first term can be associated with the weak charged current interaction, the second with the weak neutral current, and the final one with the traditional electromagnetic neutral current. To make the identification complete, we assign $e = g\sin\theta_W = g'\cos\theta_W$, where θ_W is called the weak mixing (or Weinberg) angle. [18] The neutral fields Z_μ and A_μ are distinct operators and correspond to the Z^0 and photon particles with definite mass, and the charged $W_\mu^\pm$ fields correspond to the charged $W^\pm$ bosons.

We summarize our discussion of the unbroken electroweak theory by showing the interaction between the gauge bosons and the fermions. First, we explicitly write out

the fermionic field components of the conserved currents: (here for leptons)

$$\begin{aligned} J_\mu^- &= \overline{\nu}_L \gamma_\mu e_L \\ J_\mu^+ &= \overline{e}_L \gamma_\mu \nu_L \\ J_\mu^3 &= \frac{1}{2}\overline{\nu}_L \gamma_\mu \nu_L - \frac{1}{2}\overline{e}_L \gamma_\mu e_L \\ J_\mu^Y &= -2\overline{e}_R \gamma_\mu e_R - \overline{e}_L \gamma_\mu e_L - \overline{\nu}_L \gamma_\mu \nu_L \end{aligned} \tag{2.4}$$

We can then generalize to include quarks and write for the general Dirac spinor ψ,

$$\begin{aligned} J_\mu^\pm &= \sqrt{2}\overline{\psi}\gamma_\mu \mathbf{T}^\pm \psi \\ Z_\mu &= \overline{\psi}\gamma_\mu [T_3 - \sin^2\theta_W Q]\psi \\ J_\mu^{e.m.} &= \overline{\psi}\gamma_\mu Q \psi \end{aligned} \tag{2.5}$$

where $\mathbf{T}$ is the weak isospin operator (vanishing on ψ_R and having representation $\mathbf{T} = \frac{1}{2}\tau$ with the Pauli matrices τ for the isodoublets ψ_L), and Q is the electric charge. All the fermion fields, now, can interact with the gauge fields.

The gauge bosons are still massless in this description, however (as are the fermions). We accommodate massive gauge bosons by introducing the *Higgs field* Φ, an SU(2) doublet of scalar fields. We arrange for the potential to have a ground state with non-vanishing expectation value in the physical vacuum state. Spontaneously breaking this symmetry generates masses while maintaining renormalizibility, but leaves the gauge symmetry of the Lagrangian hidden; we have introduced a preferred direction in weak isospin-hypercharge space. Without going into details, the procedure is to add a potential to the Lagrangian for the new Higgs field:

$$V = \mu^2|\Phi|^2 + \lambda|\Phi|^4 \qquad \mathcal{L}_\Phi = |D_\mu \Phi|^2 - V(|\Phi|^2)\ , \tag{2.6}$$

where we have introduced the new weak isodoublet of fields

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix}\ . \tag{2.7}$$

We require $\lambda > 0$ to bound the potential as $\Phi \to \infty$, and consider the "interesting" case where $\mu^2 < 0$. The potential V has a minimum along $\Phi^\dagger\Phi = -\frac{\mu^2}{2\lambda}$, a locus of points that is invariant under rotations in the SU(2) space. We study the particle spectrum, which requires perturbation theory near the minimum. We must choose a particular direction in which to expand, however. This spontaneously breaks the SU(2) symmetry.

The original field Φ had four degrees of freedom, one for each Φ_i. The Standard Model choice for Φ is $\Phi_1 = \Phi_2 = \Phi_4 = 0, \Phi_3^2 \equiv v^2$, bringing us to

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} . \tag{2.8}$$

We have eliminated three degrees of freedom by choosing the *unitary* gauge such that three of the four fields have been absorbed. This is "legal" since the Lagrangian is symmetric under local gauge transformations. However, these three degrees of freedom will later emerge as mass terms for the gauge fields themselves.

The requirement of local gauge symmetry requires the introduction of a massless vector field and associated covariant derivatives, and so with the addition of Φ, we modify the Lagrangian appropriately. We obtain Higgs-$\mathbf{W}_\mu$ and Higgs-B_μ interactions from the covariant derivative, after substituting in the form of Φ as chosen above:

$$|D_\mu \Phi|^2 \sim |(-ig\frac{\tau}{2} \cdot \mathbf{W}_\mu - i\frac{g'}{2}B_\mu)\Phi|^2 \tag{2.9}$$

$$= \frac{1}{8} \left| \begin{pmatrix} gW_\mu^3 + g'B_\mu & g(W_\mu^1 - iW_\mu^2 \\ g(W_\mu^1 + iW_\mu^2 & -gW_\mu^3 + g'B_\mu \end{pmatrix} \begin{pmatrix} 0 \\ v \end{pmatrix} \right|^2 . \tag{2.10}$$

Working through the math and using the mixed fields A_μ and Z_μ that we already deduced in the electroweak unification, we arrive at the conclusions for the masses of

the photon, Z^0, and $W^{\pm}$:

$$m_A = 0 \tag{2.11}$$

$$m_Z = \frac{1}{2}\sqrt{g^2 + g'^2} \tag{2.12}$$

$$m_W = \frac{1}{2}vg \tag{2.13}$$

$$m_h = \sqrt{-2\mu^2} \tag{2.14}$$

The physical states with mass are the full set of final bosons: the charged particles, $W^{\pm}$, and the neutrals Z^0 and A_μ (the photon), and of course, the Higgs. The result of the Higgs mechanism is that vacuum everywhere can emit or absorb (colorless) quanta of the Higgs field that carry weak isospin and hypercharge. The fermions, W, and Z bosons which can couple to the quantum field effectively gain mass, but particles such as the photon and gluon which do not interact with it remain massless. In summary, the Lagrangian has remained gauge invariant, but the vacuum has not.

2.1.3 The CKM Matrix

Fermion masses can be generated in the spontaneous symmetry breaking mechanism if we include Yukawa interactions between the physical states and Higgs field Φ (using coupling parameters g and the conjugate isodoublet $\tilde{\Phi}$) of the form

$$g_d^{ij}(\overline{u}_i, \overline{d}_i)_L \tilde{\Phi}^\dagger d_{j\,R} + g_u^{ij}(\overline{u}_i, \overline{d}_i)_L \Phi^\dagger u_{j\,R} + h.c.\,, \tag{2.15}$$

where we have just used the quark doublets $(u_i, d_i)_L$ and $i, j = \{1..3\}$ corresponding to the three generations. However, the weak interactions actually operate on the isospin doublets, $(u_i, d'_i)_L$, which is, in general, a different set of basis vectors. We are forced to introduce the unitary mixing matrix V, the so-called CKM matrix [4] which rotates the quark mass eigenstates into the weak eigenstates. (In short, the

mass eigenstates are not weak eigenstates.)

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.16}$$

The modified form of the charged current, for instance, is

$$J^{+}_{L\mu} = \overline{(u,c,t)}_L \gamma_\mu V \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L . \tag{2.17}$$

Note that this mixing occurs only in the quark sector since we have assumed that the neutrinos are massless; such mixing is therefore "invisible" in the lepton sector. By convention, the mixing is assigned completely to the "down-type" quarks.

The CKM matrix V is a complex, unitary matrix. As such, it has at most 9 independent real parameters since the unitarity conditions offer 9 constraints on the original set of 18 parameters.

$$V^\dagger V = V V^\dagger = 1 \quad \Rightarrow \quad \sum_j V^*_{ji} V_{jk} = \sum_j V_{ij} V^*_{kj} = \delta_{ik} \tag{2.18}$$

Furthermore, we can absorb a relative phase in each left-handed field, eliminating 5 more degrees of freedom. This leaves us with 4 physically independent parameters describing the CKM matrix [‖], and as such, can be envisioned as three Euler angle rotations and one phase insertion. [**]

The charged current portion of the electroweak Lagrangian for the quark fields

[‖] In general, the CKM matrix for N generations has $(N-1)^2$ parameters.

[**] Recall that the general $N \times N$ unitary matrix has N^2 independent parameters. The fundamental representation of SO(N) uses $N \times N$ orthogonal matrices which have $(N^2 - N)/2$ independent parameters, where each matrix represents a rotation around a real axis. Finally, we compare the number of free parameters in the CKM matrix to the number of the SO(N) matrix: $(N-1)^2 - (N^2 - N)/2 = \frac{1}{2}(N-1)(N-2)$. This difference is the excess of parameters which cannot be associated with rotations about any of N available axes. $N > 2$ is required to have a positive excess, and hence to enjoy *CP* violation in the CKM matrix.

becomes

$$\mathcal{L}_{\rm EW}^{cc} = \frac{g}{\sqrt{2}} \left\{ \bar{u}_i^L \gamma^\mu W_\mu^+ \mathsf{V}_{ij} d_j^L + \bar{d}_i^L \gamma^\mu W_\mu^- \mathsf{V}_{ij}^* u_j^L \right\} \tag{2.19}$$

with u_i^L representing the vector of up-type quarks and d_i^L representing the down-type quarks. Applying the *CP* operator to the Lagrangian, one obtains:

$$CP(\mathcal{L}_{\rm EW}^{cc}) = \frac{g}{\sqrt{2}} \left\{ \bar{d}_i^L \gamma^\mu W_\mu^- \mathsf{V}_{ij} u_j^L + \bar{u}_i^L \gamma^\mu W_\mu^+ \mathsf{V}_{ij}^* d_j^L \right\} \tag{2.20}$$

which is identical to Eq. 2.19 for the complex conjugation of V if we reverse of order of the summands. Thus, if we could find a basis for which the elements of V are real, then *CP* is a good symmetry. Based on our parameter counting earlier, we saw that in general, the CKM matrix will have a complex phase, which would generate *CP* violation as described here if we can show it to be non-zero.

The unitarity conditions on the CKM matrix are often used to construct so-called "triangles" in the complex plane. We list the six vanishing ones here for completeness (the diagonal products yield unity):

$$V_{cd}V_{ud}^* + V_{cs}V_{us}^* + V_{cb}V_{ub}^* = 0 \tag{2.21}$$

$$V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* = 0 \tag{2.22}$$

$$V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* = 0 \tag{2.23}$$

$$V_{us}^*V_{ud} + V_{cs}^*V_{cd} + V_{ts}^*V_{td} = 0 \tag{2.24}$$

$$V_{ub}^*V_{us} + V_{cb}^*V_{cs} + V_{tb}^*V_{ts} = 0 \tag{2.25}$$

$$V_{ub}^*V_{ud} + V_{cb}^*V_{cd} + V_{tb}^*V_{td} = 0 \tag{2.26}$$

The last unitarity condition listed is traditionally the one used to construct the "Unitarity Triangle," a geometrical way of depicting the vanishing sum of three complex numbers; see Fig. 2.5.

The relative sizes of the CKM elements are key to understanding electroweak interaction rates. The most popular parameterization uses four values (λ, A, ρ, η) to

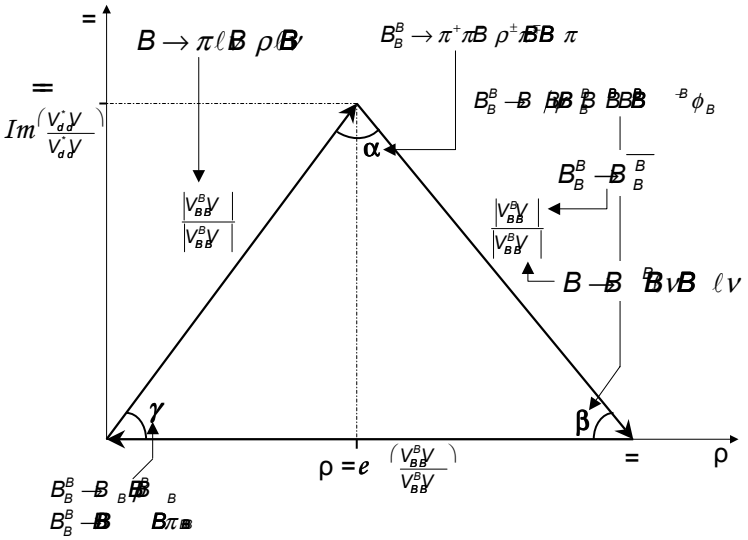


Figure 2.5: The standard CKM triangle expresses the unitarity condition between the first and third columns.

describe the matrix in the following way (known as the Wolfenstein parameterization [19]):

$$V_{CKM} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.27}$$

We observe that the diagonal elements are indeed the largest, and furthermore that some of the entries are very small (the ranges are the experimental bounds):

$$|V^{ij}_{CKM}| = \begin{pmatrix} 0.9742 - 0.9757 & 0.219 - 0.226 & 0.002 - 0.005 \\ 0.219 - 0.225 & 0.9734 - 0.9749 & 0.037 - 0.043 \\ 0.004 - 0.014 & 0.035 - 0.043 & 0.9990 - 0.9993 \end{pmatrix}. \tag{2.28}$$

Perspective

Where has this taken us? We have described the weak component of the Standard Model Lagrangian, and as such we have understood that it can be responsible for

flavor-changing charged current interactions. And because of the $SU(2)_L$ symmetry of weak isospin, the "down-type" quarks can mix according to the charged current prescription. In particular, we see that the the b quark can change flavor to a c quark via a charged current interaction. Later we'll see that it is possible for a b quark to transition into a d quark, via second order processes (*i.e.,* loops). Finally, we've seen two examples of the idea of "mixing": the mass eigenstates of the quarks are mixed to form the weak eigenstates, and the two neutral components of the electoweak gauge field were mixed to form the neutral and electromagnetic currents. The lesson is that if there's more than one interaction in a theory, the natural basis for one interaction will not necessarily be the states distinguished by another, *i.e.,* different dynamical processes usually have distinct sets of eigenstates that will be linear combinations of each other.

2.1.4 Lifetime

We next give meaning to the lifetime of a particle, since the measurement of the lifetime of the B^0 meson is one of the two goals of this study. By lifetime, we refer to the length of time between the "birth" and "decay" of a particle. The decay of any particular particle is a random process, but in an ensemble of many particles we can describe the average evolution by observing that the ensemble population decreases at a constant fractional rate. The decay process is stochastic or "memoryless," which means that the chance of decay for a particle is independent of how long it has "lived" so far. These observations lead to the following mathematical description. Introducing the decay rate Γ, we write the differential equation relating the fractional decay rate to the population at time t:

$$\frac{dN}{dt} = -\Gamma \cdot N(t) \tag{2.29}$$

$$\therefore N(t) = N(0) \exp^{-\Gamma t} \ . \tag{2.30}$$

The solution is the familiar negative exponential, and we define the lifetime $\tau \equiv 1/\Gamma$ to be the time that it takes for $(1 - 1/e) \sim 64\%$ of the population to decay. We call Γ

the decay rate in general, and in particle physics in particular, we also call it the *decay width*. At any point in time, a surviving particle has a $1/e$ chance of surviving for another lifetime. Again, the lifetime for any particular particle is a random variable, but in an ensemble we observe that the distribution of many particles' lifetimes follows a negative exponential law.

Most composite particles decay quickly. It is only in "rare" circumstances where special symmetries are at work that composite particles have an extended lifetime. Particles can decay via the strong, weak or electromagnetic forces, which operate on characteristically different timescales. Typical lifetimes are generally determined by the force mediating the decay, but the kinematics of the decay and various symmetries can also greatly affect the result. The typical lifetime for particles decaying strongly is 10^{-23} sec, electromagnetically 10^{-16} sec, and decaying weakly is about 10^{-8} sec.

The decay of any particle can be factorized into two parts: one part representing the pure kinematics of the decay (the so-called *phase space* contribution) and the other part the matrix element between the initial and final states. Note that the heavier mesons will have larger phase space available for decays, making them generally shorter-lived. The matrix element carries the information about the physics involved in the particular decay.

For the decay $X \rightarrow 1 + 2 + ... + n$: [9], the differential decay rate is

$$d\Gamma = \frac{1}{2E_X}|\mathcal{M}|^2 \frac{d^3p_1}{(2\pi)^3 2E_1} ... \frac{d^3p_n}{(2\pi)^3 2E_n} (2\pi)^4 \delta^{(4)}(p_X - p_1 - ... - p_n) \tag{2.31}$$

where $\mathcal{M} = \langle f|\mathcal{H}|i\rangle$ is the matrix element sandwich about the interaction Hamiltonian, $\mathcal{H}$. The matrix element represents the overlap between the initial and final states, as described by the physics at work, and is really a measure of the connection between them. Angular dependence can be delivered by the matrix element $\mathcal{M}$, amongst other differential properties, which is why we write Eq. 2.31 as the differential decay rate. The net decay width to a specific final state i, after integrating over all differentials is the partial decay width Γ_i. The total decay rate of a particle is the

sum of all individual channels:

$$\Gamma_{total} = \sum \Gamma_i \ . \tag{2.32}$$

This in turn means that the lifetime of a particle is determined by the details of all its available decay channels, both in number and in kind. This is a powerful statement: we cannot speak of the "lifetime for the decay to a particular final state'," although the partial decay widths control the overall lifetime. * For instance, the difference in lifetimes for the B^0-$\overline{B}^0$ system is controlled by the difference in available number of decay modes. We'll learn more about the B^0 lifetime, τ_{B^0}, in Chapter 3.2.

2.1.5 Mixing

We now introduce the second parameter which this analysis measures, the mixing frequency of the B^0-$\overline{B}^0$ meson system, Δm_d. We've already seen some examples of mixing when we discussed the electroweak sector of the Standard Model. The key idea is that mixing occurs because the flavor eigenstates are not equivalent to the mass eigenstates; *i.e.,* one cannot measure both the mass and the flavor of the particle simultaneously. As such, time evolution (according to the Hamiltonian governing the system) will rotate the flavor eigenstates as a function of time while it preserves the mass eigenstates. In the Standard Model, the four pairs of neutral mesons that decay weakly share the mixing diagrams shown in Fig. 2.1.5. Such a process is the result of second-order charged current interactions involving box diagrams containing virtual quarks.

The details of the mixing, of course, depend on the meson pair under consideration. Because of the different CKM matrix elements in question (as mentioned before, these mixing diagrams rely on the charged current interaction), mixing occurs at different rates in each of the neutral meson conjugate pair systems. For instance, K^0-$\overline{K}^0$ mixing relies on the CKM matrix elements $V_{is} \times V_{id}$ which is of order λ, but the matrix elements of the B^0-$\overline{B}^0$ system are instead of $\mathcal{O}(\lambda^3)$ and significantly suppressed.

*Quantum mechanics is also at work here: we don't know how the particle will decay until we actually see that it does.

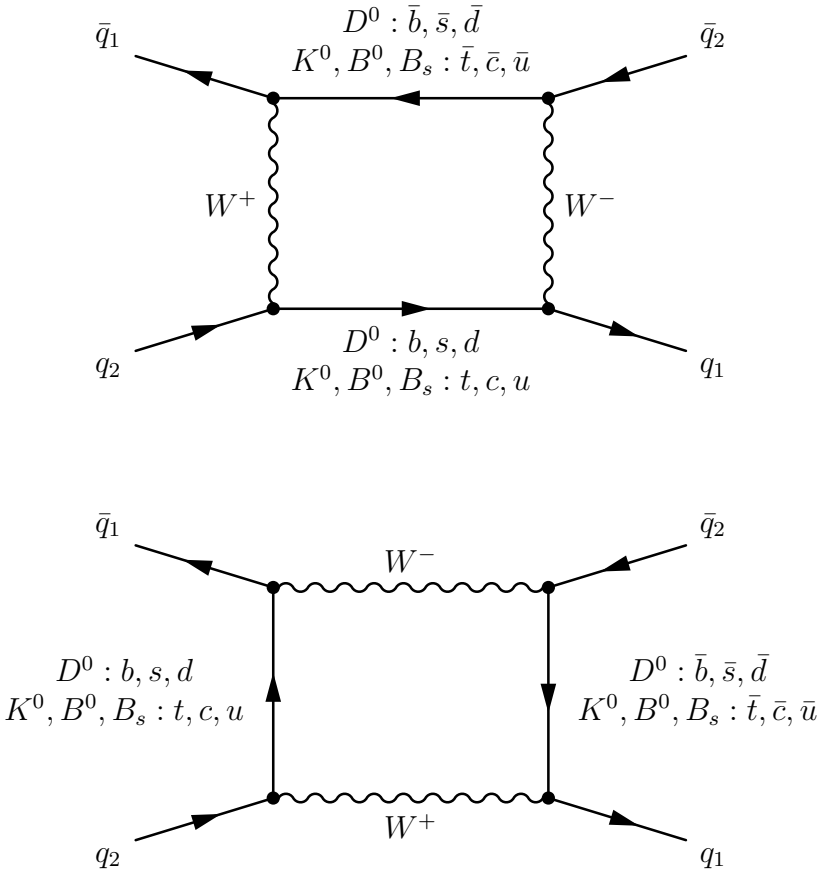


Figure 2.6: The two mixing diagrams for K^0-$\overline{K}^0$, D^0-$\overline{D}^0$, B^0-$\overline{B}^0$ and B_s-$\overline{B}_s$ neutral mesons systems. The upper Feynman graph depicts the W-exchange diagram, the lower one the box diagram. At each vertex, we accumulate a factor of the corresponding CKM matrix element.

The mass eigenstates of the full Hamiltonian are linear combinations of the flavor eigenstates. Let X^0 and $\overline{X}^0$ be the so-called flavor eigenstates of definite quark content that are most useful in understanding particle production and decay, and let X_H and X_L be the states of definite mass and lifetime which propagate through space in a definite fashion. We then write the general superposition as:

$$\begin{aligned} |X_H\rangle &= p_H|X^0\rangle - q_H|\overline{X}^0\rangle \\ |X_L\rangle &= p_L|X^0\rangle + q_L|\overline{X}^0\rangle. \end{aligned} \tag{2.33}$$

The two states X_H and X_L have different time evolution, so the flavor of the state evolves in time as well. The mass splitting Δm of X_H and X_L is dominated by virtual (off-shell) transitions while real intermediate states (*i.e.,* the imaginary parts of the box diagram, here via real c and u quarks) contribute to the lifetime difference $\Delta\Gamma$. [†] In general, virtual transitions are suppressed by the GIM mechanism [9], but these cancellations aren't perfect due to differences in the various quark masses. In the B^0 mixing system, for instance, the relevant products of CKM matrix elements are of the same order, and the mass hierarchy of the quarks comes into play, making the t quark dominant. In the K^0 system, however, the t quark transitions are suppressed, leaving the c quark dominant.

In preparation for later discussions, let us note that:

- *CPT* invariance would require that $p_H = p_L$ and $q_H = q_L$
- T invariance would require that $|q_L/p_L| = 1$ and $|q_H/p_L| = 1$
- *CP* invariance (in mixing) would require $p_L = p_H = p$, $q_L = q_H = q$, and that $|p/q| = 1$.

We'll discuss the structure of the time evolution more formally in Chapter 3. From the present discussion, however, we can already see that the time dependence of the system will depend on both Δm and $\Delta\Gamma$. We call Δm the mixing frequency when it dominates. A useful figure-of-merit in this scenario is the ratio of the mixing frequency to the average lifetime, which describes "how many times the meson can mix before it decays." For the B^0 system, $\frac{\Delta m}{\Gamma} \sim 0.7$, while for the B_s system, $\frac{\Delta m}{\Gamma} \sim 14.0$. In the K^0 system, the lifetime difference is large, but the equivalent figure of merit might be the mass splitting divided by the shorter K^0_S lifetime component, yielding $\frac{\Delta m}{\Gamma} \sim 0.6$. (These values have not yet been established for the D^0 system.)

For these neutral mesons, a difference in lifetimes between the elements of the pair can only arise from decay channels in common to both mesons. The best way to think about this is in the case of *CP* conservation ($|p| = |q|$): only when the final

[†]Recall from ordinary quantum mechanics that degenerate states will undergo "splitting" when an interaction Hamiltonia is introduced that couples them.

state f is accessible to X^0 and $\overline{X}^0$ (*i.e.,* a *CP* eigenstate perhaps) will the different admixture of X and $\overline{X}^0$ in X_H and X_L become apparent and the partial widths to f will be different. [‡] For the heavier mesons, the lifetime difference is expected to be smaller than the average lifetime since the branching ratios to the states accessible to both of the pair is small, and therefore the "opportunity" for final states in common is reduced. CKM factors come can come into play, however, to create additional suppression. For instance, when comparing the expected width difference $\Delta\Gamma$ in the B_s and B_d systems, note that such decays would involve $b \to c\bar{c}q$ where $q = d$ or s. These decays are Cabibbo-suppressed if $q = d$ and Cabibbo-allowed if $q = s$; we therefore expect $\Delta\Gamma/\Gamma$ to be as much as 20% in the B_s system, but less than 1% ($\lambda^2\Delta\Gamma_s/\Gamma_s$) in the B^0 system.

2.2 Goals of this Study

The goal of this study is to understand the life cycle of the B meson by measuring the parameters Δm_d and τ_{B^0} with high precision. In doing so, we will pioneer the technique for measuring these parameters simultaneously, and will establish a deep level of understanding of so-called "time dependent analyses" that will be critical for many other measurements at *BABAR* and elsewhere.[§]

We motivate our study of B^0-$\overline{B}^0$ lifetime by observing that the interesting part of a particle's decay rate comes from the contribution from the matrix element (see Eq. 2.31) since this piece encapsulates the complex physics. There is much interest in this term because the b quark is so much heavier than the d (and u and s) quark that the phenomenologist typically assumes that the light quark's flavor doesn't matter. The approximation that the dynamics of the meson are dominated by the heavy quark is called the *spectator* model. [20] Empirically however, we know this model isn't very accurate in the D system. By studying the B^0 lifetime, we seek to improve

[‡] If *CP* violation occurred, $\Delta\Gamma$ might even be exacerbated because, for example, both B^0 and $\overline{B}^0$ can decay to D^+ D^-, but if the particular rate for $B^0 \to D^+D^-$ were to be greater than $\overline{B}^0 \to D^+D^-$, (*CP* violating) the B^0 would decay more quickly due to the larger partial width.

[§] Other time-dependent analyses include those measuring the CKM angles $\sin2\beta$, $\sin2\alpha$, and $\sin2\gamma$ (see Fig. 2.5).

our understanding of the relationship between the light companion quark and the much heavier b quark.

Our study of the B^0 mixing frequency is motivated by two goals: (1) the loops in the box diagrams may be sensitive to the introduction of new physics, and (2) because Δm_d depends on V_{td}, a relatively poorly measured element of the CKM matrix, we can improve our estimate of V_{td}. By precisely comparing the measured value of Δm_d with that predicted by the Standard Model, we can, with a certain degree of confidence, exclude the presence of new physics in the loop processes.

We are also interested in B^0-$\overline{B}^0$ time evolution from a larger perspective. Because the system includes neutral meson mixing and decay, it is an excellent environment for testing the symmetries of the Standard Model.

2.3 Methods of this Study

Exploring the B system and its potential impact on the Standard Model requires copious production of B mesons, accurate measurement of the B flight path and flavor, and reasonably low backgrounds for reconstruction. This section introduces the key experimental principle used to achieve this. Note that we use the term *flavor* to label the flavor of the b quark within the B meson.

2.3.1 Collisions at the $\Upsilon(4S)$

In e^+e^- interactions, there is a resonant state $\Upsilon(4S)$ near $E_{cm} = 10.58\,\mathrm{GeV}$ that decays almost exclusively to b-$\overline{b}$ pairs. As the b quarks hadronize, lighter quarks are pulled from the vacuum, creating two B mesons that separate. Depending on the flavor of the spontaneously generated light quarks, the B mesons are either charged or neutral. By isospin symmetries, this implies 50% B^+B^- and 50% $B^0\overline{B}^0$ pairs. The $\overline{b}$ production cross-section near the $\Upsilon(4S)$ peak (see Fig. 2.7) is about 1 nb, which is about one-third as much as other continuum processes, and comparable to the $\mu^+\mu^-$, $\tau^+\tau^-$ pair production rates as well. The dominant physics, of course, is e^+e^- Bhabha scattering.

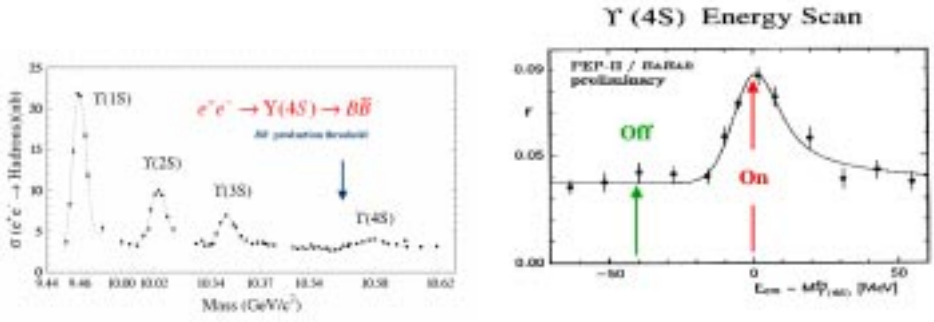


Figure 2.7: Left: Plot of the Υ resonances as a function of the e^+e^- center of mass energy. The $\Upsilon(4S)$ is right above the threshold for $\overline{b}$ quark production, and is a short-lived bound-state resonance. Right: Closeup of the $\Upsilon(4S)$ peak where the radiative tail is clear. The two arrows mark the E_{cm} energies where we collect collision data, "on" and "off" the resonant peak.

An ideal production mechanism for $B\overline{B}$ pairs is to create $\Upsilon(4S)$ particles and study their decay products which are essentially exclusively B (see Fig. 2.8). However, the energy release in the $\Upsilon(4S)$ decay is so small that the B decay products are nearly at rest in $\Upsilon(4S)$ frame. In short, the b quark mass is $m_b = 5.28\,\text{GeV}$, which is almost exactly half the mass of the $\Upsilon(4S)$ resonance leaving little extra for kinetic energy. This makes the observer's job difficult: the B mesons don't travel very far in the lab since they have such little momentum; in this case, it becomes extremely difficult to separate the B production point and its decay point. The technique proposed by Pier Odonne [21] is to produce the $\Upsilon(4S)$ so that it is traveling at moderate velocity in the lab frame. [¶] If the boost of the $\Upsilon(4S)$ is chosen wisely, modern particle detectors can then resolve the B production and decay vertices.

As described in detail later, we collide electron and positrons together at the PEP-II facility at SLAC [‖]. The novelty of the PEP-II design is the asymmetric energies of the circulating beams: the electrons are accelerated to 9.0 GeV and the positrons to 3.1 GeV. The center-of-mass is therefore boosted in the lab, travelling in the direction of the e^- beam with a boost of $\gamma\beta \sim 0.55$. The center-of-mass energy is slightly more

[¶] The boost in the lab frame also provides a small time dilation, but it is only a 10% effect.
[‖] Stanford Linear Accelerator Center, Menlo Park, California, USA

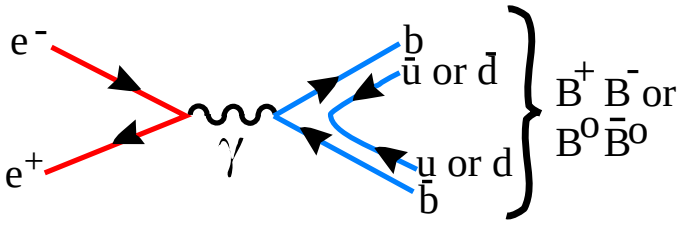


Figure 2.8: Feynman diagram for the $B\overline{B}$ production mechaniasm at the $\Upsilon(4S)$ resonance via a virtual photon.

than 10.5 GeV, near the production threshold of the $\Upsilon(4S)$ resonance. The $\Upsilon(4S)$ decay products (the $B\overline{B}$ pairs) "fly" in the lab, travelling hundreds of μm before their decay.

An additional benefit of working with the $\Upsilon(4S)$ decay is the correlated time evolution of the daughter B^0-$\overline{B}^0$ mesons. Akin to the EPR paradox with correlated photons, when the $\Upsilon(4S)$ decays and produces the B^0-$\overline{B}^0$ pair, we don't know which B is the $\overline{B}^0$ and which B the B^0; and because of their mixing, either can be either at any time. However, since the $\Upsilon(4S)$ is a vector, and the B mesons pseudoscalar bosons (*i.e.,* spin 0), we know that the B-$\overline{B}$ state must have $L = 1$ to conserve the angular momentum of the $\Upsilon(4S)$. This means that the B and $\overline{B}$ are in antisymmetric p-wave state. Because of mixing between the neutral B mesons, we might expect to occasionally have B^0-B^0 or $\overline{B}^0$-$\overline{B}^0$; this is forbidden however because that would leave us with identical bosons in an antisymmetric state, violating Bose statistics. We conclude therefore, that after the $\Upsilon(4S)$ decay, there can be at most one meson of each flavor.

2.3.2 Decays of the $\Upsilon(4S)$

We now have the ingredients for a time-dependent analysis. As shown in Fig. 2.9, the e^- and e^+ collide and form the resonant $\Upsilon(4S)$ state which immediately decays into a $B\overline{B}$ pair, either B^+B^- or $B^0\overline{B}^0$ (about half and half). If neutral, the pair time-evolves in a correlated fashion. At time t_1, one of the B's decays, which "starts the clock." At that exact instant, the surviving B must be of the opposite flavor,

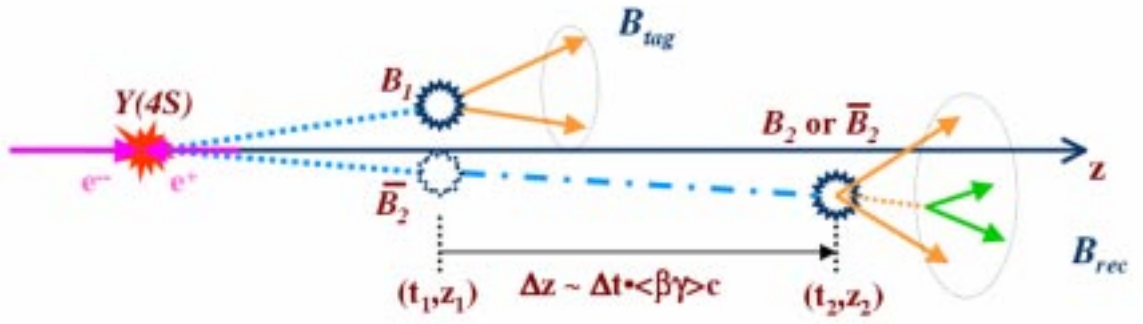


Figure 2.9: An artist's conception of the experiment. The $\Upsilon(4S)$ decays to two B mesons which later decay themselves. The decay products can be used to identify the vertex and flavor of each B at decay.

as explained above; it subsequently decays at some later time t_2. Because mixing is periodic and lifetime stochastic, the same physics governs B_2 between t_1 and t_2 as between its true origin (at the decay of the $\Upsilon(4S)$) and its final decay at t_2. We can use the decay products of each B to determine its decay vertex, and its flavor at decay. Because the B mesons are boosted in the lab frame along the z-axis, we can assume that most of their travel is along that axis (*i.e.,* the B mesons travel single file and along $\hat{z}$).

Our study of B^0 lifetime and mixing analyzes a large set of these $B\overline{B}$ pair events. We look for events in which we can exclusively reconstruct one B meson's decay in the semileptonic channel $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ channel. ** As in Fig. 2.9, the reconstructed B is denoted $B_{\rm rec}$, and we claim to understand it very well. The other B is called $B_{\rm tag}$, because it "tags" the flavor and location of our $B_{\rm rec}$ at the time t_1. To repeat: we know when and into what $B_{\rm rec}$ decayed because we have fully reconstructed it from its decay products; and we use the more inclusive information from $B_{\rm tag}$ to get "another data point" on the life-cycle (trajectory) of $B_{\rm rec}$. (Keep in mind that while the figure suggests that $B_{\rm rec}$ decays after $B_{\rm tag}$, this ordering is arbitrary – the argument still holds.)

**We imply charge conjugate modes wherever possible to reduce the notation, unless explicitly stated otherwise.

- We call a $B\overline{B}$ event *mixed* if $B_{\rm rec}$ has the opposite flavor at decay as it had at the tagging point. Consequently, $B_{\rm tag}$ will be measured as having the same flavor as $B_{\rm rec}$ at decay because it's the opposite of the opposite.

- Likewise, we call an event *unmixed* if $B_{\rm rec}$ is found to have the same flavor at decay and at the tagging point. (And so $B_{\rm tag}$ will be measured to have the opposite flavor.)

By measuring the separation in z between the two decay vertices, and using our knowledge of the boost in the lab frame, we can estimate Δt, the time between the two decays by using the relation:

$$\Delta t = \frac{\Delta z}{\langle \beta_z \gamma \rangle c} \tag{2.34}$$

This study accumulates measurements of mixing status (via flavor identification) and Δt for events with reconstructed $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ decays, analyzes them to characterize and remove backgrounds (where possible), and then uses an unbinned maximum-likelihood fit to extract the time evolution parameters of the B^0 meson.

The $B_{\rm rec}$ is reconstructed in the decay chain, $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$, $D^{*-} \to \overline{D}^0\pi^-$, and $D^0 \to K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^-\pi^+$, or $K^0_S\pi^+\pi^-$ (summing to about 25% of the total D^0 branching fraction). 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 $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ 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.

We reconstruct the D^* in the mode $D^{*-} \to \overline{D}^0\pi^-$, which has a branching fraction of 68%; because of the small mass difference, the momentum transfer is only 39 MeV. The emitted pion is therefore slow-moving in the D^* frame, and is often called the *soft* or *slow pion*. Due to the small energy release, accurate reconstruction of the soft pion is an experimental challenge.

2.3.3 Time Dependent Analysis

We are now ready to examine B^0-$\overline{B}^0$ time evolution as it might appear in *BABAR*. The strategy is to describe the $\Delta t_{\rm meas}$ distributions with a probability density function, build a full likelihood for the entire dataset, and then use an unbinned maximum likelihood fit to determine the parameters which best fit the data. Before describing the technique in subsequent chapters, we consider the realistic smearings and uncertainties that complicate that Δt distributions. In this study, we will model and parameterize these effects with coefficients extracted from the Data simultaneously with the physics parameters τ_{B^0} and Δm_d.

Based on the *mixed* and *unmixed* Δt distributions, we define the *mixing asymmetry* to be

$$A(\Delta t) = \frac{N_{\rm unmixed}(\Delta t) - N_{\rm mixed}(\Delta t)}{N_{\rm unmixed}(\Delta t) + N_{\rm mixed}(\Delta t)} \, , \tag{2.35}$$

The $\Delta t_{\rm true}$ distributions for events are shown in Fig 2.10. Note the complete absence of mixed events at $\Delta t_{\rm true} = 0$ because no time has elapsed for the B^0 to mix to $\overline{B}^0$. Also note that mixed distribution dominates in the intermediate range of $\Delta t_{\rm true}$. Finally, the mixing asymmetry has unit amplitude and a period of $2\pi/\Delta m_d$.

The flavor identification algorithms at *BABAR* are not perfect, however, and we typically incorporate a *mistag rate* which parameterizes the rate at which we assign the $B^0\overline{B}^0$ event the wrong mixing status. (In the final analysis, we even allow the mistag rates for B^0 and $\overline{B}^0$ mesons to be different.) In this situation, we start trading events between the two curves in $\Delta t_{\rm meas}$ space, and we arrive at the distributions described in Fig. 2.11. Note that the mixing asymmetry amplitude has been *diluted* to $1 - 2\omega$ where ω is the mistag probability. Ideally, one could extract the mistag rate by measuring the amplitude of the mixing asymmetry at $\Delta t_{\rm meas} = 0$.

The final complication in imaging $\Delta t_{\rm true}$ in the experiment is due to finite resolution effects of the detector. $\Delta t_{\rm meas}$ will, in general, be a biased and smeared estimate of $\Delta t_{\rm true}$, which can have dramatic effects on the measured Δt distributions as depicted in Fig. 2.11. To accommodate this, we include a resolution function in our model which maps the physics in $\Delta t_{\rm true}$ space into the real world of $\Delta t_{\rm meas}$.

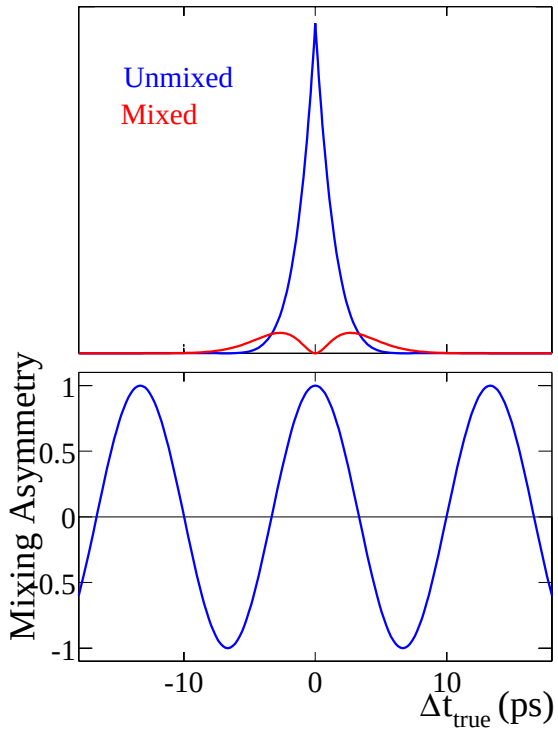


Figure 2.10: Upper: The distributions in $\Delta t_{\rm true}$ for mixed and unmixed $B^0\overline{B}^0$ events. Lower: The time dependent mixing asymmetry as defined in Eq. 2.35.

As we have suggested, the key principle in this time-dependent analysis is careful characterization of the Data while maximizing the statistical power of the event sample itself. That is, we try to minimize assumptions and external inputs in the analysis by adding components to the model (*i.e.,* mistag probabilities, background fractions, and resolution parameters) which will be determined simultaneously with the physics results in the final fit.

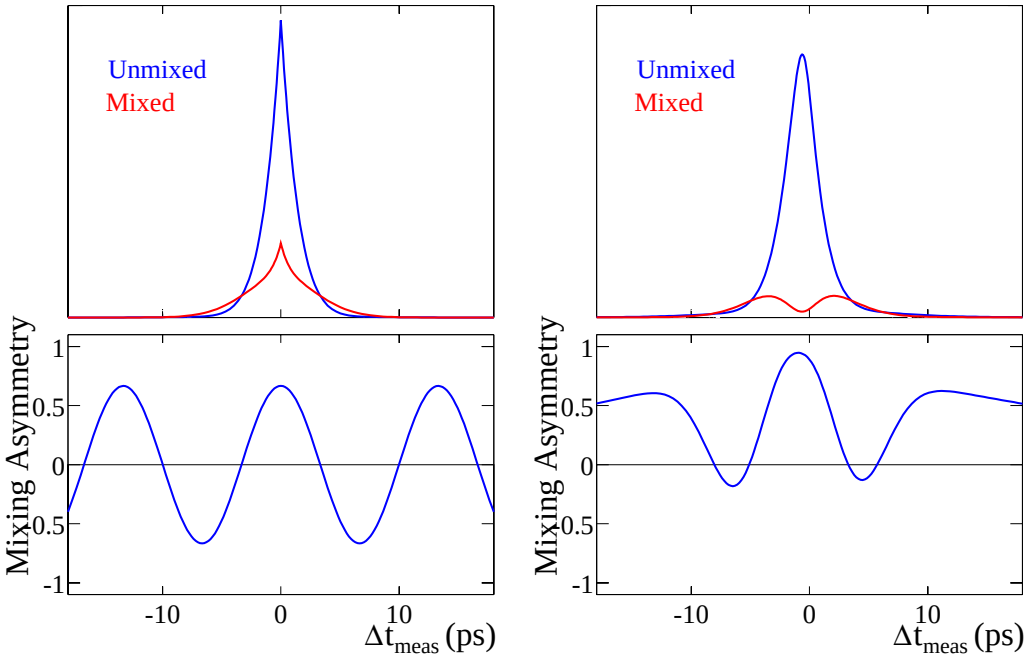


Figure 2.11: Upper Left: The distributions in $\Delta t_{\rm meas}$ for mixed and unmixed $B^0\overline{B}^0$ events with a mistag probability of 20%. Lower Left: The time dependent mixing asymmetry. Upper Right: The distributions in $\Delta t_{\rm meas}$ for mixed and unmixed $B^0\overline{B}^0$ events with an exaggerated resolution. Lower Right: The time dependent mixing asymmetry.

Chapter 3

Theory and Phenomenology

The framework for describing B meson time evolution is analogous to that of a 2-level system in quantum mechanics. Because the weak eigenstates are not identical to the physical eigenstates, a generic state has time-varying components of both B^0 and $\overline{B}^0$. The physical eigenstates have a mass splitting Δm_d and a lifetime difference $\Delta\Gamma$, where the latter is expected to be small in the Standard Model. Theoretical predictions for the mixing frequency Δm_d and the average lifetime τ_{B^0} still inherit significant uncertainties compared to the experimental precision.

The semileptonic decay mode $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ is a three-body decay whose neutrino we do not reconstruct. We approximate the B center of mass frame as that of the $\Upsilon(4S)$, and construct angular variables which measure the consistency with the $D^{*-}\ell^+\nu_\ell$ decay.

3.1 Time Evolution Formalism

Neutral B mesons are produced in states of definite flavor, *i.e.,* flavor eigenstates. As suggested in Section 2.1.5, these states are not eigenstates of the weak interaction; we therefore expect the flavor to change under action of the electroweak Hamiltonian. We will first consider the form of an effective Hamiltonian to describe this process, find the new eigenstates and eigenvalues, and then consider their general time evolution.

3.1.1 Mixing of Neutral B Mesons

We borrow the formalism from that of a quantum mechanical two-level system where transitions are introduced by an interaction Hamiltonian, $\mathcal{H}$. The general state $|\Psi\rangle$ can be written as a linear combination of the flavor eigenstates, as they form a set of basis vectors in this subspace.

$$\begin{aligned} |B^0\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad\qquad |\overline{B}^0\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ |\Psi\rangle &= a|B^0\rangle + b|\overline{B}^0\rangle \\ &= \begin{pmatrix} a \\ b \end{pmatrix} \end{aligned} \tag{3.1}$$

We can decompose $\mathcal{H}$ in the flavor basis as:

$$\begin{aligned} \mathcal{H} &= \mathit{M} - \frac{i}{2}\mathit{\Gamma} \\ &= \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \end{aligned} \tag{3.2}$$

where M is the "dispersive" term representing virtual intermediate states and $\mathit{\Gamma}$ is the "absorptive" term representing real intermediate states. Conservation of probability (*i.e.,* unitarity) requires $\mathcal{H}$ to be unitary, which forces M and $\mathit{\Gamma}$ to be Hermitian matrices (the diagonal elements must be real, and the off-diagonal elements complex conjugates of each other, *e.g.,* $M_{12} = M_{21}^*$).

For the moment, we will also impose *CPT* symmetry, which requires that $M_{11} \equiv M_{22} = M$ and $\Gamma_{11} = \Gamma_{22} \equiv \Gamma$. Using the Wigner-Weisskopf approximation that the weak interaction Hamiltonian does not couple the final states to each other, [22] we write the time-independent Schrödinger equation, and wish to find the (energy) eigenvalues and eigenstates:

$$\begin{gathered} \mathcal{H}|\Psi\rangle = \lambda|\Psi\rangle \\ \begin{pmatrix} M - \frac{i}{2}\Gamma & M_{12} - \frac{i}{2}\Gamma_{12} \\ M^*_{12} - \frac{i}{2}\Gamma^*_{12} & M - \frac{i}{2}\Gamma \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \lambda \begin{pmatrix} a \\ b \end{pmatrix} \end{gathered} \tag{3.3}$$

where we assume $|\Psi\rangle$ is an energy eigenstate. Solving the equation $det[\mathbf{M} - \frac{i}{2}\mathbf{\Gamma} - \lambda] = 0$ yields a quadratic equation with two roots, the energy eigenvalues λ_+ and λ_-:

$$\lambda_\pm = M - \frac{i}{2}\Gamma \pm \sqrt{(M_{12} - \frac{i}{2}\Gamma_{12})\ (M^*_{12} - \frac{i}{2}\Gamma^*_{12})}\ . \tag{3.4}$$

Substituting back in, we arrive at the rather cryptic forms for the new eigenstates:

$$|B_{1,2}\rangle = \left(\sqrt{(M_{12} - \frac{i}{2}\Gamma_{12})\ (M^*_{12} - \frac{i}{2}\Gamma^*_{12})}\ |B^0\rangle \ \pm\ (M^*_{12} - \frac{i}{2}\Gamma_{12})\ |\overline{B}^0\rangle \right)\ , \tag{3.5}$$

where we have not included the normalization factors. By comparing the real and imaginary parts of the two eigenvalues, we realize that the two energy eigenstates have different masses ($m_{1,2} = \mathcal{R}e\ \lambda_\pm$) and different lifetimes ($\Gamma_{1,2} = -2\ \mathcal{I}m\ \lambda_\pm$); we associate the differences with the so-called *mass* and *(decay) width splittings*: *

$$\begin{aligned} \Delta m_d = m_1 - m_2 &= \mathcal{R}e\ (\lambda_+ - \lambda_-) \\ \Delta\Gamma = \Gamma_2 - \Gamma_1 &= 2\ \mathcal{I}m\ (\lambda_+ - \lambda_-)\ . \end{aligned} \tag{3.6}$$

*Note that we are using mixed notation here: τ_{B^0} for the B_d average lifetime, Δm_d for the mass splitting, and $\Delta\Gamma$ for the lifetime different. Apologies...the correctly labelled set of parameters should probably be Δm_d, $\Delta\Gamma_d$, and τ_{B_d}.

and deduce the following relationships

$$\begin{aligned}(\Delta m_d)^2 - \frac{1}{4}(\Delta\Gamma)^2 &= 4\ |M_{12}|^2 - |\Gamma_{12}|^2 \\ \Delta m_d\ \Delta\Gamma &= -4\ \mathcal{R}e\ (M_{12}^*\ \Gamma_{12})\ .\end{aligned} \tag{3.7}$$

Let us now update our notation so that we can easily discuss the new states, but still requiring CPT symmetry:

$$\begin{aligned}p &= \sqrt{(M_{12} - \frac{i}{2}\Gamma_{12})(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)} \\ q &= M_{12}^* - \frac{i}{2}\Gamma_{12}^* \\ \frac{q}{p} &= \frac{\sqrt{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}}{\sqrt{M_{12} - \frac{i}{2}\Gamma_{12}}} \\ \lambda_\pm &= (M - \frac{i}{2}\Gamma) \pm \frac{q}{p}\ (M_{12} - \frac{i}{2}\Gamma_{12})\ .\end{aligned} \tag{3.8}$$

This allows us to write the energy eigenstates as

$$\begin{aligned}|B_L\rangle &= p|B^0\rangle + q|\overline{B}^0\rangle \\ |B_H\rangle &= p|B^0\rangle - q|\overline{B}^0\rangle \\ \Delta m_d &= m_H - m_L = -2\mathcal{R}e\ (\frac{q}{p}(M_{12} - \frac{i}{2}\Gamma_{12})) \\ \Delta\Gamma &= \Gamma_L - \Gamma_H = -4\mathcal{I}m\ (\frac{q}{p}(M_{12} - \frac{i}{2}\Gamma_{12}))\ .\end{aligned} \tag{3.9}$$

(We use the convention that $CP|B^0\rangle = +|\overline{B}^0\rangle$ and that the real part of q/p is positive.) As an interesting sidenote, we introduce δ as the complex CP-violating parameter (in mixing):

$$\delta \equiv \langle B_L|B_H\rangle = |p|^2 - |q|^2 = \frac{-2\ \mathcal{I}m\ (M_{12}^*\Gamma_{12})}{(\Delta m_d)^2 + |\Gamma_{12}|^2}\ , \tag{3.10}$$

which measures the amount to which the energy eigenstates are equally composed of B^0 and $\overline{B}^0$. (We also note that in the presence of CP violation, then, the energy eigenstates are not orthogonal.) We can explicitly observe the effect of CP-violation

in mixing as modifications to Δm_d and $\Delta\Gamma$:

$$\begin{aligned}(\Delta m_d)^2 &= \frac{4\ |M_{12}|^2 - \delta^2\ |\Gamma_{12}|^2}{1+\delta^2} \\ (\Delta\Gamma)^2 &= \frac{4\ |\Gamma_{12}|^2 - 16\ \delta^2\ |M_{12}|^2}{1+\delta^2}\ .\end{aligned} \tag{3.11}$$

In the limit of exact *CP* invariance ($\delta = 0$) the mass eigenstates coincide with the *CP* eigenstates, $CP|B_H\rangle = -|B_H\rangle$ and $CP|B_L\rangle = +|B_L\rangle$ and the mass difference and width difference are given by $\Delta m_d = 2|M_{12}|, \Delta\Gamma = 2|\Gamma_{12}|$. Thus, as stated earlier, the lifetime and mass differences of the physical eigenstates in B^0-$\overline{B}^0$ system can (and will) manifest themselves without *CP* violation.

3.1.2 Time Evolution of Neutral B Mesons

We now invoke the time-*dependent* Schrödinger equation to study the time evolution of the again-generic state $|\psi(t)\rangle$:

$$i\frac{\partial}{\partial t}\begin{pmatrix} a(t) \\ b(t)\end{pmatrix} = (\mathit{M} - \frac{i}{2}\mathit{\Gamma})\begin{pmatrix} a(t) \\ b(t)\end{pmatrix}\ . \tag{3.12}$$

The time-dependent generic state can just be written in terms of an admixture of the energy eigenstates with time-evolving amplitudes for each component (with appropriate choices for $a_{H,L}(0)$ to match the initial conditions):

$$\begin{aligned}|\psi(t)\rangle &= a_H(t)\ |B_H\rangle + a_L(t)\ |B_L\rangle \\ &= a_H(0)\ e^{-\lambda_+\ t}\ |B_H\rangle + a_L(0)\ e^{-\lambda_-\ t}\ |B_L\rangle\ .\end{aligned} \tag{3.13}$$

We can then substitute in the expressions for $|B_H\rangle$ and $|B_L\rangle$ in terms of the flavor eigenstates. We can construct a state $|B^0_{phys}(t)\rangle$ which is created at time $t = 0$ as purely B^0, and likewise for $\overline{B}^0$ as $|\overline{B}^0_{phys}(t)\rangle$. The general time evolution is summarized

as:

$$\begin{aligned} |B^0_{phys}(t)\rangle &= f_+(t)\ |B^0\rangle + \frac{q}{p}\ f_-(t)\ |\overline{B}^0\rangle \\ |\overline{B}^0_{phys}(t)\rangle &= f_+(t)\ |\overline{B}^0\rangle + \frac{p}{q}\ f_-(t\ |B^0\rangle\ , \end{aligned} \tag{3.14}$$

where the notation $f_\pm$ is simply

$$f_\pm = \frac{e^{-i\lambda_+\ t} \pm e^{-i\lambda_-\ t}}{2} = \frac{1}{2} e^{(-im_1+\Gamma_1/2)\ t}\left[1 \pm e^{(-i\Delta m_d+\Delta\Gamma/2)\ t}\right]. \tag{3.15}$$

We can find the transition probability, or oscillation rate, for the physical states to oscillate from one flavor to the other by computing the projection of the $|B^0_{phys}\rangle$ onto the flavor eigenstates $|\langle B^0|B^0_{phys}\rangle|^2$ and $|\langle\overline{B}^0|B^0_{phys}\rangle|^2$. Using the fact that the flavor eigenstates are orthogonal, and after some algebra of complex numbers, [†] we arrive at

$$\begin{aligned} |\langle B^0|B^0_{phys}\rangle|^2 &= |\langle B^0|f_+|B^0\rangle + \langle B^0|\frac{q}{p}f_-|\overline{B}^0\rangle|^2 \\ &= |f_+|^2 = \frac{e^{-\Gamma t}}{2}\left[\cosh(\frac{\Delta\Gamma}{2}t) + \cos(\Delta m_d\ t)\right] \\ |\langle\overline{B}^0|B^0_{phys}\rangle|^2 &= |f_-|^2 = \frac{e^{-\Gamma t}}{2}\left[\cosh(\frac{\Delta\Gamma}{2}t) - \cos(\Delta m_d\ t)\right]\ , \end{aligned} \tag{3.16}$$

At time $t = 0$ there is zero probability for finding the *mixed* state, where the $|B^0_{phys}\rangle$ has oscillated into a $|\overline{B}^0\rangle$, but afterwards, the probability exhibits a cosine time dependence with a frequency governed by Δm_d. [‡] This time-dependent mixing probability directly expresses the fact that since the flavor eigenstates are not the eigenstates of the interaction Hamiltonian, a generic state's projection on the flavor-space basis vectors will in general evolve in time. This phenomenon is called *mixing*.

In the B^0 system, we note that $\Delta\Gamma \ll \Gamma$ because of the small number of final states accessible to both B^0 and $\overline{B}^0$ (see Section 2.1.5). Additionally, experimental measurements report that $x_d = \frac{\Delta m_d}{\Gamma} = 0.73 \pm 0.05$. [23] Together these suggest that

[†] Pardon this sentence construction, but we wouldn't want to suggest that algebra itself is *complex*!

[‡] Note that these probabilities are bounded in time because $\Delta\Gamma$ is, by definition, bounded by $2 \times \Gamma$.

$\Delta m_d \ll \Delta\Gamma$ so we can reasonably collapse the cosh term to unity. In the study described here, we will assume the $\Delta\Gamma = 0$, but please see Section 15.3.2 for some comments about relaxing this assumption.

3.1.3 Time Evolution From $\Upsilon(4S)$ Decays

One important lesson from quantum mechanics is that multi-particle systems must be treated collectively, in order to properly account for their potentially interfering amplitudes. The observer can really only make observations about the state of the system. At the $\Upsilon(4S)$, B mesons are produced in pairs, and they evolve coherently in an Einstein-Podolsky-Rosen correlated fashion; in particular, when they are neutral, they oscillate in phase. As such, then, we comment here on the so-called *coherent formalism* used to describe the time evolution of the $B\overline{B}$ system and its eventual decay to a set of final states $f_1,\ f_2$.

Following the description in Ref. [24], we write the general state containing two B mesons in the mass eigenstate basis as

$$|B\overline{B}\rangle = c_1|B_L B_H\rangle + c_2|B_H B_L\rangle \tag{3.17}$$

where $c_1,\ c_2$ are in general complex numbers satisfying the normalization condition $|c_1|^2+|c_2|^2 = 1$. This state time-evolves just as the combination of the isolated states,

$$|B\overline{B}\rangle = |(B\overline{B})_{phys}(t_1,t_2)\rangle = c_1\, e^{-i\lambda_+ t_1} e^{-i\lambda_- t_2}|B_L B_H\rangle + c_2\, e^{-i\lambda_- t_1} e^{-i\lambda_+ t_2}|B_H B_L\rangle\ , \tag{3.18}$$

where $\lambda_\pm$ are the energy eigenvalues as before. We can use two arguments, as listed below, to derive constraints on the coefficients $c_1,\ c_2$.

Argument 1

The $\Upsilon(4S)$ is a vector meson, with $J = 1$, and $C \times P$ eigenvalue of $-1 \times -1 = +1$. Because the $\Upsilon(4S)$ decays via the strong interaction to the $B\overline{B}$ meson system, we expect CP to be conserved. Therefore, the daughter $B\ B$ system's wavefunction must also have $CP = +1$. Consider the wavefunction as $\Psi_{flavor}\Psi_{space}$. The $B\overline{B}$ system

must have one unit of orbital angular momentum, $L = 1$, since the B mesons are pseudoscalar bosons (*i.e.,* spin-0) and we need to conserve the angular momentum $J = L + S$ of the $\Upsilon(4S)$. The space part of the wavefunction Ψ_{space} is therefore negative under CP. To obtain an overall CP eigenvalue of $+1$, we are forced to choose the antisymmetric combination of the B and $\overline{B}$ to form the flavor part: $\Psi_{flavor} = |B\overline{B}\rangle - |\overline{B}B\rangle$ which also has $CP = -1$.

Argument 2

We can reach a similar conclusion by invoking Bose symmetry on the general $B\overline{B}$ mass eigenstate formalism. If both B mesons, in the $L = 1$ antisymmetric p-wave, were to decay simultaneously to identical bosonic final states f, then we would have two identical bosons in an antisymmetric state. This violates the requirements of Bose statistics. This has an immediate consequence for the coefficients $c_1,\ c_2$:

$$\begin{aligned} |B\overline{B}\rangle &= c_1|B_L B_H\rangle + c_2|B_H B_L\rangle \\ \langle ff|B\overline{B}\rangle &= 0 \\ c_1\langle ff|B_L B_H\rangle + c_2\langle ff|B_H B_L\rangle &= 0 \\ c_1\langle ff|B_L B_H\rangle &= -c_2\langle ff|B_H B_L\rangle \\ \text{but}\quad \langle ff|B_L B_H\rangle &= \langle ff|B_H B_L\rangle \\ \therefore c_1 &= -c_2 \end{aligned} \tag{3.19}$$

Recall the normalization condition $|c_1|^2 + |c_2|^2 = 1$, so we see that up to an overall phase, we can set $c_1 = 1, c_2 = -1$. This conclusion that the two-meson state is antisymmetric in the mass eigenstate basis is equivalent to a statement about the system in the flavor eigenstate basis.

We therefore conclude that the $|B\overline{B}\rangle$ meson state is antisymmetric in flavor, which means that the projection on the set of states with two B^0 or two $\overline{B}^0$ mesons is vanishing; this is why we write the state as $|B\overline{B}\rangle$.

In the experimental program as outlined in Section 2.3.3, we noted that the time-evolution is recorded in terms of the variable Δt, the decay time difference, $\Delta t \equiv$

$t_2 - t_1$. We can change variables in Eq. 3.18 to achieve:

$$t \equiv \frac{t_1 + t_2}{2} \quad , \quad \Delta t \equiv t_2 - t_1 \tag{3.20}$$

$$|B\overline{B}(t, \Delta t)\rangle = e^{-i2\lambda t} \left\{ e^{+i\Delta\lambda\Delta t}|B_L B_H\rangle - e^{-i\Delta\lambda\Delta t}|B_H B_L\rangle \right\} \tag{3.21}$$

where we have introduced more notation,

$$\lambda \equiv \frac{\lambda_+ + \lambda_-}{2} = M - \frac{i}{2}\Gamma \quad , \quad \Delta\lambda \equiv \frac{\lambda_+ - \lambda_-}{2} = \frac{1}{2}\left(\Delta m_d - \frac{i}{2}\Delta\Gamma\right) . \tag{3.22}$$

We'll examine the case where one B decays to f_1 and the other to f_2 to make one observation which will complete our discussion. Taking the first B decay to f_1, we can write the *partially projected* state (borrowing again from Ref. [24]) as

$$\langle f_1|B\overline{B}(t, \Delta t)\rangle = e^{-i2\lambda t} \left\{ e^{+i\Delta\lambda\Delta t}\langle f_1|B_L\rangle \cdot |B_H\rangle - e^{-i\Delta\lambda\Delta t}\langle f_1|B_H\rangle \cdot |B_L\rangle \right\} . \tag{3.23}$$

We next write the matrix elements in terms of the flavor eigenstates (the basis in which the physics of the actual decay is most easily described), introducing A_i as the quantum mechanical transition amplitude to a state i:

$$\begin{aligned} \langle f_1|B_L\rangle &= p\langle f_1|B^0\rangle + q\langle f_1|\overline{B}^0\rangle = pA_1 + q\overline{A}_1 \\ \langle f_2|B_H\rangle &= p\langle f_2|B^0\rangle - q\langle f_2|\overline{B}^0\rangle = pA_2 - q\overline{A}_2 . \end{aligned} \tag{3.24}$$

Combining this notation with that of the $|B^0_{phys}\rangle$, we can write the state after decay to f_1 at t_1:

$$\langle f_1|B\overline{B}(t_1, \Delta t)\rangle = e^{-2i\lambda t_1} \left[|B^0_{phys}(\Delta t)\rangle \cdot \overline{A}_1 - |\overline{B}^0_{phys}(\Delta t)\rangle \cdot A_1 \right] . \tag{3.25}$$

The exponential factor out front is representative of the lifetime associated with the decay probability. If we choose the state f_1 to be accessible only to B^0, (*e.g.,* $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$), then we have $A_1 = 1, \overline{A}_1 = 0$, which identifies the decay as that of a B^0. (The state actually collapses to the time evolution for an isolated $\overline{B}^0$ meson after t_1.) If the other B decays to a state f_2 at time t_2, we can write the full decay

amplitude as

$$\langle f_1 f_2 | B\overline{B}(t_1, \Delta t)\rangle = e^{-2i\lambda t_1}\left[\langle f_2|B^0_{phys}(\Delta t)\rangle \cdot \overline{A}_1 - \langle f_2|\overline{B}^0_{phys}(\Delta t)\rangle \cdot A_1\right] , \quad (3.26)$$

which we can use to compute the corresponding rate by taking the modulus-squared. Because of the B^0-$\overline{B}^0$ correlation, an interference term is introduced,

$$\begin{aligned}\left|\langle f_1 f_2 | B\overline{B}(t_1, \Delta t)\rangle\right|^2 \propto \\ e^{-2\Gamma t_1} \cdot \Big\{ \left|\langle f_2|B^0_{phys}(\Delta t)\rangle\right|^2 \cdot \left|\overline{A}_1\right|^2 + \left|\langle f_2|\overline{B}^0_{phys}(\Delta t)\rangle\right|^2 \cdot |A_1|^2 \\ - 2\mathcal{R}e\ \left(\langle f_2|B^0_{phys}(\Delta t)\rangle\langle f_2|\overline{B}^0_{phys}(\Delta t)\rangle^* \cdot \overline{A}_1 A_1^*\right)\Big\} . \quad (3.27)\end{aligned}$$

The interference term is only non-zero for states f_2 to which both B^0 and $\overline{B}^0$ can couple, and is a direct consequence of the correlated time-evolution.

Finally, let us make a comment on the general structure of the decay rate when we substitute back in the Δm_d-$\Delta\Gamma$ form of the energy eigenvalues. The algebra is beyond the scope of this document, so we quote from Ref. [24]:

$$\begin{aligned}\left|\langle f_1 f_2 | B\overline{B}(t, \Delta t)\rangle\right|^2 \propto \\ e^{-2\Gamma t}\left[c_1 \cosh(\frac{\Delta\Gamma\Delta t}{2}) + c_2 \cos(\Delta m_d \Delta t) + c_3 \sinh(\frac{\Delta\Gamma\Delta t}{2}) + c_4 \sin(\Delta m_d \Delta t)\right] ,\end{aligned} \quad (3.28)$$

where the coefficients c_i obey a normalization condition and depend on the decay amplitudes A_1, A_2, their conjugates and the symmetries required of the theory. [§] If we assume $\Delta\Gamma \equiv 0$, the sinh term vanishes and the cosh term collapses to unity.

3.2 Comments on τ_{B^0}

One goal of this analysis is the measurement of τ_{B^0}, the average lifetime of the neutral B meson, often referenced as $1/\Gamma$. We comment here on some of the theoretical

[§]Note that (a) we can integrate out the t dependence at fixed Δt to achieve the overall decay time difference distribution and that (b) I have hidden some of the relationships between the c_i.

framework for Standard Model predictions of the lifetime.

Calculation of the average B lifetime can be estimated to first order through use of the *spectator model*, in which we assume the the lighter quark in the meson(here, the $\overline{d}$) is not involved in the decay dynamics. In fact, a spectator model would predict that the heavy b quark dictates the entire decay process, and in fact, would predict the B^+ and B^0 lifetimes to be identical, as governed by the primitive equation below. The B meson decay width is dominated by $b \to c$ (λ^2) transitions, bringing in V_{cb}, and $b \to u$ contributions are suppressed by V_{ub} (λ^3) which we can neglect: [25]

$$\Gamma = \frac{1}{\tau} \sim \frac{G_F^2 \, m_b^5}{192\pi^3} \cdot |V_{cb}|^2 \times \rho(E, \vec{k}) \; , \tag{3.29}$$

where $\rho(E, \vec{k})$ is the appropriate set of phase-space factors. Experiments have measured that $\tau_{B^+}/\tau_{B^0} = 1.062 \pm 0.029$ [26], which indicates a significant deviation from the naive expectations of the spectator model. This is understood to be due to complications such as W-exchange or weak annihilation and Pauli interference diagrams.

Applying the optical theorem, the inclusive decay width of a hadron H_b containing a b quark can be written as the forward matrix element of the imaginary part of the transition operator and a few Wilson coefficients which take into account QCD corrections. Since the energy release in the decay of b quark is relatively large, its possible to construct an Operator Product Expansion (OPE) for the decay rate. The forward matrix elements in the OPE are expanded in inverse powers of the b-quark mass, following the prescriptions of heavy-quark effective theory (HQET). [27] These calculations predict differences from unity in the lifetime ratio of up to 10%, consistent with the observed value. ¶ Measurement of the B lifetime can be an important test for models of spectator quark effects, and can be used to extract the semileptonic decay constant, f_{B_d}, discussed in the next section.

The lifetime τ_{B^0} is a critical input to the extraction of other parameters of the Standard Model, such as measuring V_{cb} using semi-leptonic decays. τ_{B^0} is also needed

¶The b quark is the heaviest quark with which we can test these heavy flavor models, as the top quark is so heavy it decays weakly before it even hadronizes.

in the study of *CP*-asymmetries since it is used to characterize the signal in time-dependent analyses.

3.3 Comments on Δm_d

Study of the mixing frequency Δm_d is motivated by two reasons:

- Measurements of Δm_d can be used to extract the CKM matrix element V_{td} and thus determine the length of one of the sides of the unitarity triangle.
- In principle, since neutral meson mixing involves loop contributions, Δm_d is sensitive to new physics because interactions at higher scales can produce changes in the local operators at lower scales.

The mixing frequency Δm_d corresponds to the off-diagonal element M_{12} of the Hamiltonian $\mathcal{H}$ which we discussed above, and is related to the dispersive part of the transition amplitude between B^0 and $\overline{B}^0$. [‖] The important corrections from QCD are implemented with the help of the operator product expansion. [5] The top-quark contribution dominates and we can describe M_{12} by the so-called "local $\Delta B = 2$" Hamiltonian below the electroweak symmetry breaking scale as [28]

$$\Delta m_d = \frac{G_F^2}{8\pi^2} M_W^2 \, |V_{td} V_{tb}^*|^2 \, \eta_B S_0(x_t) C_B(\mu) \, \frac{|\langle \overline{B}_d | \mathcal{O}_d^{\Delta B=2}(\mu) | B_d \rangle|}{2 M_{B_d}} \, , \tag{3.30}$$

where $x_t = m_t^2/M_W^2$, $S_0(x_t) \simeq 0.784\, x_t^{0.76}$ (to better than 1%) is the relevant Inami-Lim function [29], μ the renormalization scale, $\mathcal{O}_d^{\Delta B=2}$ the four-quark operator, [**] and $\eta_B = 0.55$ and $C_B(\mu)$ are short-distance coefficients. The renormalization-scale dependence of $C_B(\mu)$ and the hadronic matrix element cancel at the level of perturbative analysis. Typically, we are after $|V_{td} V_{tb}^*|^2$, and calculate $S_0(x_t)$, η_B and b_B perturbatively. Because V_{tb} is known to be close to unity to good accuracy, Eq. 3.30 enables extraction of V_{td}. The precision is limited at present by the precision of lattice

[‖] Ironically, although M_{12} is the real part of the amplitude from the box diagram, it describes virtual transitions.

[**] $[\overline{b}\gamma^\mu(1-\gamma^5)d]$ $[\overline{b}\gamma_\mu(1-\gamma^5)d]$

QCD efforts to compute the hadronic matrix element. That matrix element is often parameterized as

$$\langle \overline{B}_d | \mathcal{O}_q^{\Delta B=2}(\mu) | B_d \rangle = \frac{8}{3} M_{B_d}^2 f_{B_d}^2 B_{B_d}(\mu) \ , \tag{3.31}$$

where f_{B_d} is the meson decay constant and B_{B_d} the so-called bag parameter. This matrix element is the subject of much debate, as different theoretical estimates using lattice QCD techniques give different results, although all agree that $f_{B_d}\sqrt{B_{B_d}} \sim 200\,\mathrm{MeV}$. It becomes easier to consider the ratio of mixing frequencies for B_d and B_s in which many of the uncertainties cancel [††]:

$$\frac{\delta m_s}{\Delta m_d} = \frac{m_{B_s}}{m_{B_d}} \zeta^2 \left| \frac{V_{ts}}{V_{td}} \right| \ , \tag{3.32}$$

where $\zeta = (f_{B_d}\sqrt{B_{B_d}})$ $((f_{B_s}\sqrt{B_{B_s}}) \sim 1$ as estimated from lattice QCD at the level of 5-6%, significantly reducing the hadronic uncertainties. [26] It is in this direction that current efforts are focused. However, Δm_s hasn't been measured experimentally yet, so the utility of this technique is still limited.

The phenomenological and theoretical interest in Δm_d is high, and when combined with expected results in the B_s system, offers a significant battery of tests for the Standard Model.

3.4 The Decay $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$

As suggested earlier, we need to unmistakably reconstruct a B candidate in the event to make it useful. Ideally, we would make a list of techniques and algorithms and use them to process every event, maximizing our chances of reconstructing every possible B-decay mode known today. This would certainly be the most efficient way to find B_{rec} candidates, but as Wm. McDonough has observed, "We must seek to be effective, not just highly efficient." [30] By maximizing our efficiency, we inevitably sacrifice our purity; that is, we include candidates in our sample that are not true B decays. It becomes difficult to understand the backgrounds (mistakes) in our sample

[††] Assuming unitarity, of course

when we strive to include as much signal as possible.

It is for these reasons that our analysis focuses on one set of $B_{\rm rec}$ decay modes. We look for semileptonic decays of the B, which are distinguished experimentally by a high momentum lepton and a missing neutrino: $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$. The branching fraction for this decay channel is relatively high, about 4.6% [5]. This means that for every 100 decays of the B, about 5 of them will involve a D^* and an e, and another about 5 will yield a D^* and a μ. We discard all other events with other types of B^0 decay. However, the semileptonic features of the decay allow the collection of a sample with high efficiency and purity, as this decay has the largest branching fraction of any other exclusive B decay mode.

We define the *signal* for this analysis to be events in which at least one B decays into a final state:

- which contains a $D^{*\pm}$ and either an electron or a muon of the opposite charge produced by $W^\pm$ decay,
- where the D^* and ℓ originate from a common vertex, within the detector resolution,
- and which has the same time-dependent decay structure as the mode $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$.

Under this definition, the D^* and ℓ must either be direct decay products of a parent B^0 via $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$, or else be secondary decay products of short-lived resonances. The known resonances which we expect could contribute are radially and orbitally excited D states, [31] $Y_c = D_1^+, D_1'^+, D_2^{*+}$:

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

By requiring the D^* and lepton to originate from a common vertex within the detector resolution, we exclude leptons from intermediate τ or D^0 resonances. We require the lepton to be produced by $W^\pm$ decay, which excludes short-lived electromagnetic decays, such as $B^0 \to D^{*-}\omega\pi^+, \omega \to \pi^0 e^- e^+$.

Due to imperfections in our event selection, we also select some $D^*\ell\nu$ decays which have charged B parents, via the reaction:

$$B^\pm \to D^{*0}\ell^\pm\nu_\ell X.$$

These decays, due to detector resolution, can appear consistent with the signal B^0 decay kinematics and can thus contaminate our sample. We remedy this situation in the final analysis by incorporating their expected Δt structure into the model and explicitly measuring the contamination fraction.

The $D^{*-}\ell^+\nu_\ell$ decay is a three-body decay which means that the angular structure of the decay products is not uniquely specified by the kinematics. The variables $\hat{E}_\ell$ (the lepton energy in the B center of mass frame) and $\hat{Q}^2$ (the $W^\pm$ invariant mass equal to $m_X^2 - 2m_{B^0}\hat{E}_{D^*}$, where $\hat{E}_{D^*}$ is the hadron energy in the B frame) parameterize the phase space of the decay

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

and so are useful for disentangling the effects of phase space and dynamics. Fig. 3.1 compares the distribution of decays in the plane of these two variables [‡‡] for three-body phase space and $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ decays.

The variables $\hat{p}_{D^*}$ (magnitude of the hadronic momentum in the B frame) and $\cos\hat{\theta}_{D^*,\ell}$ (angle between the directions of the lepton and the D^* in the B frame) are completely correlated with the Dalitz variables $\hat{Q}^2$ and $\hat{E}_\mathcal{L}$:

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

Because they are also not independent, cuts applied to more than two of these four variables are not independent ($\hat{Q}^2$, $\hat{E}_\ell$, $\hat{p}_{D^*}$, $\cos\hat{\theta}_{D^*,\ell}$). Figure 3.2 shows the distribution of these variables. Figure 3.2(a) shows that the hadronic momentum has a sharp

[‡‡] Yes, a Dalitz plot!

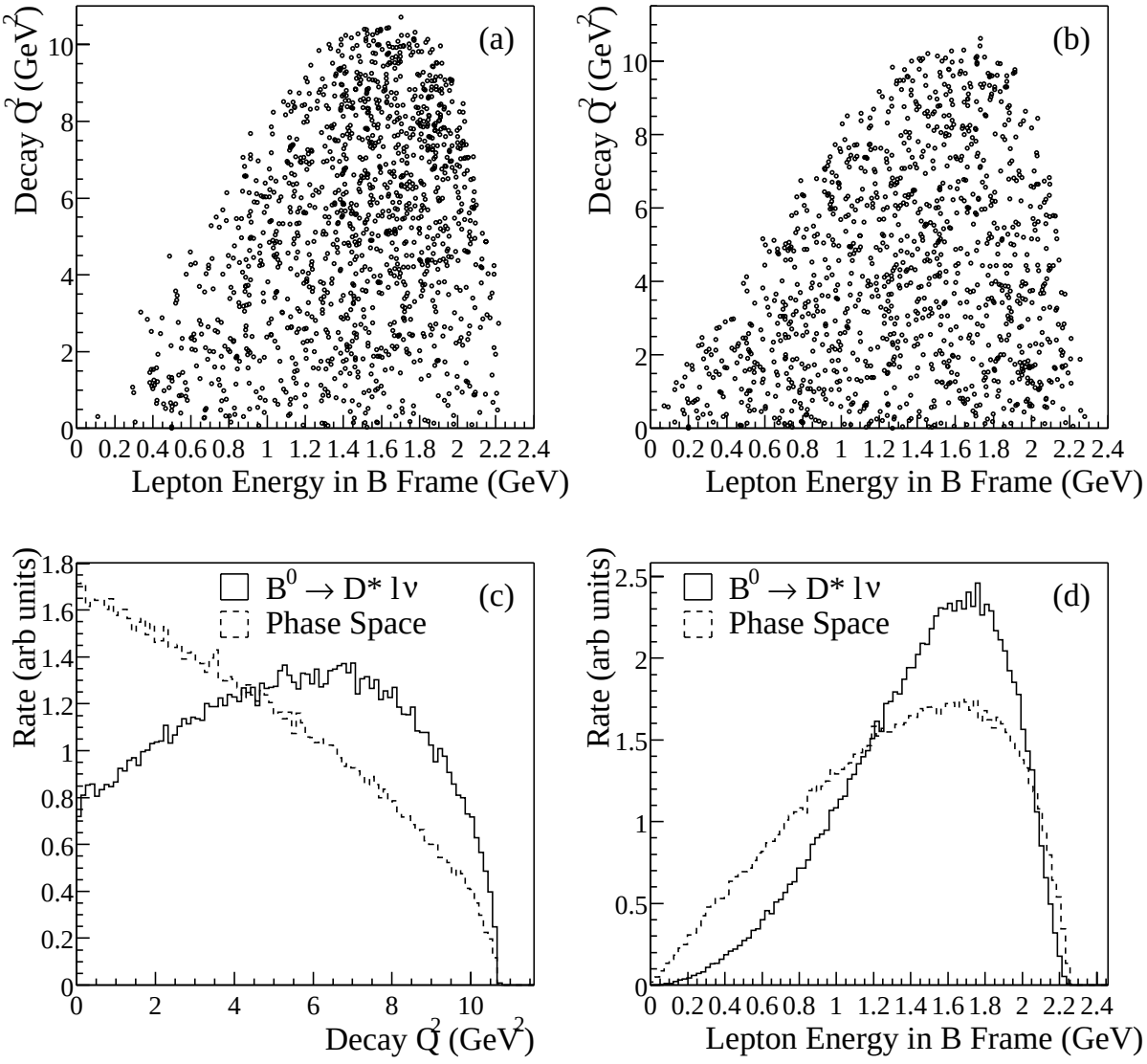


Figure 3.1: Dalitz distribution of $B^0 \to D^{*-}\ell^+\nu_\ell$ using (a) the form factors and angular-momentum correlations, or (b) three-body phase space. The lower plots show the normalized projections of these variables: (c) $\hat{Q}^2$, and (d) $\hat{E}_{\mathcal{L}}$.

edge at the upper kinematic limit, although the spectrum is significantly softened by the form factors and angular-momentum constraints (because of the $V - A$ structure of the weak interaction). Figure 3.2(b) shows that the D^* and lepton tend to be

back-to-back in the B^0 decay frame, and that this effect is mostly due to phase space considerations.

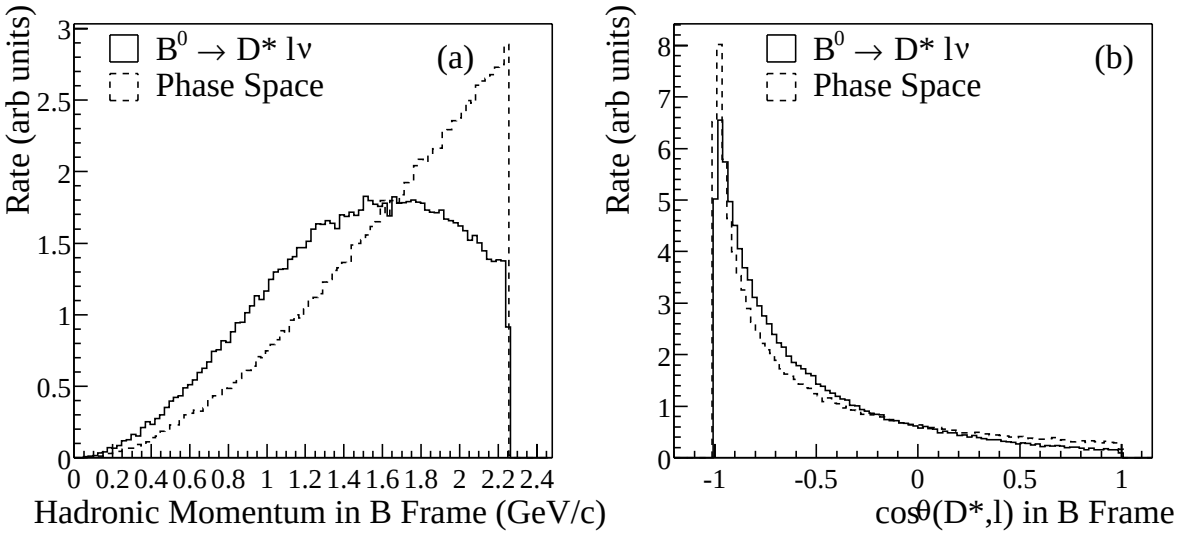


Figure 3.2: Normalized distributions of (a) the hadronic momentum, $\hat{p}_{D^*}$, and (b) the angle between the hadronic and lepton momenta, $\cos\hat{\theta}_{D^*,\ell}$, measured in the B^0 decay frame. The plots compare the distributions calculated using either three-body phase space, or else including the full decay dynamics.

Reconstruction of exclusive semileptonic decays is often more difficult than those which terminate in hadronic final states because of the undetected neutrino. Because of that deficit, there is no quantitative prescription for combining reconstructed D^* candidates with a lepton candidate to form a signal $D^*\ell$. (That is, we cannot build a B candidate because we don't have the daughter neutrino with which to apply the usual kinematic constraints.) At the $\Upsilon(4S)$, however, we can constrain the energy of the $B_{D^*\ell}$ candidate because we know the beam energy to high precision, so that $E_B = E_{beam}$. By comparing the sum of the four-vectors of the D^* and ℓ candidates with this constraint, one can test for consistency of the reconstructed D^* and ℓ system with a single missing neutrino (assumed to be massless). We can examine the missing

invariant mass:

$$\begin{aligned} p_{D^*\ell} &= p^*_D + p_\ell \\ p^2_{miss} &= p^2_\nu = (p_B - p_{D^*\ell})^2 \\ &= m^2_B + m^2_{D^*\ell} - 2E_B E_{D^*\ell} + 2|\mathbf{p}_B||\mathbf{p}_{D^*\ell}| \cos\theta_{B,D^*\ell} \, . \end{aligned} \tag{3.33}$$

If the kinematics of the measured D^* and ℓ are really consistent with a missing neutrino, we expect the missing mass as expressed above to vanish. While this quantity is a Lorentz invariant, we note that it depends on p_B (magnitude and relative direction with respect to the $D^*\ell$) which, because of our lack of complete knowledge of the final state, is unknown. At the $\Upsilon(4S)$, $|\mathbf{p_B}| = 330\,\mathrm{MeV}$ which suggests that we could simply neglect the third term in the equation, effectively smearing the p^2_{miss} distribution by the magnitude of the cross-term (about $0.5\,\mathrm{GeV}^2$). Another option is to assume the hypothesis is correct by setting $p^2_{miss} = 0$, isolating the angular cross-term, and requiring that $\cos\theta_{B,D^*\ell}$ be in the physical range -1 to $+1$ to preferentially select candidates consistent with the hypothesis. This is the approach we use in this study.

Finally, we note that since the B momentum in the $\Upsilon(4S)$ frame is small, the two rest-frames are nearly co-moving. Assuming we know the beam energy and momentum very well, we often replace quantities in the B rest frame with quantities measured in the $\Upsilon(4S)$ frame because of the improved precision (and knowledge) in the $\Upsilon(4S)$ rest frame.

Chapter 4

Other Measurement Techniques

The study of the B meson lifetime and mixing frequnecy is not new; it's been a business for about two decades now. Dozens of other measurements of the B^0 lifetime τ_{B^0} and the mixing frequency Δm_d have been made. The large variety of techniques extract the same physics, but they differ in the strategies used to (a) produce and identify candidate B decays, (b) measure the flight length of the B, and (c) as needed for mixing, identify the b quark flavor. Additionally, most previous efforts have concentrated on measuring only one aspect of the time evolution. Time-dependent analyses, which use measurements of the decay time of each B candidate, have emerged as the preferred method for achieving high precision results.

4.1 Overview

In approaching the general task of studying B meson time dependence, it is important to understand the restrictions imposed by the "production" method: B physics can be studied at e^+e^- or hadron-hadron colliders, at threshold for B production or far above it. The three major groupings are

- hadron-hadron collisions: Experiments such as CDF and D0 * at the Tevatron rely on this mechanism in which protons and anti-protons are collided to produce $b\overline{b}$ quark pairs form parton-parton interactions. Hadronization yields uncorrelated B mesons, typically at high momentum and surrounded by a jet of other hadrons.

- e^+e^- interactions at the Z^0 pole or some other "higher" energy (30 – 90 GeV): Experiments at LEP, SLC, and PEP used this technique to create $b\overline{b}$ quark pairs from virtual Z^0 or γ interactions. The resulting B mesons are, again, heavily boosted in the lab frame, but the production vertex is known with good accuracy because of the small beamspot. However, the B mesons are still typically embedded in hadronic jets.

- e^+e^- interaction at threshold, the $\Upsilon(4S)$: As described earlier, the $\Upsilon(4S)$ resonance is just above the kinematic threshold for $b\overline{b}$ quark pair production, which provides a correlated pair of B mesons nearly at rest in the center-of-mass system. The subsequent decay of the B meson is relatively isolated, and "clean." CLEO uses this mechanism. The B production point is less well-known, however. A variation on this technique is the asymmetric B factory, which boosts the center-of-mass system in the lab frame. Both *BABAR* and Belle employ the asymmetric technique at the $\Upsilon(4S)$ resonance.

*And UA1

4.2 Lifetime

The general strategy for measuring particle lifetimes is to produce the particle directly and then reconstruct its decay length and momentum to extract the decay time. The decay time distribution is described with a negative exponential (lifetime) distribution, often convolved with some smearing or resolution function, usually Gaussian in nature. The analysis relies on correctly identifying events that contain B mesons, and cleanly measuring the production and decay points. One then needs an accurate model for converting measured flight distances to decay times — a momentum distribution resolution model. As such, lifetime measurements are inherently time-dependent analyses. Typically, the B meson system is boosted in the lab frame, increasing the flight path enough for imaging by precision tracking detectors. (This is the main reason that CLEO does not have a B lifetime measurement; the B decay vertex is too close to the production point for separation and distinction. They have, however, made some contributions by extracting the lifetime ratio from measurements of the B^+ to B^0 semileptonic branching ratios. [32])

The first measurements of the B^0 lifetime were reported in 1983 by the collaborations MAC [33] and MARK II [34] at the PEP facility. Their technique took advantage of the relatively large B semileptonic branching fraction by inclusively identifying b and c events with high momentum leptons from e^+e^- collisions at 30 GeV. The impact parameter of the lepton with respect to the interaction point was computed, and distribution was fit to find the average decay length of the B meson. Interestingly, the Mark II measurement included both leptons from charm and leptons from bottom as two separate samples in their analysis, and simultaneously estimated the c-quark meson and the b-quark meson lifetimes. [†] Note, however, that these measurements did not distinguish the type of B hadron produced, and so the results are actually average B hadron lifetime measurements. [‡]

[†]This use foreshadows the simultaneous technique using background control samples that is a key feature of this analysis.

[‡]At that time, B mesons were expected to have roughly identical lifetimes as dictated by the spectator quark model.

Measurements of τ_{B^0} have not changed significantly since these first efforts. Modern techniques use decay time measurements rather than projected impact parameters due to the advent of larger data samples that permit the "luxury" of exclusive reconstruction, and the development of silicon detectors for precise secondary vertex determination. Candidate reconstructions takes advantage of sophisticated particle identification technology to more efficiently select a purer sample of B decay candidates. Three different approaches are popular today, each having to do with how the B sample is collected, with tradeoffs between sample size and measurement precision.

- *Full Reconstruction:* Reconstructing the full decay chain of the B meson, typically using hadronic final states. This provides a good determination of the decay vertex and the boost, and incorporates little dependence on simulation. However, reconstruction efficiencies are low, typically 0.1%, so the statistical power is limited by small sample sizes. Both CDF and the asymmetric B-factories have capitalized on this technique.

- *Limited/Partial Reconstruction:* This approach uses the semileptonic decay mode of the B meson to identify candidates, either by identifying exclusive modes such as $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ (but misses the neutrino), or by identifying the $D^{*-} \to \overline{D}^0\pi^-$ portion of the decay. This method is more efficient at B finding, and still allows for good decay vertex determination, but the incomplete knowledge of the event (even if only of just the neutrino) degrades the boost/momentum determination. There is a somewhat larger dependence on simulation, especially in the partially reconstructed case, where one needs to understand backgrounds in the selected sample. Almost every modern experiment has used this technique with some variations.

- *Inclusive Reconstruction:* The final approach is to inclusively identify B mesons, using minimal information about the details of the event. The cost of high signal efficiency is significantly larger backgrounds. At the B-factories, this is implemented as an inclusive dilepton analysis in which two high momentum leptons are used to identify candidate events and their distance of closest approach to the beamspot is used to estimate the decay length. At the higher $\sqrt{s}$

experiments, inclusive reconstruction can take advantage of *topological vertex reconstruction*, which Delphi has used with great success. Topological reconstruction of the vertex tries to assign all tracks in an identified b jet to either the primary or secondary vertex. The charge of the B is simply the sum of the charges of the tracks assigned to the secondary vertex. These methods (except at the B-factories where the higher mass B hadrons are not produced) suffer from contamination by B_s and Λ_B, however.

As of 2002, the most precise measurements come from the B-factories, closely followed by results from OPAL and SLD using the inclusive strategies. The industry of measuring τ_{B^0} will probably become secondary in the future, taking a back seat in the race to measure more sophisticated properties of B time evolution, *i.e.,* τ_{B^0} will be continue to be updated, but only as the byproduct of other analyses for which it is yet another floating parameter in the fit. However, if lattice QCD calculation converge on a precise prediction, interest in τ_{B^0} for its own sake might flare up.

4.3 Mixing Frequency

Strategies for measuring the neutral B mixing frequency come in two flavors [§], so-called "time-integrated" and "time-dependent" analyses. Time-integrated techniques seek to measure the mixing probability χ_d, where time-dependent analyses actually map out the oscillations of the B^0 mesons by recording the initial and final state flavor of the B^0 at the decay time. These techniques are used at all types of experiments,

The parameter χ_d is the integral of the $B^0 \to \overline{B}^0$ mixing probability (akin to Eq. 1.2), and if we assume *CP* invariance (and $\Delta\Gamma$ *equiv*0),:

$$\begin{aligned} \chi_d &\equiv \int \frac{\Gamma}{2} e^{-\Gamma t}(1-\cos(\Delta m_d t)) \\ &= \frac{1}{2}\left[\frac{(\Delta m_d/\Gamma)^2}{1+(\Delta m_d/\Gamma)^2}\right] \\ &= \frac{N_{mixed}}{N_{unmixed}+N_{mixed}} \end{aligned} \tag{4.1}$$

[§]Pun intended!

The time-integrated technique, then, relies only on detecting and accurately counting the number of mixed and unmixed B^0 events (assuming the efficiencies and misidentification probabilities for mixed and unmixed events are understood). With knowledge of the lifetime, one can immediately extract Δm_d. The earliest technique employed for identifying mixed and unmixed B^0 events uses the charge correlations of the lepton from B^0 semileptonic decays. The sign of the lepton's charge indicates the charge of the W, which in turn labels the flavor of the initial quark as b or $\overline{b}$. The key element counting the number of like-sign $B\overline{B}$ decays to measure the number of mixed events.

$B^0 - \overline{B}^0$ mixing was first reported in 1987 by ARGUS [1] (generally credited with the discovery) and UA1 [35]. ARGUS operated at the $\Upsilon(4S)$ resonance at the DORIS e^+e^- storage ring, and observed mixing in three different ways: by completely reconstructing one mixed event, by reconstructing B mesons in a set of hadronic and semileptonic final states and identifying the flavor of the recoiling B using a high momentum lepton, and by measuring the fraction of like-sign lepton $B\overline{B}$ decays. All three observations gave convincing evidence that B^0 was mixing into $\overline{B}^0$. CLEO has also reported precision measurements of χ_d, using similar time integrated techniques [36]. They expand upon the ARGUS technique by requiring only partial reconstruction of one B. The time-integrated technique is in principle the easiest, and perhaps most direct method, but it suffers from systematic uncertainties. Both CLEO and ARGUS ran at the $\Upsilon(4S)$ resonance where semileptonic decays of $B^{\pm}$, $c\overline{c}$, and $B^0\overline{B}^0$ daughters can dilute the sample. As of Summer 2002, the combined time-integrated measurement of the $B^0 - \overline{B}^0$ mixing frequency has an overall precision of 6.5%. [5]

The alternative to measuring the mixing probability is to observe the mixing dynamically, using a *time-dependent* analysis where better sensitivity is obtained. In this approach, the flavor of the candidate B meson is measured at two specific points in time. (Typically, this is the production and decay times. ¶) The task reduces to the identification of signal B^0 candidates, *flavor-tagging* the initial and final states, and measuring the proper decay time. Flavor-tagging is the process of identifying the flavor of the b quark, often using charge correlations with the B decay daughters. Because most experiments produce $b\overline{b}$ pairs, such as $\Upsilon(4S) \to b\overline{b}$ or $Z^0 \to b\overline{b}$, the

¶But no at the B factories.

flavor of the B decaying on one side of the event is correlated with the flavor of the B on the other side. This is often used to provide the mixing status information. The proper time is estimated from the distance between the production point (usually the beamspot) and the B decay vertex (often relying on the precision tracking capabilities of silicon vertex detectors) and the B momentum. The decay time distributions are then used in a fit to extract the mixing frequency.

Different analyses employ different techniques to reconstruct the B meson whose decay time they measure. Exclusive reconstruction of final states is difficult in the higher energy experiments because of the multi-track jetty environments, so partial and inclusive reconstruction techniques have been preferred. An interesting note is that the LEP and Tevatron experiments use time evolution of uncorrelated B mesons; since the b and $\overline{b}$ quarks hadronize independently, one B meson can be charged and the other neutral. Different analyses also employ different techniques to measure the initial and final state flavors of the B. The algorithms can become quite elaborate, ranging from the use of high-momentum leptons, jet charge, reconstructed D^* charge, so-called charge dipole separation, and even the polar angle of the B candidate at SLD where the colliding beams were polarized.

The asymmetric B-factories follow this same time-dependent approach, but as described in Section 2.3.2, the decay time difference measurement is used, where the initial measurement of the B is not made at the production time, but rather at the time of the decay of the (correlated) other B. Time-dependent studies of B^0 mixing are typically limited by sample contamination, knowledge of the b-hadron lifetimes, and flavor-identification mistakes known as *mistags*. Before the B-factories, the world average on Δm_d using time-dependent techniques had an overall precision of about 2.6%, [5] significantly better than the time-integrated extractions discussed above.

4.4 Comparisons to This Measurement

As mentioned, this study is a time-dependent analysis, meaning that we directly observe the B meson's life-cycle. A novel feature of this study is the use of the asymmetric B Factory, which, by virtue of the asymmetric-energy head-on collisions,

produces the B mesons at a moderate velocity along a single direction in the lab frame. The important advantage of this technique is that the B mesons travel a significantly longer distance in the lab (on average) before decaying, as compared to symmetric energy collisions at the $\Upsilon(4S)$. The decay distance difference becomes great enough to be resolved with modern particle detectors.

A key difference between the measurement reported here and those described above is the simultaneous extraction of Δm_d and τ_{B^0} using the same analysis on the same data sample. To date, there is only one other measurement of this (simultaneous) type in the B system, by OPAL at LEP, published in 2000 [37]. The OPAL result also uses the $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ decay channel, where the Z^0 is approximated as a two-body decay: $Z^0 \to b\overline{b}$ and the D^0 decay is inclusively reconstructed amidst the b-jet. The proper time is reconstructed by measuring the energy of the B^0 and the decay length (separation of the e^+e^- beamspot and the combined lepton and $\pi^-_{\rm soft}$ tracks). They include energy resolution effects and a triple-Gaussian resolution function for the estimated proper decay time. The decay flavor of the B^0 candidate is determined by the sign of the lepton's charge, and the flavor at production is measured by studying the properties of the other b hadron. [‖] The event is divided into two hemispheres by a plane perpendicular to the thrust axis, one containing the candidate b jet, and the other opposite it. By considering the net charge of the other jet, reconstructed secondary vertices, and high momentum leptons in the the opposite hemisphere, an estimate of the other b flavor is made. The final fit uses the Δt distributions, resolution functions, $m(D^*) - m(D^0)$ spectra, and four different background models to simultaneously measure Δm_d and τ_{B^0} from a sample of about $80,000$ events with an overall precision better than 7% and 3% respectively. They also observe a statistical correlation between τ_{B^0} and Δm_d of -14%.

There are several reasons why we choose to fit the B^0 lifetime and oscillation frequency simultaneously. Traditionally, they have been studied separately, and the reasoning has been that (a) the lifetime is better measured without using the flavor identification information since biases and inefficiencies can be introduced by the

[‖] In the case where the other b hadronizes as a B^0, there will be some dilution due to its own mixing.

flavor-tagging algorithms, (b) at the true physics level, the lifetime and mixing parameters are completely uncorrelated, and (c) the world average value for the observable τ_{B^0} was typically far better than could be achieved in a single measurement. Our reasoning to the contrary is listed here:

1. The expected precision for both τ_{B^0} and Δm_d in this analysis is comparable to that of the year 2000 world average (an overall precision of 2% and 4%, respectively). Therefore, there is no benefit from fixating on the world average measurement of either. Depending on other measurements for parameter input (such as τ_{B^0}) introduces systematic uncertainties as well. The resolution on Δt leads to a correlation between τ_{B^0} and Δm_d. We can study the correlations between τ_{B^0} and Δm_d.

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} by about 15%.

   (b) Since $B^0\overline{B}^0$ and B^+B^- 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.

3. The use of untagged events (actually, "poorly tagged events") even for the mixing measurement increases our sensitivity to the resolution function. As will be described later, we use the tagging information from all events in the fit for lifetime and mixing.

4. Future measurements (see Chapter 15) which could extract $\Delta\Gamma$ and study the C, P, and T properties of B time evolution will need to re-evaluate the average values τ_{B^0} and Δm_d in the event that these additional parameters are found. (For instance, τ_{B^0} is measured assuming that $\Delta\Gamma$ is zero. If $\Delta\Gamma$ is non-zero (as expected), it will become more important as the precision of the τ_{B^0} measurement increases.)

Chapter 5

The BaBar Experiment

This study uses data collected with the BABAR detector at the PEP-II asymmetric-energy B-factory at SLAC, the Stanford Linear Accelerator Center in Menlo Park, California. The PEP-II accelerator collides 9.0 GeV *electrons and 3.1* GeV *positrons to produce* $B\overline{B}$ *pairs which decay quickly into sets of charged and neutral particles. These decay products in turn pass out of the beampipe collision volume into the BABAR detector, which consists of a series of concentric cylindrical detectors, each of which focuses on certain types of measurements. The momenta of charged particles are measured with a combination of a 40-layer drift chamber (DCH) and a five-layer silicon vertex tracker (SVT) in a 1.5-T solenoidal magnetic field. A detector of internally-reflected Cherenkov radiation (DIRC) is used for charged hadron identification. Kaons are identified with a neural network based on the likelihood ratios calculated from* $\mathrm{d}E/\mathrm{d}x$ *measurements in the SVT and DCH, and from the observed pattern of Cherenkov light in the DIRC. A finely-segmented CsI(Tl) electromagnetic calorimeter (EMC) is used to detect photons and neutral hadrons and to identify electrons. Electron candidates are required to have a ratio of EMC energy to track momentum, an EMC cluster shape, DCH* $\mathrm{d}E/\mathrm{d}x$*, and DIRC Cherenkov angle all consistent with the electron hypothesis. The instrumented flux return (IFR) contains resistive plate chambers for muon and neutral hadron identification. Muon candidates are required to have IFR hits located along the extrapolated DCH track, a suitable IFR penetration length, and an energy deposit in the EMC consistent with the muon hypothesis.*

The overarching goal of the *BABAR*-PEP-II facility is to study physics processes at the $\Upsilon(4S)$ resonance with high precision, including B, c, and τ physics. The primary focus is on decays of the neutral B meson, however. As suggested in Chapter 6,the B meson has a lifetime of only 1.5 ps, and at the $\Upsilon(4S)$ resonance, this corresponds to flight path in the lab too short to measure. By "boosting" the decay in the lab frame, we achieve better spatial separation of the decay and its products so that we can examine the physics more closely.

5.1 The PEP-II Facility

The PEP-II storage ring is the main engine of the experiment. PEP stands for *Positron Electron Project*, highlighting the fact that the system collides electrons and their antiparticles, positrons. The PEP-II facility is an upgraded version of the original PEP complex from around 1980. Both detector and accelerator personnel worked long and hard to bring all parts of the project online on-schedule and on-budget in May, 1999.

We list several important features of PEP-II here.

- PEP-II stores and maintains the energy of circulating 9.0 GeV electrons and 3.1 GeV positrons in two stacked storage rings. The asymmetric beam energies are necessary to produce center-of-mass collisions at the $\Upsilon(4S)$ resonance of 10.58 GeV, and boost the the $\Upsilon(4S)$ in the lab frame with $\beta\gamma = 0.55$. We refer to the higher-energy electron beam direction as the *forward* direction.

- The high beam currents and strong focusing of PEP-II are necessary to achieve the high luminosity required, in excess of $3 \times 10^3 3/\,\mathrm{cm}^2$ s. The PEP-II design was aggressive in its magnitude and scale, and it has performed above and beyond expectations. To provide final focusing as the beams come into collision, several accelerator components are actually inside the *BABAR* detector. The SLAC Linear Accelerator provides a high rate injector for the storage rings. Compared to other e^+e^- colliders, PEP-II has significantly higher beam currents due to the hundreds of so-called "bunches," the buckets of concentrated charge

in the ring.

- The large data samples needed to achieve the physics goals of *BABAR* depend not only on the instantaneous luminosity of the accelerator, but also the *up-time*; that is, the fraction of each day that stable collisions are delivered. The *BABAR*-PEP-II team focuses on minimizing downtime in order to maximize the integrated luminosity.

Fig. 5.1 depicts the PEP-II rings and their relationship to the SLAC Linear Accelerator. For details of the PEP-II design parameters please see Ref. [38].

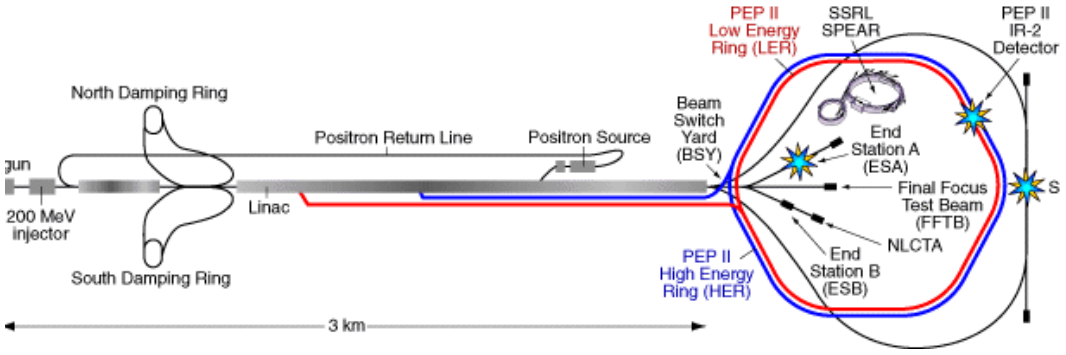


Figure 5.1: Schematic of the the PEP-II facility and SLAC. The 2 mile long Linear Accelerator provides intense beams of electrons and positrons which are injected into the PEP-II storage rings and come into collision at region IR-2.

The interaction region which includes the final focusing elements and the *BABAR* detector is a technical feat of modern engineering. We include an illustration of the final approach of the two beams in Fig. 5.2. The storage ring is, of course, not 100% efficient, as there are small losses around the ring where circulating charge is ejected from the nominal orbit. Primarily because of the high beam currents, these losses can sum to substantial radiation backgrounds in the *BABAR* detecotr, degrading not only instantaneous response (due to the presence of extra tracks, triggers, and noise) but also long-term performance through radiation damage. We refer the reader to the Appendices for several notes on these matters.

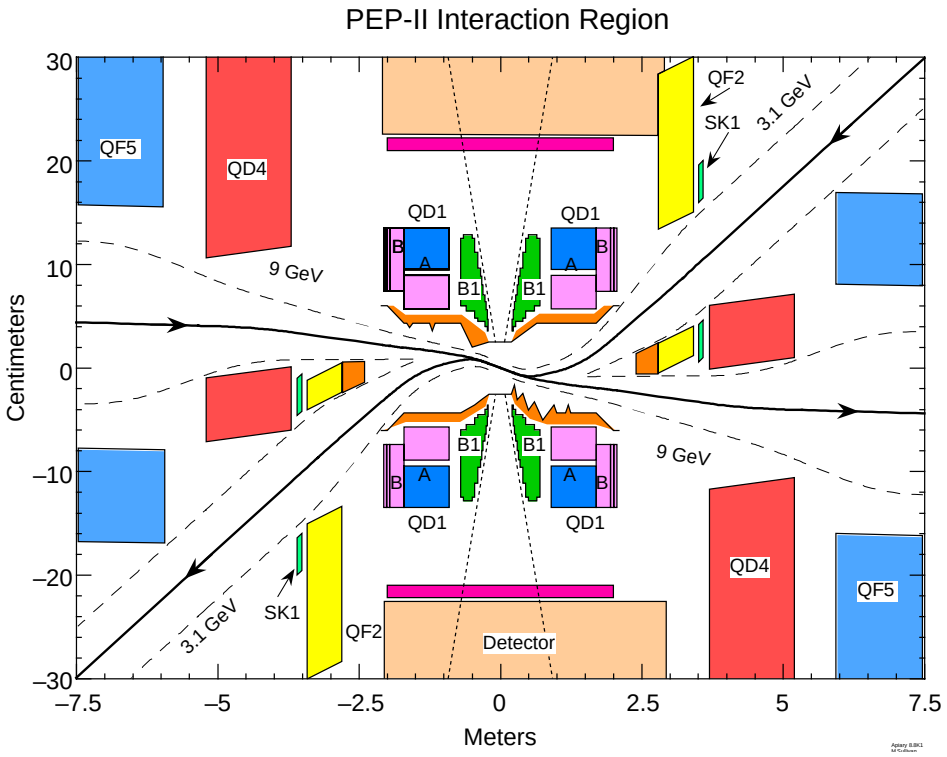


Figure 5.2: Plan view of the PEP-II interaction region, showing the focusing of the two beams in the magnetic elements. The beams collide head-on and are then separated. The strong focusing is achieved by the final doublets of quadropole magnets, Q1, Q2 and Q3, Q4. The internal B1 dipole magnets are permanent ferromagnets and perform the final bends to place the beams in (and out of) collision. (Diagram courtesy Ref. [39].)

5.2 The *BABAR* Detector

The *BABAR* detector is designed to quickly and accurately image the $B\overline{B}$ decay products as they travel outward from the interaction point. It is roughly a series of

concentric cylinders surrounding the near-IR beamline. Fig. 5.3 shows a side- and end-view of the detector, illustrating its symmetries. In this Section, we'll touch upon each of the *sub-systems* of the detector, each of which consists of a separate cylindrical layer specialized in some functionality with specific types of instruments.

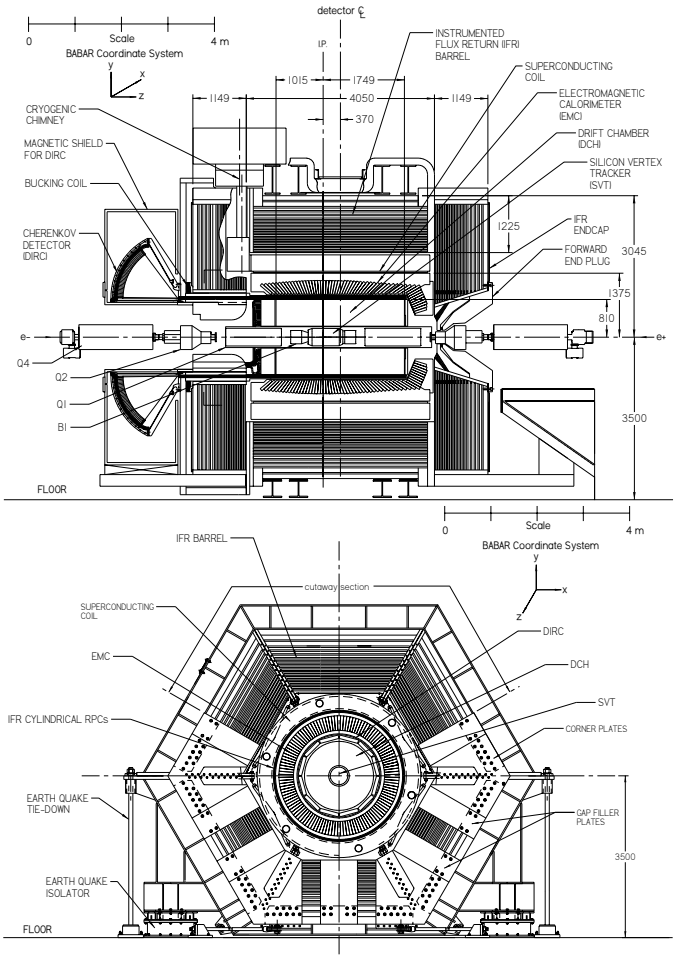


Figure 5.3: Side- and end-views of *BABAR*, illustrating the set of cylindrical layers that comprise the detector. The e^+e^- interaction takes place inside the innermost volume, an evacuated beampipe through which the beams travel. Each collision typically has enough transverse momentum to spray the particles outward and forward (due to the boost) into the detector volumes.

To achieve its physics goals, the following design attributes were required for *BABAR*:

- Maximum possible acceptance in the center of mass system
- Excellent vertex resolution (this requires not only good track momentum and angle measurements, but also an inner tracking system as close to the interaction point as reasonably achievable)
- Optimal tracking over a wide range of momenta
- Efficient neutral particle reconstruction
- Excellent particle identification over a wide kinematic range.

These requirements are met through the combination of six specialized systems, layered in increasing radius:

1. The Silicon Vertex Tracker (SVT)
2. The Drift Chamber (DCH)
3. The Detector of Internally Reflected Cherenkov Light (DIRC)
4. The Eletromagnetic Calorimeter
5. A superconducting solenoid magnet to immerse the tracking volumes in a uniform 1.5T magnetic field
6. The Instrumented Flux Return (IFR).

These systems are each complex and fine-tuned, but they also collaborate seamlessly, as particles must travel through all the inner detectors before reaching the next outer one. The goal is, then, to make each sub-detector nearly "invisible" to the particles in all aspects except the one mechanism used to detect their passage.

The *BABAR* coordinate system places $+z$ in the forward, e^- beam direction, and uses cylindrical coordinates to specify a polar angle θ measuring from the horizontal and the azimuthal angle ϕ.

5.2.1 Silicon Vertex Tracker

The innermost *BABAR* subsystem is the *Silicon Vertex Tracker*, known as the SVT. The SVT consists of 5 cylindrical layers of double-sided silicon wafers arranged around the beampipe. The inner radius is about 3.2 mm from the beamline; the SVT extends outward to about 144 mm. Fig. 5.4 shows a side- and end-view of the SVT to illustrate its geometry. Note that it is asymmetric between the fore and aft; this is because of the asymmetric energies: the substantial boost forces decay products to be flung forward, such that soft (low transverse momentum to be precise) tracks will be at a low polar angle. The design of the SVT accommodates this by placing all the cooling and readout components in the backward direction where they will not occlude the detector. The bend in the arch modules increases the solid angle coverage and avoids very large track incidence angles. The design of the SVT was heavily constrained by the accelerator components in the interaction region, including the aforementioned B1 dipole magnets.

We note several key features of the SVT here:

- The principle of operation is the creation of electron-hole pairs caused by the passage of charged particles through the silicon semiconductor substrate. The $e - h$ pairs flow under the influence of an applied electric field, and the image charge's movement is noted at AC-couple charge-integrating contacts. (See Fig. 5.5.) Since $\mathrm{d}E/\mathrm{d}x$ is about 3.8 MeV/cm for a minimum ionizing particle, and the average energy to create an $e - h$ pair in silicon is 3.6 eV, the $300\,\mu$m thick wafers are capable of producing about $24,000$ electron-hole pairs per minimum ionizing track.

- Each side of the silicon wafer-sensor is divided into strips, and the two sides of each sensor have strips running in perpendicular directions. Strips that run parallel to the beam axis integrate along their length and measure the ϕ position of incident tracks, while strips running circumferentially around the detector measured the z-position. The SVT sensors sum to almost a square meter of silicon, and provide about $150,000$ channels of information to *BABAR*.

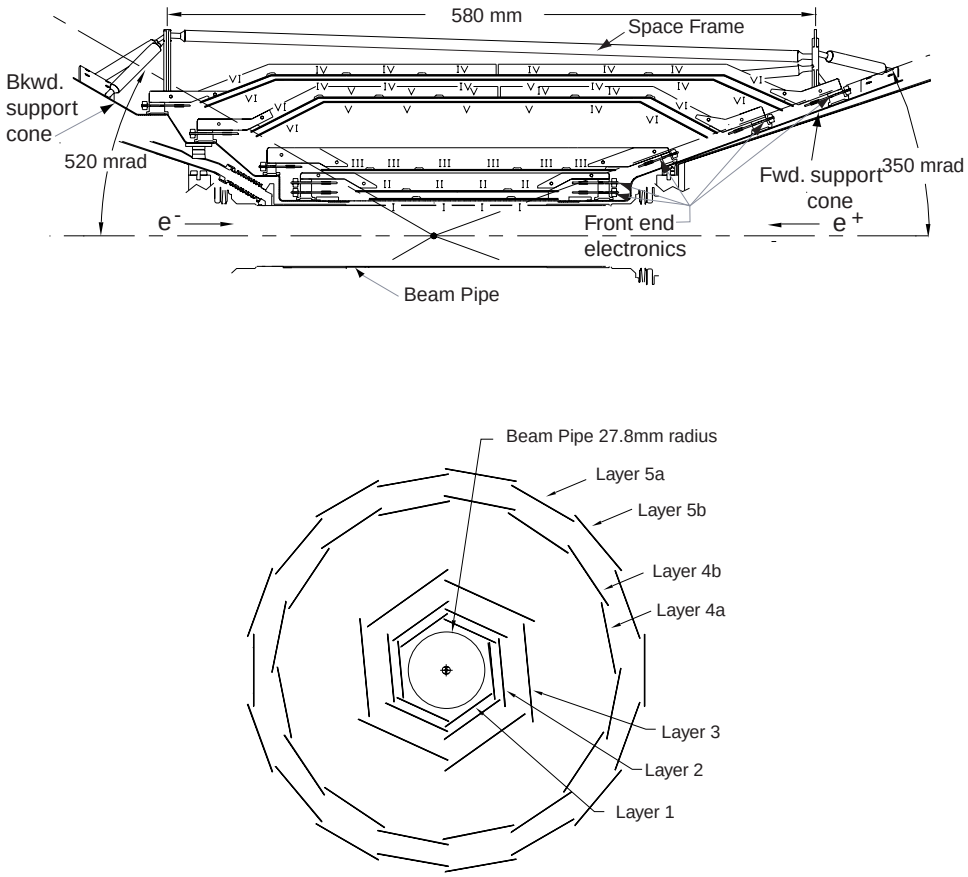


Figure 5.4: Side- and end-views of the SVT. In the longitudinal view, the lower half of the detector has been suppressed for clarity. Note the arch-shape of the outer layers, as referenced in the text.

- The SVT dominates the measurement of the decay time difference, while the other subsystems contribute to reconstruction and particle identification. As such, it's hit resolution is critical, but limited by multiple scattering, which is of order 10-15 μm in the inner layers. Taking just the readout pitch into consideration, for an inner layer, the best single resolution possible is about 10 μm, so the SVT design is well-balanced. As indicated in Fig. 5.6, the single

hit resolution of the SVT in Layer-1 is better than 20 μm in ϕ and 40 μm in z. *

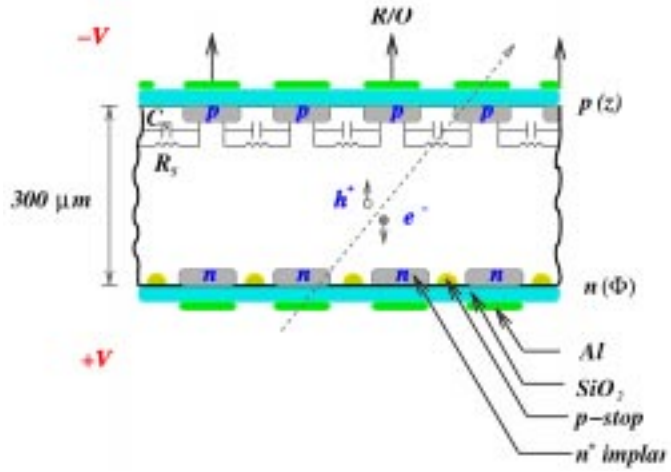


Figure 5.5: A simplified portrayal of an inner layer SVT wafer showing the readout strips which sense the movement of charge in the active volume. The implants serve to isolate the channels.

5.2.2 Drift Chamber

Traveling outward in radius, the next detector is the precision tracking chamber, the *Drift Chamber* or DCH. The DCH provides up to 40 coordinate measurements per charged track for particles with momenta greater than approximately 100 MeV. In order to minimize multiple scattering losses, the DCH has been constructed of low mass materials. It, too, is fore/aft asymmetric, as all the readout and control electronics are on the rear bulkhead, outside of which there are no external sub-detectors. It is 280 cm long and extends from 23.6 cm in radius out to 80.9 cm. † The Drift Chamber is perhaps the most "conventional" particle detector employed in *BABAR*.

We mention a few keys aspect of the DCH here.

*The hit resolution is calculated in Data by measuring the distance between the single hit and the reconstructed track without using that hit.

†Between the DCH and the SVT is the *Support Tube*, a mechanical component which suspends the SVT and gimbals it off the magnet rafts which support the B1 dipole magnets.

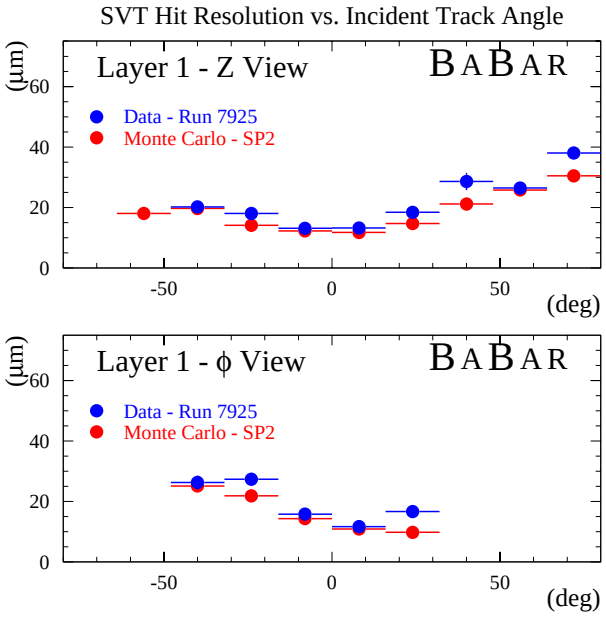


Figure 5.6: Single hit resolution for Layer-1 of the SVT in Data and and Monte Carlo as a function of incident track angle.

- The principle of operation for a drift chamber is the ionization of gas due to energy deposited by a charged particle traversing the detector. An array of high voltage wires strung the length of the cylindrical chamber causes the ionized electrons to drift and then avalanche. Nearby sense wires detect the movement of charge and are connected to the readout electronics. So-called *time-to-distance relations* relate the arrival time of the charge pulses to the initial distance from the wire of the particle's energy deposition.

- The 40 "layers" are organized into 10 superlayers, some of which are purely axial (the wires run parallel to the axis of the chamber) and some are slightly rotated, meaning that they connect to points on the front and back endplates of different (r, ϕ). [‡] The stereo layers alternate in positive and negative stereo

[‡]This, in fact, is what limits the amount of stereo angle: the effective "waistening" in the chamber center must be small enough to prevent the wires from one layer from touching the wires of another.

angles, and the "intersection" helps provide information about the z-coordinate of tracks. Each layer is organized into cells, which consist of a single sense wire surrounded by six field shaping wires ($\sim 1900V$). There are a total of $7,104$ cells.

- A particle must have a minimum transverse momentum of about 50 MeV to reach the DCH, but accurate track reconstruction requires more than one hit; the quality cutoff is typically around 120 MeV.

- Information from the DCH is used for particle identification. Fig. 5.7 shows the energy loss per unit distance ($\mathrm{d}E/\mathrm{d}x$) for different particles of the same momentum. DCH $\mathrm{d}E/\mathrm{d}x$ is most useful at lower momenta.

- At full design voltage, the average tracking efficiency is better than 98%. The momentum resolution is parameterized as fractionally sensitive to the transverse momentum of the charged track:

$$\frac{\sigma_{p_T}}{p_T} = (0.13 \pm 0.01)\% p_T + (0.45 \pm 0.03)\% \tag{5.1}$$

5.2.3 Detector of Internally Reflected Cherenkov Light

BABAR uses a novel ring-imagine Cherenkov light detector for particle identification, the *Detector of Internally Reflected Cherenkov* Light, known as the DIRC. The DIRC consists of a thin layer of highly polished quartz bars arranged regularly around the outside of the DCH, as shown in Fig. 5.8.

As a charged particle passes through matter, the surrounding atoms polarize and subsequently depolarize, and a weak electromagnetic wave spreads out from the instantaneous position of the particle. For a particle traveling more slowly than the local speed of light, wave-fronts originating at different times can never meet, and no interference is possible. For a particle travelling faster than light, the wave-fronts do overlap, and constructive interference is possible, leading to a significant, observable signal. This constructive interference produces light known as Cherenkov radiation,

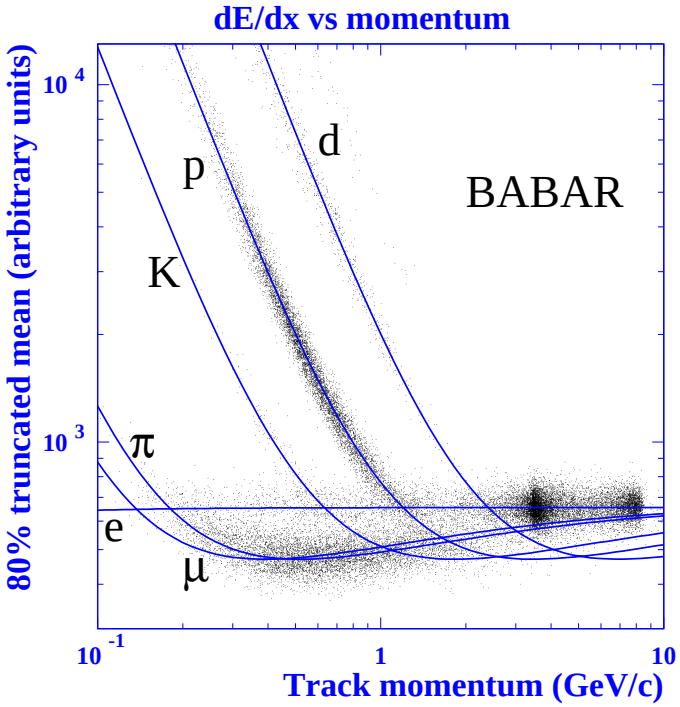


Figure 5.7: Measurement of $\mathrm{d}E/\mathrm{d}x$ in the DCH as a function of track momentum. The Data includes beam background triggers to enrich the sample in protons. The curves are the Bethe-Bloch predictions for the different particle masses.

and the effect is velocity dependent. If the momentum of the particle is known, as perhaps provided by the DCH, one can use the DIRC information to extract the mass and therefore identify the particle species.

The DIRC is specially constructed so Cherenkov photons are emitted in the quartz bars; the highly polished sides have excellent internal reflectivity so that the photons bounce back and forth as they travel down the length of the bar and into the water-filled detection volume. The Cherenkov photons produce a ring of light on the back wall, which is instrumented with phototubes. A measurement of the diameter of the ring identifies the original Cherenkov photon angle, when the incident track parameters are also known.

We note here a few features associated with the DIRC. [40]

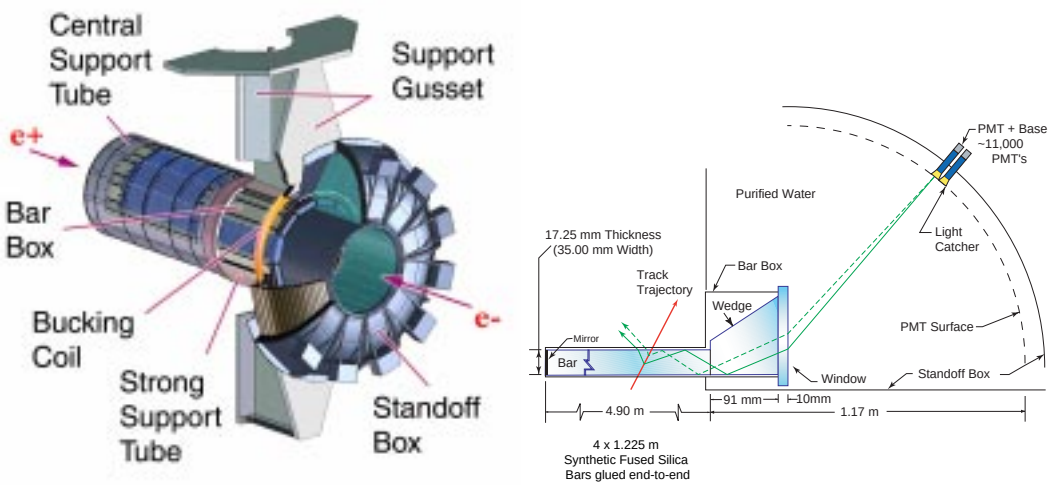


Figure 5.8: Sketch of the DIRC subsystem, including the strong support tube which holds the large water-detection volume carpeted with phototubes. Note again that we make use of the asymmetry at *BABAR*: the detection volume is mounted off the rear. The right-hand figure shows the principle of DIRC operation: total internal reflection, and then "expansion" of the Cherenkov ring onto the imaging rear face of the water-volume.

- Good $K^{\pm}/\pi^{\pm}$ separation is achieved with the SVT and DCH $\mathrm{d}E/\mathrm{d}x$ measurements up to momenta of about 0.7 GeV. The DIRC offers $K^{\pm}$-$\pi^{\pm}$ separation of about 4-σ at 3 GeV.

- The error on the Cherenkov angle is controlled by the photon counting statistics. There are about 28 Cherenkov photoelectrons detected per $\beta = 1$ normally incident particle. The average resolution on the Cherenkov angle for a single track is 2.5 mrad.

- The particle's mass can be related to the Cherenkov angle and momentum by

$$m = p\sqrt{n_{quartz}^2 \cos^2\theta_C - 1} \tag{5.2}$$

  where n_{quartz} is 1.473, the index of refraction for quartz.

- The DIRC has excellent timing resolution, as it must be able to separate photons

from different beam crossing interactions, and all of the "internals" operate at the speed of light (as opposed to drift velocities, etc.).

- The reflection coefficients of the quartz bars must be better than 0.9992, because photons can bounce as many as 365 times down the length of the bar.

5.2.4 Electromagnetic Calorimeter

So far all the subsystems discussed are sensitive only to the passage of charged particles. The *Electromagnetic Calorimeter* (or EMC for short) is the first detector which is sensitive to neutral particles (the photon being one of the most important) and, in fact, is used to detect photons and separate electrons and positrons from charged hadrons. Because half of the photon energies from generic B decay are below 200 MeV, the EMC is designed to reconstruct low energy photons with high efficiency and excellent resolution.

The (roughly) cylindrically symmetric EMC is shown in Fig. 5.9. The EMC consists of $6,580$ crystals, the majority of which are arranged in rings of ϕ to form the barrel part of the calorimeter. The remaining 12% are distributed in a forward endcap, where again we take advantage of the preferred boost direction do not even instrument the backward direction. The crystals are Cesium-Iodide, doped lightly with Thalium. In comparison with other scintillating crystals commonly used, CsI(Tl) has high density, excellent light output and long decay constant (the latter being a drawback).

We note here several aspects of the EMC. [40]

- Again, to minimize material in front of the subdetector, all services and electronics are located behind the crystals.

- Charged particles produce scintillation in the crystals, and the photons are detected by redundant photodiodes on the backs of crystal. Phototubes are inappropriate due to the ambient 1.5T magnetic field. The average light yield is $7,300$ photoelectrons/ MeV.

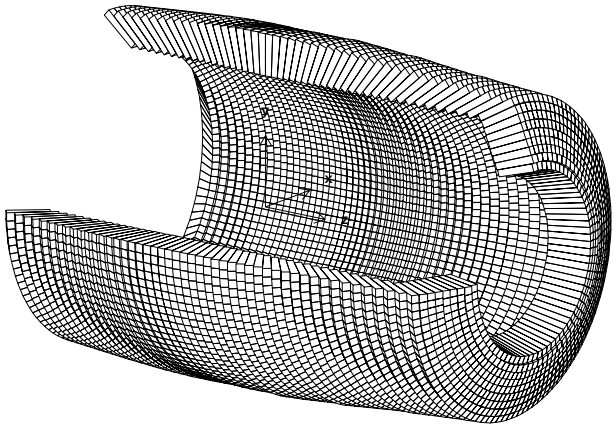


Figure 5.9: Plan drawing of the assembly of CsI(Tl) crystals that form the' active regions of the EMC subdetector.

- The EMC is more than 96% efficient for photons with energy greater than 20 MeV, where this threshold is set by beam backgrounds and material in front of the crystals. The energy resolution for photons is measured (from a combination of radioactive source calibrations and Bhabha scattering) to follow:

$$\frac{\sigma(E)}{E} = \frac{(2.32 \pm 0.30)\%}{E^{1/4}} \oplus (1.85 \pm 0.12)\% \; / \tag{5.3}$$

  The first term measures the primarily photon counting statistics; the second term (dominant at $E > 1\,\mathrm{GeV}$ arises from non-uniformities in light collection and uncertainties in calibration. The angular resolution is determined by the transverse crystal size and the average distance to the interaction point, and has been measured from an analysis of π^0 and η_0 decays to two photons to be

$$\sigma_\phi = \sigma_\theta = \frac{(3.87 \pm 0.07)}{\sqrt{E(\,\mathrm{GeV})}} \oplus (0.00 \pm 0.04)\; mrad\,. \tag{5.4}$$

5.2.5 Instrumented Flux Return

A superconducting solenoidal magnet surrounds the EMC and provides a uniform 1.5T field throughout its interior. The provided field is uniform in the tracking

chamber (the DCH) to within 2%. Final focusing of the PEP-II beams in the domain of the solenoidal field is tricky, since it can couple some of the beam modes. Sextupoles on either side of the interaction region, when finely tuned, remedy this effect.

The significant flux of the magnet is returned through a hexagonal, instrumented detector of iron, the *Instrumented Flux Return* (IFR). The iron acts as a yoke, and is segmented into sandwiches of active detector volume and iron plates, as shown in Fig. 5.10. This sequence of absorber-detector-absorber serves as a long-lived neutral hadron and muon detector. Each sandwich is designed as a resistive plate chamber with high voltage Bakelite cathodes coated with graphite strips separated by a 2 mm gap. Muons with sufficient momenta (greater then 1 GeV) pass through the chamber and cause an ionization discharge across the gap, losing energy each time in the iron regions. Long-lived neutral hadrons such as K_L^0 pass through the IFR chambers and, due to interactions in the iron, cause small "showers" of charged daughters that traverse the gaps and cause the quenched discharge. A key difference is that the charged muon candidates will have an associated track leading up to the IFR chambers, whereas the K_L^0 (or others) will be "invisible" until forced to interact with the very dense iron and steel. Additionally, the expected number of interactions for muons and K_L^0 is different.

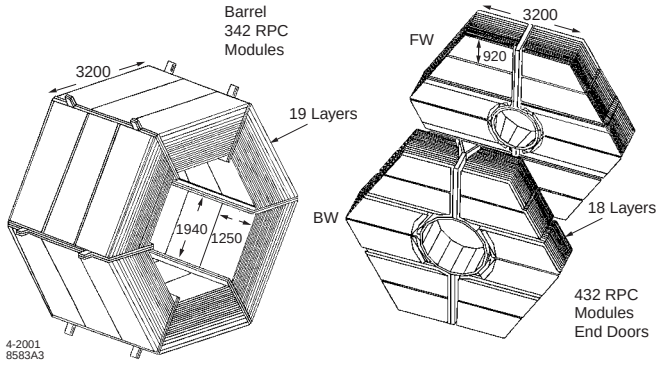


Figure 5.10: Two different views of the IFR detector, showing the hexagonal shaped barrel region and the two endplates. All are instrumented with the resistive plate chamber sandwiches described in the text.

Unfortunately, the resistive plate chambers suffer from efficiency problems which

are apparently aggravated by temperature and humidity. As a result, the *BABAR* muon identification has been steadily decreasing as the quality of the chambers has been degrading. During the 1999-2000 run, over 75% of the chambers were better than 90% efficient.

5.2.6 Trigger and Data Acquisition

Beam crossings occur at several hundred MHz, depending on the pattern of colliding bunches in the PEP-II rings. Not every collision results in an interesting interaction of course, and as mentioned earlier, not every collision contains B physics. *BABAR* uses a two level *trigger* system to filter out machine background events, Bhabha scatterings, and other unwanted physical processes. The trigger is designed to gather categorical and global event information to quickly identify the structure of an event, and then to make a decision about whether information gathered in the entire detector should be stored. At design luminosity, the $b\overline{b}$ production cross-section leads to about 3 $B\overline{B}$ pairs produced per second, which is "plenty of time" for modern electronics and the physical processes at work inside each subdetector to achieve equilibrium and/or deposit signal. As such, *BABAR* often refers to itself as a nearly deadtime-less detector, as relatively little (fractional) time is spent reading out the entire detector.

The key purpose of the filter, then, is reject events associated with the e^+e^- interaction that are unwanted, and to preferentially retain those events that are consistent with interesting physics. The tricky part is making the accept/reject algorithm unbiased. *BABAR* accomplishes this is by making the trigger relatively loose. Table 5.1 shows the performance of the first stage of the trigger (the so-called *Level 1* decision) for several different physics processes. The Level 1 decision is made using charged track candidates in the DCH, showers in the EMC, and tracks detected in the IFR. Because of the need for minimum latency, the Level 1 algorithms are implemented in hardware.

Events accepted by Level 1 are then passed to Level 3 which is designed to accept events at a maximum rate of around 100 Hz. Level 3 performs additional selection and filtering to help reduce backgrounds, and runs in software. Once an event is

Event type	Cross section (nb)	Production Rate (Hz)	Level 1 Trigger Rate (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

Table 5.1: Cross sections, production and trigger rates for the principal physics processes at 10.58 GeV for a luminosity of $3 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The e^+e^- cross section refers to events with either the e^+, e^-, or both inside the EMC detection volume. (This tabulation courtesy Ref. [40].)

accepted by Level 3, the detector freezes. Each subsystem stores sets of recent event information in a *pipeline memory* which buffers the events for later readout. A Level 3 accept causes the *online data flow* software to take over which slowly pulls information from each subsystem and passes it the event processing routines. The event processing is handled by a many-node computer farm which organizes and collates the Data for permanent storage on the local filesystem.

For more information on the *BABAR* event store, particularly its implementation, please see Appendix C.

5.3 Performance and Data Samples

The *BABAR* experiment has been running successfully since May, 1999. As of July 2002, the experiment has recorded more than $90\,\mathrm{fb}^{-1}$ of Data, and maintained an average efficiency of more than 90%. Fig. 5.11 shows the phenomenal success of the *BABAR*-PEP-II team effort. The *BABAR* shift crew consisted of 4 people on duty 24-hours a day: the Shift Leader, the Data Quality Shifter, the Data Acquisition Shifter, and the *BABAR*-PEP-II Liaison.

The analysis described in this study uses the 1999-2000 dataset which includes

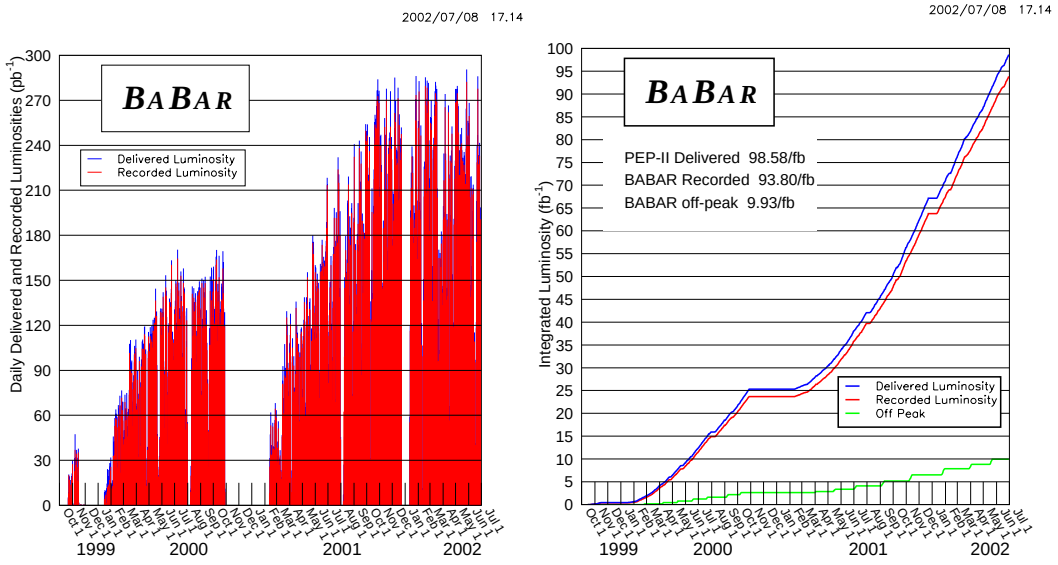


Figure 5.11: These publicity plots highlight the tremendous startup of the *BABAR*-PEP-II facility. Design capacity was achieved early on by reaching the goal of 120 p pb^{-1} er day.

$20.6\,fb^{-1}$ of on-resonance data and $2.6\,fb^{-1}$ of off-resonance data. [§]

5.4 Monte Carlo Data Samples

A key toolset in successfully extracting physics from the *BABAR* data is the complementary simulated data samples, the so-called *Monte Carlo* data. "Monte Carlo" refers to the core technique for generating the simulated data — given probability distributions and analytic constraints, the Monte Carlo method is used to quickly generate a random set of events that are distributed according to the input probability density functions. The simulated data is computer-generated, and can only exhibit patterns or relationships inasmuch we have parameterized and modeled them. In short, the output events are only as realistic as the input physics. These simulated datasets are powerful tools, since they provide the physicist with two features.

[§]This particular dataset has been "re-processed" since then (improving calibrations and track-finding algorithms), although we elected not to start over by switching to the new version of the sample.

1. An alternative dataset that is sufficiently complex and general enough to provide an analysis strategy test and development environment, and

2. A dataset whose "truth" can be revealed, since the original information about what events were generated is also stored in the file.

In this study, we do not rely on simulated data to make our final measurements [¶] but we do use it in order to convince ourselves that our techniques are extracting the information we expect.

MC event simulation begins with the generation of a selected B decay (parented by the $\Upsilon(4S)$), including event kinematics and decay chains. The algorithm then tracks the $B\overline{B}$ meson pair and their decay products as they traverse the detector volume, depositing simulated energy in simulated front-end detector elements. This map of the simulated tracks of the particles from the original $B\overline{B}$ pair is stored in a format nearly identical to that of the real data provided by the true detector. In this way, the offline candidate reconstruction software can process either simulated data or real data, yielding the best estimate about what physics was in the event, independent of "where" the data really came from.

The trickiest part of creating a simulated data sample is ensuring that it is systematically consistent with (and even faithful to) actual event data. [‖] There have been four significant cycles of centrally-produced simulated data for *BABAR* (so far); the set set of Simulation Production data used in this analysis is called SP3. The simulated data is produced centrally so that collaborators can study and benchmark to the same set of standards. There are two broad classes of simulated data: "signal Monte Carlo data" (signal MC) and "generic Monte Carlo data" (generic MC). A signal MC dataset is one that was specifically generated for one or two specific decay modes of the original B meson, in our case, this is $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$. By generating only events of this type, and passing them through the reconstruction software, we can focus specifically on what happens to true signal events in the presence of detector reconstruction effects, event selection effects, and other inefficiencies. The other

[¶] Unlike, perhaps, current students at the LHC.

[‖] Of course, it won't be exactly like the Data, since if we knew how to precisely reproduce real Data, we wouldn't need to do the experiment to measure the true physics.

class of MC data is generic Monte Carlo, where the generated B meson pairs are allowed to decay to any of many different final states. In this way, we simulate the true conditions in the Data — we must search through an entire dataset of many different processes to identify the events that specifically contain $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ signal candidates. Generic Monte Carlo data is used, therefore, to examine how natural (and understood) competing physics processes can influence the analysis; *i.e.,* generic MC allows us to study backgrounds in the sample.

The key to understanding Monte Carlo data is that although it is "man-made" and computer-generated, it is also an elegant technique to produce realistic responses from very complex systems, whose internal dynamics are driven by random processes. We can often learn from a proper physics-based Monte Carlo about effects which were not explicitly encoded in the inputs. (That is, a good Monte Carlo simulation can expose correlations between measured quantities.)

For this study, we use all available SP3 $B^0\overline{B}^0$, B^+B^-, $c\overline{c}$, and $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ signal Monte Carlo samples generated and reconstructed with Release 8.8.x. ** The generic Monte Carlo samples each correspond to about $15\,\mathrm{fb}^{-1}$ equivalent luminosity for the $B^0\overline{B}^0$ and B^+B^- samples, and about $7.6\,\mathrm{fb}^{-1}$ for the $c\overline{c}$.

In order to increase the size of our generic Monte Carlo sample, an additional sample of "filtered generics" was produced. In this "filtered" sample, a true lepton with center-of-mass momentum greater than 1 GeV/c is required at the generator level. These samples provide an additional 10.4 fb^{-1} $B^0\overline{B}^0$ and 16.4 fb^{-1} B^+B^- equivalent luminosity. [41]

The signal Monte Carlo samples correspond to events in which one B decays to D^{*-} ℓ^+ ν_ℓ with $D^{*+} \to D^0\pi^+$. Separate samples are created for the four different D^0 decay modes under study. The samples consist of 424k $K\pi$ ($161.3\,\mathrm{fb}^{-1}$), 436k $K\pi\pi\pi$ (84.9 fb^{-1}), 654k $K\pi\pi^0$ (68.7 fb^{-1}) and 102k $K^0_S\pi\pi$ (55.2 fb^{-1}). Since the effective luminosities for the Monte Carlo samples are different for each D^0 decay mode, we must re-weight them in any combined study in order to realistically model the contributions from each in real Data.

** "SP3" and "Release 8.8.x" are labels for specific versions of the software used to produce the dataset. These versions correspond to the releases used in the real Data processing as well.

The value generated in the Monte Carlo across all samples for τ_{B^0} is 1.548 ps and for Δm_d is 0.472 ps^{-1}. In generic Monte Carlo, the ratio of mixed to unmixed events (parameterized by χ_d) is consistent with these values of τ_{B^0} and Δm_d. However, signal Monte Carlo samples were generated with a value of $\chi_d = 0.160$, which is inconsistent with the value of $\chi_d = 0.174$ that corresponds to the generated values of τ_{B^0} and Δm_d. [††] We have corrected this effect for all studies described in this document by rejecting a precise fraction (9.5785%) of the unmixed events[‡‡].

We rely on Monte Carlo "truth-matching" in some proof-of-principle studies for this analysis. Truth-matching refers to looking at the event's generated characteristics and identifying which reconstructed objects most likely correspond to the underlying true candidates. This matching process is often hampered by the processes where particles interact in the detector volume (in simulation as in reality) and change form. We often refer to the "correctly reconstructed signal MC sample" which is the set of signal MC events whose $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ signal candidates pass the Monte Carlo truth matching test. This sample therefore has realistic tagging-side performance characteristics, but the best possible resolution on the reconstructed side.

5.5 Candidate Reconstruction

The reconstruction stage of event processing is the final one performed centrally at *BABAR*. *Online Prompt Reconstruction* processes each run and builds so-called "candidate lists" in the event and stores *tagbits* which record (typically) binary information about what types of candidates were found in the event. Again, the dataflow model at *BABAR* is that the general user has access to the entire Data sample, but each event has already been examined by a variety of "loose selectors" that have recorded whether or not basic particle candidates were found (such as K^0_S, or $D^0 \to K^-\pi$, etc.). The analyst then needs only to "assemble" the basic composite candidates to fit their particular decay sequence, and retain those events in which candidates

[††] *i.e.,* There was a typo in the decay file. χ_d, Δm_d, and Γ are related by $\chi_d = \frac{(\Delta m_d/\Gamma)^2}{2((\Delta m_d/\Gamma)^2+1)}$.

[‡‡] A special `perl` script (asciiChi_d.pl) was prepared to remove these candidates randomly from the signal Monte Carlo datasets.

are successfully built. In this section we'll touch briefly on some of elements of that process.

5.5.1 Track Reconstruction

Track measurements are important for the extrapolation to the DIRC, EMC, and IFR. Most critical are the angles at the DIRC, because the uncertainties in the charged particle track parameters add to the uncertainty in the measurement of the Cherenkov angle. At lower momenta the DCH measurements are more important, while at higher momenta the SVT measurements dominate.

Charged particle tracks are reconstructed using information from both the DCH and the SVT. Tracks are found by fitting the expected charged-particle-in-a-magnetic-field helices to the sequences of hits. Once a track is successfully seeded, hits will be added if they consistent with the track. Tracks are parameterized by five parameters, $(d_0, z_0, \phi_0, \omega, \tan\lambda)$. These parameters are measured at the point of closest approach to the z-axis; d_0 and z_0 are the distances of this point from the origin of the coordinate system in the x–y plane and along the z-axis, respectively. The angle ϕ_0 is the azimuth of the track, λ the dip angle relative to the transverse plane, and $\omega = 1/p_T$ is its curvature. d_0 and ω are signed variables depending on the charge of the track. The track finding and the fitting procedures make use of a Kalman filter algorithm [42] that takes into account the detailed distribution of material in the detector and the full map of the magnetic field.

The particles that are actually observed in the detector are only pions, electrons, muons, kaons, protons and photons. All other particles are reconstructed from these basic building blocks, as they have typically decayed in flight leaving just the longer-lived daughters. The candidates (neutral clusters and charged tracks) are grouped into the following lists depending on the quality and characteristics of the track/cluster [31].

- *ChargedTracks*: All reconstructed tracks from DCH or SVT. The pion mass hypothesis is the default.
- *GoodTracksVeryLoose*: subset of *ChargedTracks*, must pass:

1. a distance of closest approach to the per-event beam spot of $|\Delta z| < 10$ cm, and $\sqrt{\Delta x^2 + \Delta y^2} < 1.5$ cm
2. a maximum track momentum measured in the lab frame of 10 Gev
3. a minimum number of DCH + SVT track hits ≥ 5

- *GoodTracksLoose*: subset of *GoodTracksVeryLoose*, must pass:

1. a minimum transverse momentum of 100 MeV
2. a minimum number of 12 track hits recorded in the DCH

- *GoodTracksTight*: subset of *GoodTracksLoose*, must pass:

1. a distance of closest approach to the per-event beam spot of $|\Delta z| < 3$ cm, and $\sqrt{\Delta x^2 + \Delta y^2} < 1$ cm
2. a minimum number of 20 track hits recorded in the DCH

- *GoodPhotonLoose*

1. a minimum calorimeter energy of 30 MeV
2. a minimum number of EMC crystals hit > 0
3. LAT (an energy deposit shape variable) < 0.8

- *GoodNeutralLoose*

1. a minimum calorimeter energy of 30 MeV
2. a minimum number of EMC crystals hit > 0
3. no LAT cut

These lists serve as a pool of candidates for creating composite particles such as the K_S^0 or D^0, especially when combined with particle identification information.

5.5.2 Particle Identification

One of the most powerful tools in *BABAR* reconstruction is "particle identification." This refers to the ability to identify the species of a particle in addition to its momentum, which can be provided by the tracking system. Aside from muons which are a somewhat special case, particle identification typically exploits physical processes whose rate varies with the particle's velocity, thereby providing access to its identity. As expected, then, particles whose masses are near each other are often difficult to distinguish (provided they are similarly charged, etc.) Bear in mind that one can also deduce a particle's mass by separately measuring its energy and momentum. *BABAR* uses the following techniques to identify charged particles.

1. Momentum-dependent rate/angle of Cherenkov radiation for particles passing through specific radiator materials,

2. Mass-dependent rate of energy loss when passing through material (so-called "dE/dx" measurements,

3. Lateral shower shape of energy deposition in the EMC .

Particle identification is used in this analysis to identify the kaon used when reconstructing D^0 mesons, to identify the lepton candidates for the ℓ in $D^*\ell$, and in the inclusive determination of the $B_{\rm tag}$ flavor.

Without going into too many details, we discuss some of the primary methods of particle identification. Each class of particle type in *BABAR* has a set of associated selectors which will select candidate tracks from any of the aforementioned track lists in the event. Selected tracks are consistent with the particle identification hypothesis.

The kaon selector is based on forming combined likelihood ratios using quantities from three *BABAR* subdetectors: SVT, DCH, and DIRC. For low momenta, the discriminating quantity is the ionization loss, dE/dx, from the SVT and the DCH. For higher momenta, the Cherenkov angle θ_C and the number of photons, N_γ, detected by the DIRC are added. The calculated probability for dE/dx from the SVT and DCH, and for the DIRC Cherenkov angle θ_C, is Gaussian. For the number of observed photons in the DIRC, N_γ, the probability includes a Poisson term. The total

likelihood $\mathcal{L}$ for a given particle type is a product of the likelihoods from all three subdetectors. The average efficiency is 82% for the very tight selector and the average pion mis-identification rate is 2%. Pion identification is performed in the veto mode: pions are only required to not be identified by the tight kaon selector as kaons.

The variables used for electron identification are [43]:

1. The ratio of the energy E deposited in the EMC to the track momentum p, E/p. An ideal calorimeter will have $E/p = 1$ for an electron.

2. The ionization loss dE/dx .

3. A lateral energy distribution variable, LAT, defined as

$$\text{LAT} = \frac{\sum_{i=3}^{n} E_i r_i^2}{\sum_{i=3}^{n} E_i r_i^2 + E_1 r_0^2 + E_2 r_0^2}, \quad E_1 \geq E_2 \geq ... \geq E_n, \tag{5.5}$$

   where n is the number of crystals in a shower, r_0 is the average distance between two crystal front faces ($\approx$ 5 cm), E_i is the energy deposited in the i^{th} crystal, and r_i is the distance between the i^{th} crystal and the shower center. Electromagnetic showers are typically confined to only one or two crystals, so the corresponding LAT is smaller than that for hadronic showers.

4. The Zernike moment, A_{42} which measures the shower irregularities, which is akin to an expansion in angular moments, defined as

$$A_{42} = \sum_{r_i \leq R_0}^{4} \frac{E_i}{E} \cdot f_{42}\left(\frac{r_i}{R_0}\right) \cdot e^{-2i\phi_i}, \tag{5.6}$$

   where $R_0 = 15$ cm, and

$$f_{42}\left(\rho_i \equiv \frac{r_i}{R_0}\right) = 4\rho_i^4 - 3\rho_i^2 \ , \tag{5.7}$$

   with r_i as the distance from the crystal to the shower center. This variable is useful because the electromagnetic showers tend to be more regular than hadronic ones.

5. The measured DIRC Cherenkov angle, $\theta_{\rm C}$, of the e^+ or e^- candidate, which is compared to the expected Cherenkov angle, $\theta_{\rm C}^{\rm exp}$ calculated assuming an electron hypothesis.

The efficiency plateau for the very tight selector is 88.1% for tracks with momenta $0.5 < p < 2.0$ GeV/c, and the pion fake rate is 0.15%.

The following quantities are combined for muon identification: [44]

1. The energy deposited in the EMC, E, when available.

2. The number of IFR layers in a cluster matched to the muon candidate.

3. The measured and expected interaction lengths, λ and $\lambda_{\rm exp}$, respectively, traversed by the track in the entire detector. The expected interaction length is calculated assuming the muon hypothesis. *

4. The variable $\chi^2_{\rm trk}$, which measures how closely the hit IFR strips in a cluster match the track extrapolation. It is computed using the coordinates of each hit in an IFR layer and the coordinates of the intersection point of the track extrapolated to the same layer. $\chi^2_{\rm fit}$ is the χ^2 of a fit of all the hits in a given cluster to a 3$^{\rm rd}$ order polynomial. The discriminating power of both $\chi^2_{\rm trk}$ and $\chi^2_{\rm fit}$ comes from the fact that a pion or a kaon can interact in the material and produce a hadronic shower, leading to a more spread out distribution of hits, while a muon is expected to match the track extrapolation reasonably well.

5. The average multiplicity of hit strips per layer, $\overline{\rm m}$, and its standard deviation, $\sigma_{\rm m}$.

6. $\rm T_c$ describes the continuity of the track in the IFR, *i.e.,* the fraction of layers with recorded hits between the first and last observed interaction. This variable peaks at 1 for continuous tracks in an ideal detector (with a 100% detector efficiency and geometric acceptance).

*An "interaction length" is defined as the mean free path of a particle before undergoing inelastic collisions. This pair of variables (measured and expected interaction lengths) is really comparing the penetrating power of muons versus the stoppage of pions. For muons, we expect the expected number of interaction lengths to be nearly identical.

7. Very tight muons are required not to be selected by the tight kaon selector.

The plateau efficiency is 70.2%, with the average pion (kaon) mis-identification rate 2.3% (0.7%). The efficiency has been determined using Data control samples with $\mu\mu ee$ and $\mu\mu\gamma$ final states, and the mis-identification rate has been determined using pions from three-prong τ decays and $K^0_S \to \pi^+\pi^-$. (See Fig. 5.12.) Decays in flight constitute $\sim 2\%$ of the pion mis-identification rate.

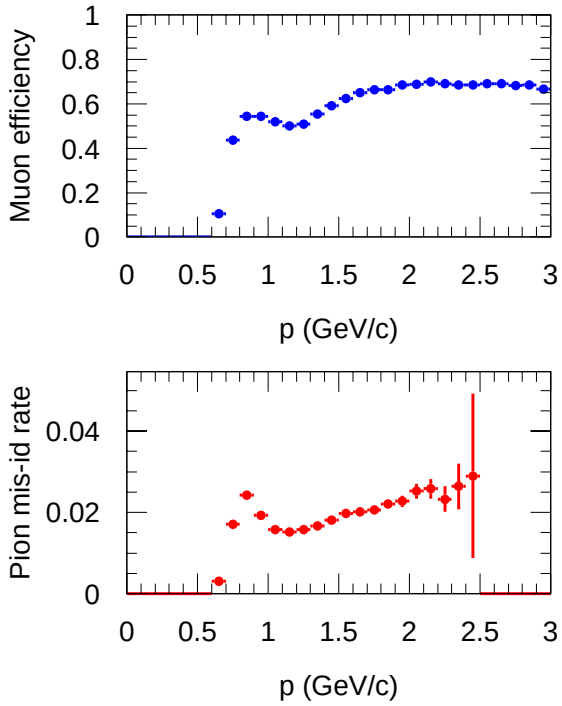


Figure 5.12: Upper: Muon identification efficiency as a function of lab momentum. Lower: Pion mis-identification rate as a function of lab momentum.

Chapter 6

Decay-time Measurement

The decay-time difference Δt between B decays is determined from the measured separation $\Delta z = z_{D^*\ell} - z_{\rm tag}$ along the z axis between the $D^*\ell$ vertex position ($z_{D^*\ell}$) and the flavor-tagging decay $B_{\rm tag}$ vertex position ($z_{\rm tag}$). This measured Δz is converted into Δt according to the relation $\Delta t = \Delta z/(\beta\gamma c)$ with the use of the known $\Upsilon(4S)$ boost, determined for each run.

In this section, we describe the method used to determine the decay-time difference Δt and its uncertainty $\sigma_{\Delta t}$ for each event. We define Δt as the difference between the decay time of the fully reconstructed $B \to D^* \ell \overline{\nu}_\ell$ candidate and the decay time of the inclusively-reconstructed tagging B:

$$\Delta t = t_{\rm rec} - t_{\rm tag} = t_{D^*\ell} - t_{\rm tag}.$$

For reference, we also define the residual as $\delta\Delta t = \Delta t_{\rm meas} - \Delta t_{\rm true}$, a signed quantity that compares the measured Δt to the true value. (This quantity is only explicitly available in the Monte Carlo datasets where the "truth" is known, of course.)

6.1 $B_{\rm rec}$ Vertex

The momentum and position vectors of the D^0, the soft π^- from $D^{*-} \to \overline{D}^0 \pi^-$, and ℓ candidates, and the run-averaged position of the e^+e^- interaction point (called the beam spot) in the plane transverse to the beam are used in a constrained fit to determine the best-fit position of the $D^*\ell$ vertex. The beam-spot constraint is of order $100\,\mu$m in the horizontal direction and $30\,\mu$m in the vertical direction, corresponding to the RMS size of the beam in the horizontal direction and the approximate transverse flight path of the B in the vertical direction. The constrained fit improves the resolution on $z_{D^*\ell}$ by about 30%. The RMS spread on the difference between the measured and true position of the $D^*\ell$ vertex as measured in Monte Carlo is about $80\,\mu$m (0.5 ps).

6.1.1 The $B_{D^*\ell}$ refitting algorithm

The $B_{D^*\ell}$ refitting algorithm deserves further comment. It uses a constrained vertex fitting technique that combines information from the ℓ, beamspot, and the D^* together with its daughters simultaneously to find the best vertex. The algorithm explicitly incorporates a flight distance of the daughter D^0, correctly considers the correlated errors in the fit, and provides an improved estimate of $m(D^*) - m(D^0)$ and

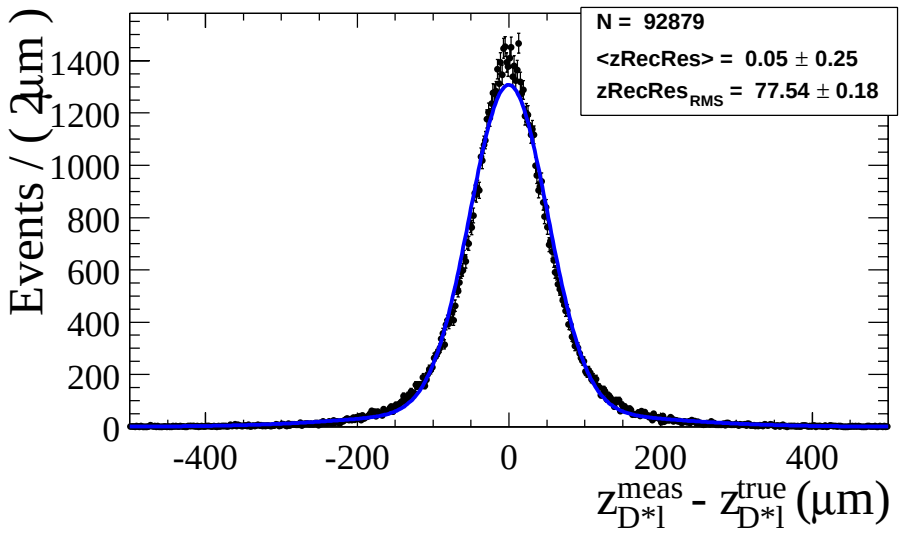


Figure 6.1: The difference between the reconstructed z position and the true z position $B_{D^*\ell}$ vertex, in μm, as measured in correctly reconstructed signal Monte Carlo. The overlaid curve is the fit resultfor a double Gaussian, yielding a resolution of 77 μm. (Recall the conversion 166 μm/ps.)

the $D^*\ell$ vertex of the B.* In short, when using this algorithm, the $m(D^*) - m(D^0)$ resolution improves by 40%, and the $z_{D^*\ell}$ vertex resolution improves by about 30%.

The soft π^- candidate is poorly measured, especially when only SVT-hits are available. Our knowledge of the soft π^- (particularly the polar angle) is also highly correlated with the D^0 momentum and D^0 decay length. Only by fitting for the D^0 vertex and D^* vertex simultaneously can the correlations be taken properly into account. Second, the D^0 vertex serves as a new data point for determining the momentum of the soft pion, which helps improve the entire fit. The result is the $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ vertex with improved results for the D^* vertex, D^0 vertex, and slow pion momentum which, in turn, provides a better measure of $m(D^*) - m(D^0)$.

*The source code for this strategy is primarily located in the **VtxFitter/VtxB0_DstarlnuAlgorithm** class. As initially prepared, the algorithm did not support the triply-constrained fit of the B vertex; we provided this extension and conclusively demonstrated the benefits. See [45] for further details.

We show next that including the beamspot constraint in the simultaneous fit improves the average vertex resolution, and directly contributes to an improvement in the Δt residuals, while introducing negligible correlations between the reco- and tag-side vertex estimates. For this study, we use $D^0 \to K^-\pi$ signal Monte Carlo and four different reco-side vertexing algorithms.

- (raw) - The basic vertexing of the D^* and ℓ candidates as used in the default composition sequence. [†] No explicit treatment of the daughter D^0 flight length.
- (D^*, ℓ) - Employs the **VtxB0_DstarlnuAlgorithm** which forces a separation between the (D^*,ℓ) and the D^0 decay vertex in the simultaneous refitting of the entire D^* tree and the ℓ.
- (D^*, bs) - Same as above, except that the lepton is not used in the vertex fit; rather the beamspot is applied as constraint on the D^* decay vertex.
- (D^*, ℓ, bs) - Same as above, except the lepton candidate is also considered in the vertex fit.

Naïvely, we expect the most constrained refitting, the (D^0, ℓ, bs), to perform the best on average. Because these algorithms adjust the D^* 4-momentum and position, we also might expect to see a dependence of the the $\delta m = m(D^*) - m(D^0)$ resolution on the choice of algorithm.

The technical implementation of this study involved producing new ntuples from the SP3 $D^0 \to K^-\pi$ signal Monte Carlo kanga files which contained four different estimates (corresponding to the four different reco-side vertexing algorithms) of each candidate's refitted δm, $z_{D^*\ell}$ and their corresponding error estimates. To maintain the integrity of the study, we only compare candidates that pass the final selection with all four algorithms. This is because of the slight difference observed in selection efficiency as shown in Table 6.1. (It is primarily the $\sigma_{\Delta t}$ cut and convergence of the refitting algorithm that causes this difference.)

The expected correlation between the δm resolution, the reco-side vertex algorithm, and angular features of the candidates is not supported in the dataset due to

[†]This is implemented by the **GeoKin** vertexing operator as used in **CompositionTools**.

Vtx Algorithm	Candidates Selected	% Change
raw	84780	0.0
(D^*, ℓ)	83347	-1.7
(D^*, bs)	81810	-3.5
(D^*, ℓ, bs)	85433	+0.8

Table 6.1: Crude comparison of the efficiency of the different reco-side vertexing algorithms for signal MC events

other challenging correlations. (We define the the resolution of δm for an algorithm to be the combined width of a double gaussian fit to the peak of the mass difference distribution.) For instance, one might expect the power of the (D^*, bs) algorithm to depend on the azimuthal angle of the soft pion (from the D^* decay) so that when the soft pion is in the azimuthally "horizontal direction" it will be most affected by the tight y-constraint imposed by the beamspot. Alternatively, we might think that the resolution (improvement with respect to the (raw) algorithm) of the (D^*, ℓ) algorithm might depend on the lab opening angle between the D^* and ℓ candidates. Due to a variety of correlations, however, these expectations are not borne out in data. For instance, the momentum of the soft pion appears to be empirically correlated with the center-of-mass opening angle of the candidate due to shared kinematics.

We conclude that the three refitting algorithms improve the resolution of the δm estimate, but that based on this alone, we cannot distinguish between them. Figures 6.3- 6.6 illustrate our results.

We ultimately wish to choose the the $B_{\rm rec}$ vertexing algorithm that optimizes our Δt resolution, and in that dimension, we can draw distinctions between the three refitting algorithms. For instance, the distribution of per-event estimates for this sample is different for each of the algorithms (see Figure 6.5). We now focus only on correctly reconstructed signal candidates in this Monte Carlo sample, and again the (D^*, ℓ, bs) algorithm is most efficient. We compare the measured z-position of the reco-side vertex to that from Monte Carlo truth association to form a z-residual on just the reco-side. This distribution ultimately describes the Δt performance of the different algorithms. Figure 6.6 shows these residual distributions, and the

combined (D^*, ℓ, bs) algorithm is singled out as the most accurate and precise estimate of the z-position of the $D^*\ell$ candidate. The pull distribution supports this conclusion, although are harder to use to distinguish the algorithms because each algorithm returns a very good estimate of its own error.

An important consideration is whether the application of the beamspot constraint separately to the $B_{\rm rec}$ vertex introduces an undesirable correlation with the tag-side vertex (which also uses the beamspot in the inclusive vertex fit). To evaluate this, we measure the correlation coefficient between measured $z_{D^*\ell}$ and $z_{\rm tag}$ and compute the combined $\sigma_{\Delta z}$ including the individual vertex resolutions. For the sample in question, we obtain the results shown in Table 6.2, which indicate that the change in correlation between the two vertices is negligible compared to the improvement in Δz resolution. (In this sample, we measure the average resolution of the $z_{\rm tag}$ estimate to be 130.9 μm.)

Reco Vtx Algorithm	$\rho(z_{D^*\ell}, z_{\rm tag})$	RMS $\sigma_{\Delta z}$ [μm]
raw	0.00276	155.4
(D^*, ℓ)	0.00211	155.5
(D^*, bs)	-0.00341	192.4
(D^*, ℓ, bs)	-0.00449	147.3

Table 6.2: Comparison of correlation between $z_{\rm tag}$ and $z_{D^*\ell}$ for the different reco-side vertexing algorithms, and resulting RMS of $\sigma_{\Delta z}$ using signal MC candidates.

A final cross-check uses the residual distribution in a fit to determine a triple gaussian resolution function. As expected, the (D^*, ℓ, bs) algorithm returns the best RMS resolution on Δt of 0.746 ps, a 33% improvement over the (raw) algorithm. In conclusion, we select the multiply constrained simultaneous refitting algorithm (D^*, ℓ, bs) for $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ candidates since it (a) offers the best $z_{D^*\ell}$ vertex resolution with negligible change in the correlation between the tag- and reco-side vertices, and (b) measurably improves the δm resolution.

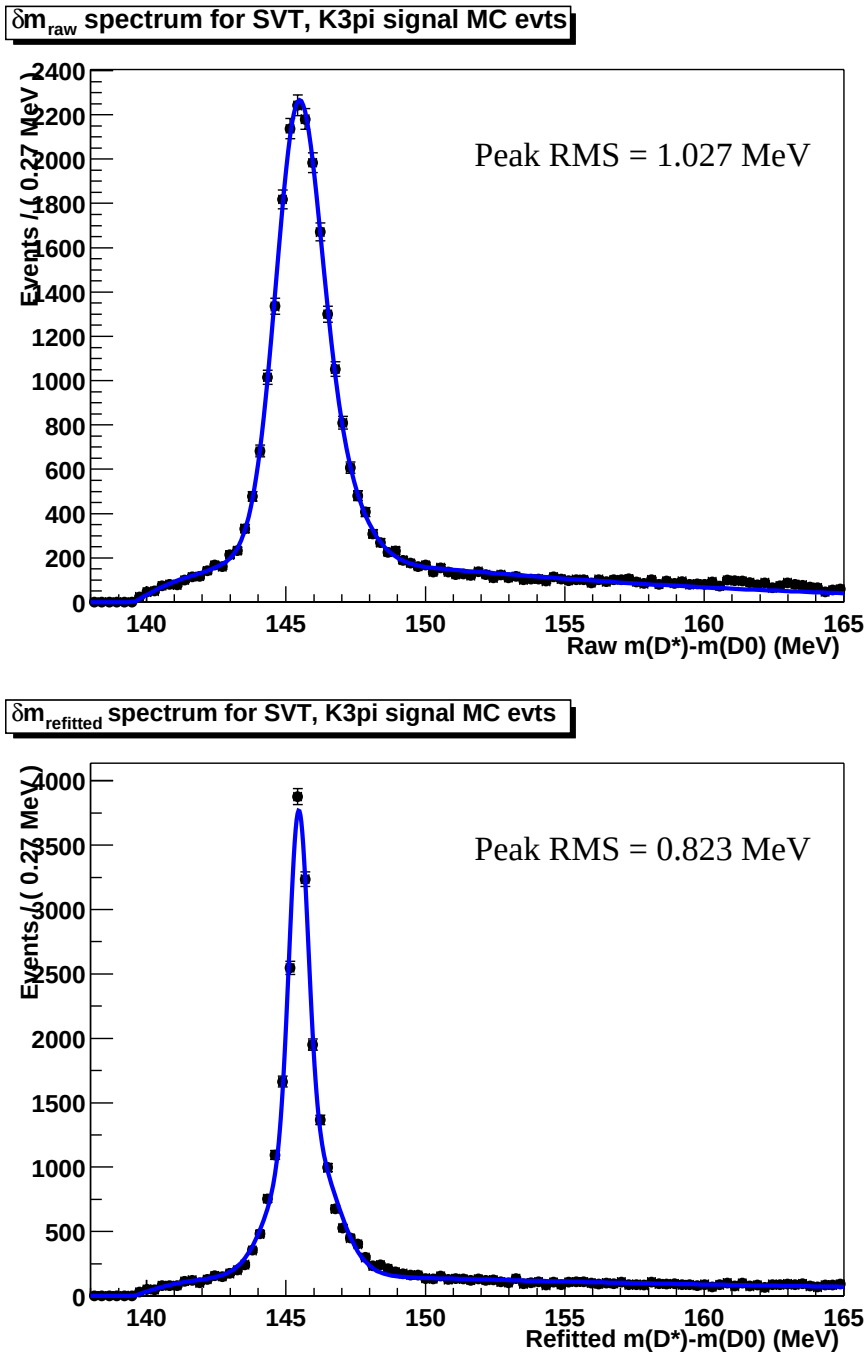


Figure 6.2: This pair of plots shows the improvement in δm resolution when the constrained refitting algorithm is used. The effective rms of the peak (where the signal is concentrated) is improved by more than 40%.

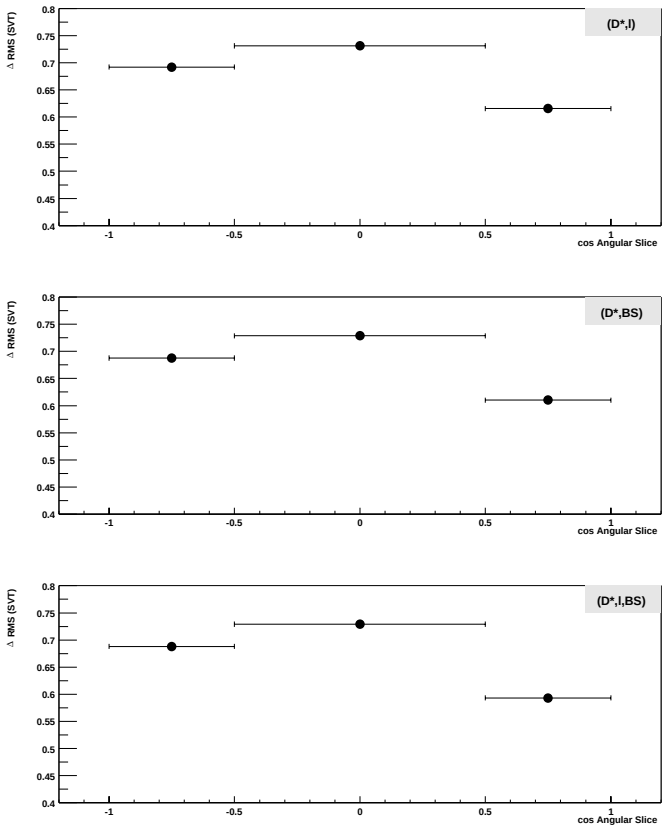


Figure 6.3: Each of these three figures shows the improvement in δm resolution for the reco-side vertexing algorithm shown in the upper right corner as compared to the (raw) algorithm. (The sample of candidates in each plot is identical, except for the reco-side vertex quantities, and is drawn from signal MC.) The x-axis indexes the slice in angular region of the lab quantity, $\cos\theta_{D^*,\ell}$: {[-1,-0.5],(-0.5,0.5],(0.5,1],}. If indeed the resolution of the (D^*,ℓ) vertexing algorithm were dependent on the lab opening angle of the D^*-ℓ candidate, the upper plot would show the largest improvement in the center bin, while the other two plots would exhibit little structure.

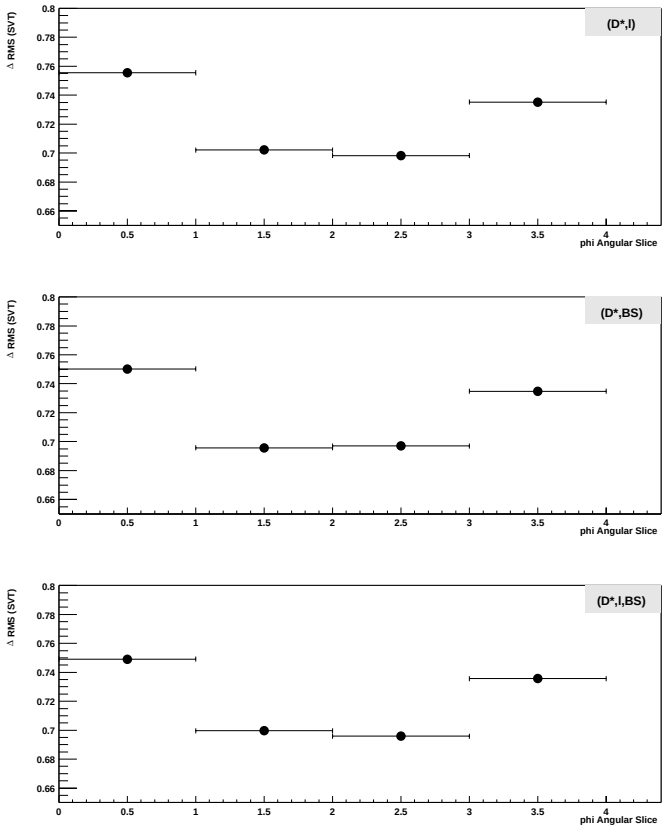


Figure 6.4: Each of these three figures shows the improvement in δm resolution for the reco-side vertexing algorithm shown in the upper right corner as compared to the (raw) algorithm. (The sample of candidates in each plot is identical, except for the reco-side vertex quantities, and is drawn from signal MC.) The x-axis indexes the slice in angular region of the lab quantity, $\phi_{\pi_{soft}}$: $\{[-\pi/2,\pi/2),[\pi/2,3\pi/2),[3\pi/2,-3\pi/2),$ $[-3\pi/2,-\pi/2)\}$ or more simply {right, up, left, down}. If indeed the resolution of the (D^*, bs) vertexing algorithm were dependent on the azimuthal angle of the D^*-ℓ candidate's soft pion, the middle plot would show the largest improvement in the first and third bins, while the other two plots would exhibit little structure.

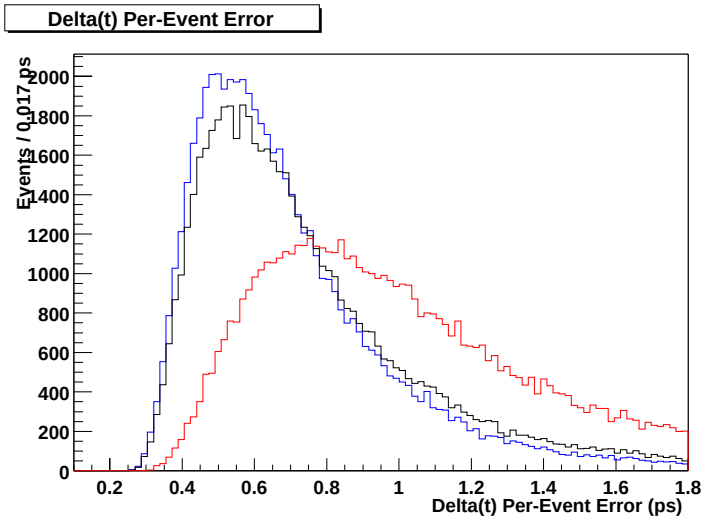


Figure 6.5: Distribution of $\sigma_{\Delta t}$ for the signal Monte Carlo sample described in the text. The blue curve (upper) corresponds to the distribution as reported by (D^*, ℓ, bs) algorithm, the black (middle) to the (D^*, ℓ), and the red curve (lower) to the (D^*, bs) algorithm.

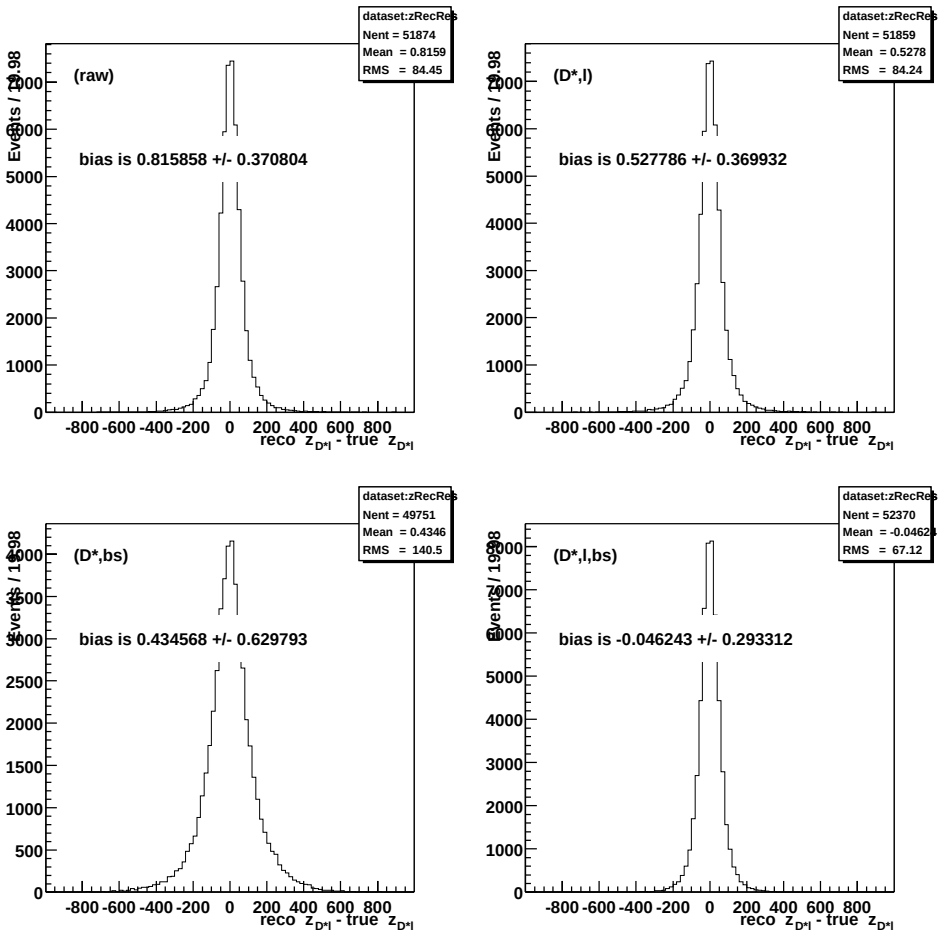


Figure 6.6: Residual distributions for the z-position of correctly reconstructed $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ candidates in the signal Monte Carlo sample described in the text. Each of the four plots corresponds to a different reco-side vertexing algorithm used to estimate the z-vertex. The bias reported on each plot is the mean of the distribution, and the error is computed from the RMS of the distribution.

6.2 $B_{\rm tag}$ Vertex

To maintain high efficiency, we inclusively determine the vertex of the other B, $B_{\rm tag}$. We determine the position of the $B_{\rm tag}$ vertex from all charged tracks in the event except the daughters of the $D^*\ell$ candidate, using K^0_S and Λ candidates in place of their daughter tracks, and excluding tracks that are consistent with being due to photon conversions. The same beam-spot constraint applied to the $B_{D^*\ell}$ vertex is also applied to the $B_{\rm tag}$ vertex. To reduce the influence of charm decay products, which bias the determination of the vertex position, tracks with a large contribution to the χ^2 of the vertex fit are iteratively removed until those remaining have a reasonable fit probability or only one track remains. The RMS spread on the difference between the measured and true position of the $B_{\rm tag}$ vertex is about 160 μm (1.0 ps) as measured in signal Monte Carlo. For implementation details of the $B_{\rm tag}$ vertex fit, please see Ref. [46] [‡].

In Fig. 6.7 we show the residual distribution from signal Monte Carlo events for the $B_{\rm tag}$ z vertex. We observe a positive bias in the residual of 34 μm, which corresponds to a systematic overestimate of $z_{\rm tag}$. This is because the vertex-finding algorithm is easily biased by $B_{\rm tag}$ daughters which leave behind slightly detached decay vertices. (See the cartoon in Fig. 6.8.) These daughters are typically D mesons or other charm-containing particles, which have significant lifetimes in the lab, on the order of 1 ps. Because Δt is linearly dependent on Δz, the overall Δt resolution will be dominated by the poorer z resolution of the tag vertex position.

6.3 Extracting the Decay-Time Difference

The two B's are boosted in the z direction because of the asymmetric collision with $\beta_z\gamma \approx 0.55$. The momentum in the boost direction (6 GeV/c) is much larger than the motion of the B^0 in the $\Upsilon(4S)$ rest frame ($p^*_B = \frac{1}{2}\sqrt{s - 4m^2_{B^0}} \simeq 320$ MeV/c).

[‡]We use the algorithm **VtxTagBtaSelFit** described in that document.

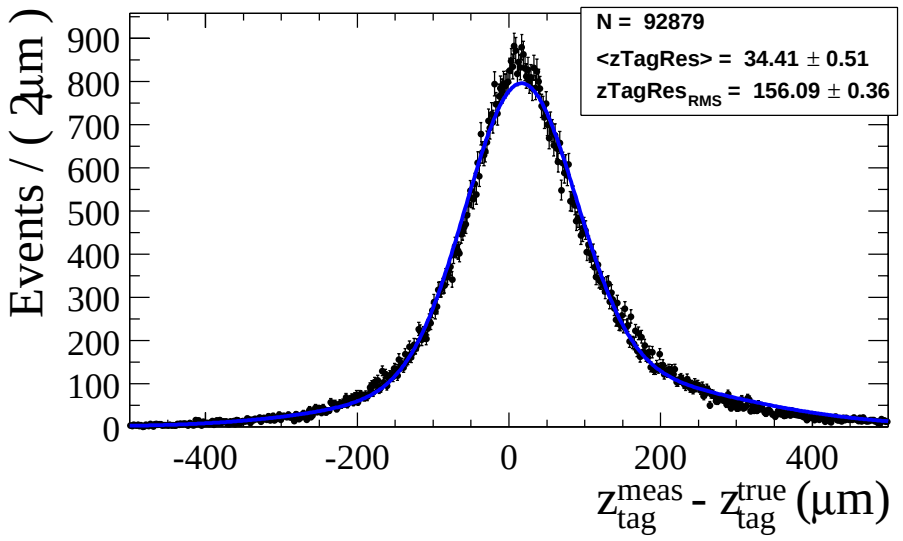


Figure 6.7: The difference between the reconstructed z position and the true z position of the $B_{\rm tag}$ vertex, in μm, as measured in correctly reconstructed signal Monte Carlo. (Recall the conversion 166 μm/ps.) The overlaid curve is the fit result for a double gaussian. (Refer to Fig. 6.6 for the corresponding figure for $B_{\rm rec}$.) Note the significant bias and width of this residual distribution.

Δt is defined as the (proper) decay time difference of the two B mesons:

$$\Delta t = t_1 - t_2 = m_B \left(\frac{z_1}{p_1^z} - \frac{z_2}{p_2^z} \right) . \tag{6.1}$$

To a good approximation, though, the two B's are traveling purely in the z direction with a lab-velocity determined by the boost. The time difference between the two decays is then given by

$$\begin{aligned} \Delta z &= z_{\rm rec} - z_{\rm tag} \\ \Delta t &= \Delta z / c\beta_z\gamma \end{aligned} \tag{6.2}$$

where γ is the boost factor for the $\Upsilon(4S)$ in the lab frame and β_z its velocity projected on the z-axis of the detector.

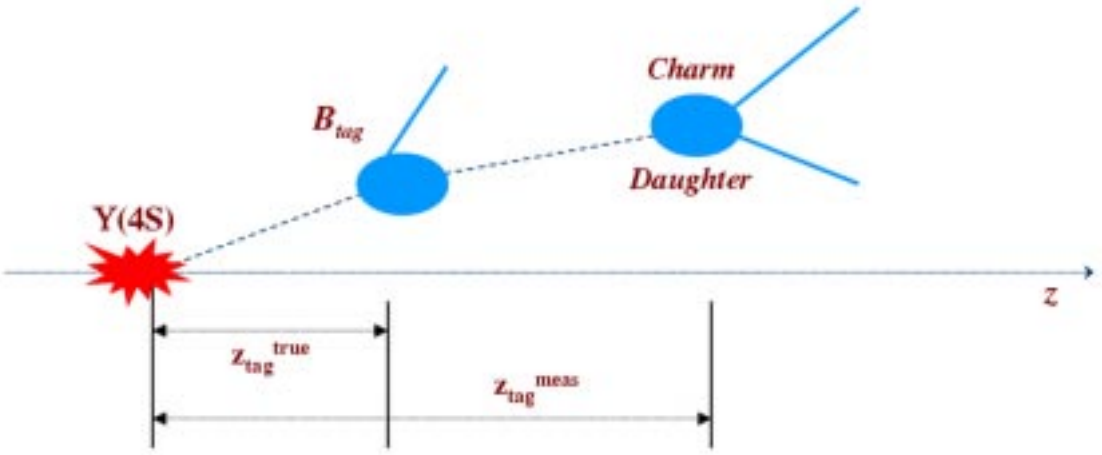


Figure 6.8: Cartoon description of the mechanism which biases the estimated $z_{\rm tag}$ vertex position high. The daughters of the $B_{\rm tag}$ have tracks which point back to a detached decay vertex; the vertex-fitting algorithm can mistakenly combine tracks from these vertices when estimating the $B_{\rm tag}$ vertex. (This cartoon is exaggerated, of course.)

The average separation between the two B's is given by the average lifetime times velocity:

$$\langle \Delta z \rangle = \langle \beta\gamma c\tau_{B^0} \rangle \quad \approx \; 250\mu m \tag{6.3}$$

The boost from the center-of-mass to the lab frame due to the asymmetric-energy collision is important, because without it, the B's would be travelling in the lab frame with only their momentum from the $\Upsilon(4S)$ decay:

$$\beta\gamma = \frac{p_B E_B}{m_B^2} = \frac{E_B\sqrt{E_B^2 - m_B^2}}{m_B^2} \approx 0.0647 \tag{6.4}$$

In this case, each B would travel $\approx 30\,\mu$m, for a total separation of only $60\,\mu$m. The experimental resolution is about $130\,\mu$m so the asymmetric boost is critical to the detector's ability to resolve the time difference between the B decays. The ability to accurately measure Δt depends critically on how well the separate $B_{\rm tag}$ and $B_{\rm rec}$

decay vertices can be determined.

6.3.1 Boost Approximation

Equation 6.2 ignores three effects. The most important is the finite momentum of the B meson in the $\Upsilon(4S)$ rest frame, mentioned previously and quantified below. Secondly, the $\Upsilon(4S)$ momentum is smeared due to the momentum spread of the colliding beams. This smearing is about 6 MeV/c, which corresponds to 0.1% smearing of $\beta\gamma$. The factor of β_z accounts for the third observation, the fact that the direction of the colliding beams is tilted by $\theta \approx 20$ mrad with respect to the detector z-axis. The boost along the z-axis is reduced by 1.2 MeV (0.02%) due to this effect. There is also an effective boost of about 118 MeV/c 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 [46]

$$\Delta z = c\gamma\beta\gamma^*(t_1 - t_2) + c\gamma\gamma^*\beta^* \cos\theta^*(t_1 + t_2)\,, \tag{6.5}$$

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. 6.2 is only biased by a factor γ^*, which is approximately 1.002, since $\langle\cos\theta^*\rangle = 0$.

In general we don't know $t_1 + t_2$ on an 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. 6.5, we obtain

$$\Delta z = c\gamma\beta\gamma^*\Delta t + c\gamma\gamma^*\beta^* \cos\theta^*(\tau_B + |\Delta t|)\,. \tag{6.6}$$

which is known as the "average τ_B approximation."

In this analysis, we cannot measure the B decay direction because we haven't fully

reconstructed the neutrino. We take $\Delta z = c\gamma\beta\gamma^*\Delta t$ as a further approximation. The effect of the second term in Eq. 6.5 is to smear the exponential distribution of Δz. The RMS of the error due to this final approximation is

$$c\gamma\gamma^*\beta^*\sqrt{\langle\cos^2\theta^*\rangle\,\langle(t_1+t_2)^2\rangle} \simeq 21\cdot\sqrt{6}\,\tau_B\sqrt{\langle\cos^2\theta^*\rangle}\ (\mu\mathrm{m}) \tag{6.7}$$

where we have used $\beta^* \simeq 0.0606$, $\gamma \simeq 1.144$ and the expectation value of $\langle(t_1+t_2)^2\rangle = 6\tau_B^2$. The angular distribution for $\Upsilon(4S) \rightarrow\ B\overline{B}$ 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. 6.7. Therefore the smearing introduced by using the approximation $\Delta z/(\gamma\beta\gamma^* c)$ can be written as a function of $|\Delta t|$:

$$\frac{\beta^*}{\sqrt{5}\beta}\sqrt{\tau_B^2+(\tau_B+|\Delta t|)^2} \simeq 0.0557\sqrt{\tau_B^2+(\tau_B+|\Delta t|)^2}\,. \tag{6.8}$$

We can illustrate this effect by plotting the RMS of $\Delta z_{\mathrm{truth}}/(\gamma\beta\gamma^* c) - \Delta t_{\mathrm{truth}}$ in bins of $|\Delta t_{\mathrm{truth}}|$. Figure 6.9 shows this plot for 80k signal MC events that pass all our final signal cuts, and a fit to a function of the form given in Eq. 6.8: $p_0\cdot\sqrt{p_1^2+(p_1+|\Delta t|)^2}$. The quality of the fit result suggests that we have correctly parameterized the dependence.

This smearing effect is even visible at the level of reconstructed Δt. Consider Fig. 6.10 which shows the RMS of several Δt-related quantities as a function of $|\Delta t_{\mathrm{true}}|$. The top plot in the figure shows the RMS of the residual; there is clear evidence of structure, reminding us that the real residual of the Δt_{meas} does depend on Δt_{true} because of the boost approximation. The middle plot shows the effect of the boost approximation in unadulterated form: we compare the Δt estimate from the

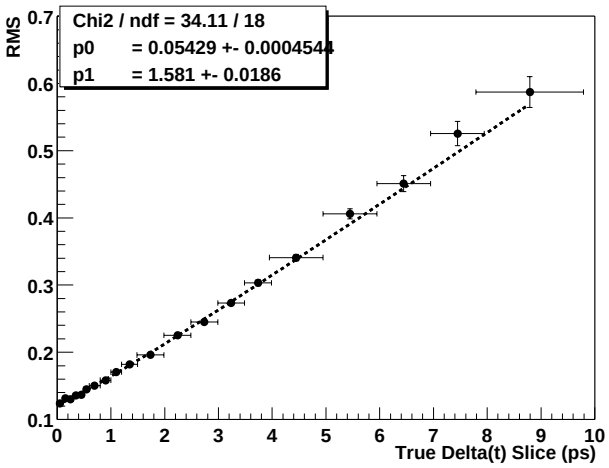


Figure 6.9: RMS of the distribution of residuals $\Delta z_{\rm truth}/(\gamma\beta\gamma^* c) - \Delta t_{\rm truth}$ in bins of $|\Delta t_{\rm truth}|$ for 80k signal MC events that pass all event selection criteria. A linear fit to the function $p_0 \cdot \sqrt{p_1^2 + (p_1 + |\Delta t|)^2}$ is superimposed. The $\Delta z \to \Delta t$ conversion factor used here is 166.87μm/ps.

boost-approximation using $z_{\rm true}$ ($\Delta t_{boost-approx}$) with the $\Delta t_{\rm true}$ and show the RMS of that difference. The bottom plot is the bin-by-bin difference in quadrature between the top and middle histograms. The observed "V" structure is almost entirely absent, suggesting that the boost approximation is the major source of the dependence of resolution on $\Delta t_{\rm true}$.

Finally, we can observe some of this effect in the difference between measuring the lifetime from a fit to $\Delta t_{\rm true}$ in signal Monte Carlo, and from a fit to $\Delta t_{boost-approx} = \Delta z_{\rm truth}/(\gamma\beta\gamma^* c)$ using the Monte Carlo $z_{\rm true}$ information. This (partially) isolates any event selection bias from any bias introduced by the boost approximation. The shift in fit results for the two techniques is about 0.0047 ps for τ_{B^0} and 0.00015ps^{-1} for Δm_d, as shown in Fig. 6.11. In the final analysis with Data, the introduction of a Δt resolution function (described in the next section) absorbs these effects and makes them negligible compared to the precision expected from the full dataset.

6.3.2 Per-event error

We calculate the uncertainty on Δz due to uncertainties on the track parameters from the 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, errors in associating tracks with the B vertex, or the effects of misalignment within and between the tracking devices. The calculated uncertainties will also be incorrect if our assumptions for the amount of material in the tracking detectors or the beam-spot size or position are inaccurate. We use parameters in the Δt resolution model, measured with data, to account for uncertainties and biases introduced by these effects. We assume, for this analysis, that the per-event error estimate is completely uncorrelated with any other event characteristics.

The left-hand plot in Fig. 6.13 shows the distribution of $\sigma_{\Delta t}$ values observed in signal MC. The right-hand plot in Fig. 6.13 shows that the pull distribution observed in signal MC 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%.

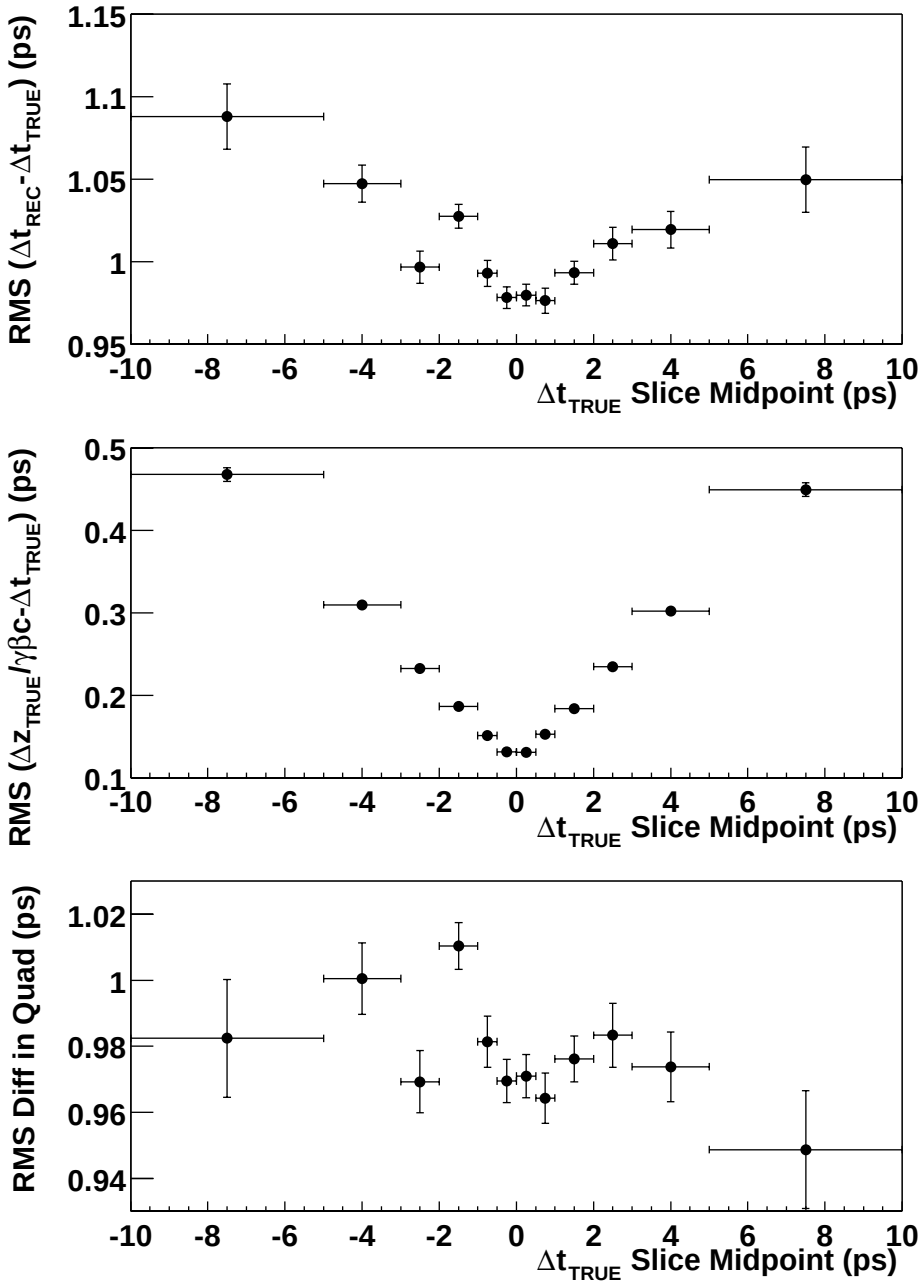


Figure 6.10: Top: the RMS of the residual $\delta\Delta t$ in bins of $\Delta t_{\rm true}$ for Monte Carlo simulation. Middle: same plot with Δt replaced by $\Delta z_{true}/\gamma\beta\gamma^* c$ in the calculation of the RMS of the residual. Bottom: bin-by-bin difference in quadrature between the top two histograms. Note the factor of 10 difference in the vertical scales of the top and bottom histograms.

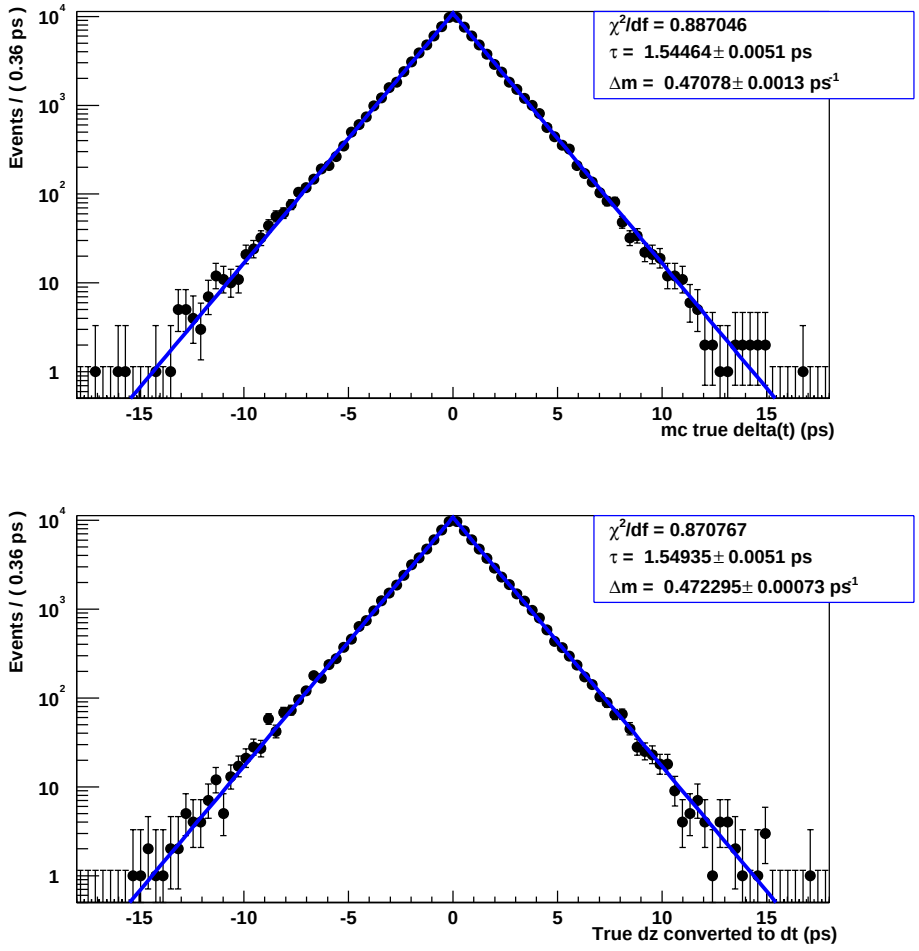


Figure 6.11: Fits for B^0 lifetime (and mixing) using two-types of truth information for signal Monte Carlo events. Upper: $\Delta t_{\rm true}$, Lower: computing Δt using Δz_{true} and the boost approximation.

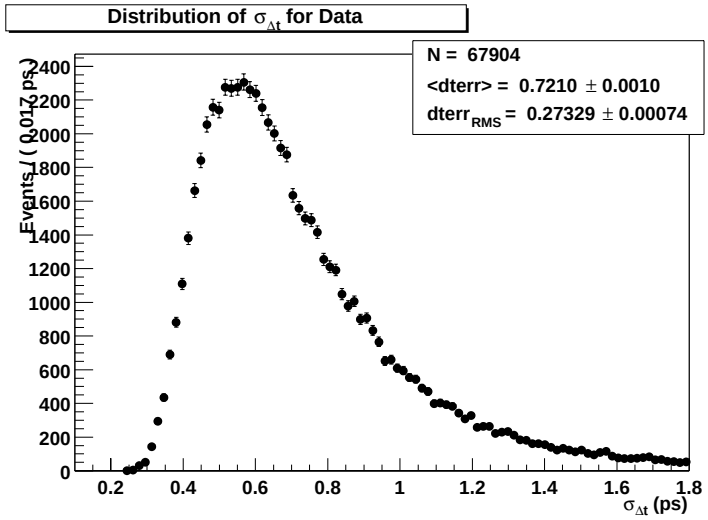


Figure 6.12: The distribution of $\sigma_{\Delta t}$ for Data. The minimum value is set by the single track error resolution.

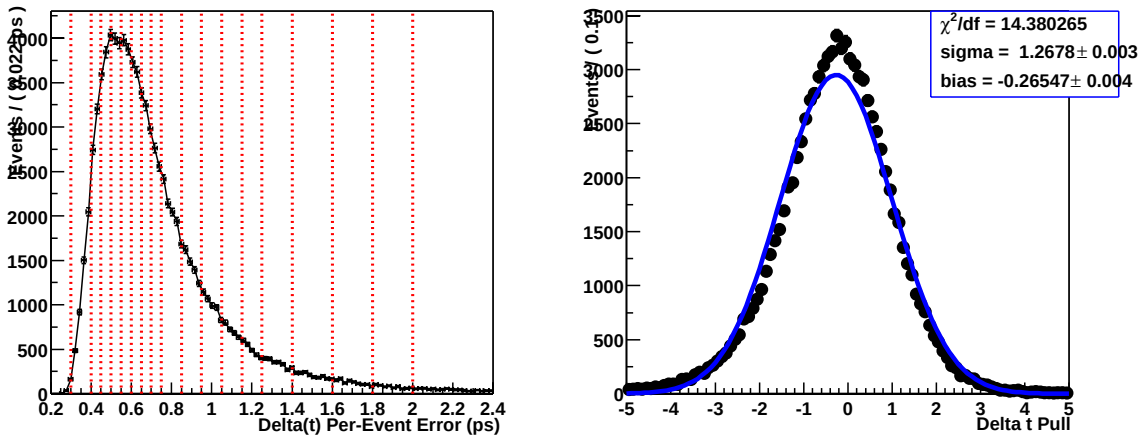


Figure 6.13: Distribution of calculated vertex errors $\sigma_{\Delta t}$ (left side) and pulls (right side) obtained from correctly reconstructed signal MC. Vertical dashed lines in the $\sigma_{\Delta t}$ distribution indicate the slices used to select subsamples with approximately equal calculated errors. The solid curve superimposed on the pull distribution is the result of a single Gaussian fit.

6.4 Resolution

6.4.1 Nature of the Resolution Model

As we have discussed, the variable we measure is $\Delta t_{\rm meas}$, while the true physics of lifetime and mixing involves the variable $\Delta t_{\rm true}$. We construct a convolution integral (the resolution function $\mathcal{R}$) which connects the two correlated spaces in terms of the residual $\delta\Delta t$:

$$P(\Delta t_{\rm meas}) = \int G_{sig}(\Delta t_{\rm true}) \cdot \mathcal{R}(\cdot\sqcup_{\rm meas} - \cdot\sqcup_{\rm true}) \lceil (\cdot\sqcup_{\rm true})$$

A resolution model $\mathcal{R}$, then, 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 are assuming, then, that resolution is independent of $\Delta t_{\rm meas}$, but expect a strong correlation with $\sigma_{\Delta t}$; *i.e.,* $\mathcal{R} = \mathcal{R}(\delta\Delta t, \sigma_{\Delta t})$.

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 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

$$G(x;\, x_0, \sigma) \equiv \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \exp\bigl(-(x - x_0)^2/(2\sigma)^2\bigr) \ . \tag{6.9}$$

6.4.2 Dependence on $\sigma_{\Delta t}$

Based on the grouping of events by common $\sigma_{\Delta t}$ as suggested by the vertical lines in Fig. 6.13, we measure the mean and RMS of the residual distribution in each $\sigma_{\Delta t}$ bin. We observe a clear dependence of both the mean and the RMS of the residual distribution with $\sigma_{\Delta t}$, as shown in Fig. 6.14. The RMS of the residuals scales with the calculated error as expected — this just reflects the desirable feature that the estimated Δt error is highly correlated with the true error (the residual). The slope

of the correlation is not quite unity, so we introduce a scale parameter which multiplies $\sigma_{\Delta t}$ to compensate.

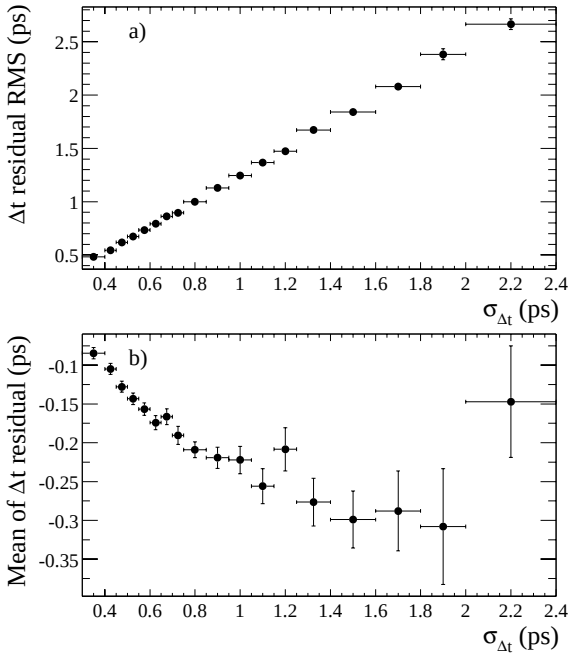


Figure 6.14: RMS (a) and mean (b) of the residual distributions obtained from correctly reconstructed signal MC, in slices of $\sigma_{\Delta t}$.

The dependence of the mean of the residual distribution on $\sigma_{\Delta t}$ merits some explanation. The correlation is due to the fact that, in B decays, 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 calculated uncertainty on the vertex position in z for the $B_{\rm tag}$ candidate. In addition, the flight length of the D in the z direction is correlated with its flight direction. Therefore the bias in the measured $B_{\rm tag}$ position due to inclusion of D decay products in the vertex estimate is correlated with the D flight direction. Taking into account both of these correlations, we conclude that D mesons that have a flight direction perpendicular to the z axis in the laboratory frame will have the best z resolution and will introduce the least bias in a measurement of $z_{\rm tag}$, while D mesons that travel forward in the lab frame will introduce a larger bias in the corresponding measurement. Fig. 6.15

illustrates this argument geometrically. Fig. 6.16 uses signal Monte Carlo truth to isolate the source of the $z_{\rm tag}$ bias as longer-lived particles containing c-quarks. The dependence turns over at larger $\sigma_{\Delta t}$ due to competing effects such as the dependence of the D momentum on the lab polar angle. [47]

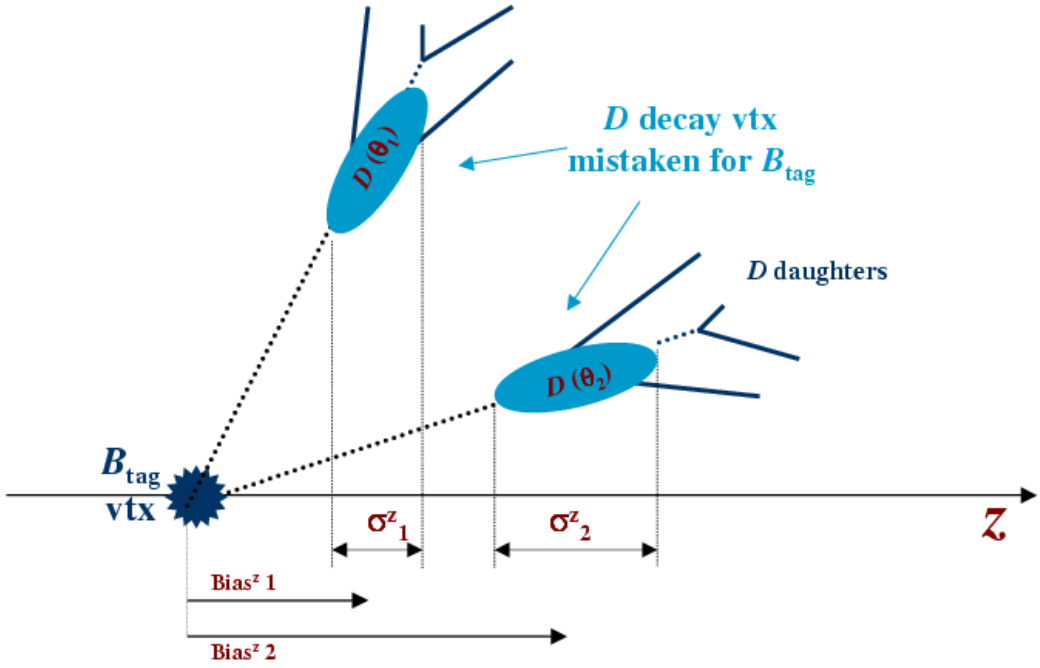


Figure 6.15: Diagram of the mechanism by which longer-lived daughters of the $B_{\rm tag}$ meson introduce a correlation between vertex error estimate $\sigma_{\Delta t}$ and the vertex estimate $z_{\rm tag}$ itself. Consider two different situations where the $B_{\rm tag}$ charm daughter (here, a D meson) is emitted at two different angles in the lab frame: θ_1 and θ_2. We observe that the projection of the D vertex error ellipse on the z axis introduces a correlation with the bias on $z_{\rm tag}$ as well.

6.4.3 Choice of Resolution Model

We summarize the observations about Δt resolution and motivate our choice for the final resolution model:

- Although the true resolution function is not independent of $\Delta t_{\rm true}$, we shall

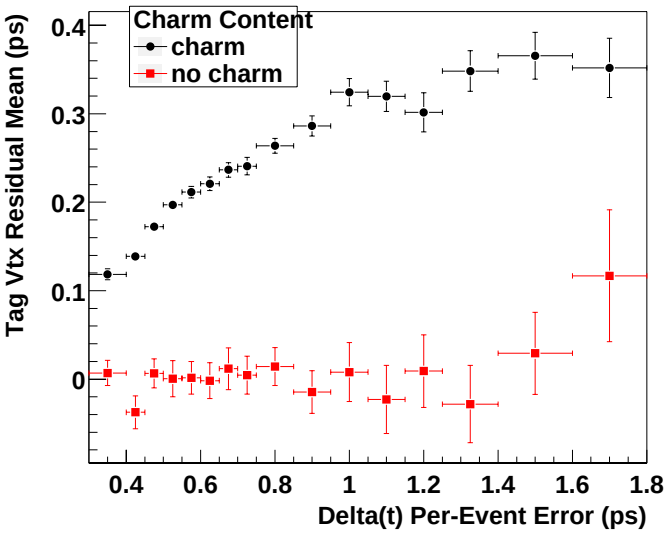


Figure 6.16: Mean residual of $z_{\rm tag}$ (scaled by $\gamma\beta c$) in slices of $\sigma_{\Delta t}$ for signal MC events in which at least one track used for tag-vertexing is produced from charm-containing or τ-containing $B_{\rm tag}$ daughters (black) and for events in which no track used for tag-vertexing is produced from charm-tau (red). We observe that the correlation vanishes for non-charm/τ-containing $B_{\rm tag}$-vertex events.

assume it is

- The per-event error estimate $\sigma_{\Delta t}$ for Δt is typically underestimated, but is proportional to the residual $\delta\Delta t = \Delta t_{\rm meas} - \Delta t_{\rm true}$
- Because $z_{\rm tag}$ is typically overestimated, Δz is biased low
- The residual exhibits a correlation with the per-event error such that the mean and RMS of the residual distribution are positively correlated with $\sigma_{\Delta t}$

In order to accommodate these effects, we consider two different models for the "core" of the resolution function (loosely defined as $|\delta\Delta t| < 5\,\sigma_{\Delta t}$, as used in Fig. 6.13. The first core model consists of a double Gaussian whose biases and widths are scaled by the measured $\sigma_{\Delta t}$:

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

This model is similar to the one used for the CP [48] and hadronic mixing [49] analyses, except for the scaling of the bias that is clearly present in signal Monte Carlo (Fig. 6.14).

The second model is the sum of a single Gaussian and the same Gaussian convolved with a one-sided ($\delta\Delta t < 0$) exponential $E(\delta\Delta t;\, \kappa)$:

$$\begin{aligned}\mathcal{R}_{\rm GExp}(\delta\Delta t, \sigma_{\Delta t};\, \kappa, s, f) \equiv\, & f \cdot G(\delta\Delta t;\, 0, s\,\sigma_{\Delta t}) + \\ & (1-f) \cdot \frac{1}{2\,\kappa\sigma_{\Delta t}} \int_{-\infty}^{0} du\, \exp\bigl(+u/(\kappa\,\sigma_{\Delta t})\bigr) \cdot G(u - \delta\Delta t;\, 0, s\,\sigma_{\Delta t})\ . \quad (6.10)\end{aligned}$$

The integral in this definition can be performed analytically to give

$$\begin{aligned}\mathcal{R}_{\rm GExp}(\delta\Delta t, \sigma_{\Delta t};\, \kappa, s, f) = & \, f\, G(\delta\Delta t;\, 0, s\,\sigma_{\Delta t}) + \\ & (1-f)\,\frac{1}{2\,\kappa\sigma_{\Delta t}}\, \exp\left(\frac{s^2 + 2\,\delta\Delta t\cdot\kappa/\sigma_{\Delta t}}{2\kappa^2}\right)\, \mathrm{erfc}\left(\frac{s^2 + \delta\Delta t\cdot\kappa/\sigma_{\Delta t}}{\sqrt{2}\,s\,\kappa}\right)\ .\end{aligned} \quad (6.11)$$

This model was also used in the hadronic lifetime analysis[50].

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

$$\langle\delta\Delta t\rangle_{\rm G+G} = \bigl(f b_1 + (1-f) b_2\bigr) \cdot \langle\sigma_{\Delta t}\rangle \quad (6.12)$$

$$\langle(\delta\Delta t)^2\rangle_{\rm G+G} = \bigl(f(s_1^2 + b_1^2) + (1-f)(s_2^2 + b_2^2)\bigr) \cdot \langle(\sigma_{\Delta t})^2\rangle\ , \quad (6.13)$$

from which we calculate the G+G RMS squared

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

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

$$\langle \delta\Delta t \rangle_{\rm GExp} = -(1-f)\,\kappa \cdot \langle \sigma_{\Delta t} \rangle \tag{6.14}$$

$$\langle (\delta\Delta t)^2 \rangle_{\rm GExp} = \left(s^2 + 2\,(1-f)\,\kappa\right) \cdot \langle (\sigma_{\Delta t})^2 \rangle \;, \tag{6.15}$$

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

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

and the predicted slopes for the RMS are

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

Figure 6.17 shows the results of G+G and GExp fits to each of the slices shown in Figure 6.13. We observe the expected scaling of the fit parameters at least out to 1.8 ps. Since events with calculated uncertainties $\sigma_{\Delta t} > 1.8$ ps represent only about 1.5% of our sample and contribute even less to our statistical sensitivity, we will apply an additional vertex quality requirement of $\sigma_{\Delta t} < 1.8$ ps. [§]

Because of the strong influence of the charm daughters on the $B_{\rm tag}$ vertex, we prefer the GExp-based resolution model as it explicitly parameterizes this effect in the $\Delta t_{\rm true}$ to $\Delta t_{\rm meas}$ transformation with the "Exp" convolution. Additionally, it has been shown [51] that the G+G-based resolution model can be biased when fitting for the lifetime in a sample of limited size.

[§]This is tighter than the cut at 2.4 ps applied in the *CP* and hadronic mixing analyses.

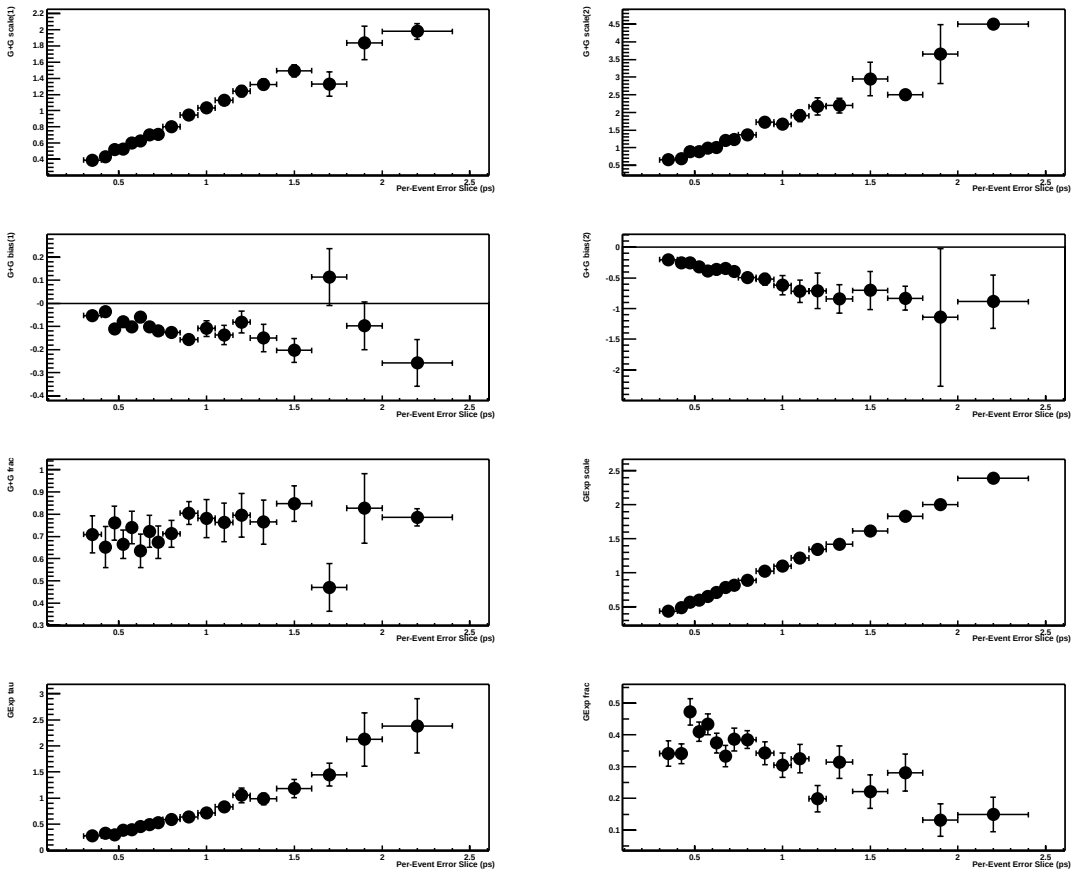


Figure 6.17: Parameters obtained from G+G and GExp fits to the residual distributions in slices of $\sigma_{\Delta t}$ for signal MC. The first 5 plots are the G+G parameters s_1, s_2, b_1, b_2, and f. The last 3 plots are the GExp parameters s, κ, and f.

Fig. 6.18 illustrates the final form of resolution model as determined by the best fit to well-reconstructed signal Monte Carlo. Each of the components of the GExp model are explicitly shown as well. Note the overall asymmetry and bias, which accommodates the residual distribution behavior. (See Section 11.2.1 for more details.)

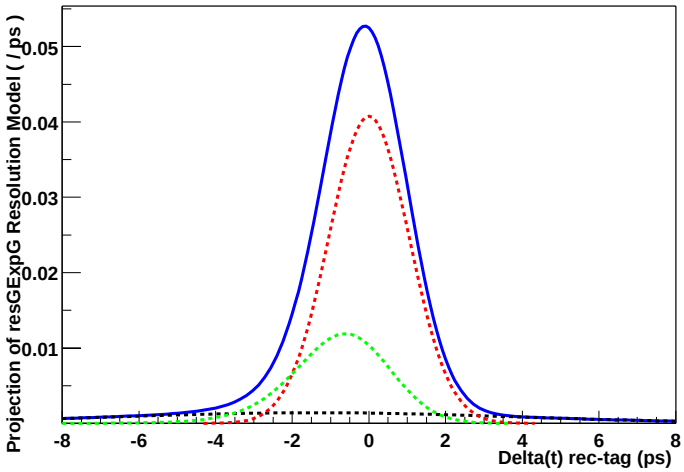


Figure 6.18: A projection of the GExp resolution model onto residual space for signal Monte Carlo. The outer curve in solid blue is the sum of all components; the inner red dashed curve is the wide Gaussian contribution; the lower green dotted curve is the G⊗Exp core mode; the smallest dashed black curve is the outlier term, but the fraction has been multiplied by 10 in order to make it visible.

6.5 Outliers

In addition to the core of the resolution function, we model the contribution of "outliers": loosely, 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 measured.) Outliers represent about 1% of the candidates selected in signal Monte Carlo. Outliers are a critical concern when measuring the lifetime because they have large residuals but small $\sigma_{\Delta t}$, making them hard to identify. The outlier population is expected to be uniform in $\Delta t_{\rm meas}$, and since signal events are distributed according to a negative exponential (*i.e.,* lifetime), the outliers are fractionally dominant at large $|\Delta t|$. If we mistakenly assume that the outliers events at large Δt are actually well-measured signal events, we grossly overestimate the lifetime in order to accommodate them.

Using Monte Carlo truth, we can continue the definition of outlier events as those events with $pull = \delta\Delta t/\sigma_{\Delta t}$ greater than 5 units. Keep in mind, however, that this definition is arbitrary and we employ it only as a means of isolating events that

have high potential of being outliers. From truth-matching, we learn that outlier candidates are usually one of three types of events:

- Lepton from other B: The D^* is correctly reconstructed but the lepton candidate is from the other B.
- Both B mesons are misreconstructed: The event is completely misreconstructed. A mix of charged tracks from both B candidates is used in each vertex.
- Other B misreconstructed: The $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ candidate is reconstructed correctly but the $B_{\rm tag}$ is misreconstructed either because the daughter charm vertex is found, or a mixture of primary B tracks and daughter charm tracks is used in the vertex.

The first two sources of outliers will tend to decrease $|\Delta t|$ but will not bias the average value of Δt. The third source will lead to a negative bias in Δt since $\Delta t = t_{D^*\ell} - t_{\rm tag}$ and $t_{\rm tag}$ will be biased in a positive direction.

In the top plots in Fig. 6.19, we show the distribution of Δt for all signal Monte Carlo candidate events and for outliers. The distribution for outliers is significantly broader as expected. The third plot shows the distribution of residuals $\delta\Delta t = \Delta t_{\rm meas} - \Delta t_{\rm true}$ for outliers. The final plot shows the fraction of events that are outliers, as a function of Δt. Although a very significant fraction of events at large Δt are outliers, this is partly due to the definition of outliers ($|\delta\Delta t/\sigma_{\Delta t}| > 5$) combined with the earlier cut on maximum Δt error: $\sigma_{\Delta t} < 2.4$ ps. We have decided not to change the existing cut on $|\Delta t| < 18$ ps. Instead, we will study the effect of varying the Δt and $\sigma_{\Delta t}$ cuts on the final physics results to estimate systematic uncertainties. As described elsewhere [47], we also find the events with less than two tracks used in the $B_{\rm tag}$ vertex are more likely to be outliers as well. To minimize our outlier contamination, we will require two or more tracks on the tag-side vertex.

We model outliers with a single Gaussian in residual space whose width and bias are not scaled by the calculated error because, by their definition, $\sigma_{\Delta t}$ is not a useful estimator:

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

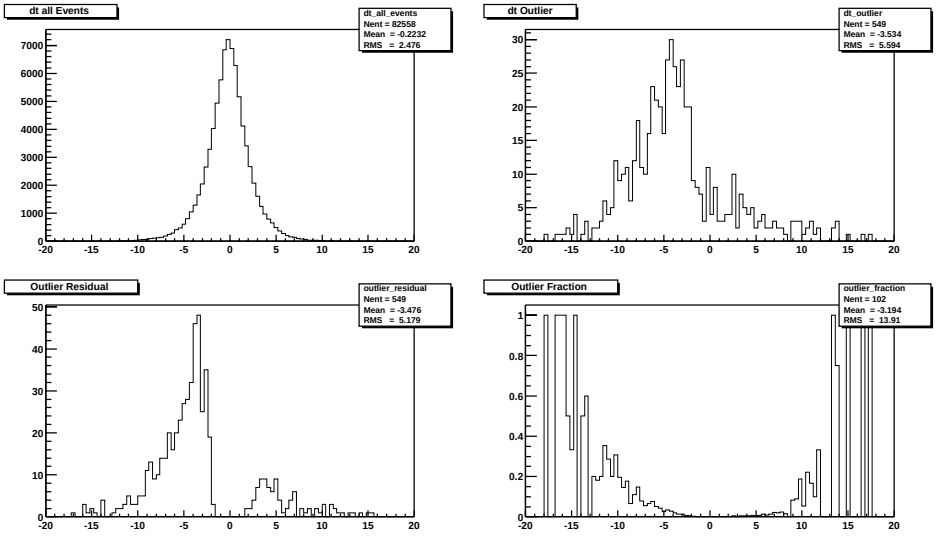


Figure 6.19: Distribution of Δt for all signal Monte Carlo $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ candidate events (top left) and for outliers (top right). The bottom left plot shows the distribution of residuals $\delta\Delta t = \Delta t_{\rm reco} - \Delta t_{\rm true}$ for outliers. The bottom right plot shows the fraction of events that are outliers, as a function of Δt. (Assuming a definition of *pull* ≥ 5 for outliers!)

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:

$$\begin{aligned}\mathcal{R}_{\rm G+G+G}(\delta\Delta t, \sigma_{\Delta t};\, b_1, b_2, b_{\rm out}, s_1, s_2, s_{\rm out}, f, f_{\rm out}) \equiv \\ f \cdot G(\delta\Delta t;\, b_1\sigma_{\Delta t}, s_1\sigma_{\Delta t}) + \\ (1 - f - f_{\rm out}) \cdot G(\delta\Delta t;\, b_2\sigma_{\Delta t}, s_2\sigma_{\Delta t}) + \\ f_{\rm out} \cdot R_{\rm out}(\delta\Delta t;\, b_{\rm out}, s_{\rm out})\ ,\end{aligned} \tag{6.16}$$

and

$$\begin{aligned}\mathcal{R}_{\mathrm{GExp+G}}(\delta\Delta t, \sigma_{\Delta t};\, \kappa, s, f, b_{\mathrm{out}}, s_{\mathrm{out}}, f_{\mathrm{out}}) \equiv \\ & f \cdot G(\delta\Delta t;\, 0, s\,\sigma_{\Delta t}) + \\ & (1 - f - f_{\mathrm{out}}) \cdot \frac{1}{2\,\kappa\sigma_{\Delta t}} \int_{-\infty}^{0} du\, \exp\bigl(+u/(\kappa\,\sigma_{\Delta t})\bigr) \cdot G(u - \delta\Delta t;\, 0, s\,\sigma_{\Delta t}) + \\ & f_{\mathrm{out}} \cdot R_{\mathrm{out}}(\delta\Delta t;\, b_{\mathrm{out}}, s_{\mathrm{out}})\ .\end{aligned} \tag{6.17}$$

We will return to the subject of outliers in Section 12.1.1.

Chapter 7

Flavor Identification

"Which tag of the tagging tagbits should I use to tag the candidate?" – Anonymous

All charged tracks in the event, except the daughter tracks of the identified $D^*\ell$ candidate, are used to determine whether the $B_{\rm tag}$ decayed as a B^0 or a $\overline{B}^0$. This is called flavor tagging, and is in turn used to classify the event as mixed or unmixed. We use five different types of flavor tag, or tagging categories, in this analysis. The first two tagging categories rely on the presence of a prompt lepton, or one or more charged kaons, in the event. The other three categories exploit a variety of inputs with a neural-network algorithm.

7.1 Definitions

We have defined the mixing state (or status) of a candidate $B\overline{B}$ event to be:

- *mixed* - The b quark in $B_{\rm rec}$ has changed flavor between $t_{\rm tag}$ and $t_{\rm rec}$, or
- *unmixed* - The b quark in $B_{\rm rec}$ has the same flavor at $t_{\rm rec}$ as it had at $t_{\rm tag}$.

In order to determine the mixing status of the event, we need to determine the b quark flavor of the $B_{\rm rec}$ and $B_{\rm tag}$ mesons. Because we have exclusively reconstructed the $B_{\rm rec}$ as $B_{D^*\ell}$, the flavor of this candidate is trivial (a negative lepton indicates $\overline{b}$).

We also define the *mistag rate* or (*probability*) to be the chance of mis-identifying the flavor of the $B_{\rm tag}$ meson. (For all intents and purposes, the mistag rate for $B_{\rm rec}$ is negligible, since we do account for the chance the $B_{\rm rec}$ is not well-reconstructed by assigning it a per-event signal probability.) We assume that the mistag probability is independent of Δt, however.

The time dependent mixing asymmetry can be reformulated to include the realistic effect of non-zero mistag probabilities. Consider a tagging algorithm that correctly tags a B^0 with probability ϵ and $\overline{B}^0$ with probability $\overline{\epsilon}$. Because the algorithm is 100% efficient, events tagged incorrectly are merely assigned the opposite (incorrect) flavor. Assume a sample of N events which contain true $B^0\overline{B}^0$ decays with $N_{B^0\to\overline{B}^0}+N_{\overline{B}^0\to B^0}$ mixed events and $N_{B^0\to B^0}+N_{\overline{B}^0\to\overline{B}^0}$ unmixed events. The number of measured mixed and unmixed events is then:

$$\begin{aligned} N^{meas}_{mixed} &= N^{meas}_{B^0\to\overline{B}^0} + N^{meas}_{\overline{B}^0\to B^0} \\ &= \overline{\epsilon}N_{B^0\to\overline{B}^0} + (1-\epsilon)N_{B^0\to B^0} + \epsilon N_{\overline{B}^0\to B^0} + (1-\overline{\epsilon})N_{\overline{B}^0\to\overline{B}^0} \\ N^{meas}_{unmixed} &= N^{meas}_{B^0\to B^0} + N^{meas}_{\overline{B}^0\to\overline{B}^0} \\ &= \epsilon N_{B^0\to B^0} + (1-\epsilon)N_{\overline{B}^0\to\overline{B}^0} + \overline{\epsilon}N_{\overline{B}^0\to\overline{B}^0} + (1-\epsilon)N_{B^0\to B^0} \end{aligned} \tag{7.1}$$

and likewise for unmixed events. We can combine these observed quantities to form the measured mixing asymmetry as below, where we have suppressed the dependence on Δt, and where we have identified in advance that the denominator is unchanged

(flavor tagging does not alter the overall number of events):

$$\begin{aligned} A^{meas} &= \frac{N^{meas}_{\rm unmixed} - N^{meas}_{\rm mixed}}{N^{meas}_{\rm unmixed} + N^{meas}_{\rm mixed}} \\ &= \frac{\epsilon N_{B^0\to B^0} + (1-\epsilon)N_{\overline{B}^0\to\overline{B}^0} + \overline{\epsilon} N_{\overline{B}^0\to\overline{B}^0} + (1-\epsilon)N_{B^0\to B^0}}{N_{unmixed} + N_{mixed}} \quad - \\ &\quad \frac{\overline{\epsilon} N_{B^0\to\overline{B}^0} + (1-\epsilon)N_{B^0\to B^0} + \epsilon N_{\overline{B}^0\to B^0} + (1-\overline{\epsilon})N_{\overline{B}^0\to\overline{B}^0}}{N_{unmixed} + N_{mixed}} \\ &= \frac{(2\epsilon-1)(N_{B^0\to B^0} - N_{\overline{B}^0\to\overline{B}^0}) + (2\overline{\epsilon}-1)(N_{\overline{B}^0\to B^0} - N_{B^0\to\overline{B}^0})}{N_{unmixed} + N_{mixed}} \ . \end{aligned} \tag{7.2}$$

But from Eq. 2.35, we recognize that this is really just a linear combination of the true mixing asymmetry:

$$\begin{aligned} A^{meas} &= \frac{1}{2}(2\epsilon-1)A + \frac{1}{2}(2\overline{\epsilon}-1)A \\ &= \left[\frac{1}{2}(2\epsilon-1) + \frac{1}{2}(2\overline{\epsilon}-1)\right]\cos(\Delta m_d\ \Delta t) \\ &= \left[\frac{1}{2}(1-2\omega) + \frac{1}{2}(1-2\overline{\omega})\right]\cos(\Delta m_d\ \Delta t) \\ &= \left[\frac{D}{2} + \frac{\overline{D}}{2}\right]\cos(\Delta m_d\ \Delta t) \\ &= \langle D\rangle \cos(\Delta m_d\ \Delta t)\ , \end{aligned} \tag{7.3}$$

where we have made the substitution $\omega = 1-\epsilon$ in the last steps and defined $D \equiv 1-2\omega$ to be the *dilution* and $\langle D\rangle$ to be the average dilution. We see that, as claimed earlier, a non-zero mistag probability merely reduces the amplitude of the mixing asymmetry and does not interfere with the the time-dependence. Because of this, it is straightforward to extract both the mixing frequency and the mistag probabilities when fitting the observed time dependence.

In the event that different tagging categories have different mistag rates, we can construct a separate diluted mixing asymmetry for each of them (using their respective mistag rates ω_i) and use the population-weighted sum to describe the measured asymmetry in the Data.

7.2 Algorithms

B meson flavor-identification at *BABAR* is broken into several components. The first stage is to remove the $B_{\rm rec}$ and its identified daughters from the event, in a similar way as was done for $B_{\rm tag}$ vertexing. The "rest of the event" is then subjected to several different algorithms whose outputs are ranked and then combined to form the best estimate of the $B_{\rm tag}$ flavor. This analysis uses the so-called *Elba Tagger* which assigns a flavor status and one of five categories to each event; the category is representative of the information used in making the final decision.

BABAR employs two very different but complementary strategies: one is a cut-based algorithm (NOT - Non Optimal Tagging) while the second is a neural network based approach (NetTagger). NOT uses identified electrons, muons, kaons and slow pions from the decay $D^{*+} \to D^0\pi^+$ to determine the $B_{\rm tag}$ flavor. The algorithm uses all charged tracks not used in the $B_{\rm rec}$ candidate and the various PID hypotheses assigned to each track (as discussed in Chapter 5.5.2). The neural network allows non-linear responses to linear combinations of input variables, and therefore greatly enhances sensitivity to correlations. The ultimate approach is a hybrid which retains the best qualities of both methods:

- `LTag`: This is the tagging category with the lowest mistag fraction. It uses the correlation between the charge of the primary lepton and the flavor of the b quark: $b \Leftrightarrow \ell^-$. A momentum cut in the $\Upsilon(4S)$ rest frame of 1.1 GeV for electrons and 1.0 GeV for muons (lower to recover efficiency) helps to reduce contamination due to secondary leptons from semileptonic charm decays. (see Fig. 7.1).

- `KTag`: If the sum of charges of all identified kaons is nonzero, the event is assigned a `KTag` tag, with the flavor determined by $b \Leftrightarrow \Sigma Q_K > 0$, using the correlation between the charge of the kaon and the b quark flavor in the most probable decay chain $b \to c \to s$ in a neural network. * This category is the

*On average there are 0.8 K mesons per B decay, and about 15% of these are so-called *wrong-sign* which means that the charge of their contained s quark does not match that of the originating b quark. See Chapter 7.4 for more details.

most efficient.

- `NT1`: Events for which the absolute value of the NetTagger neural net output is greater than 0.5.
- `NT2`: Events for which the absolute value of the neural net output is between 0.2 and 0.5.
- `NT3`: Events for which the absolute value of the neural net output is less than 0.2. (least certain)

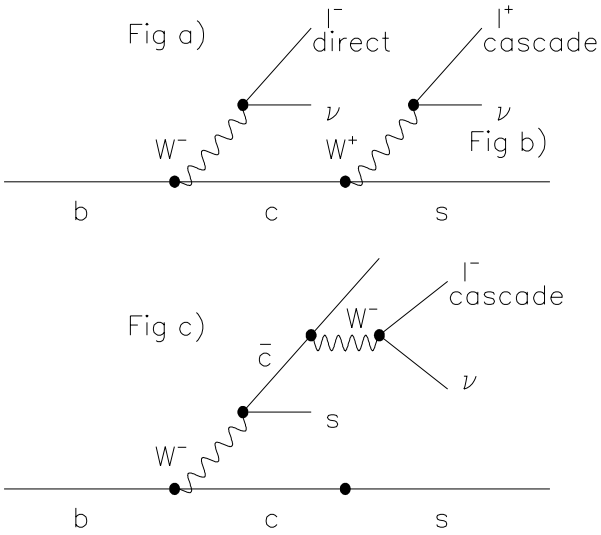


Figure 7.1: Feynman diagram showing the production mechanism for direct and cascade leptons. The region labelled (a) shows a direct lepton from the b to c transition, while the region labelled (b) shows a cascade (or secondary) lepton from the subsequent c to s decay. Finally, region (c) shows another cascade lepton diagram, this time with the same sign lepton charge as that of the b quark.

The first two categories contain more than 45% of the events. The final three tagging categories involve the multi-variate analysis based on a neural network (NetTagger), which is trained to identify primary leptons, kaons, and soft pions, and the

momentum and charge of the track with the maximum center-of-mass momentum. Depending on the output of the neural net, events are assigned to an `NT1` (most certain), `NT2`, or `NT3` (least certain) tagging category (see Fig. 7.2). About 30% of events are in the `NT3` category, which has a mistag rate close to 50%. Therefore, these events do not carry much sensitivity to the mixing frequency, but they increase the sensitivity to the B^0 lifetime.

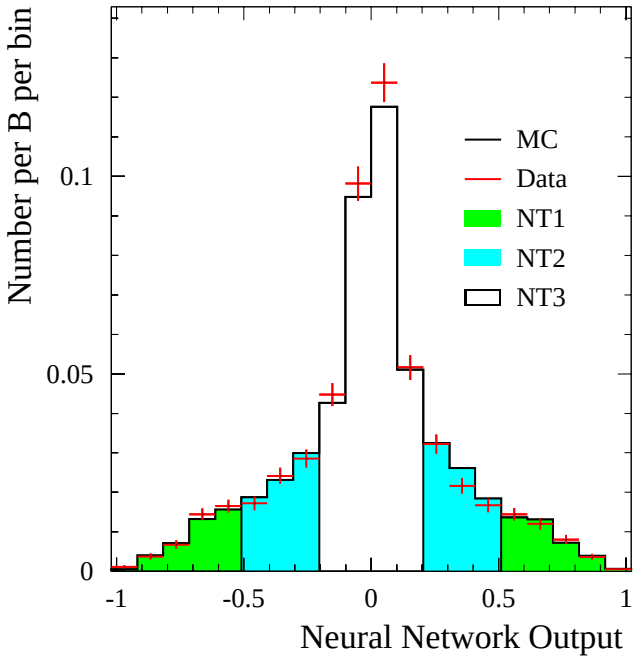


Figure 7.2: Distribution of neural network responses for events not claimed by the `LTag` and `KTag` categories. The histogram is from simulation, the points from fully reconstructed B events. An output value of $+1$ corresponds to a B^0 flavor assignment, and -1 to $\overline{B}^0$.

Tagging categories are mutually exclusive due to the hierarchical use of the tags. Events with a `LTag` tag and no conflicting `KTag` tag are assigned to the `LTag` category. If no `LTag` tag exists, but the event has a `KTag` tag, it is assigned to the `KTag` category. Otherwise events are assigned to corresponding neural network categories.

7.3 Performance

Because we use the full spectrum of output from the neural network, the flavor-tagging algorithm is 100% efficient in this analysis. Using Monte Carlo data and its corresponding truth information, we can measure the absolute efficiencies of each tagging category, as well as the *power* $Q = \epsilon(1-2\omega)^2$. Q measures the effectiveness, combining the efficiency with the error rate. Using Data, we can only measure the mistag rates per tagging category in the full fit. (Please refer to Table 12.1 for the final mistag measurements in Data.)

Category	ϵ_i	ω_i	$Q_i = \epsilon_i(1-2\omega_i)^2$
`LTag`	13.2	.063	10.1
`KTag`	35.1	.157	16.5
`NT1`	8.3	.193	3.1
`NT2`	14.5	.338	1.5
`NT3`	28.9	.465	0.14
Total	100		29.9

Table 7.1: Tagging performance on signal Monte Carlo events. [52]

7.4 Vertexing-Tagging Correlations

A correlation of about 0.12 ps^{-1} is observed between the mistag rate and the Δt resolution for `KTag` tags. Fig. 7.3 shows the computed mistag rate (using Monte Carlo truth) in increasing bins of $\sigma_{\Delta t}$, and the correlation is strongest in the `KTag` category. This effect is modeled in the resolution function for signal as a linear dependence of the mistag rate on $\sigma_{\Delta t}$, as shown here:

$$\omega_{\text{kaon}} = m_{\text{kaon}} \cdot \sigma_{\Delta t} + \omega_{\text{kaon}}^{\text{offset}} . \tag{7.4}$$

We briefly describe the source of this correlation [53], but please see Ref. [54], [55] for more details. We find that both the mistag rate for `KTag` tags and the calculated error on Δt depend inversely on $\sqrt{\Sigma p_t^2}$, where p_t is the transverse momentum with

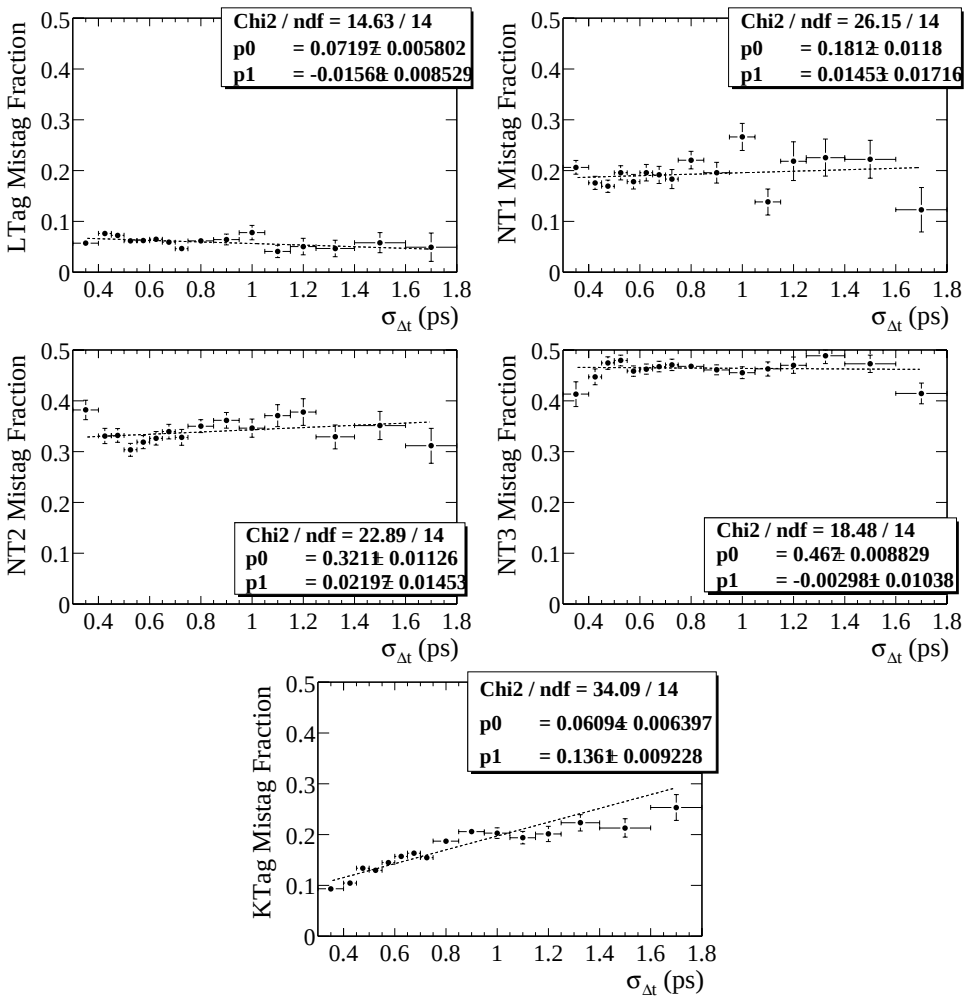


Figure 7.3: From signal Monte Carlo, computed mistag fractions in increasing bins of $\sigma_{\Delta t}$ for the 5 tagging categories described in the text. The KTag category exhibits a significant correlation.

respect to the z axis of tracks from the $B_{\rm tag}$ decay (see Fig. 7.4). Correcting for this dependence of the mistag rate removes most of the correlation between the mistag rate and $\sigma_{\Delta t}$ as demonstrated in Fig. 7.5.

The mistag rate dependence originates from the kinematics of the physics sources for wrong-charge kaons. The three major sources of mistags are wrong-sign D^0 mesons

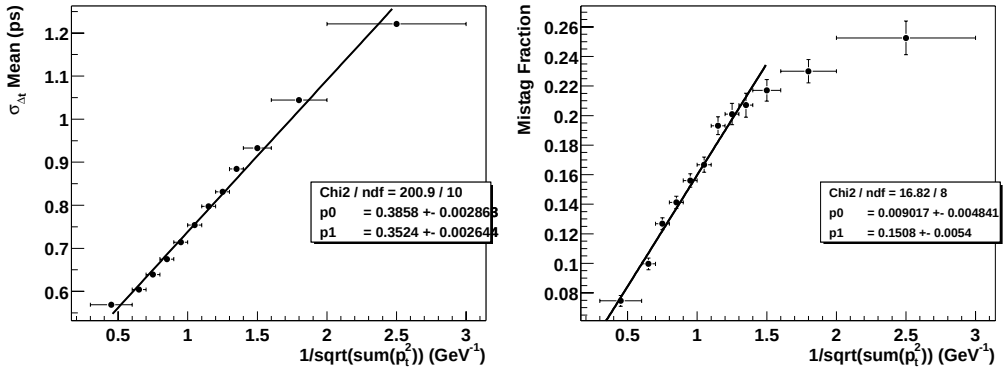


Figure 7.4: Observation of common dependence of $\sigma_{\Delta t}$ (left) and mistag rate ω (right) on $\sqrt{\Sigma p_t^2}$ from Monte Carlo.

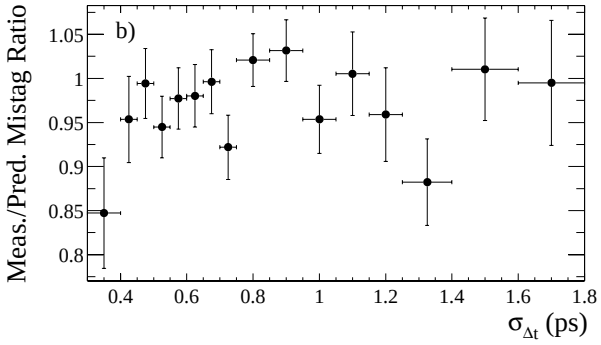


Figure 7.5: Dependence of mistag rate in `KTag` category on $\sigma_{\Delta t}$ after scaling by $\sqrt{\Sigma p_t^2}$. The flat slope indicates that most of the correlation between mistag rate and $\sigma_{\Delta t}$ is removed after correcting for the common dependence shown in Fig. 7.4.

from B decays to double charm, wrong-sign kaons from D^+ decays, and kaons produced directly in B decays, as illustrated in Fig. 7.6. All these sources produce a spectrum of charged tracks that have smaller $\sqrt{\Sigma p_t^2}$ than B decays that produce a correct tag. The $\sigma_{\Delta t}$ dependence originates from the $1/p_t^2$ dependence of σ_z for the individual contributing tracks.

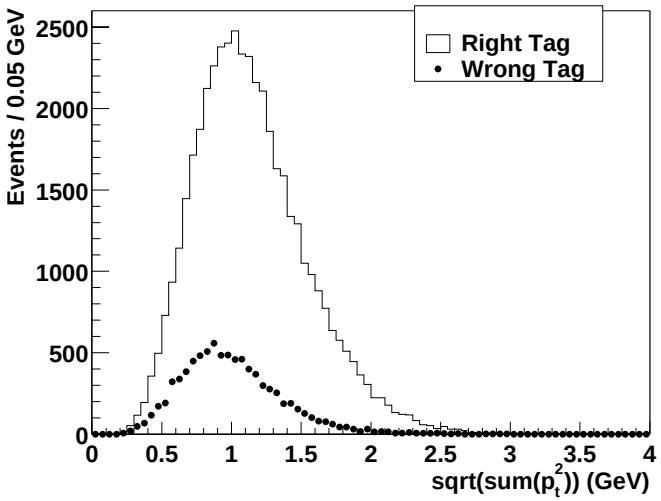


Figure 7.6: The $\sqrt{\sum p_t^2}$ spectra for correctly and incorrectly tagged events by `KTag` in signal Monte Carlo. As shown elsewhere [54], $B_{\rm tag}$ decays terminating in "wrong-sign" K mesons typically have a softer p_T spectrum due to a combination of effects such as momentum losses in double-charm decays and charged track multiplicities in D^+ decays.

Chapter 8

Event Sample Selection

The Data used in this analysis were recorded with the BABAR detector [40] at the PEP-II storage ring [38] in the period October 1999 to December 2000. The total integrated luminosity of the data set is equivalent to 20.6 fb^{-1} collected near the $\Upsilon(4S)$ resonance and 2.6 fb^{-1} collected 40 MeV below the $\Upsilon(4S)$ (off-resonance data). The corresponding number of produced $B\overline{B}$ pairs is estimated to be about 23 million.

We select events containing a fully reconstructed D^{*-} and an identified oppositely-charged electron or muon. This $D^*\ell$ pair is then required to pass kinematic cuts that enhance the contribution of semileptonic $B \to D^*\ell\overline{\nu}_\ell$ decays. In addition to the signal sample, we select several control samples that are used to characterize the main sources of background.

8.1 Overview

Not every event recorded by the *BABAR* detector is relevant for this analysis. The physicist's job is to look through the entire Data sample and reduce it to a smaller (private) sample by filtering out "uninteresting" events. We focus on semileptonic decays of the B meson: in particular, the mode $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$. Using the tagbit technology defined in Chapter 5, we scan through the entire set of *BABAR* Data taken from 1999–2000 and select those events which appear most promising as $D^*\ell$ candidates. The integrated luminosity considered is $20.6\,\mathrm{fb}^{-1}$ on the $\Upsilon(4S)$ peak, and $2.6\,\mathrm{fb}^{-1}$ off-resonance.

We build $D^*\ell$ candidates on an event-by-event basis, starting by filtering on global event properties and then reconstructing the in the mode $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$, starting to build parent particles, working backwards up the decay chain. We select events by requiring them to:

- be consistent with global event topology of B decay,
- contain a reconstructed D^0 candidate in one of the four decay modes used: $D^0 \to K\pi$, $D^0 \to K\pi\pi\pi$, $D^0 \to K^0_S\pi\pi$, $D^0 \to K\pi\pi^0$, (as appropriate, a reconstructed K^0_S is required),
- contain a reconstructed soft pion from the $D^{*-} \to \overline{D}^0\pi^-$ decay, π^-_{soft},
- contain a reconstructed D^* candidate using the D^0 and π^-_{soft},
- contain a lepton candidate that associates with the $D^*\ell$ candidate,
- pass B_{tag} side criteria such as vertexing,
- contain a candidate $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ candidate in a suspected $B\overline{B}$ event.

This amounts to an overall branching fraction of 1.8% summing over both lepton types. In the following sections, we'll discuss each of these criteria with the ultimate goal of efficiently selecting $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ events in a way that incorporates non-signal in a known and parameterizable fashion; that is, following a strategy whose "mistakes" we can identify and accommodate. (See Appendix B for more information.)

We define here the following classification of the sources of signal and background that we expect to contribute to this sample. The nomenclature shown in italics will be used throughout this paper to define signal and all possible types of background. This classification scheme will become more important as we move through the analysis.

1. Events with a correctly reconstructed D^{*-} candidate:
   (a) Events that originate from $B\overline{B}$ events:
      i. Events with a correctly identified lepton candidate:
         A. *Signal* – $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ (X) decays.
         B. *Uncorrelated-lepton background* – ($B \to D^{*-}X$, other $B \to \ell^+X$) or ($B \to D^{*-}X$, $X \to \ell^+Y$)
         C. *Charged B background* – $B^+ \to D^{*-}\ell^+\nu_\ell X$.
      ii. *Fake-lepton background* – events with a misidentified lepton candidate.
   (b) *Continuum background* – $c\overline{c} \to D^{*-}X$.
2. *Combinatoric background* – events with a misreconstructed D^{*-} candidate.

The careful accounting and labelling of events allows us to disentangle the background (*i.e.,* non-signal) contributions in our sample.

8.2 Signal Reconstruction

8.2.1 Daughter Reconstruction

The short-lived particles * used in this selection are: π^0, K^0_S, D^0, and $D^{*\pm}$. We identify candidates for each of these composites by combining entries in the reconstruction lists, and possibly other composite lists, and requiring that the resulting combinations pass loose kinematic cuts (applied to four-vector sums, without any refitting to include geometric or kinematic constraints).

*By which we refer to the particles that decay within the detector volume after production in the B decay.

π^0 Selection

π^0 candidates are reconstructed from two photons with invariant mass within 15.75 MeV of the nominal π^0 mass. The mass of the photon pair is then constrained to the π^0 mass and the photon pair is kept as a π^0 candidate if the χ^2 probability of the fit is greater than 1%. This refitting improves the energy resolution of the π^0 candidates from about 3.0% to 2.5%. [56] Combined with the geometrical acceptance of the calorimeter, the overall efficiency is 65-70% at moderate momentum.

K Selection

K_S^0 candidates are reconstructed from a pair of charged particles with invariant mass within 15 MeV of the K_S^0 mass. The pair of tracks is retained as a K_S^0 candidate if the χ^2 probability that the two tracks form a common vertex is greater than 1%. Charged kaon candidates satisfy loose kaon criteria for the $K\pi$ mode and tighter criteria for the $K\pi\pi\pi$ and $K\pi\pi^0$ modes.

D^0 Selection

The D^0 candidate is reconstructed in the modes $K^-\pi^+$, $K^-\pi^+\pi^-\pi^+$, $K^-\pi^+\pi^0$ and $K_S^0\pi^-\pi^+$. D^0 candidates must have measured invariant mass within 17 MeV of the D^0 mass for the $K\pi$, $K\pi\pi\pi$, and $K_S^0\pi\pi$ modes, and within 34 MeV for the $K\pi\pi^0$ mode. The invariant mass of the daughters is then constrained to the D^0 mass, and the tracks are constrained to a common vertex in a simultaneous fit. The D^0 candidate is retained if the χ^2 probability of the fit is greater than 0.1%. For the $K\pi\pi^0$ and $K_S^0\pi\pi$ modes, a likelihood is calculated as the square of the decay amplitude in the Dalitz plot for the three-body candidate, based on measured amplitudes and phases. [57] The candidate is retained if the likelihood is greater than 10% of its maximum value across the Dalitz plot. Fig. 8.1 shows the Dalitz distribution from Monte Carlo truth for $D^0 \to K\pi\pi^0$, *i.e.,* before convolving with resolution functions. This decay has several resonances, the largest being $D^0 \to K^-\rho^+$. The Dalitz decay amplitude criteria is generally more effective than a traditional selection based on $\pi^+\pi^0$ invariant mass and helicity analysis. [31], [58]

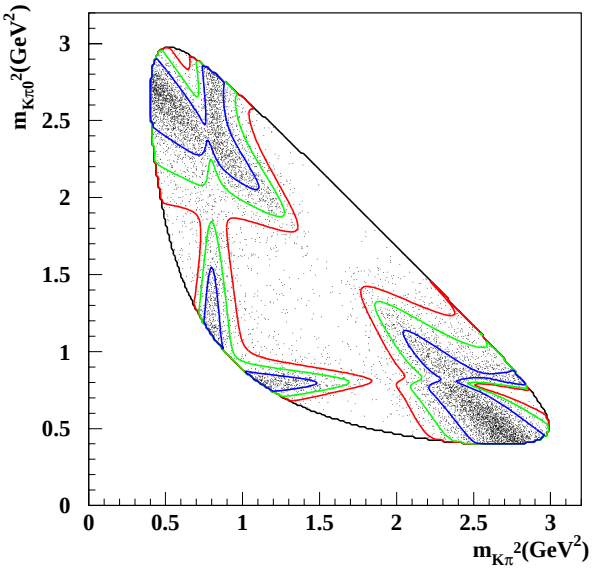


Figure 8.1: $D^0 \to K^-\pi^+\pi^0$ Dalitz distribution and contours according to parameters found in Ref. [57], before convolving with resolution functions. The values on the contours are 0, 10, 20, and 50, out of approximately maximum 357.

D^* Selection

D^* candidates are selected in the decay mode $D^{*+} \to D^0\pi^+$. The low-momentum pion candidates for the $D^{*+} \to D^0\pi^+$ decay are selected from a list of loosely † selected tracks with total momentum less than 450 MeV in the $\Upsilon(4S)$ rest frame and momentum transverse to the beam line greater than 50 MeV. (See Fig. 8.2.) The momentum of the D^* candidate in the $\Upsilon(4S)$ rest frame is required to be between 0.5 and 2.5 GeV, which helps discriminate against continuum-generated D^* mesons, as shown in Fig. 8.3. Continuum events typically have more energetic D^*'s because there is more energy available for momentum transfer to the D^* daughter, since the c quark is lighter than the b. Because of differences in resolution on the soft pion, we separate D^* candidates into those whose soft pion candidate has DCH hits in addition

†GoodTracksVeryLoose

to the SVT ones, and those that have only SVT hits on track.

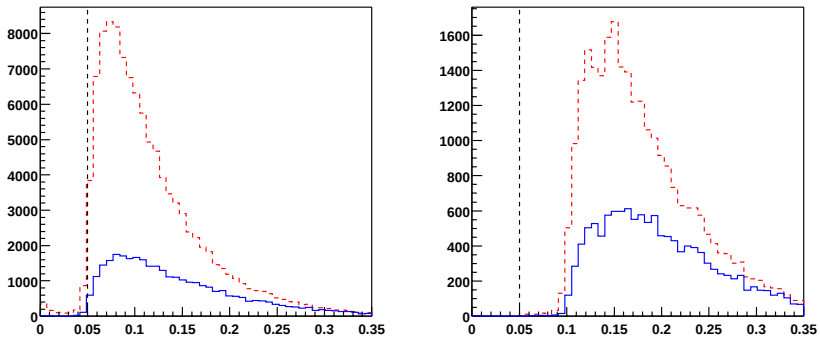


Figure 8.2: Left: Distribution of transverse momentum for the $\pi^-_{\rm soft}$ provided that it is an SVT-only candidate. Right: Distribution of $\pi^-_{\rm soft}$ momentum for candidates with at least hits in the drift chamber tracking system. The red histograms are the distribution from signal Monte Carlo, the blue from the truth-matched subset. The dashed line shows the lower cut.

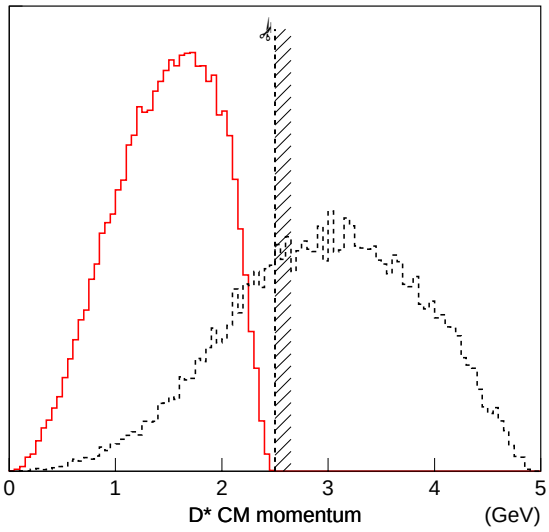


Figure 8.3: Distribution of D^* center-of-mass momenta for signal candidates (left histogram in red) and continuum generated D^* (right histogram in black), at the Monte Carlo truth level. The selection window prefers the softer signal D^* candidates.

Lepton Selection

Lepton candidates are defined as charged tracks from the tight track list [‡] with momentum in the $\Upsilon(4S)$ rest frame greater than 1.2 GeV, which helps separate primary from secondary leptons (see Fig. 8.4). For the D^*e samples, the electron candidate passes selection criteria with a corresponding electron identification efficiency of about 90% and hadron misidentification less than 0.2%. For the $D^*\mu$ samples, the muon candidate passes selection criteria with a corresponding muon identification efficiency of about 70% and hadron misidentification between 2% and 3%. For the fake-lepton control sample, $D^*\ell$ candidates are accepted if the lepton *fails* both electron and muon selection criteria looser than those required for lepton candidates.

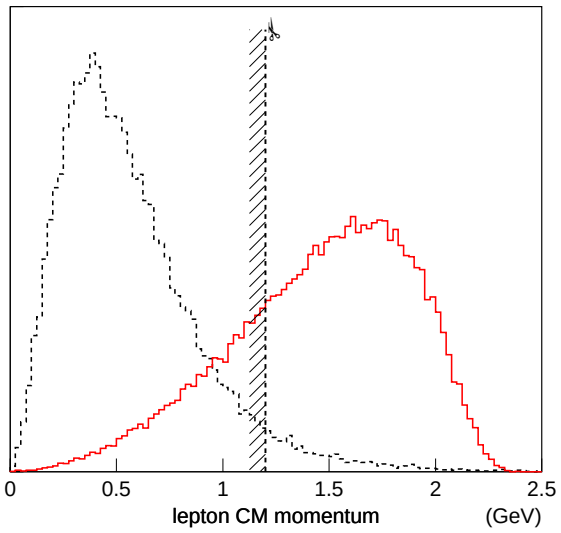


Figure 8.4: Distribution of ℓ center-of-mass momentum for signal candidates (right histogram in red) and secondary leptons (left histogram in black), at the Monte Carlo truth level. The high-momentum requirements preferentially selects leptons from the B decay.

[‡]GoodTracksTight

8.2.2 $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ Selection

The $D^*\ell$ candidate satisfies $|\cos\theta^*_{\rm thrust}| < 0.85$, where $\theta^*_{\rm thrust}$ is the angle between the thrust axis of the $D^*\ell$ candidate and the thrust axis of the remaining charged and neutral particles in the event. $D^*\ell$ candidates are retained if the χ^2 probability that the daughter tracks form a common vertex is greater than 1% and $\delta m = m(D^*) - m(D^0)$ is less than 165 MeV, where $m(D^*)$ is the candidate $\overline{D}^0\ \pi^-$ mass calculated with the candidate $\overline{D}^0$ mass constrained to the true D^0 mass, $m(D^0)$. We also then require that the triply constrained $(D^*, \ell, beamspot)$ refitting algorithm described in Chapter 6.1.1 converges.

Finally, events are retained if: (a) at least two tracks are used to determine the decay point of the other B, (b) the fit that determines the distance Δz between the two B decays along the beamline converges, (c) the time between decays (Δt) calculated from Δz is less than 18 ps, and (d), the calculated error on Δt ($\sigma_{\Delta t}$) is less than 1.8 ps.

8.3 Final Data Sample

We've described everything but the final step of $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ event selection. Because selection criteria are efficient but impure, we inevitably end up including various backgrounds in the Data sample (where we've tried to isolate the signal). In part because of the overall goal of high precision for this analysis, we need to be able to account for these background events in the sample when we perform the final measurement. We employ a (standard) technique for characterizing backgrounds in the signal sample by "inverting" the "last few" signal selection criteria to select a *background control sample* which is mostly background and poor in signal. By studying the behavior of this sample, we can extrapolate (by analogy) to describe the behavior of the background candidates in the signal sample. The novel feature of this analysis is that not only do we study the background control samples separately, but we include them *as much as possible* in the final fit for the time dependence. That is, to minimize the chance for bias and systematic error, we perform the final

measurement on a sample that is more than 75% background!

As described above, then, we have a set of $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ candidates where we have not applied the following requirements, which would select the purest subsample of signal:

- restrict to the peak region $143\,\mathrm{MeV} < \delta m < 148\,\mathrm{MeV}$ of the $m(D^*) - m(D^0)$ spectra (minimizing the number of candidates with "fake" (combinatoric) D^*'s)
- require strict e or μ particle identification of the lepton (minimizing the number of candidates with fake leptons)
- require $D^*\ell$ kinematics and angular correlations consistent with a semileptonic decay and a missing neutrino.

We apply combinations of these requirements to the raw $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ sample to form the final Data sample. We further divide the sample into several subgroups in an effort to distinguish their physics based on several different characteristics. A trivial example of this is the background control sample for combinatorics. We "invert" the signal selection criteria to use the δm sideband ($\delta m > 155\,\mathrm{MeV}$) as a relatively pure sample of combinatoric candidates.

We define two angular quantities for each $D^*\ell$ candidate to classify them into a sample enriched in $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ signal events, in which the D^* and lepton candidates are on opposite sides of the event, and a sample enriched in *uncorrelated-lepton background* events, in which the D^* and lepton candidates are on the same side of the event. The first angle is $\theta_{D^*,\ell}$, the angle between the D^* and lepton candidates in the $\Upsilon(4S)$ rest frame. The second angle is $\theta_{B^0,D^*\ell}$, the angle between the direction of the B^0 and the vector sum of the D^* and lepton candidate momenta, calculated in the $\Upsilon(4S)$ rest frame. Fig. 8.5 shows the distributions of these angles for signal and background (some of each are included in the Data sample). We also calculate the same angle with the lepton momentum direction reflected through the origin in the $\Upsilon(4S)$ rest frame: $\cos\theta_{B,D^*(-\ell)}$.

The B^0 is a spin-0 particle, so by angular momentum conservation, the total spin of the $D^{*-}\ell^+\nu_\ell$ system must also be zero. The D^* is a spin-1 meson, so the

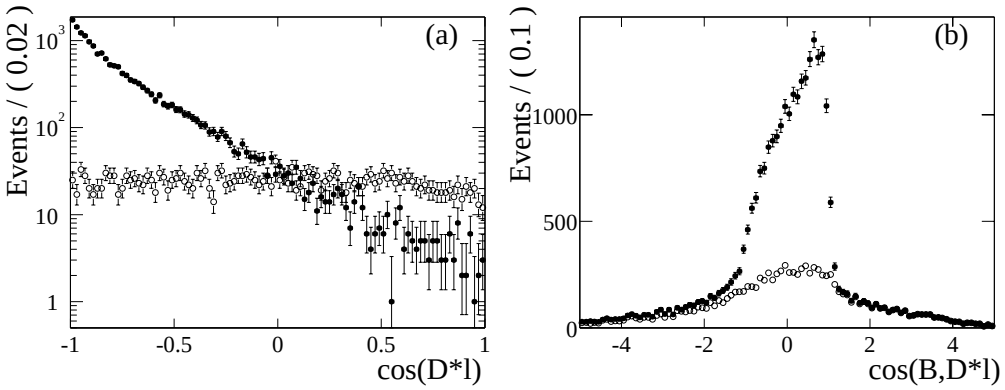


Figure 8.5: Distribution of angular quantities for $D^*\ell$ candidates in the signal Monte Carlo sample. Signal behavior is shown in solid black circles, and backgrounds in open circles. The left plot shows the preferred "opposite-side" nature of signal $D^*\ell$ candidates; the right plot shows $\cos\theta_{B,D^*\ell}$, which leaks beyond the $[-1,+1]$ physical region due to reconstruction effects.

lepton-neutrino system must align its total angular momentum to cancel out the total angular momentum projected on the z axis. Since the helicity of the neutrino is fixed (left-handed neutrinos, right handed anti-neutrinos) the lepton tends to come out back-to-back from the D^* with a harder spectrum than the neutrino. Hence, the $\cos\theta_{D^*,\ell}$ distribution for signal will be strongly peaked at -1. A number of backgrounds unfortunately have the same topology; $c\overline{c}$ events will also have a real D^*-ℓ system that behaves this way. However, *uncorrelated lepton* backgrounds can be distinguished using this variable, as they will tend to be flat across the spectrum.

Since we do not know the direction of the B^0, we calculate the cosine of $\theta_{B^0,D^*\ell}$ from the following equation, in which we assume that the only B decay particle missed in the reconstruction is a massless neutrino:

$$\cos\theta_{B,D^*\ell} = \frac{-(m_{B^0}^2 + m_{D^*\ell}^2 - 2E_B E_{D^*\ell})}{2|\vec{p}_B||\vec{p}_{D^*\ell}|} \,. \tag{8.1}$$

All quantities in Eq. 8.1 are defined in the $\Upsilon(4S)$ rest frame. The energy and the magnitude of the momentum of the B are calculated from the e^+e^- center-of-mass energy and the B^0 mass. For true $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ events, $\cos\theta_{B,D^*\ell}$ will lie in the

physical region $[-1,+1]$, aside from detector resolution effects. Backgrounds lie both inside and outside the range $[-1,+1]$.

With these definitions in hand, we select a sample enhanced in $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ signal events (called the *opposite-side* sample (OS))with $D^*\ell$ candidates with $\cos\theta_{D^*,\ell} < 0$ and $|\cos\theta_{B,D^*\ell}| < 1.1$. We additionally divide this sample by considering the lepton identification criteria, whether the candidates satisfy the criteria for an electron, a muon or a fake-lepton. The first two samples are the signal samples, and the latter is the *fake-lepton* control (sub)sample (of the OS sample).

From Fig. 8.5 we observe that there is significant *uncorrelated background* in the signal region, where $\cos\theta_{D^*,\ell} < 0$. Following the prescription above, we might consider selecting a complementary background control sample by requiring $\cos\theta_{D^*,\ell} > 0$ and $|\cos\theta_{B,D^*\ell}| < 1.1$. However, according to Monte Carlo simulation, [§] the distribution of $\cos\theta_{B,D^*\ell}$ in the background control sample is systematically different than that of the uncorrelated-lepton background in the signal sample. We observe, though, that the distribution of $\cos\theta_{B,D^*(-\ell)}$ in this control sample is similar to the distribution of the appropriate $\cos\theta_{B,D^*\ell}$, as indicated in Fig. 8.6. We can use $\cos\theta_{B,D^*(-\ell)}$ to form the background control sample, but still require $\cos\theta_{D^*,\ell} > 0$ to maintain independence from the signal OS sample. An additional background control sample, representative of the uncorrelated-lepton background and called the *same-side (SS)* sample, is composed of $D^*\ell$ candidates with $\cos\theta_{D^*,\ell} \geq 0$ and $|\cos\theta_{B,D^*(-\ell)}| < 1.1$. We also divide this sample by considering the three different lepton identification criteria (electron, muon, fake).

Approximately $68,000$ candidates pass the above selection criteria, and are distributed over the two signal samples (e and μ) and ten background control samples defined by the following characteristics: whether the Data was recorded on or off the $\Upsilon(4S)$ resonance (two choices); whether the candidate lepton is *same-side* (SS) or *opposite-side* (OS) to the $D^*\ell$ candidate (two choices); and whether the lepton candidate passes the criteria for an electron, a muon, or a fake lepton (three choices). Of the ten control samples, five are primary samples which we use to directly study

[§]Here we rely on the current parameterizations of the appropriate form factors in Monte Carlo, and therefore we do acquire some systematic uncertainty.

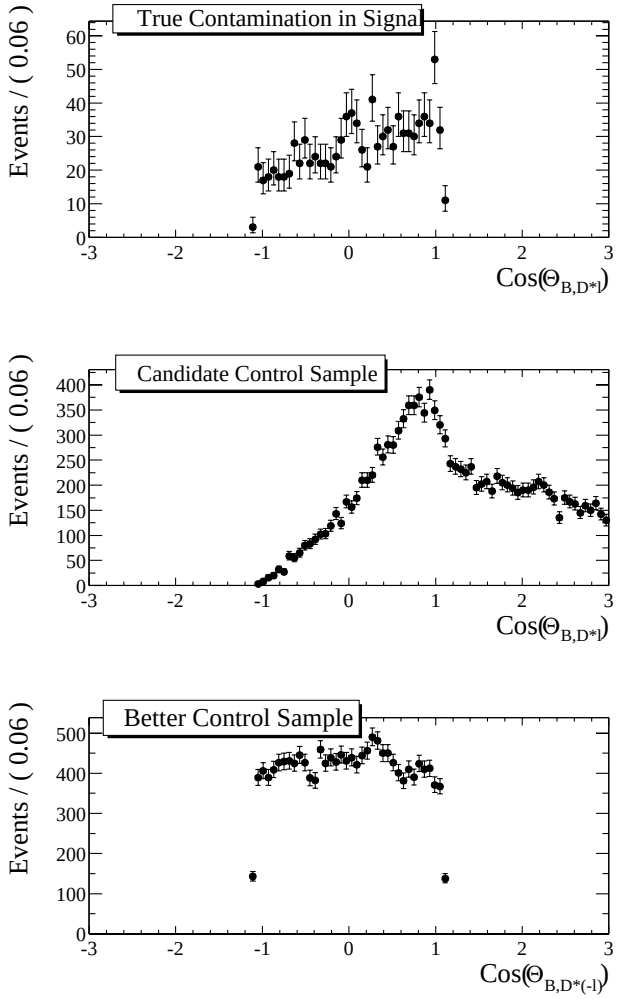


Figure 8.6: Distribution of $\cos\theta_{B,D^*\ell}$ for various subsamples of Monte Carlo. The upper plot shows the distribution for uncorrelated lepton background candidates as identified by Monte Carlo truth-matching. (This is the distribution we want to match.) The middle plot shows the $\cos\theta_{B,D^*\ell}$ distribution for the sample $\cos\theta_{D^*,\ell} > 0$, which is not a good match for the distribution of the background candidates which contaminate the signal sample. The bottom plot shows the $\cos\theta_{B,D^*\ell}$ distribution for the control sample which is computed using the flipped-lepton momentum, and therefore also $\cos\theta_{D^*,\ell} > 0$. This new sample represents the signal pollutants much better. NOTE: The distribution in the upper and lower plots are after the restriction to the range ± 1.

backgrounds that contribute to the signal samples. The other five are secondary subsamples which we use to understand the cross-contamination between the primary background control samples. Table 8.1 enumerates these samples, and their relationships.

Sample	AngCut	OnOffRes	Lepton ID	Sample enriched in...
1	OS	On	e	electron signal
2	OS	On	μ	muon signal
3	OS	On	!e,!μ	fake bkg to samples 1 and 2
4	OS	Off	e	continuum bkg to sample 1
5	OS	Off	μ	continuum bkg to sample 2
6	SS	On	e	uncorrelated bkg to sample 1
7	SS	On	μ	uncorrelated bkg to sample 2
8	OS	Off	!e,!μ	fake bkg to samples 4 and 5
9	SS	On	!e,!μ	fake bkg to samples 6 and 7
10	SS	Off	e	continuum bkg to sample 6
11	SS	Off	μ	continuum bkg to sample 7
12	SS	Off	!e,!μ	fake bkg to samples 10 and 11

Table 8.1: 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.

A key feature of this analysis is the inclusion of such large background samples in the Dataset used in the final fit. This is necessary if we wish to understand the background components in the signal samples (D^*e and $D^*\mu$) when our background samples are impure. (For instance, the $D^*\mu$ signal sample contains a significant subpopulation of fake-μ candidates. The fake-μ control sample, however, also contains combinatoric, continuum, and uncorrelated lepton candidates that will distort the overall behavior of the control sample unless we separate them out.) In the coming Chapters, we will use the $m(D^*) - m(D^0)$ distribution and and other information to evaluate the signal and background quality of each event. The cartoon in Fig. 8.7 illustrates our approach: by simultaneously considering all the subsamples of the Data defined above, we can extract reliable estimates of the signal probability for each candidate.

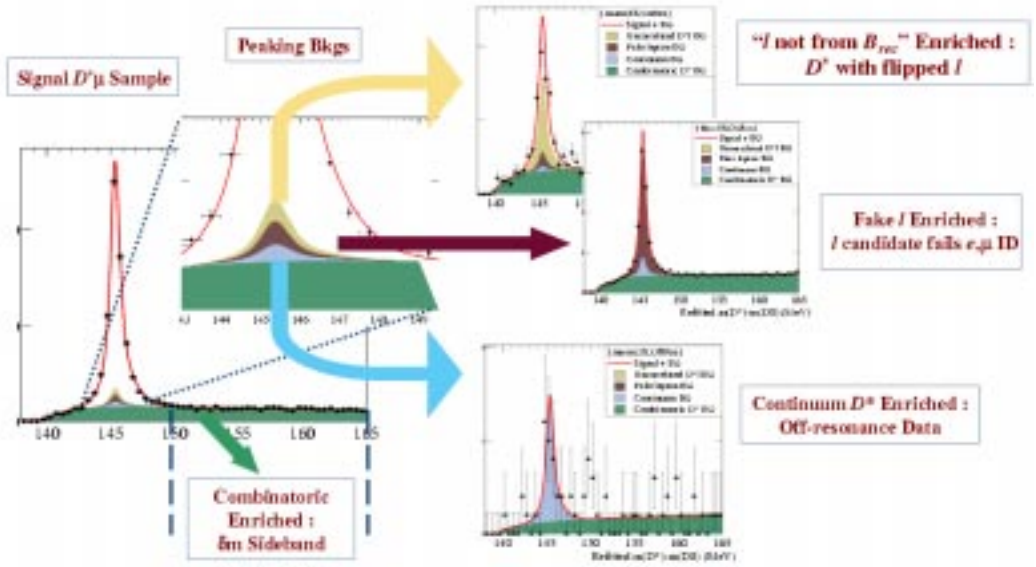


Figure 8.7: Artist's conception of the use of multiple samples to properly account for the backgrounds in the signal region. In the $m(D^*) - m(D^0)$ spectrum, the $B \to D^* \ell \overline{\nu}_\ell$ signal is concentrated in the peak region, which also includes combinatoric, continuum, fake lepton, and uncorrelated lepton backgrounds. The additional background control samples shown in the blowout figures (whose relative efficiencies we calculate in Chapter 10) clearly show the cross-contamination in the background control samples.

8.4 Implementation

For gritty details, please see Ref. [41]. In the interest of clarity, we note the following features of the technology used to select the final event sample:

- We use the `BToDstarlnuVTight` tagbit, in Stream6 of the kanga files which uses the `BToDstarlnuLoose` tagbit.
- The Data is processed in Release 8.8.0x (*i.e.,* tagbits were calculated using Release 8), with the final event reconstruction and selection occurring in `analysis-11` based executables.
- The analysis package `DstarlnuUser V00-04-00` is the primary event handling

engine for this analysis. We use `DstarlnuUser` to process the reduced collections of Data (stored in the Objectivity database) into ntuples, which were filtered to ascii files, the final format used for fitting.

Chapter 9

Analysis Strategy

We have described how the Data sample is collected and selected. We now turn to the task of extracting the most precise, unbiased measurements of the B^0 lifetime and mixing parameters. As described earlier, the events selected for inclusion in the Data sample are not all signal. In fact, as much as 20% of the selected events which satisfy nominal signal criteria are due to backgrounds from incorrect D^*'s and leptons or cross-feed from other processes. In order to properly measure the time evolution we need to characterize and then parameterize these contributions to the sample. To improve our precision in understanding these backgrounds and to reduce our dependence on systematic assumptions, our study includes the additional control samples described in the previous chapter.

The strategy for this measurement is to build up a probability density function (the master pdf) that models the distribution of all selected events (signal and background) in several key variables and then apply the method of maximum likelihood to estimate the parameters of the model which best describe the Data. At the basic level, the master pdf describes the distribution of events in the continuous variables δm and Δt.

Not all candidates are identically distributed, of course, so we divide the entire sample into subgroups where events behave similarly. Considering behavior in δm and Δt space, the Dataset is separated into 360 distinct subsamples. Each sample is not characterized independently of all others; for instance, some classes of events share Δt resolution models, but have different distributions in $m(D^*) - m(D^0)$, etc.

We use the computed per-event probabilities (extracted from fits to the δm spectra and other information) to weight events in a combined master pdf that describes the Δt distribution for each of the five physics processes (one signal plus four background sources).

To prevent the introduction of bias due to expectations of the experimenters, we insert an interface layer between the analysis software and the users that "blinds" the values of τ_{B^0} and Δm_d. The quantities used in the fitting are those preferred by the Dataset under study, but they are offset by a fixed random amount so that the user can compare shifts in the central value but not the absolute scale.

9.1 Overview

At the top level, this study measures the B^0 lifetime and mixing frequency by:

1. Selecting events that pass criteria consistent with the kinematics of semileptonic B decays

2. Reconstructing a $B_{D^*\ell}$ and inclusively identifying the $B_{\rm tag}$ candidate in each event

3. Extracting the decay-time difference Δt and computing the mixing status for each event

4. Assigning a set of per-event probabilities to each event based on the $m(D^*) - m(D^0)$ value and other candidate characteristics: signal, combinatoric background, continuum peaking background, fake lepton peaking background, and uncorrelated lepton peaking background probabilities

5. Building an overall likelihood model for the Data sample that incorporates the per-event probabilities to weight pdfs describing the Δt distribution of each signal and background source

6. Minimizing the negative log likelihood to obtain the best-fit parameter values

7. Performing a series of consistency and validation checks to substantiate the measurement.

The previous chapters of this document described Steps 1–3. We outline the strategy for Steps 4-6 here. The goal, again, is to write down a probability distribution function that describes the entire Data, including both signal and backgrounds. To preface the discussion, consider a cartoon version of the full pdf we'll introduce next.

$$\begin{aligned} Prob(\delta m, \Delta t) \sim\ & f^{\rm comb} F^{\rm comb}(\delta m) G^{\rm comb}(\Delta t) + \\ & (1 - f^{\rm comb}) F^{\rm peak}(\delta m) [f^{\rm pkgbkg} G^{\rm pkgbkg} + (1 - f^{\rm pkgbkg} G^{\rm sig}] \ . \end{aligned} \tag{9.1}$$

We see the two δm pdfs for the peak and combinatoric part of the $m(D^*) - m(D^0)$ spectrum weighting the Δt distributions. * These tell us the "quality" of the D^* candidate. † Because we cannot separate peaking background from signal based on the δm fits alone, the $F(\delta m)$ function multiplying them is common; however the background fraction $f^{\rm pkgbkg}$ weights them appropriately to form the combined peaking-δm behavior in Δt-space. (The determination of the background fractions will be described in the next chapter.) This toy likelihood is sufficient for the case where all events in each class (signal, combinatoric, or peaking background) have identical Δt probability distributions. As we shall see, this is not the case, and the situation becomes quite a bit more complex. For the moment, though, the general task is merely to:

- Determine $F^{\rm comb}$, $F^{\rm peak}$ and $f^{\rm comb}$ from fits to the $m(D^*) - m(D^0)$ spectrum,
- Determine $G^{\rm comb}$, $G^{\rm pkgbkg}$, and $G^{\rm sig}$ from fits to the Δt distributions, taking advantage of the enriched background control samples to improve our precision,
- Determine $f^{\rm pkgbkg}$ by a variety of techniques which exploit known relationships between the signal and background control sample,
- Fit the final model to the Δt distribution of the full Dataset.

9.1.1 The Master Model

Our realistic goal is to perform an unbinned fit simultaneously to events in each of the 12 signal and control samples (indexed by s) that are further subdivided into

*Recall that we can use the $m(D^*) - m(D^0)$ spectrum to evaluate the probability that we have a true D^* in the $D^*\ell$ candidate. Fake D^* candidates are called *combinatoric backgrounds*; fake $D^*\ell$ candidates with good D^*'s are called *peaking backgrounds* because they "look good" in the $m(D^*) - m(D^0)$ projection as they peak in the signal region.

†We use the δm variable to evaluate candidate quality rather than the missing invariant mass or $\cos\theta_{B,D^*\ell}$ because it is easier to parameterize and understand, and the separation of background contributions physics class is straightforward.

30 subsamples (indexed by c) using a likelihood [‡]

$$\mathcal{L} = \prod_{s=1}^{12} \prod_{c=1}^{30} \prod_{k=1}^{N(s,c)} P_{s,c}(\vec{x}_k \,;\vec{p}) \;, \tag{9.2}$$

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 $\vec{p} = (f^{\rm comb}_{s,c}, \vec{p}^{\,\rm comb}_{s,c}, \vec{p}^{\,\rm peak}_{c}, \vec{q}^{\,\rm comb}_{s,c}, f^{\rm pkg}_{s,c,1}, f^{\rm pkg}_{s,c,2}, f^{\rm pkg}_{s,c,3}, \vec{q}^{\,\rm pkg}_{s,c,1}, \vec{q}^{\,\rm pkg}_{s,c,2}, \vec{q}^{\,\rm pkg}_{s,c,3}, \vec{q}^{\,\rm sig}_{c})$ as

$$\begin{aligned} &P_{s,c}(\delta m, \Delta t, \sigma_{\Delta t}, g \,;\vec{p}) = \\ &f^{\rm comb}_{s,c} \cdot \mathcal{F}^{\rm comb}(\delta m \,;\vec{p}^{\,\rm comb}_{s,c}) \cdot \mathcal{G}^{\rm comb}(\Delta t, \sigma_{\Delta t}, g \,;\vec{q}^{\,\rm comb}_{s,c}) + \left(1 - f^{\rm comb}_{s,c}\right) \cdot \mathcal{F}^{\rm peak}(\delta m \,;\vec{p}^{\,\rm peak}_{c}) \cdot \\ &\left[\sum_{j=1}^{3} f^{\rm pkg}_{s,c,j} \cdot \mathcal{G}^{\rm pkg}_{j}(\Delta t, \sigma_{\Delta t}, g \,;\vec{q}^{\,\rm pkg}_{s,c,j}) + \left(1 - \sum_{j=1}^{3} f^{\rm pkg}_{s,c,j}\right) \cdot \mathcal{G}^{\rm sig}(\Delta t, \sigma_{\Delta t}, g \,;\vec{q}^{\,\rm sig}_{c}) \right] \;, \end{aligned} \tag{9.3}$$

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. 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. Functions labelled with $\mathcal{F}$ describe the probability of observing a particular value of δm while functions labelled with $\mathcal{G}$ give probabilities for values of Δt and $\sigma_{\Delta t}$ in category g. Parameters labelled with f describe the relative contributions of different types of events. Parameters labelled with $\vec{p}$ describe the shape of a δm distribution, and those labelled with $\vec{q}$ describe a (Δt, $\sigma_{\Delta t}$) shape.

Note that we make explicit assumptions that the δm peak shape, parameterized by $\vec{p}^{\,\rm peak}_{c}$, and the signal (Δt, $\sigma_{\Delta t}$) shape, parameterized by $\vec{q}^{\,\rm sig}_{c}$, 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

[‡]See the following chapters for a discussion of the division of the Data into subsamples. As foreshadowed, these 360 subdivisions are necessary because not all event behave identically in δm and Δt space.

are selected in the signal sample or appear as a background in a control sample.

Because of these assumptions of independence, the analysis proceeds in two stages: (1) determine the best models to describe the $m(D^*) - m(D^0)$ distribution for all events and calculate the per-event probabilities with those fit results, and (2) fit the Δt distributions of all events using the per-event weights determined from the previous step. We do not expect to have the computational power or precision to perform both the δm and Δt fits simultaneously. Note then, that the we are not using the Δt structure of the events to distinguish signal from background (except in the case of charged B semileptonic decays which contaminate the signal region).

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 10.2).

9.1.2 Sharing or Splitting Parameters

The task of Chapters 10-11 is to determine the most effective models for the δm and Δt distributions, respectively. Note that with the general pdf as defined above in Eq. 9.3, the complete space of parameters can allow for a maximum of $3,960$ [§] parameters to describe the 360 different δm distributions and per-event probabilities, and another $9,362$ [¶] to describe the potentially unique signal plus background Δt distributions in of each the 360 different categories. The real challenge, then, is to use physical argument, empirical observation, and common sense to reduce this ungodly number by "sharing" parameters between subsamples that behave similarly.

This assertion may sound preposterous, but it is the general case of a usually trivial situation. 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

[§] Assuming 5 peak plus 2 combinatoric background plus 4 independent per-event probabilities per distinct subsample

[¶] Assuming 2 overall physics parameters, and then 6 signal parameters plus 4×5 background parameters per subsample

this because we find that different subsamples have significantly different behaviors. 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 identified in Chapter 8. This approach provides the maximum statistical sensitivity, in principle, but suffers from significant book-keeping overhead and problems with fitting small samples.

We choose to compromise by "sharing" parameters between subsamples. Technically, this means that in the fitting procedure, the likelihood is re-calculated for both samples when varying a parameter that is shared between samples. In this analysis, we alternatively think of sharing parameters as the opposite of "splitting" them. For instance, if two processes share a similar Gaussian behavior, but with different widths and identical means, we could imagine performing a fit to all events but with two different "copies" of the width parameter, each restricted to the distinct subset of events to which it corresponds. In this way, the shared mean gains the statistical power of the full dataset, while the widths are allowed to be different. The methodology in this study then, is as follows (which closely follows the technical implementation as well):

1. Identify the prototype model for the full sample of events, *i.e.,* the $m(D^*) - m(D^0)$ lineshape model of double Gaussian peak plus combinatoric background term will be used to model the δm spectrum

2. Separate the full sample into subgroups (as intuition directs) that might exhibit different detailed structure in the common form, *i.e.,* events with soft $\pi^{\pm}$ tracks that reach the DCH will have better momentum resolution than those that do not

3. Fit the isolated subsamples and compare the best-fit parameters, looking for statistically significant differences, *i.e.,* the peak shape parameters for the two types of soft π candidates are separated by as much as 5σ

4. Repeat the fit to the full sample, sharing the consistent parameters across all

subsamples and "splitting" those between subsamples that behave differently (in this way, each "split" of a prototype parameter actually adds one more floating variable to the fit).

In summary, note that this procedure involves selecting a prototype model (a pdf which uniformly describes all events in the sample), and then splitting parameters to accommodate particular features of subclasses of events. Using this procedure, as illustrated in the upcoming chapters, we are able to reduce the full space of parameters from more than $13,000$ to less than $2,000$, of which only a fraction are freely floating in the δm or Δt fits at any stage in the analysis.

9.2 Notation

The final model contains more than 70 parameters. Some parameters are split according to different combinations of categories. To avoid confusion, we define the notation used in this study for parameters in the Δt models. In general, a parameter is denoted by

$$p_{\mathrm{cat}}^{\mathrm{label}},$$

where the subscript represents the category values (if any) to which this parameter is restricted, and the superscript is an extra label for distinguishing different roles of a certain type of parameter in a model, if necessary.

Recall the motivation for "splitting" a parameter: in some cases, we expect different groups of events to behave according to the same Δt model, but with different parameter values. By "splitting" a parameter according to a certain candidate property, we assign the parameter a domain of context or relevance that is the subset of data possessing a certain value of that property, and create additional instances of the parameter for the other subsets of the data with different values for that property. For instance, we often "split" resolution model parameters by tagging category. This means that each tagging category will have identical resolution model structure, but that certain parameters of that model will be unique to the LTag or KTag events, etc.

For the signal physics model ($G^{(\mathrm{sig})}$ in Chapter 11),

- τ_{B^0} and Δm_d represent B^0 lifetime and mixing frequency, respectively.
- The mistag fraction is noted as ω and is split by tagging category (`LTag`, `KTag` and `NTx` are short-handed as "`LT`", "`KT`", and "`Nx`") :
  - ω_{LT}, ω_{KT}, ω_{N1}, ω_{N2}, ω_{N3} are the mistag probabilities averaged over B and $\overline{B}$.
  - Mistag for `KTag` is a linear function of $\sigma_{\Delta t}$: $\omega_{\mathrm{KT}} = m_{\mathrm{KT}} \cdot \sigma_{\Delta t} + \omega_{\mathrm{KT}}^{\mathrm{offset}}$. The slope m_{KT} is always given in units of ps^{-1}.
  - $\Delta\omega$ is the difference in mistag probabilities for B and $\overline{B}$ (see next section).
- f_{B^+} represents the charged B fraction in signal events (see Chapter 11.2.2).

For each of the background Δt models, we use a "physics model" convoluted with a "resolution function" to describe the Δt distribution of the background events. Although for background events, the true Δt and resolution are not well-defined, we still use the names "physics model" and "resolution function" just to refer to the mathematical form of the Δt models.

A background Δt model consists of a prompt term plus either a lifetime term or an oscillating term, or a prompt term by itself. These three terms can be written as

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

where $\delta(\Delta t)$ is a δ-function, g is a sign representing mixing status, i.e., $g = +1\,(-1)$ for unmixed (mixed) events, and τ^{bkg} and Δm^{bkg} are effective lifetime and mixing frequency for this particular background. The combined physics model for a background model, then, is typically one of the following:

- $f^{\mathrm{life}} \cdot \mathcal{G}_{\mathrm{phys}}^{\mathrm{life}} + (1 - f^{\mathrm{life}}) \cdot \mathcal{G}_{\mathrm{phys}}^{\mathrm{pmt}}$

- $f^{\rm osc} \cdot \mathcal{G}^{\rm osc}_{\rm phys} + (1 - f^{\rm osc}) \cdot \mathcal{G}^{\rm pmt}_{\rm phys}$

- $\mathcal{G}^{\rm pmt}_{\rm phys}$

We use three different resolution functions for Δt models, GExp+G, G+G+G and G+G functions. The basic ingredients are a Gaussian, G (Eq. 6.9, Sec. 6.4) and the sum of a single Gaussian and the same Gaussian convoluted with a one-sided ($\delta\Delta t < 0$) exponential, $G \otimes E$.

$$\begin{aligned}
&\mathcal{R}_{\rm G+G}(\delta\Delta t, \sigma_{\Delta t}; b^1, s^1, f^1, b^{\rm out}, s^{\rm out}) \\
&\qquad = f^1 \cdot G(b^1\sigma_{\Delta t}, s^1\sigma_{\Delta t}) + (1 - f^1) \cdot G(b^{\rm out}\sigma_{\Delta t}, s^{\rm out}\sigma_{\Delta t}) \ , \\
&\mathcal{R}_{\rm G+G+G}(\delta\Delta t, \sigma_{\Delta t}; b^1, s^1, f^1, b^2, s^2, b^{\rm out}, s^{\rm out}, f^{\rm out}) \\
&\qquad = f^1 \cdot G(b^1\sigma_{\Delta t}, s^1\sigma_{\Delta t}) + (1 - f^1 - f^{\rm out}) \cdot G(b^2\sigma_{\Delta t}, s^2\sigma_{\Delta t}) + f^{\rm out} \cdot G(b^{\rm out}, s^{\rm out}) \ , \\
&\mathcal{R}_{\rm GExp+G}(\delta\Delta t, \sigma_{\Delta t}; s^1, \kappa, f^1, b^{\rm out}, s^{\rm out}, f^{\rm out}) \\
&\qquad = f^1 \cdot G(0, s^1\sigma_{\Delta t}) + (1 - f^1 - f^{\rm out}) \cdot G(0, s^1\sigma_{\Delta t}) \otimes E(\kappa\sigma_{\Delta t}) + f^{\rm out} \cdot G(b^{\rm out}, s^{\rm out}) \ ,
\end{aligned}$$

where the dependent variable $\delta\Delta t$ in G and E is not shown explicitly. The function E is defined as $E(x; a) = \frac{1}{a} \exp^{x/a}$. (Note that we deliberately do not scale the bias and width parameters of the third term, the outlier Gaussian, when present. This is because it addresses physics events which are so poorly reconstructed (mistaken, actually), that the vertex error estimate $\sigma_{\Delta t}$ has little to do with the true uncertainty.)

The same type of function may be used for different backgrounds, therefore the same symbol will represent different parameters in different background models. In principle we should add another label to distinguish them, but this will make the notation very cumbersome. Fortunately we rarely need to put different backgrounds in the same context. So we will not add this extra label unless it is necessary for clarity.

Finally, we typically represent the change in a parameter due to a variation in the fit method, data sample or some other procedure as $\delta(\text{par}) = \text{par}^{\rm changed} - \text{par}^{\rm unchanged}$.

9.3 Blinding

The physics parameters τ_{B^0} and Δm_d were kept hidden until all analysis details and the systematic errors were finalized, to eliminate experimenter's bias. This was accomplished directly by using features of the analysis technology. [59] The key is to allow the fitting application to work with the true representation of the parameter but scrambling the value by a fixed but unknown amount when showing the user. In this study, we chose to scramble τ_{B^0} by shifting it a random amount as selected from a Gaussian distribution of width 0.06 ps, while we chose a scale of 0.04 ps^{-1} for Δm_d. This "blinding" of the fit result prevented us from attaching any significance to the numerical value of the result (*i.e.,* we could not directly compare it to the world average).

However, statistical errors on the parameters and changes in the central values of the physics parameters due to variations in the analysis were not hidden. For systematic studies then, we list the difference in the central values of the fit results so the blinding offset cancels out. Fits to Monte Carlo data samples were not blinded except during tests of the blinding procedure implementation in our analysis framework.

Chapter 10

Models for Per-event Probability

We use the $m(D^*) - m(D^0)$ spectra, BABAR particle identification tables, and a small amount of Monte Carlo to determine several different probabilities for each event: signal probability, combinatoric background probability, and the probability of being each of three types of peaking background. We use these probabilities to weight each event in the master pdf which then describes its behavior in Δt space. We find that the δm spectrum depends on a number of event characteristics, including:

- *the magnitude of soft pion's transverse momentum affects the quality of track reconstruction because there is a threshold for reaching the precision tracking DCH (2 choices)*
- *the D^0 decay mode affects the level of combinatoric background (number of false D^* candidates) and the resolution of $m(D^*) - m(D^0)$ peak (3 choices)*
- *the particle identification hypothesis for the lepton candidate is correlated with the combinatoric background level because of the hadron misidentification rates (3 choices)*
- *the hint about the nature of $B_{\rm tag}$ offered by the tagging category is correlated with the background level as well due to different probabilities for mixing up tracks from the two B's (4 choices).*

By subdividing the Data sample into smaller groups, we find that our description of the δm spectrum becomes more precise, and we are able to accommodate the contaminated background control samples, i.e., the background in the background.

10.1 Strategy

Given that we've selected the events described in Chapter 8, we next address the question of evaluating our confidence in the signal probability on an event-by-event basis. Because we do not reconstruct the neutrino, the most stringent test we apply to the $B_{\rm rec}$ kinematics is the $m(D^*) - m(D^0)$ value, which should be strongly peaked for signal (see Fig. 10.1.) We can use the distribution of $\delta m = m(D^*) - m(D^0)$ to evaluate the "good D^*" probability of the candidate. The *combinatoric background* due to events with a misreconstructed D^* candidate can be distinguished from events with a real D^* in a plot of the mass difference $m(D^*) - m(D^0)$, where $m(D^*)$ is the mass of the D^* candidate with the D^0 mass constrained to its true value, and $m(D^0)$ is the true D^0 mass. Our signal mode ($B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$) requires more than just a "good D^*'," however; the D^* candidate must be paired with a lepton in a manner consistent with the signal mode's dynamics. As shown in Fig. 10.1, this means that events which have δm in the peak region can actually still be backgrounds to our signal, so-called *peaking backgrounds*. These peaking backgrounds are from three main sources, and represent events which are selected in the δm "good D^*" region but do not correspond to physics of the signal mode:

- continuum: the D^* candidate is from continuum $c\overline{c}$ production
- fake lepton: the lepton associated the D^* is a falsely identified lepton (*i.e.,* a hadron, or perhaps a mis-typed lepton)
- uncorrelated lepton: the lepton associated with the D^* is either a secondary lepton from $B_{\rm rec}$, or a lepton from the $B_{\rm tag}$ decay.

Our primary tool for characterizing the backgrounds that contribute to our signal sample is, then, to fit the distribution of the $D^* - D^0$ mass difference δm and use it to evaluate the per-event reliability of each candidate. In general, our fit to the δm distribution has three ingredients:

1. a model for the shape of the peak due to candidates with a correctly reconstructed D^*,

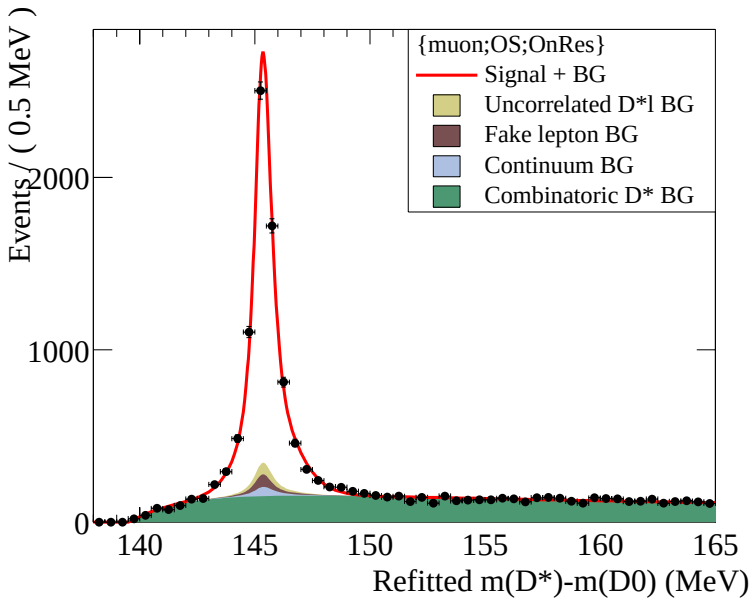


Figure 10.1: Plot of $\delta m = m(D^*) - m(D^0)$ for $D^*\mu$ candidates from Data, which we use to evaluate the quality of the D^* candidate. The colored regions represent the estimated contributions from the different background sources; that is, events which have good D^* candidates but are not proper $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ candidates. Note the three types of peaking background.

2. a model for the shape of the combinatoric background due to candidates with an incorrectly reconstructed D^*, and

3. a parameterization of the yields of peak and combinatoric components in each sample.

The general program for using these fits is to first determine the yields in different subsamples of the signal and control samples, and then to use these yields to calculate the fractions of different sources of peaking background. 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.

In studying the sample, we observe that the peak and combinatoric background shapes describing the distribution of δm depend on several different characteristics of the event. In other words, we find that it is appropriate to split the shape parameters amongst several subgroups so that we maximize our sensitivity to the $m(D^*)-m(D^0)$ information. We also find that the peaking background fractions depend on other candidate characteristics, most notably the particle identification hypothesis for the lepton candidate and the angular correlation between the D^* and ℓ. Table 10.1 summarizes the full set of candidate attributes which we might consider when subdividing the Data sample.

This allows subdivision into 360 samples. In the unbinned maximum likelihood fits to the $m(D^*)-m(D^0)$ and (Δt, $\sigma_{\Delta t}$) distributions, individual fit parameters are shared among different sets of subsamples based on physics motivation and observations in the data.

10.2 Fitting the $m(D^*) - m(D^0)$ Shapes

We perform a simultaneous fit to the $m(D^*) - m(D^0)$ distributions for all 360 subsamples. The peak due to real D^* candidates is modeled by the sum of two Gaussian distributions; the mean and variance of both the Gaussian distributions, as well as the relative normalization of the two Gaussians, are free parameters in the fit:

$$F^{\rm peak}(\delta m) = \sum_{k=1}^{2} \frac{f_k}{\sqrt{2\pi}\sigma_k} exp\left(-\frac{(\delta m - \delta m_k)^2}{2\sigma_k^2}\right) \tag{10.1}$$

where $f_1 + f_2 = 1$. We model the shape of the combinatoric background with the function

$$F^{\rm comb}(\delta m) = \frac{1}{N}\left[1 - \exp\left(-\frac{\delta m - \delta m_0}{c_1}\right)\right]\left(\frac{\delta m}{\delta m_0}\right)^{c_2}, \tag{10.2}$$

where $\delta m \equiv m(D^*) - m(D^0)$, N is a normalization constant, δm_0 is a kinematic threshold equal to the mass of the $\pi^{\pm}$, and c_1 and c_2 are free parameters in the fit. Fig. 10.2 shows the effects of varying the two free parameters in the model. The exponential term enforces the turn on at the $\pi^{\pm}$ mass and asymptotically approaches

Attribute	N_{cat}	Description
`angCut`	2	What are the angular correlations between the D^* and ℓ? (1) candidates pass the "opposite-side" (OS) cuts (2) candidates pass the "same-side" (SS) cuts
`onOffRes`	2	What center-of-mass energy was the event recorded at? (1) near the $\Upsilon(4S)$ resonance (2) below the $\Upsilon(4S)$ resonance
`leptID`	3	Is the kinematically-selected ℓ candidate lepton-like? (1) candidate passes "very-tight" electron ID (2) candidate passes "very-tight" muon ID (3) candidate fails "loose" electron and muon ID
`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 $D^0 \to K^0_S\pi\pi$
`tagCat`	5	How was the flavor of the non-$D^*\ell$ B determined? (1) using the `LTag` category algorithm, (2) using the `KTag` category algorithm, (3) using the `NT1` category algorithm, (4) using the `NT2` category algorithm, (5) using the `NT3` category algorithm

Table 10.1: Attributes used to subdivide the signal and control samples for the purposes of background characterization and extraction of lifetime and mixing parameters.

the value 1 for large values of δm. The power law factor allows the PDF to match the slope of the combinatoric tail at large values of δm.

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

$$F(\delta m) = f_{\rm peak} \cdot F^{\rm peak}(\delta m) + (1 - f_{\rm peak}) \cdot F^{\rm comb}(\delta m) \; . \tag{10.3}$$

The value of $f_{\rm peak}$ corresponds to the fraction of peak events in the entire sample

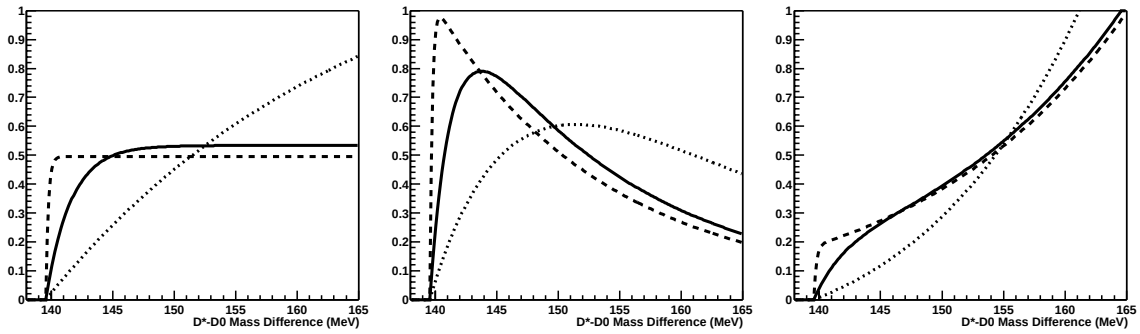


Figure 10.2: Effects of the combinatoric model shape parameters c_1 and c_2. Each plot is calculated using a different value of c_2: 0 (left), -10 (middle), +10 (right). The curves within each plot represent different values of c_1: 2 MeV (solid), 0.2 MeV (dashed), or 25 MeV (dotted). The starting values in a fit are $c_1 = 2$ MeV and $c_2 = 0$, and the allowed ranges are 0.01 MeV $< c_1 <$ 50 MeV and $-20 < c_2 < +20$.

being fit (including the δm sideband).

Once we have specified our model $F(\delta m)$, we next identify subsamples that will be fit with a single set of shape parameters. An initial fit is done to determine the shape parameters describing the peak and combinatoric background. Separate values of the five parameters describing the shape of the peak are used for the six subsamples defined by whether the π^- candidate is tracked in the SVT only or in the SVT and DCH, and the three types of D^0 decay modes. 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. Each of these six groups that share peak parameters is further subdivided into 12 subgroups that each share a common set of the two combinatoric background shape parameters. Ten of these 12 subgroups are defined by the five tagging categories for the large signal samples and for the fake-lepton control samples, in on-resonance data. The other two subgroups are defined as same-side, off-resonance samples and all off-resonance samples.

10.3 Calculating Yields

Once the peak and combinatoric background shape parameters have been determined, we fix the shape parameters and determine the peak and combinatoric background yields in each of the 360 subsamples with an unbinned extended maximum-likelihood fit. The total peak yields in the signal sample and each background control sample are then used to determine the amount of true signal and each type of peaking background in the $m(D^*) - m(D^0)$ peak of each sample as follows.

Each of the 12 samples described above are further divided into 30 subsamples according to the following characteristics that may affect the $m(D^*) - m(D^0)$ or Δt distributions.

1. The π^- from the D^* decay is reconstructed in the SVT only, or in the SVT and DCH (two choices): The $m(D^*) - m(D^0)$ resolution is worse when the π^- is reconstructed only in the SVT; the "peak" region is nominally 143-148 MeV for SVT-only candidates, and a narrower 144-147 MeV for SVT+DCH candidates.

2. The D^0 candidate is reconstructed in the mode $K\pi$ or $K\pi\pi^0$ or ($K\pi\pi\pi$ or $K^0_S\pi\pi$) (three choices): The level of contamination from combinatoric background and the $m(D^*) - m(D^0)$ resolution may depend on the D^0 decay mode.

3. The b-tagging information used for the other B ($B_{\rm tag}$) (five choices; see Chapter 7): The level of contamination from each type of background and the Δt resolution parameters may depend on the tagging information.

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 $\delta m = 147.5$ MeV (that is, near the edge of the peak region). By assuming an average shape for the peak and combinatoric background, 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 as we know that such candidates have a narrower peak (only 144-147 MeV) compared to SVT-only candidates. In short, by considering additional information about the candidate, we are able to increase the precision of our signal (well, really just peak here) versus background discrimination.

We combine our peak and combinatoric models into an unbinned extended maximum-likelihood fit using

$$\mathcal{L}(N'_{peak}, N'_{comb}) = \prod_{\text{evts } j} \left\{ \frac{N_{peak}}{N_{tot}} \cdot f_{peak}(\delta m_j) + \frac{N_{comb}}{N_{tot}} \cdot f_{comb}(\delta m_j) \right\} \cdot P(N_{tot}) \ , \quad (10.4)$$

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

$$\begin{aligned} N'_{peak} &\equiv N_{peak} \cdot \int_{\delta m_1}^{\delta m_2} f_{peak}(\delta m) d(\delta m) \\ N'_{comb} &\equiv N_{comb} \cdot \int_{\delta m_1}^{\delta m_2} f_{comb}(\delta m) d(\delta m) \ . \end{aligned} \quad (10.5)$$

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.

10.3.1 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'_{comb}, in the δm fits described above.

10.3.2 Continuum Peaking Background

We use the peak yields in off-resonance data, scaled by the relative integrated luminosity for on- and off-resonance data, to determine the continuum-background yields in on-resonance data.

10.3.3 Fake Lepton Peaking Background

Particle identification and misidentification efficiencies for the e, μ, and fake-lepton selection criteria are measured in Data as a function of laboratory momentum, polar angle, and azimuthal angle, for true electrons, muons, pions, kaons, and protons. $B^0\overline{B}^0$ and B^+B^- Monte Carlo simulations are used to determine the measured laboratory momentum, polar angle, and azimuthal angle distributions for true electrons, muons, pions, kaons and protons that pass all selection criteria for $D^*\ell$ candidates, except the lepton (or fake-lepton) identification criteria. These distributions are combined with the measured particle identification efficiencies and misidentification probabilities to determine the momentum- and angle-weighted probabilities for a true lepton or true hadron to pass the criteria for a lepton or a fake lepton in each of the $D^*\ell$ signal and background control samples. We then use these efficiencies and misidentification probabilities, and the observed number of lepton and fake-lepton candidates in data, to determine the number of true leptons and fake leptons (hadrons) in each control sample. This procedure is sophisticated, yet general, and best described in Refs. [41], [54].

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 $D^*\ell$ candidates where the lepton candidate is a true lepton or a true hadron (η_e, η_μ, η_h)

through a set of linear equations:

$$\begin{pmatrix} N_e \\ N_\mu \\ N_f \end{pmatrix} = \begin{pmatrix} \epsilon_e(e;!\mu) & \epsilon_\mu(e;!\mu) & \epsilon_h(e;!\mu) \\ \epsilon_e(!e;\mu) & \epsilon_\mu(!e;\mu) & \epsilon_h(!e;\mu) \\ \epsilon_e(!e;!\mu) & \epsilon_\mu(!e;!\mu) & \epsilon_h(!e;!\mu) \end{pmatrix} \begin{pmatrix} \eta_e \\ \eta_\mu \\ \eta_h \end{pmatrix}. \tag{10.6}$$

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

$$f_e = \eta_h \epsilon_h(e;!\mu)/N_e \tag{10.7}$$

$$f_\mu = \eta_h \epsilon_h(!e;\mu)/N_\mu \tag{10.8}$$

$$f_f = \eta_h \epsilon_h(!e;!\mu)/N_f \tag{10.9}$$

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 subsamples with the same `angCut`.

10.3.4 Uncorrelated Lepton Peaking Background

To determine the number of uncorrelated-lepton events in each sample, we use the relative efficiencies from Monte Carlo simulation for signal and uncorrelated-lepton events to pass the criteria for same-side and opposite-side samples.

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

$$\begin{pmatrix} N_{\rm OS} \\ N_{\rm SS} \end{pmatrix} = \begin{pmatrix} \epsilon_{\rm OS}^{\rm sig} & \epsilon_{\rm OS}^{\rm unc} \\ \epsilon_{\rm SS}^{\rm sig} & \epsilon_{\rm SS}^{\rm unc} \end{pmatrix} \begin{pmatrix} \eta_{\rm sig} \\ \eta_{\rm unc} \end{pmatrix}. \tag{10.10}$$

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

as

$$f_{\rm OS} = \epsilon^{\rm unc}_{\rm OS}\eta_{\rm unc}/N\ _{mathrmOS} \quad (10.11)$$

$$f_{\rm SS} = \epsilon^{\rm unc}_{\rm SS}\eta_{\rm unc}/N\ _{mathrmSS} \quad (10.12)$$

$$(10.13)$$

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. We have assumed that one efficiency matrix is valid for all SVT-only samples and one for all SVT+DCH samples.

10.4 Results

The peak yields and continuum, fake-lepton, and uncorrelated-lepton fractions of the peak yield, as well as the combinatoric fraction of all events in a $m(D^*) - m(D^0)$ signal window, are shown in Table 10.2 for the signal and background control samples in on-resonance data. The peak yields include the peaking backgrounds. The signal window is defined as (143 - 148) MeV for the calculation of combinatoric background fractions.

We illustrate the results of δm analysis with several sample plots here, but refer the reader to Appendix A for a more complete set. Fig. 10.3 shows the results of the first fit stage which estimates the peak shape parameters for the double Gaussians in each of the 6 subsamples of the Data. The following figure (Fig. 10.4) shows the results of the analysis of the combinatoric background shapes within the subsample of candidates with SVT-only soft pion candidates and reconstructed in the $D^0 \to K\pi\pi\pi$ or $D^0 \to K^0_S\pi\pi$ decay mode. Note the variations in the level and shape of the combinatoric background, making the need for separate parameterizations explicit.

Finally, we use the calculated fractions and fitted shapes of each background

Category		Peak Yield	$f_{\rm cont}(\%)$	$f_{\rm fake}(\%)$	$f_{\rm unco}(\%)$	$f_{\rm comb}(\%)$
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	0.00 ± 0.006	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

Table 10.2: Peak yields and continuum, fake-lepton, and uncorrelated-lepton fractions of the peak yield, and the combinatoric fraction of total events in a $m(D^*) - m(D^0)$ signal window for the signal and background control samples in on-resonance data. Peak yields include the peaking backgrounds. The signal window for combinatoric background is defined as (143 - 148) MeV. OS and SS refer to opposite-side and same-side samples; e, μ, and f indicate the type of lepton candidate: electron, muon or fake-lepton.

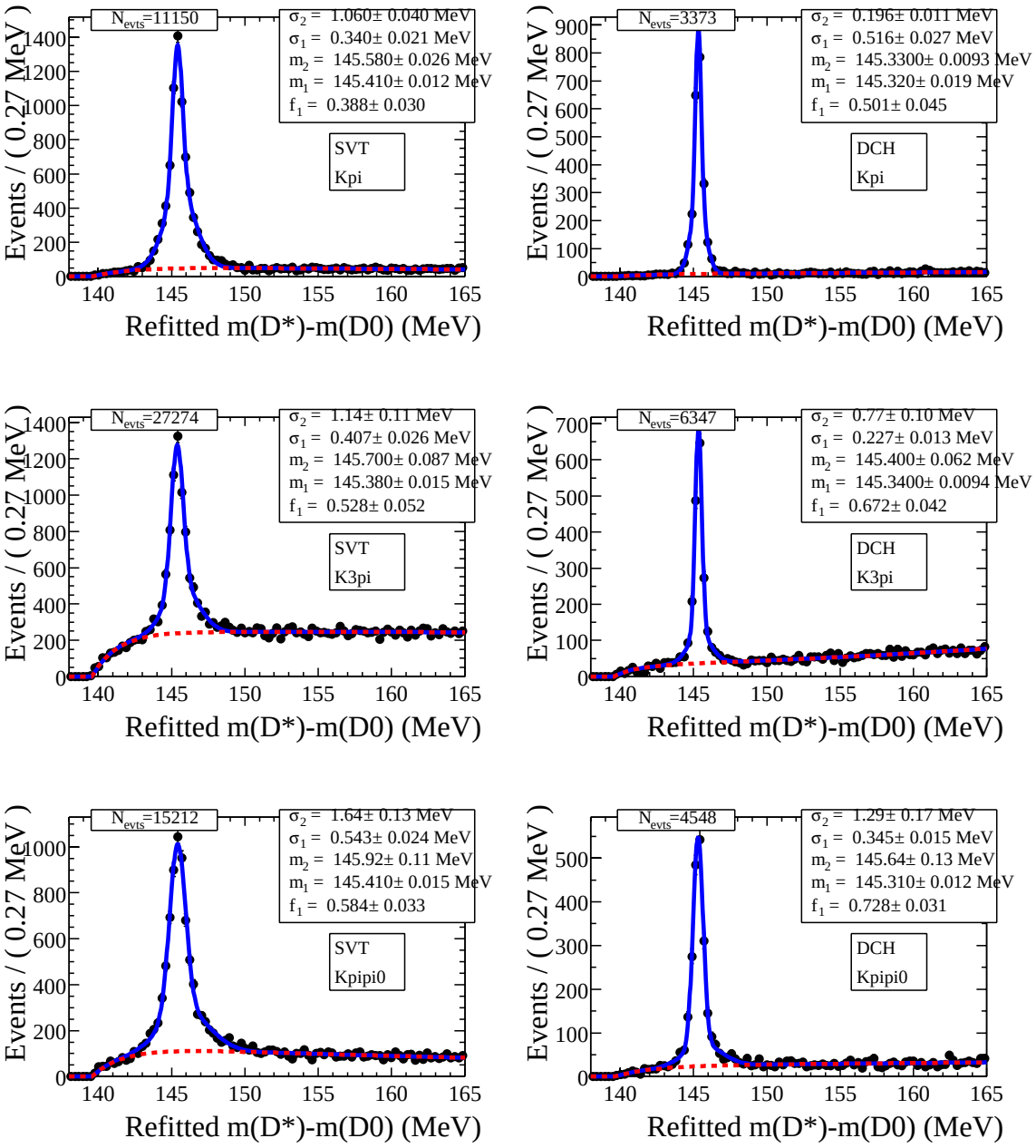


Figure 10.3: Projections of the fit result on Data for the δm peak shape fits described in the text. The 6 plots show the fit on the 6 distinct subsamples of the Data. The dashed curve shows the fitted contribution from combinatorics.

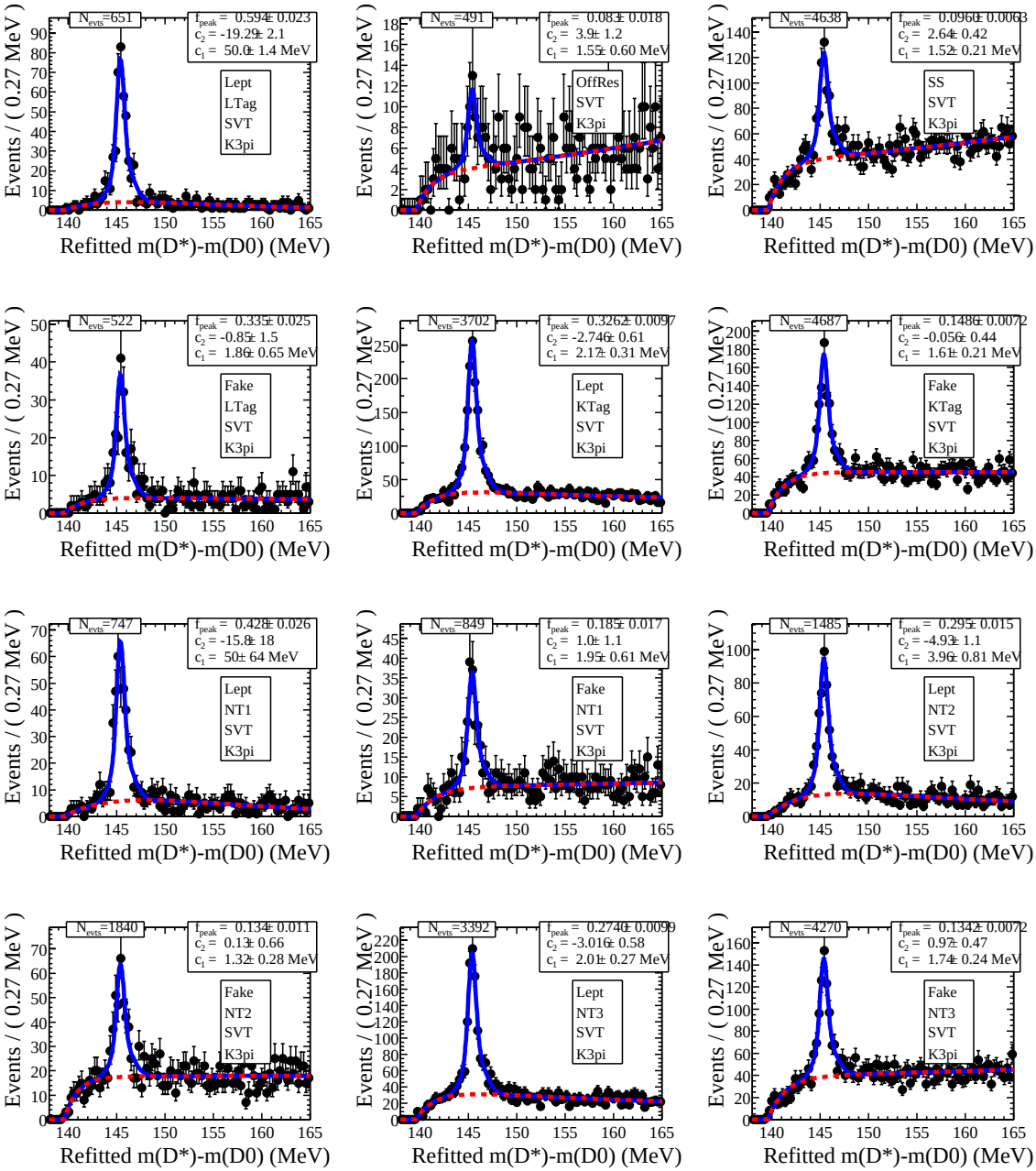


Figure 10.4: Projection of the fit result on Data for the δm combinatoric background fits described in the text. The 12 different plots show the distinct subsamples of the (SVT x ($D^0 \to K\pi\pi\pi$ + $D^0 \to K^0_S\pi\pi$) sample from Data. The dashed curve shows the fitted contribution from combinatorics.

source in each control sample to estimate the probability of each candidate to be due to signal or each type of background (combinatoric, continuum, fake-lepton, or uncorrelated-lepton) when we fit the (Δt, $\sigma_{\Delta t}$) shape to determine the lifetime and mixing parameters. We take advantage of the fact that charged and neutral B decays have different decay-time distributions (because the charged B does not mix) to determine the fraction of charged B background events in the fit to (Δt, $\sigma_{\Delta t}$).

By summing the Data subsamples and the corresponding fitted PDFs, we can form the summary plots as shown in Fig. 10.5. We observe excellent agreement between the Data and the fit result, which indicates that the technology running the individual fits down at the 360 categories level is operating successfully.

For reference, we draw histograms of the per event probabilities as determined by the sequence of fits outlined in this chapter, indicated in Fig. 10.6. These histograms show the "purity" of the sample; because of the vast δm sideband, there is a clean, easy-to-isolate combinatoric background control sample, where as the uncorrelated lepton sample is more complex.

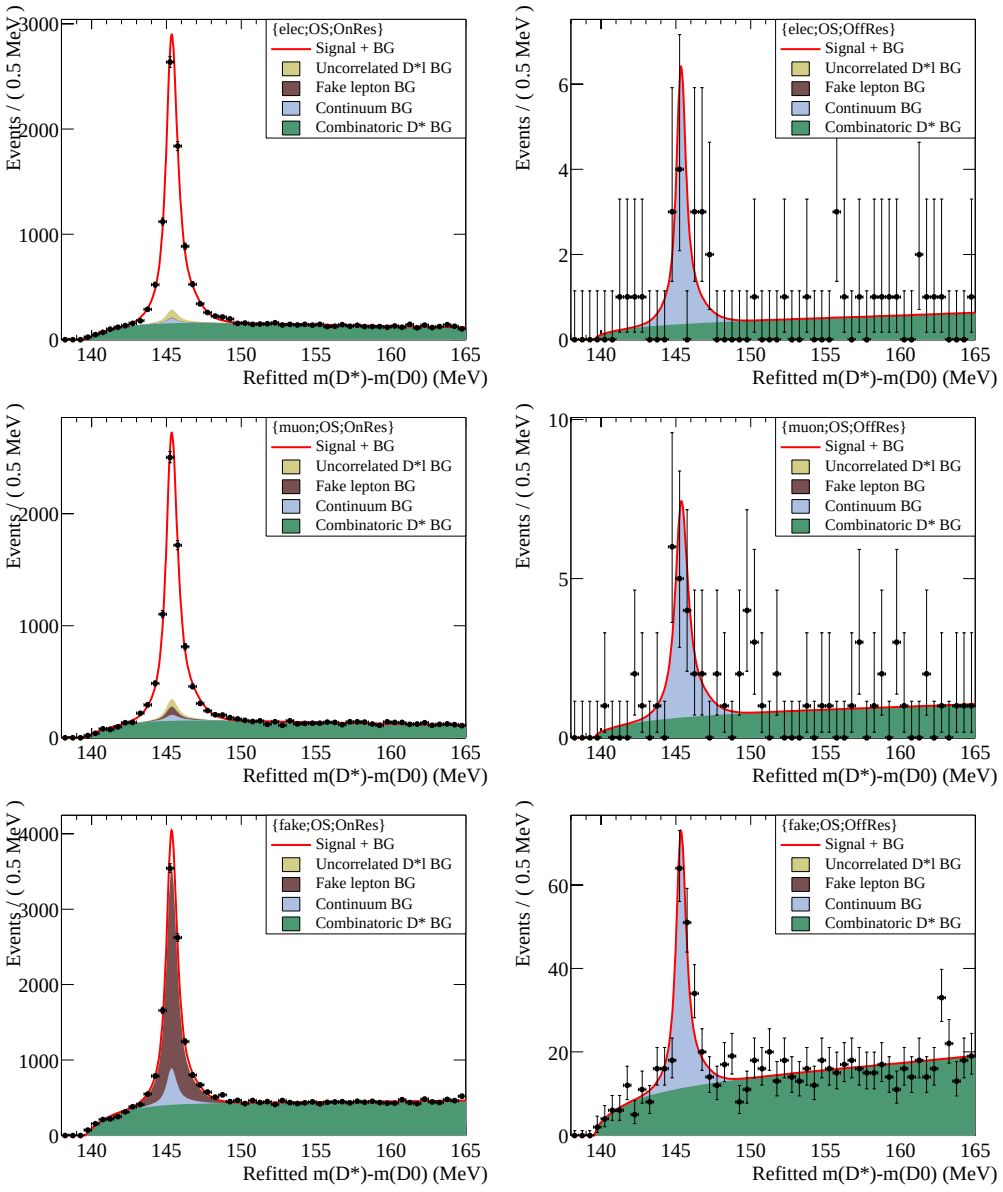


Figure 10.5: Uncorrelated $D^*\ell$, fake ℓ, continuum D^*X, and combinatoric D^* contributions for opposite-side samples. From top to bottom: electron, muon and fake control samples; left column: on-resonance and right column: off-resonance.

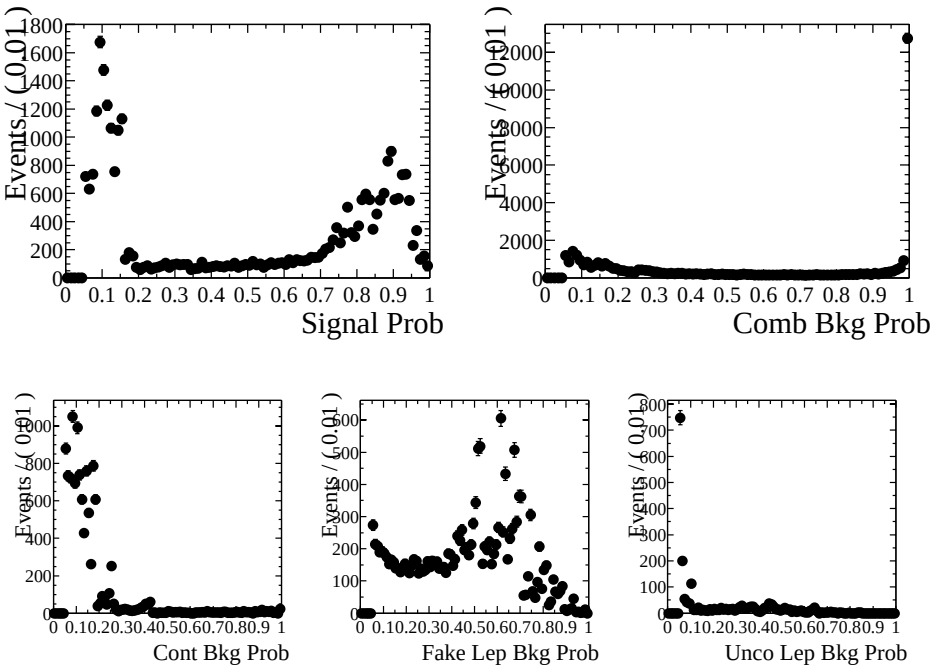


Figure 10.6: Histograms of the per-event signal probabilities as calculated for Data. The zero bin entries have been suppressed by placing a cut at 5% probability. Integrating the contents of the histograms yields the overall purities for each event class in the sample: 21.7% signal, 62.7% combinatoric, 2.8% continuum background, 11.5% fake lepton background, and 1.3% uncorrelated lepton background.

Chapter 11

Models for Time-Dependence

The final fit uses the per-event probabilities developed in the preceding chapters to weight the contributions from each of the five general classes of events (signal, combinatoric background, and the three peaking backgrounds) in the overall Δt spectrum from which we'll extract the physics. We determine the appropriate form for the Δt model for each class of event by studying Monte Carlo truth-matched samples and the control samples in Data. For each class of event, we first find the general prototype model and then determine which parameters need to split between subsamples due to statistically different behaviors. This phase is the first of a bootstrap process wherein we use smaller, enriched subsamples to predetermine behavior and performance before advancing to the full fit in the next chapter.

11.1 Strategy

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 class 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 minimizes 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 class 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 ($\approx$68k events), leaving essentially all parameters free (72, actually) in the fit.

11.2 Signal Models

11.2.1 B^0 Model

Physics

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

$$\mathcal{G}^{\rm sig}(\Delta t, \sigma_{\Delta t}, g\,; \vec{q}_c^{\rm sig}) = \left\{ \frac{1}{4\tau_{B^0}} e^{-|\Delta t_{\rm true}|/\tau_{B^0}} (1 + g(1-2\omega_c)\cos(\Delta m_d \Delta t_{\rm true})) \right\} \otimes \mathcal{R}(\delta\Delta t, \sigma_{\Delta t}; \vec{q}_c)\,, \tag{11.1}$$

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 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 (as motivated in Chapter 7.2). To account for an observed correlation between the mistag rate and $\sigma_{\Delta t}$ in the kaon category (described in Chapter. 7.4), we allow the mistag rate in the `KTag` category to vary as a linear function of $\sigma_{\Delta t}$:

$$\omega_{\rm kaon} = m_{\rm kaon} \cdot \sigma_{\Delta t} + \omega_{\rm kaon}^{\rm offset}\,. \tag{11.2}$$

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

$$\omega_{B^0/\overline{B}^0} = \omega \pm \frac{1}{2}\Delta\omega\,. \tag{11.3}$$

There are 13 free parameters in the complete physics model for all tagging categories.

Resolution

For the Δt resolution model, we use the sum of a single Gaussian distribution and the same Gaussian convolved with a one-sided exponential to describe the core part

of the resolution function, plus a single Gaussian distribution to describe the contribution of "outliers" – events in which the reconstruction error $\delta\Delta t$ is not described by the calculated uncertainty $\sigma_{\Delta t}$:

$$\begin{aligned}\mathcal{R}_{\mathrm{GExp+G}}(\delta\Delta t, \sigma_{\Delta t}; s, \kappa, f, b^{\mathrm{out}}, s^{\mathrm{out}}, f^{\mathrm{out}}) = \\ f \cdot G(\delta\Delta t; 0, s\sigma_{\Delta t}) + (1 - f - f^{\mathrm{out}}) \cdot G(u - \delta\Delta t; 0, s\sigma_{\Delta t}) \otimes E(u; \kappa\sigma_{\Delta t}) \\ + f^{\mathrm{out}} \cdot G(\delta\Delta t; b^{\mathrm{out}}, s^{\mathrm{out}}) \ , \quad (11.4)\end{aligned}$$

where u is an integration variable in the convolution $G \otimes E$. The functions G and E are defined as in Eqs. 6.9-6.11 and Chapter 6.4:

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

and

$$E(x; a) \equiv \begin{cases} \frac{1}{a} \exp(x/a) & \text{if } x \leq 0, \\ 0 & \text{if } x > 0. \end{cases}$$

Since the outlier contribution is not expected to be described by the calculated error on each event, the last Gaussian term in Eq. 11.4 does not depend on $\sigma_{\Delta t}$. However, in the terms that describe the core of the resolution function (the first two terms on the right-hand side of Eq. 11.4), the Gaussian width s and the effective decay constant κ are scaled by $\sigma_{\Delta t}$. The scale factor s is introduced to accommodate an overall underestimate ($s > 1$) or overestimate ($s < 1$) of the errors for all events. The decay constant κ is introduced to account for residual charm decay products included in the B_{tag} vertex; κ is scaled by $\sigma_{\Delta t}$ to account for a correlation observed in Monte Carlo simulation between the mean of the $\delta\Delta t$ distribution and the measurement error $\sigma_{\Delta t}$.

The mean and RMS spread of Δt residual distributions in Monte Carlo simulation vary significantly among tagging categories, as illustrated in Fig. 11.1. We find that we can account for these differences by allowing the core Gaussian fraction f to be different for each tagging category. In addition, we find that the correlations among the three parameters (b^{out}, s^{out}, f^{out}) describing the outlier Gaussian are large and

that the outlier parameters are highly correlated with other resolution parameters. Therefore, we fix the outlier bias $b^{\rm out}$ 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 instead of fitting for them (see Chapter 14.6). The resolution model then has 8 free parameters: s, κ, $f^{\rm out}$, and five fractions f_c (one for each tagging category c).

We demonstrate the performance of the $G^{\rm sig}$ model by fitting to well-reconstructed signal Monte Carlo events, as shown in the matrix of plots in Fig. 11.2. The estimates of Δm_d and τ_{B^0} are unbiased, as indicated in Table 11.1.

Signal Δt fits to signal MC				
Par	Generator	Fit to $\Delta t_{\rm true}$	Fit to $\Delta t_{\rm meas}$	Δ
Δm_d (ps^{-1})	0.4720	0.4709 ± 0.0013	0.4715 ± 0.0043	0.0006
τ_{B^0} (ps)	1.5480	1.5446 ± 0.0051	1.5523 ± 0.0091	0.0077

Table 11.1: Results of fitting $G^{\rm sig}$ to signal Monte Carlo, using either $\Delta t_{\rm true}$ and true tagging information, or $\Delta t_{\rm meas}$ (*i.e.,* using the full resolution model) and measured tagging information. The final column (Δ) shows the difference with respect to the $\Delta t_{\rm true}$ results.

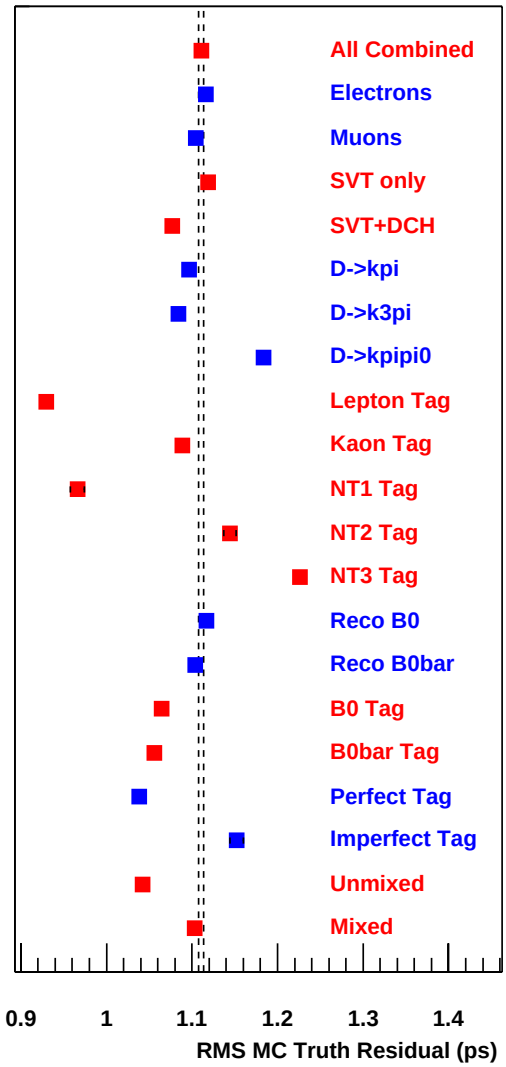


Figure 11.1: Using Monte Carlo truth information, we compute the RMS of the residual distribution when dividing the signal MC sample into various subsamples. This figure shows the spread of calculated RMS's for 8 different ways of dividing up the sample. Of key importance is the significance in residual RMS for the 5 different tagging categories.

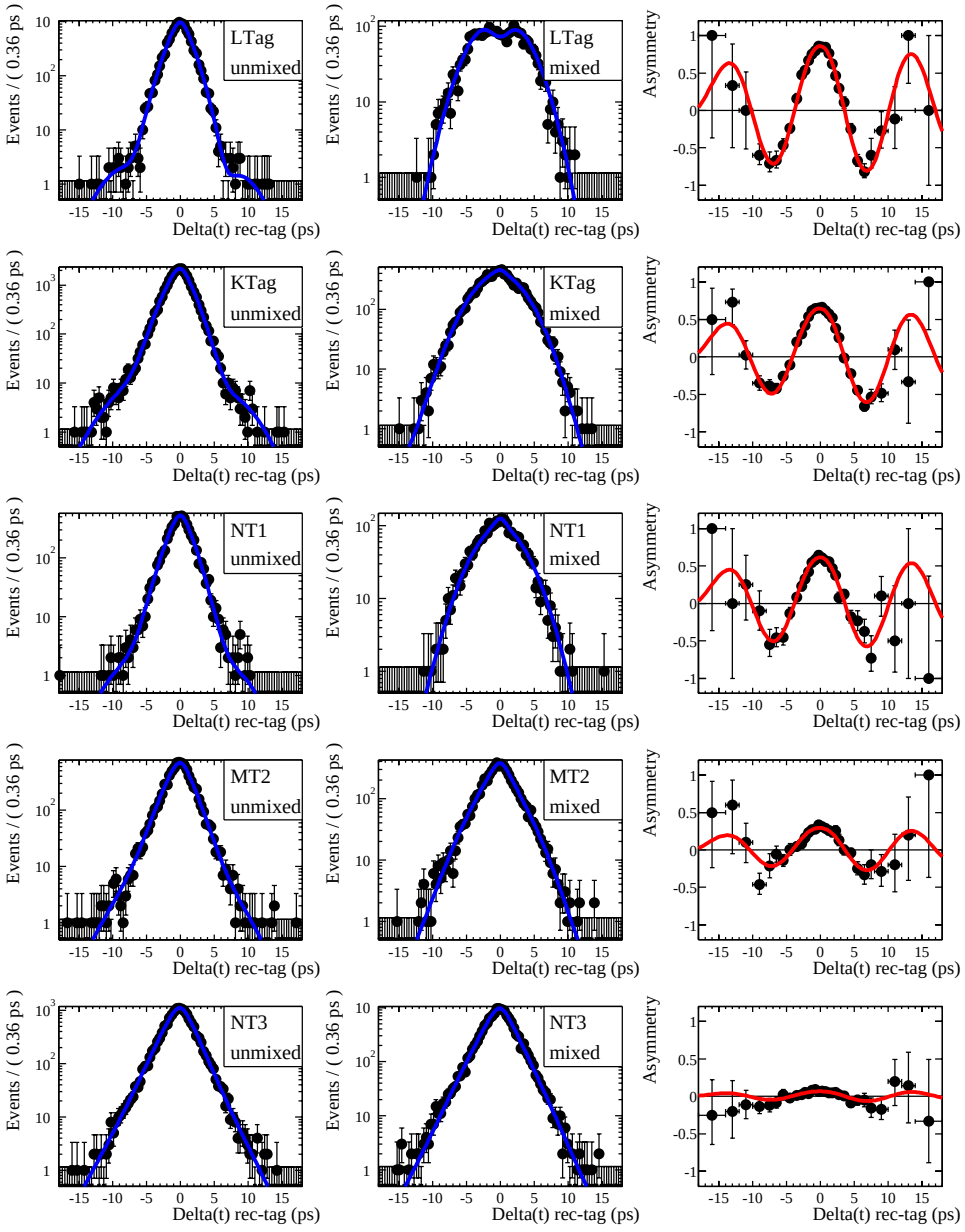


Figure 11.2: Δt distributions for unmixed and mixed events on a log scale, and the asymmetry plots for the signal Monte Carlo sample, divided into the five tagging categories. The curves are projections of the signal model from the GExp+G fit to signal Monte Carlo.

11.2.2 $B^{\pm}$ Model

The charged-B peaking background is due to decays of the type $B^{\pm} \to D^{*0}\ell^{\pm}\nu_{\ell}$. 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

$$\mathcal{G}^{\mathrm{B}^{\pm}}(\Delta t, \sigma_{\Delta t}, g\,; \vec{q}_c^{\,\mathrm{sig}}) = \frac{1}{4\tau_{B^+}} e^{-|\Delta t_{\mathrm{true}}|/\tau_{B^+}}\,(1 + g(1 - 2\omega^c_{B^+}))] \otimes \mathcal{R}(\delta\Delta t, \sigma_{\Delta t}; \vec{q}_c) \quad (11.5)$$

where $\omega^c_{B^+}$ is the mistag fraction for charged B mesons for tagging category c.

Since the ratio of the charged B to neutral B lifetime is close to 1 and the fraction of charged B mesons in the peaking sample is small, we do not have sufficient sensitivity to distinguish the lifetimes in the fit. We parameterize the physics model for the $B^{\pm}$ in terms of the lifetime ratio τ_{B^+}/τ_{B^0}, and fix this ratio to the PDG2002 world average. [5] For Monte Carlo, we use the ratio of generated lifetimes.

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 so close, the time evolution of mixed and unmixed events can be expressed approximately as

$$\begin{aligned} U(\Delta t) &\propto (1 + f_{B^+}\mathsf{D}_{B^{\pm}})/2 + ((1 - f_{B^+})\mathsf{D}_{B^0}\cos\Delta m_d\Delta t)/2 \\ M(\Delta t) &\propto (1 - f_{B^+}\mathsf{D}_{B^{\pm}})/2 - ((1 - f_{B^+})\mathsf{D}_{B^0}\cos\Delta m_d\Delta t)/2\,, \end{aligned}$$

where D is the dilution $1 - 2\omega$. Only two quantities, $f_{B^+}\mathsf{D}^{\pm}_B$ and $(1 - f_{B^+})\mathsf{D}^0_B$ are measurable (the constant term and the amplitude of the oscillating term). 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. For Monte Carlo, we fix these ratios to the results of truth-counting in the generic sample. See Tables 11.2-11.3 for these values. (The values for Data are obtained from the *BABAR* BTagging Group. [52])

Nominal B^+/B^0 Ratios for MC	
τ_{B^+}/τ_{B^0}	1.0690 ± 0.029
$\omega^{B^+}_{\rm LT}/\omega^{B^0}_{\rm LT}$	1.210 ± 0.46
$\omega^{B^+}_{\rm KT}/\omega^{B^0}_{\rm KT}$	0.724 ± 0.12
$\omega^{B^+}_{\rm N1}/\omega^{B^0}_{\rm N1}$	0.740 ± 0.26
$\omega^{B^+}_{\rm N2}/\omega^{B^0}_{\rm N2}$	0.942 ± 0.17
$\omega^{B^+}_{\rm N3}/\omega^{B^0}_{\rm N3}$	1.121 ± 0.12

Table 11.2: Values for ratios of B^+ to B^0 parameters for use in the signal Δt model which describes candidates from both charged and neutral parents. The mistag rates were determined directly from Monte Carlo truth by counting.

Nominal B^+/B^0 Ratios for Data	
τ_{B^+}/τ_{B^0}	1.083 ± 0.017
$\omega^{B^+}_{\rm LT}/\omega^{B^0}_{\rm LT}$	0.54 ± 0.10
$\omega^{B^+}_{\rm KT}/\omega^{B^0}_{\rm KT}$	0.68 ± 0.05
$\omega^{B^+}_{\rm N1}/\omega^{B^0}_{\rm N1}$	0.99 ± 0.12
$\omega^{B^+}_{\rm N2}/\omega^{B^0}_{\rm N2}$	1.05 ± 0.07
$\omega^{B^+}_{\rm N3}/\omega^{B^0}_{\rm N3}$	1.12 ± 0.12

Table 11.3: Values for ratios of B^+ to B^0 parameters for use in the signal Δt model which describes candidates from both charged and neutral parents. These values for performance on Run-1 Data were obtained from other *BABAR* studies, except for the NT3 category, which are determined by Monte Carlo truth counting.

11.2.3 Combined Signal Model

We combine the charged-B and neutral-B models with the fraction f_{B^+} since we assert that they will have identical resolution models. The combined model becomes

$$\begin{aligned}
&\mathcal{G}^{\rm sig}(\Delta t, \sigma_{\Delta t}, g\,; \vec{q}_c^{\,\rm sig}) = \\
&\quad [\frac{1-f_{B^+}}{4\tau_{B^0}} e^{-|\Delta t_{\rm true}|/\tau_{B^0}} \left(1 + g(1-2\omega^c_{B^0})\cos(\Delta m_d \Delta t_{\rm true})\right) + \\
&\quad \frac{f_{B^+}}{4\tau_{B^+}} e^{-|\Delta t_{\rm true}|/\tau_{B^+}} \left(1 + g(1-2\omega^c_{B^+})\right)] \otimes \mathcal{R}(\delta\Delta t, \sigma_{\Delta t}; \vec{q}_c)\,,
\end{aligned} \tag{11.6}$$

We use this model to fit a sample of neutral and charged B signal candidates in the

generic Monte Carlo, 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^+} 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 PDG2000 value, 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 event counting based on the truth. The fit results for Δm_d, τ_{B^0} and the charged B fraction are shown in Table 11.4, along with the results of the fit to B^0 signal only, and the fit with charged B fraction fixed at the true value. In all fits, the resolution function and mistag fractions are free parameters.

The fitted charged B fraction is consistent with the "true" value. The generic MC sample is significantly smaller than the signal MC sample, but the fitted lifetime and mixing values appear consistent with the generator values at the 1σ level. We note however, that in this truth-matched sample of Generic MC events, we are unable to measure a non-zero outlier fraction. The correlation between f^{out} and τ_{B^0} is what drives the measured lifetime higher in the Generic MC fits. This "systematic" failure is reflected in the "statistical" error which incorporates these correlations. Upon closer inspection, we see the error on Δm_d inflates when we float f_{B^+}, as expected, and because of the positive correlation, an underestimate of f_{B^+} also causes an underestimate of Δm_d, as seen by comparing columns 2, 3 and 4 of the table. We conclude that while we are unlucky with our Generic MC sample in that we cannot reliably measure the signal outlier fraction, the fit results are still correctly described by their statistical errors.

Signal Δt fits to (truth-matched) Signal & Generic Monte Carlo				
Par.	Sig MC	Gen MC $B^0\overline{B}^0$	GenMC $B^0\overline{B}^0$ &B^+B^-	Gen MC $B^0\overline{B}^0$ &B^+B^-
Δm_d (ps^{-1})	0.4715 ± 0.00427	0.4640 ± 0.0101	0.4630 ± 0.0106	0.4617 ± 0.0132
τ_{B^0} (ps)	1.5523 ± 0.00915	1.5749 ± 0.0227	1.5666 ± 0.0221	1.5656 ± 0.0227
$f^{\rm out}$	0.0044 ± 0.0015	-0.0012 ± 0.0027	-0.0005 ± 0.0029	-0.00024 ± 0.0034
f_{B^+}	0	0	0.0521	0.047 ± 0.030

Table 11.4: $G^{(\rm sig)}$ fits to MC truth-matched signal $B^0\overline{B}^0$ and B^+B^- from Generic MC (we include the result for the fit to signal MC for reference). We show the fitted results (with full statistical errors) for Δm_d, τ_{B^0}, f_{B^+}, and the signal outlier fraction $f^{\rm out}$. The third column of values represents a fit to truth-matched $B^0\overline{B}^0$ and B^+B^- Generic MC events where we fix f_{B^+} to its true value and hold it constant in the fit; in the fourth column we float this parameter and extract it from the fit. In all the fits, resolution function and mistag fractions are floating, except for the two signal outlier model shape parameters.

11.3 Background Models

One of the key features of this study is the effort to describe the four classes of background events using knowledge derived directly from the Data. Here, we elucidate models for the Δt distributions of the combinatoric, continuum, fake lepton, and uncorrelated lepton background events. We use samples of Generic Monte Carlo events and the control samples in Data to determine the functional forms, but leave almost all of the parameters floating in the final fit to the full sample. This properly treats the correlations between the Δt models because while the "true" background candidates are separable into the four distinct classes, we only have per-event probability estimates, *i.e.,* we do not have certainty about which events are absolutely combinatoric or absolutely peaking background. Finally, although the $\Delta t_{\rm true}$ and resolution on $\Delta t_{\rm true}$ are not well-defined for background events, we still describe the total Δt model as a "physics model"convolved with a "resolution function" for convenience.

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:

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

where $\delta(\Delta t)$ is a δ-function, $g = +1$ for unmixed and -1 for mixed events, and $\tau^{\rm bkg}$ and $\Delta m^{\rm bkg}$ 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, the so-called double Gaussian:

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

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 wide fraction is allowed to be different for each type of background.

11.3.1 Combinatoric background

Model

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. We have an advantage in studying this largest source of background, as the control sample (the δm sideband) is very pure in combinatorics.

An important concern about the sideband control sample for combinatorics is whether it accurately describes the behavior of combinatorics in the peaking region of δm. If it does, then there is a benefit to including it in the final analysis since it will significantly increase the number of events used to determine the combinatoric Δt portion of the model. Using generic Monte Carlo, we observe that B^+B^- combinatorics do not exhibit time-dependent mixing, but combinatoric candidates from $B^0\overline{B}^0$ parents can, as seen in Fig.11.3. The relative fraction of $B^0\overline{B}^0$ and B^+B^- events in the combinatoric background depends slightly on δm as shown in Fig. 11.4(a). Clearly the $B^0\overline{B}^0$ background dominates, and there is a tradeoff with B^+B^- sources as one nears the peak region. We might expect then, that an effective mixing frequency for combinatorics could depend on the δm value.

If we fit the generic Monte Carlo sample of combinatorics in slices of δm, we find that the effective mixing frequency $\Delta m^{\rm bkg}$ is uniform across $m(D^*) - m(D^0)$ (see Fig. 11.4(b)). No significant dependence of the parameters of the Δt model on δm

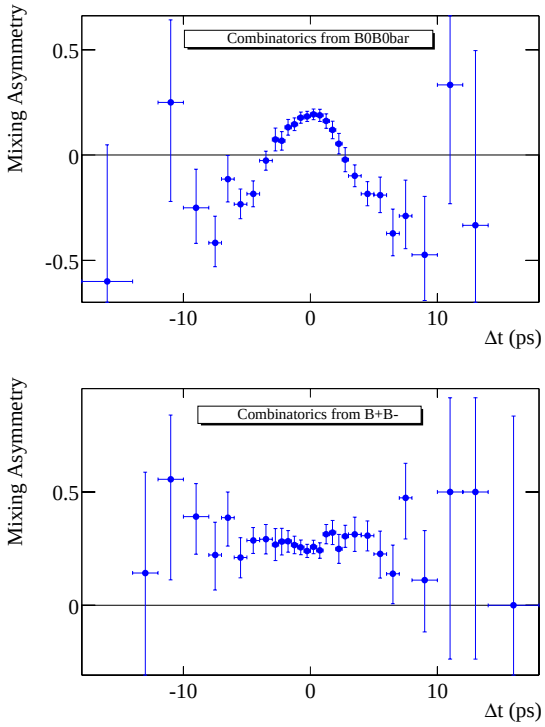


Figure 11.3: Effective mixing asymmetry for combinatoric candidates selected from $B^0\overline{B}^0$ parents and B^+B^- parents. We observe that the candidates from charged B mesons do not exhibit any mixing asymmetry, while the $B^0\overline{B}^0$ combinatorics do.

is observed in Data or the Monte Carlo simulation. We therefore choose to employ a common $\Delta m^{\rm bkg}$ for all combinatoric candidates. The data and Monte Carlo Δt distributions are described well by a prompt plus oscillatory term convolved with a double-Gaussian resolution function:

$$\mathcal{G}^{\rm comb} = [f^{\rm osc} \cdot \mathcal{G}^{\rm osc}_{\rm phys}(\Delta t_{\rm true}, g; \tau^{\rm comb}, \Delta m^{\rm comb}, \omega^{\rm osc}) + (1 - f^{\rm osc}) \cdot \mathcal{G}^{\rm pmt}_{\rm phys}(\Delta t_{\rm true}, g; \omega^{\rm pmt})] \otimes \mathcal{R}_{\rm G+G}(\delta\Delta t, \sigma_{\Delta t}; b, s, f, b^w, s^w) \; . \quad (11.7)$$

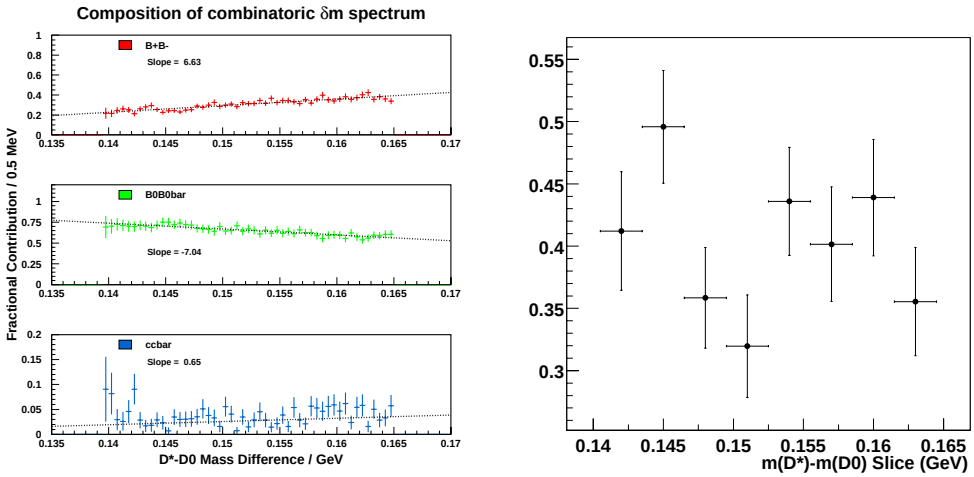


Figure 11.4: **Left:** Relative composition of the combinatoric δm spectrum from Monte Carlo. Note that $B^0\overline{B}^0$ dominates everywhere, although the relative composition between $B^0\overline{B}^0$ and B^+B^-, changes by 15% from the signal region to the far sideband. **Right:** The fitted value of $\Delta m^{\rm bkg}$ from fits to slices of the generic Monte Carlo sample in $m(D^*) - m(D^0)$. We observe that $\Delta m^{\rm bkg}$ is consistent across all eight bins, despite the changing relative fraction of contributions from $B^0\overline{B}^0$.

Parameterization

Because the sideband control sample is so pure, we develop the Δt model without using the detailed δm per-event probability information. The prototype form of $G^{\rm comb}$ above has 10 parameters. To determine the most appropriate sharing or splitting of these parameters amongst subsamples in the Data, we study the behavior of fit results for different variations of the model. For instance, we suspect that $f^{\rm osc}$ might vary between subsamples, so we consider different groupings of events to share a common $f^{\rm osc}$ and compare the fitted values. The worst case scenarios might be that all 360 subsamples have a unique fraction of oscillatory time dependence; the best might be that all combinatoric candidates share the same fraction identically. Consider the plots in Fig. 11.5. The left-hand figure shows the different values for the parameter obtained when splitting it by one type of category at a time, *i.e.,* sideband candidates with identified leptons prefer a higher level of oscillation in their time structure as compared to those with fake leptons, or combinatoric candidates

with same-side angular correlation prefer a lower oscillation fraction than those with opposite-side angular correlation. However, subdivisions like these only divide the data by the values of one category at a time. We also perform tests such as the one in the right-hand plot of Fig. 11.5 where we separate the sample into $3 \times 5 = 15$ different groups which have separate oscillation fractions (and share all other parameters). Here we observe that, as expected, the fake lepton combinatoric candidates have a small oscillation fraction, but not all: the combinatoric candidates with a fake lepton and assigned the `KTag` tagging category have an oscillation fraction more like that of identified leptons. After a complete study of different groupings, we arrive at the simplified splitting shown in Fig. 11.6. Here, the combinatoric control sample (*i.e.,* the δm sideband) is divided into 5 distinct groups, each of which has a unique oscillation fraction.

We find that the parameters $\omega^{\rm pmt}$, $\Delta m^{\rm comb}$, $\tau^{\rm comb}$, f, b^w, and s^w are best shared among all subsamples of the combinatoric background. 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 sample of events in the δm sideband is used to determine the starting values for the parameters in the final full fit to all data samples.

11.3.2 Continuum peaking background

For this and the other peaking backgrounds, we follow a similar strategy as we did for the combinatorics. For brevity, however, we will not go into as much depth. Please see Ref. [47] for more details.

Model

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

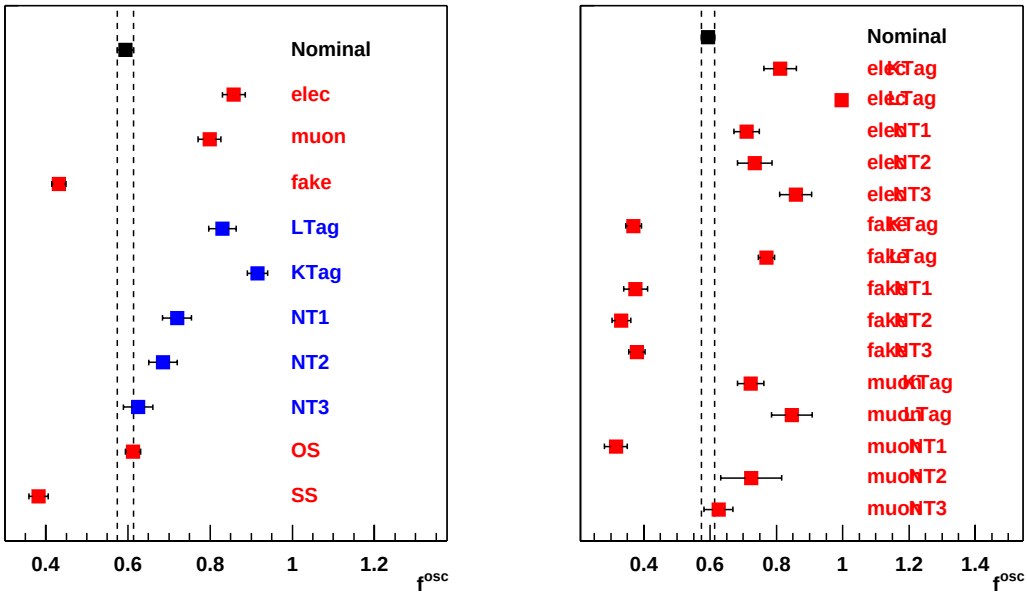


Figure 11.5: Spread of fit results for $f^{\rm osc}$ for fits to the sideband sample of Data, almost 100% pure in combinatoric events. Each region of the plot indicates a fit where the oscillation fraction was allowed to be different for each category of events labelled on the right hand side. The left-hand plot shows 4 fits with 4 different groupings of events; the right-hand plot shows one fit with the sample divided into 15 different groups for $f^{\rm osc}$.

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:

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

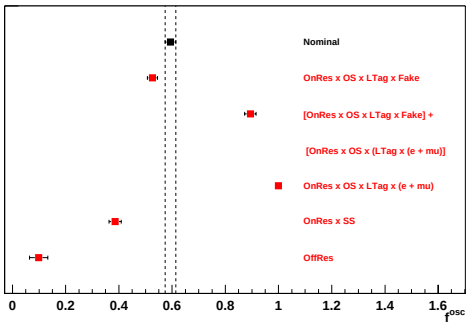


Figure 11.6: Spread of fit results for $f^{\rm osc}$ for fits to the sideband sample of Data using the final divisions of the sample. For instance, all off-resonance combinatoric candidates are non-$B^0\overline{B}^0$ parented, and so we expect a small fraction of oscillatory time dependence. We observe this in the diagram. We also see that on-resonance, same-side events have their own unique oscillation fraction. Finally, combinatoric candidates that are on-resonance, opposite-side split into three groups, each displaying a different value for $f^{\rm osc}$.

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

Parameterization

In Data, the background control samples for the continuum peaking background are the off-resonance samples. To bootstrap the continuum Δt parameters in the final fit, we first fit to a combination of combinatoric background and continuum peaking background in both data and Monte Carlo. For the Monte Carlo sample, the $c\overline{c}$ sample is artificially divided into two samples with the proportion of numbers of events being 1:7 (similar to the ratio of off- to on-resonance data). The smaller Monte Carlo sample is artificially marked as off-resonance. Both the data and Monte Carlo samples are analyzed as described in Chapter 10 to extract δm shapes and background fractions. In each case, the δm sideband events are used to obtain the parameters of the Δt model for combinatoric background. We fix all parameters except the continuum peaking background Δt model. The results are shown in Table 11.5, and directly illustrate the process of bootstrapping. The parameters from the fit to data

(right-hand column) will be used as starting values in the final full fit to data.

Par	$c\overline{c}$ D^* MC	"Off-Res" MC	Off-Res Data
$\omega^{\rm pmt}_{\rm LNTx}$	0.483 ± 0.018	0.491 ± 0.050	0.445 ± 0.055
$\omega^{\rm pmt}_{\rm KT}$	0.117 ± 0.013	0.064 ± 0.033	0.094 ± 0.034
$b^1_{\rm OS}$	0.015 ± 0.039	-0.048 ± 0.15	-0.098 ± 0.15
$b^1_{\rm SS}$	-0.177 ± 0.072	0.01 ± 0.41	-0.597 ± 0.28
s^1	1.145 ± 0.037	1.07 ± 0.25	1.16 ± 0.16
f^1	0.930 ± 0.023	0.69 ± 0.28	0.83 ± 0.11

Table 11.5: $G^{(\rm cont)}$ fit parameters for continuum D^* candidates from $c\overline{c}$ Monte Carlo with a real D^*, and combined fits to combinatoric background and continuum peaking background in "off-resonance" $c\overline{c}$ Monte Carlo (defined in text) and data. Recall that the subscripted labels on the parameters indicate the subset of the sample to which they are restricted. For instance, $b^1_{\rm SS}$ is the mean of the core Gaussian of the resolution model for candidates with so-called "same-side" angular correlation between the D^* and ℓ.

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 `KTag` 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, *i.e.,* split the parameters across the aforementioned category values. 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.

11.3.3 Fake-lepton peaking background

Model

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 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:

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

Parameterization

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.

11.3.4 Uncorrelated-lepton peaking background

Model

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 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 $B_{D^*\ell}$ 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 the "wrong" ℓ to a common vertex will converge for these events. Finally, we expect truly mixed events to have a higher probability (than unmixed events) of being reconstructed as uncorrelated-lepton events because in mixed events the charge of the primary lepton(s) on the tagging side is opposite that of D^*, and can therefore be consistent with the $B_{D^*\ell}$ hypothesis. These expectations are confirmed in the Monte Carlo simulation. For instance, Table 11.6 shows the calculated χ_d parameter for uncorrelated-lepton events identified in the generic Monte Carlo sample. We observe that this sample is enriched in mixed events, as expected (that is, mixed events are more likely to become uncorrelated lepton candidates than unmixed events).

Tagging Category	% Sample	χ_d
`LTag`	13	0.68
`KTag`	43	0.53
`NT1`	7	0.59
`NT2`	11	0.43
`NT3`	26	0.55

Table 11.6: Ratio of the number of truly mixed events to the total number of events, χ_d, for the pure uncorrelated lepton sample from generic $B\overline{B}$ Monte Carlo. Note that the sample has a high average value of χ_d, and is relatively enriched in `LTag` events.

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:

$$\mathcal{G}^{\rm uncor} = \left[f^{\rm life} \cdot \mathcal{G}^{\rm life}_{\rm phys}(\Delta t_{\rm true}, g; \tau^{\rm uncor}, \omega^{\rm life}) + (1 - f^{\rm life}) \cdot \mathcal{G}^{\rm pmt}_{\rm phys}(\Delta t_{\rm true}, g; \omega^{\rm pmt}) \right] \otimes \mathcal{R}_{\rm G+G}(\delta\Delta t, \sigma_{\Delta t}; b, s, f, b^w, s^w) \; . \tag{11.8}$$

Parameterization

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.

Chapter 12

Results

The final model is fitted to the full Data sample, and we report the fit results. To improve the robustness of the result, we fix the two shape parameters of the signal outlier shape after surveying the likelihood values in that two-dimensional space. We extract the following (uncorrected) values for lifetime and mixing:

$$\tau_{B^0} = 1.545 \pm 0.023\,\mathrm{ps}$$

and

$$\Delta m_d = 0.512 \pm 0.018 ps^{-1}$$

and

$$\rho(\Delta m_d, \tau_{B^0}) = -0.22\ .$$

We observe non-linearities in the likelihood space around the minimum in the τ_{B^0} direction, so we correct the upper statistical error on τ_{B^0} by 6%. We also study the source of the correlation between τ_{B^0} and Δm_d and find that it is largely due to the signal resolution model and the charged B fraction.

12.1 Comments on Implementation

12.1.1 Signal Outlier Model

The outlier portion of the signal 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 gross errors in tagside vertex reconstruction. Since there is little reason to expect the "mean" and "width" of the signal outlier Gaussian to scale with the per-event error, the outlier resolution function is parameterized as a Gaussian in $\delta\Delta t$ ($= \Delta t_{\rm meas} - \Delta t_{\rm true}$).

We found that the fit to Data (and Generic Monte Carlo) is not 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. 12.4. 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 repeat 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 12.1 shows the negative-ln-likelihood surface as a function of outlier bias and width. The negative-ln-likelihood has been offset so that the minimum value over the plot is zero. For fixed width, the minimum in negative-ln-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-ln-likelihood is less than 0.5 for most of the values of bias and width scanned in this study.

A scatter plot of the resulting values of Δm_d and τ_{B^0} is shown in Fig. 12.2. 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. 12.1).

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

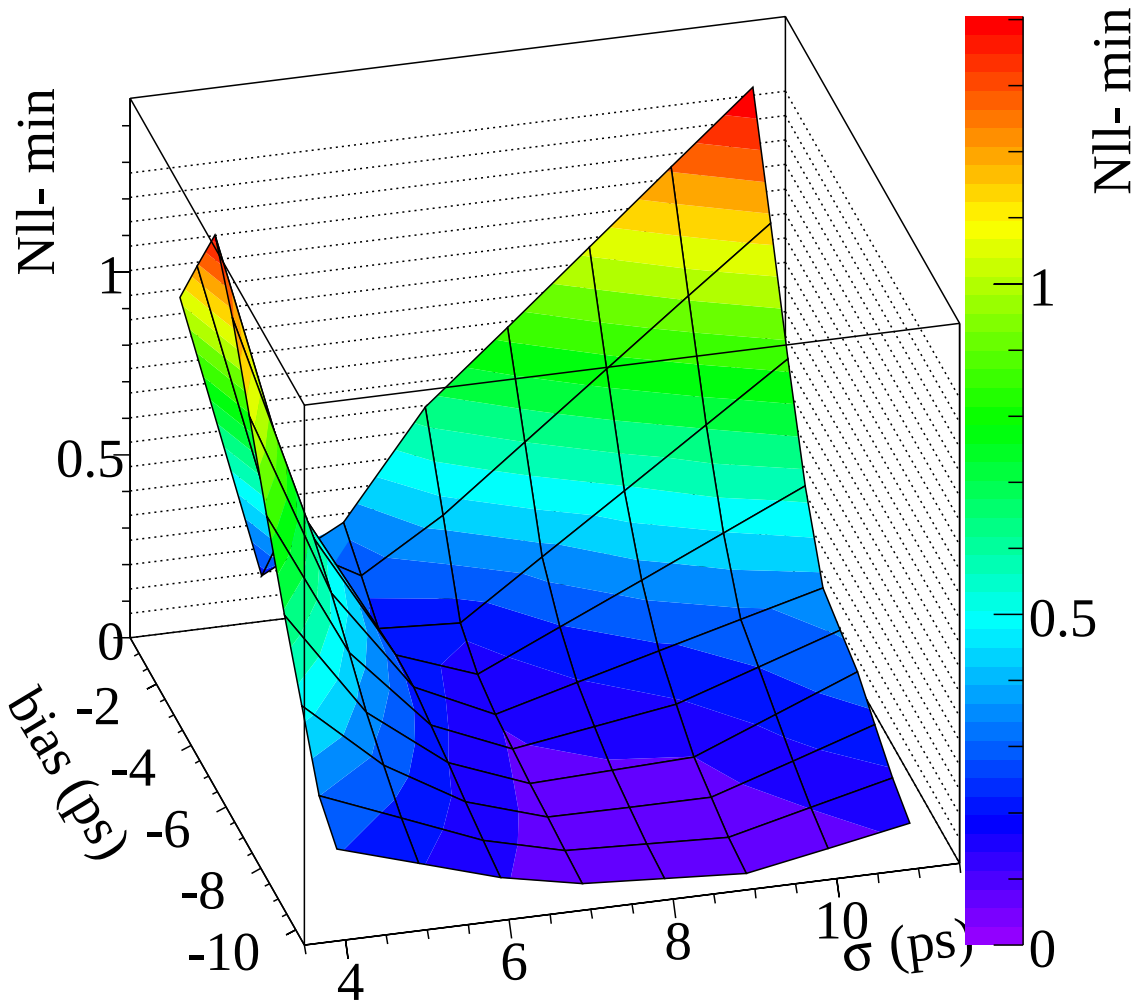


Figure 12.1: The negative-log-likelihood surface in the space of outlier bias and outlier width for the outlier Gaussian of the signal resolution model in Data. The vertical scale has been offset so that minimum is at 0. The surface is constructed from 36 fits to the Data with different fixed values of the signal outlier parameters.

(Fig. 12.2) and use the corresponding bias and width. We selected a bias of -5 ps and a width of 6 ps.

To summarize, we fix the bias and width in the outlier Gaussian of the resolution

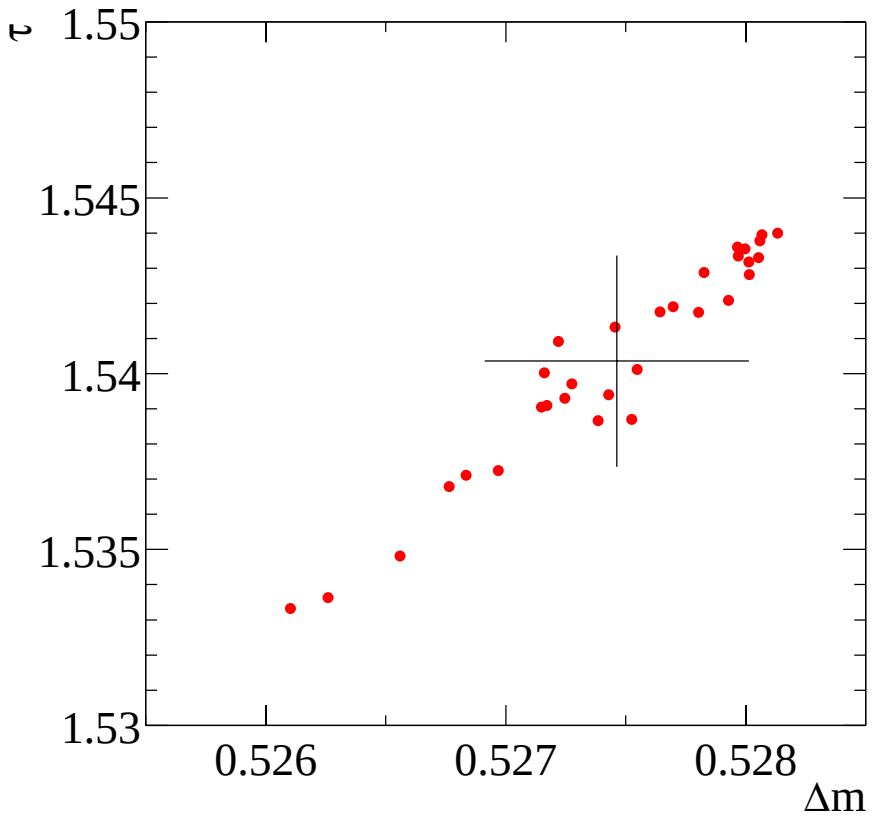


Figure 12.2: Scatter plot of the set of (blinded) Δm_d, τ_{B^0} values obtained from data for 36 fits with different fixed values of the outlier bias and width in the range -1 to -10 ps for bias and 4 to 11 ps for width. The cross indicates the mean and RMS of the Δm_d, τ_{B^0} distribution. Note that the range of the τ_{B^0} (y) axis is slightly less than one statistical error bar, and that the range of the Δm_d (x) axis is less than 1/5 of a statistical error bar.

function to -5 ps and 6 ps, respectively. We note that it is not unusual to fix outlier parameters in lifetime and mixing fits in *BABAR*— no other lifetime and mixing analysis in *BABAR* has allowed the outlier parameters to be free. (In fact, this detailed studied of the likelihood surface in the outlier shape parameter space is most likely a first.)

12.1.2 Performance

The technology used to define the probability density functions, build the likelihood and perform the fit consists of two general toolkits, **RooFit** [60] and **MINUIT** [61] both running under the umbrella environment of **ROOT** [62]. Because of the complexity of the likelihood, the computational power required to perform the minimization is impressive. A typical fit to the Data set, starting from good initial values, takes 8-9 hours on a Pentium III with 500 MB of memory running at 800 MHz to find the minimum, and another 8-9 hours to calculate the covariance matrix. In order to make study of the measurement more tenable, we typically conducted studies without requesting the covariance matrix unless it was absolutely necessary (thereby halving the time of each fit.)

We also found that the covariance matrix calculation was sometimes unstable, reporting that a non-positive-definite matrix was found (*i.e.,* the minimum appeared to be a saddle point). By optimizing the precision of the calculation and feeding the more detailed information from the minimization step to the covariance-calculation stage, we were able to substantially improve the reliability of the error matrix computation.

12.2 Fit Results

12.2.1 Final Parameter Values

The total number of free parameters in the final fit is 72: 22 in the signal model, 24 in the combinatoric background model, and 26 in peaking background models. The fitted signal Δt model parameters are shown in Table 12.1. The fitted combinatoric background model parameters are shown in Table 12.2 and the fitted peaking background parameters are listed in Table 12.3.

The statistical correlation coefficient between τ_{B^0} and Δm_d is $\rho(D^-, \tau_{B^0}) = -0.22$. 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 12.4. We draw the error ellipse for the τ_{B^0}-Δm_d correlation in Fig. 12.3, where the contour is estimated from the covariance matrix.

Table 12.1: Results of full fit to data — signal model and resolution function parameters. A small correction, described in Chapter 13.3.3, will be applied to τ_{B^0} and Δm_d.

Signal Model and Δt Resolution Function Parameters					
parameter	value	parameter	value	parameter	value
Δm_d (ps^{-1})	0.512 ± 0.018	f_{B^+}	0.082 ± 0.029	s	1.201 ± 0.063
τ_{B^0} (ps)	1.545 ± 0.023	$\omega_{\rm lepton}$	0.071 ± 0.015	κ	0.86 ± 0.17
-	-	$\omega_{\rm kaon}^{\rm offset}$	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	$f^{\rm out}$	0.0027 ± 0.0017
-	-	$\Delta\omega_{\rm kaon}$	-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 charged B fraction in data from this final fit is $(8.2 \pm 2.9)\%$. Although the selection criteria are not exactly the same, it is interesting to compare this result with the charged B fraction found for the $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ sample described in the Ref. [12]: $(4.5 \pm 0.3 \pm 2.2)\%$. This result was determined from a fit to the $\cos\theta_{B,D^*\ell}$ distribution in data, with shapes taken from Monte Carlo. Given that the errors on the two results are not correlated, the difference is $(3.7 \pm 3.7)\%$. Therefore, the two results are consistent.

12.2.2 Plots of Δt Projections

We can review the performance of the fit by projecting the master PDF onto Δt separately for mixed and unmixed events, overlaying the Data as well. A subsample of known purity can be selected by computing the average per-event probability as predicted from the δm analysis. For instance, in Fig. 12.4 we plot the fit result projected onto a sample of 80% signal purity by selecting events with per-event signal

Final Fit to Data, Combinatoric Δt Model Parameters			
parameter	value	parameter	value
$\Delta m^{\rm comb}$	0.422 ± 0.020	$b^1{}_{\rm LNT12}$	-0.056 ± 0.021
$\tau^{\rm comb}$	1.234 ± 0.024	$b^1{}_{\rm KNT3}$	-0.104 ± 0.012
$\omega^{\rm osc}_{\rm Fake;LTag}$	0.434 ± 0.026	$f^{\rm osc}_{\rm Off}$	0.099 ± 0.035
$\omega^{\rm osc}_{\rm Fake;KTag}$	0.128 ± 0.017	$f^{\rm osc}_{\rm SS}$	0.385 ± 0.023
$\omega^{\rm osc}_{\rm Fake;NT1}$	0.443 ± 0.032	$f^{\rm osc}_{\rm NoLept}$	0.526 ± 0.019
$\omega^{\rm osc}_{\rm Fake;NT2}$	0.507 ± 0.024	$f^{\rm osc}_{\rm OneLept}$	0.894 ± 0.022
$\omega^{\rm osc}_{\rm Fake;NT3}$	0.589 ± 0.020	$f^{\rm osc}_{\rm TwoLept}$	1.000
$\omega^{\rm osc}_{\rm Lept;LTag}$	0.156 ± 0.023	$s^1{}_{\rm OffSS}$	1.337 ± 0.024
$\omega^{\rm osc}_{\rm Lept;KTag}$	0.297 ± 0.010	$s^1{}_{\rm On;OS}$	1.259 ± 0.021
$\omega^{\rm osc}_{\rm Lept;NT1}$	0.377 ± 0.025	f^1	0.9666 ± 0.0071
$\omega^{\rm osc}_{\rm Lept;NT2}$	0.401 ± 0.016	$b^{\rm out}$	-0.98 ± 0.24
$\omega^{\rm osc}_{\rm Lept;NT3}$	0.491 ± 0.011	$s^{\rm out}$	4.60 ± 0.42
$\omega^{\rm pmt}$	0.3960 ± 0.0093	-	-

Table 12.2: Default fit result for Data — Δt model parameters for combinatoric background, $G^{(\rm comb)}$.

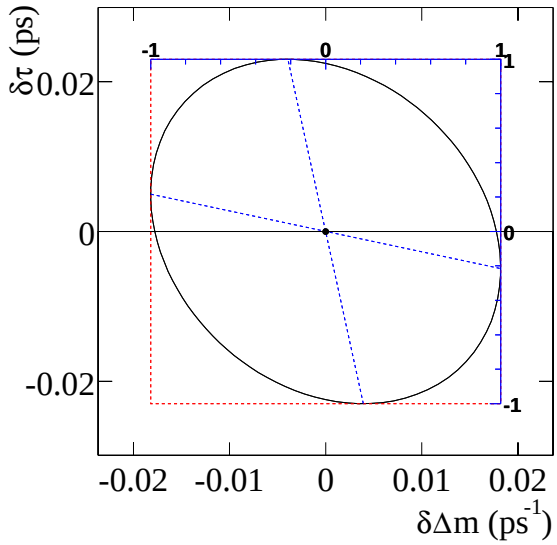


Figure 12.3:]
The one-sigma contour for τ_{B^0} *vs.* Δm_d from the full fit to Data. The correlation coefficient is -0.22.

Final Fit to Data, Peaking BKG Model Parameters					
cont. par.	value	fake par.	value	unco. par.	value
ω_{KT}	0.083 ± 0.028	Δm^{fake}	0.444 ± 0.031	τ^{unco}	1.07 ± 0.17
ω_{LNT}	0.457 ± 0.040	τ^{fake}	1.416 ± 0.060	$\omega^{\mathrm{life}}_{\mathrm{KNT}}$	0.71 ± 0.13
$b^1{}_{\mathrm{OS}}$	0.04 ± 0.11	$\omega^{\mathrm{osc}}_{\mathrm{LT}}$	0.248 ± 0.038	$\omega^{\mathrm{life}}_{\mathrm{LT}}$	0.9998
$b^1{}_{\mathrm{SS}}$	-0.08 ± 0.13	$\omega^{\mathrm{osc}}_{\mathrm{KT}}$	0.142 ± 0.029	ω^{pmt}	0.00000 ± 0.00064
s^1	1.300 ± 0.090	$\omega^{\mathrm{osc}}_{\mathrm{NT1}}$	0.342 ± 0.044	f^{life}	0.877 ± 0.085
f^1	0.909 ± 0.040	$\omega^{\mathrm{osc}}_{\mathrm{NT2}}$	0.377 ± 0.039	b^1	-0.30 ± 0.18
-	-	$\omega^{\mathrm{osc}}_{\mathrm{NT3}}$	0.512 ± 0.042	f^1	1.000
-	-	$\omega^{\mathrm{pmt}}_{\mathrm{KNT3}}$	0.46 ± 0.14	s^1	1.34 ± 0.24
-	-	$\omega^{\mathrm{pmt}}_{\mathrm{LNT12}}$	0.64 ± 0.15	-	-
-	-	f^{osc}	0.852 ± 0.046	-	-
-	-	$b^1{}_{\mathrm{KNT3}}$	-0.208 ± 0.062	-	-
-	-	$b^1{}_{\mathrm{LNT12}}$	-0.078 ± 0.065	-	-
-	-	f^1	0.944 ± 0.023	-	-
-	-	s^1	1.159 ± 0.091	-	-

Table 12.3: Default fit result for Data — Δt model parameters for peaking backgrounds. The outlier bias (μ_{out}) and scale factor (σ_{out}) for these peaking background are shared with combinatoric background parameters.

Table 12.4: 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
$\rho(\tau_{B^0}, \sigma^1_{\mathrm{sig}})$	-0.49
$\rho(\tau_{B^0}, f^{\mathrm{out}}_{\mathrm{sig}})$	-0.26

probability ≥ 0.4, yielding about $15,900$ events. (Recall Chapter 10.4.) Likewise we construct the mixing asymmetry by combining the projections in Fig. 12.4 to form the appropriate ratio. The asymmetry in Data between mixed and unmixed events per unit time is shown in Fig. 12.5 along with the projection of the fit result.

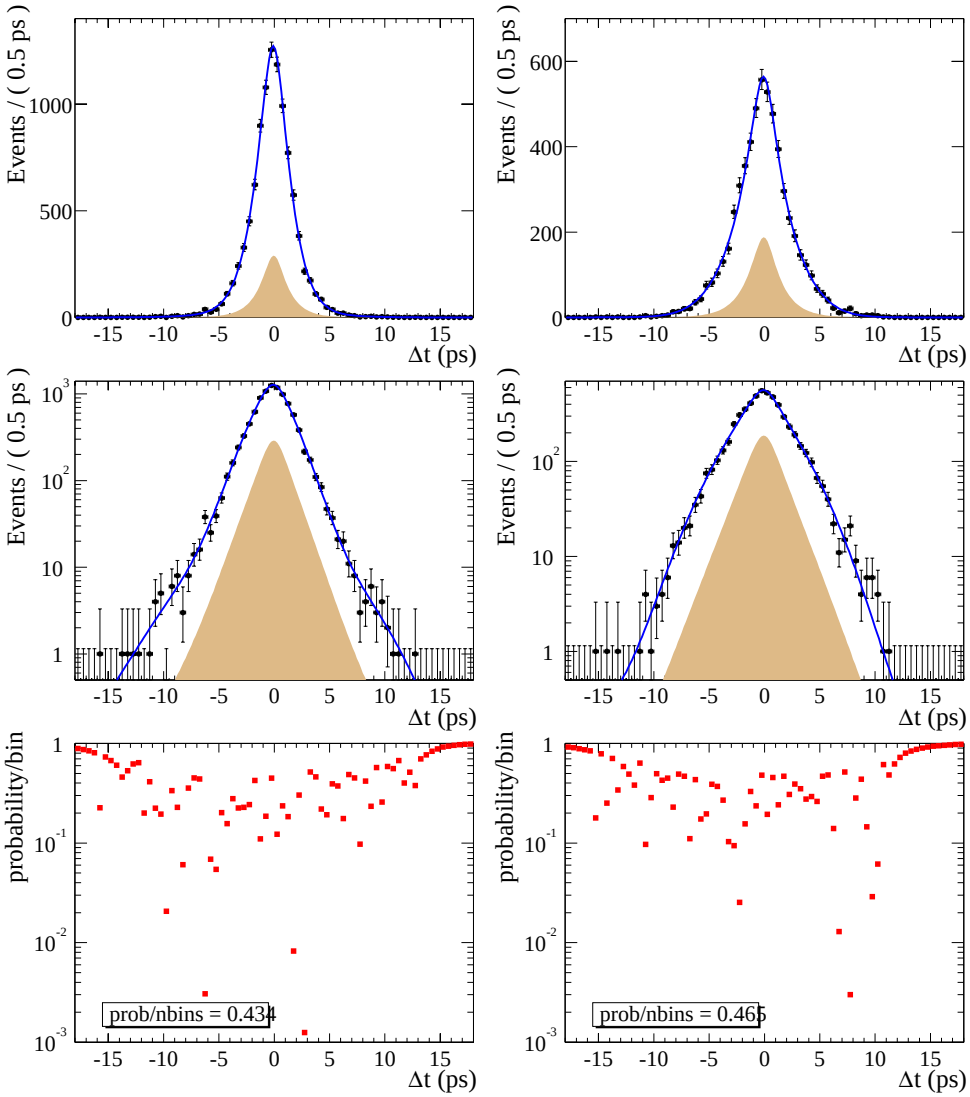


Figure 12.4: The Δt distribution of unmixed and mixed events in an 80% pure signal subsample of Data, and the projection of the model. The left hand plots are for unmixed events, the right for mixed events; the middle row is simply a $\log y$ plot. The superimposed shaded areas on the upper plots show 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 $n > \mu$.

12.3 Discussion

Parameter Errors and Correlations

Since we float many parameters in the model, it is useful to see how the errors on τ_{B^0} and Δm_d, and their correlation change when different parameters are free in the

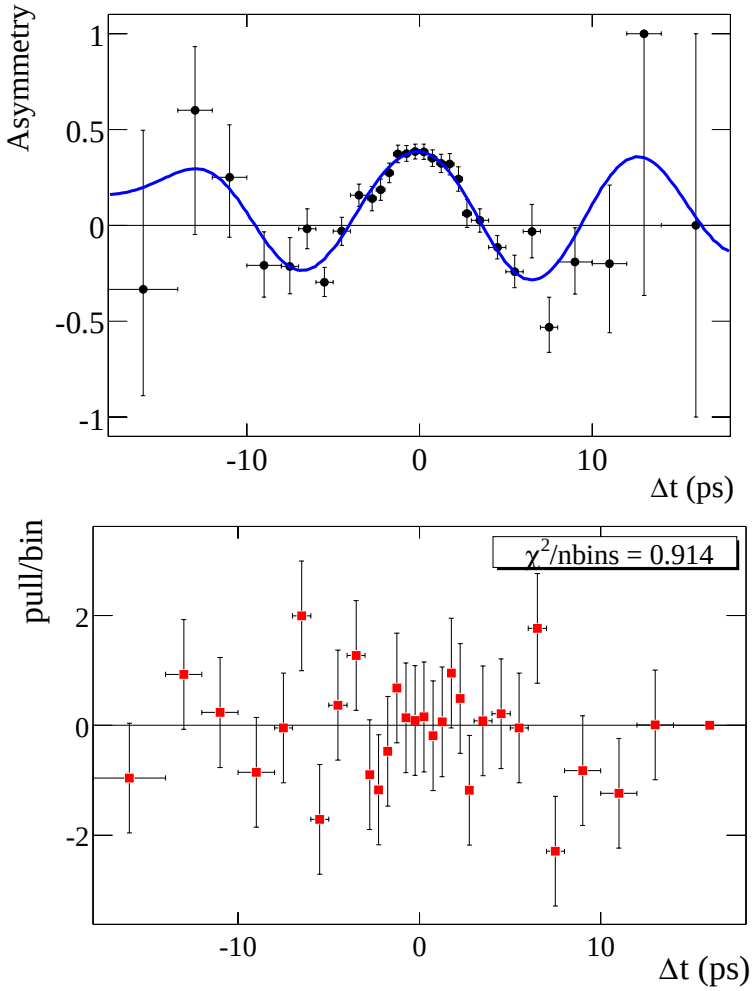


Figure 12.5: The asymmetry plot for the 80% pure signal sample of Data, and the projection of the model fit result from Data. The lower plot shows the bin-by-bin difference normalized by the error estimated from data on each bin.

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. 12.6.

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.

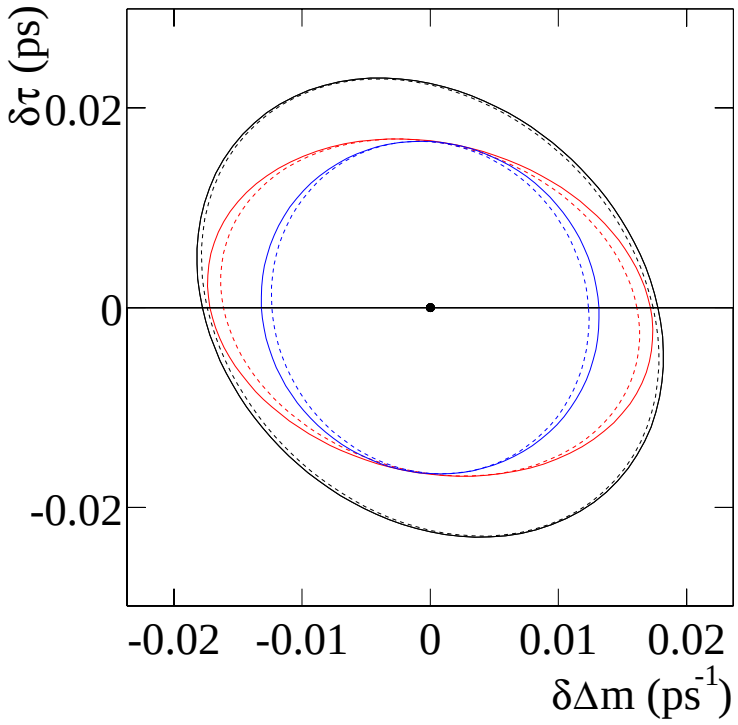


Figure 12.6: Comparison of one-sigma error ellipses in the Δm_d-τ_{B^0} plane for fits in which different sets of parameters are free. From the innermost to the outermost ellipse, the floating parameters are (Δm_d, τ_{B^0}), (Δm_d, τ_{B^0}, mistag fractions), (Δm_d, τ_{B^0}, f_{B^+}), (D^-, τ_{B^0}, f_{B^+}, mistag fractions), all signal Δt parameters, and the default fit (72 floating parameters).

The covariance matrix we report above is estimated from the partial derivatives of the likelihood surface near the minimum and assuming locally parabolic behavior. However, the log-likelihood function may not be symmetric and the real error (deviation from the minimum such that log-likelihood changes by 0.5 units) may not

be the same as that estimated from the curvature. The more correct technique involves actually stepping along each single parameter at a time and re-minimizing the negative-log likelihood at each step until it changes by required 0.5 units.* Due to the large number of floating parameters, it is impractical to measure the true errors on every parameter, however, because this would involve repeating the full fit at least 72 times. However, since we observe that the statistical error contribution from the background Δt parameters to τ_{B^0} and Δm_d is minimal, we can hold them constant and explore the reduced parameter space. We find that the parabolic error estimates for the signal model parameters (including τ_{B^0} and Δm_d) are accurate and symmetric at the 0.1% level.

To estimate the statistical errors at higher accuracy while including the contribution from the background parameters, we scan the log-likelihood function along the $1-\sigma$ contour on the Δm_d-τ_{B^0} plane predicted by the approximate covariance matrix. That is, we fix Δm_d and τ_{B^0} to the values on the $1-\sigma$ contour in Fig. 12.3 and repeat the full fit with all other parameters floating. If the errors from the are correct, the log-likelihood should be greater than the minimum by exactly 0.5 units.

The resulting values of eight points on the one-sigma contour are shown in Fig. 12.7. The percentages shown in parenthesis are how much the error should increase in that particular direction if a parabolic curve is assumed. The largest difference from 0.5 is in the positive τ_{B^0} direction. We therefore correct the positive τ_{B^0} statistical error by 6.5%. For Δm_d and negative τ_{B^0} errors, the correction is negligible.

To more directly understand these conclusions from the strategy outlined in Chapter 9, we visually portray the correlation matrix for the 72 parameters in the final Δt fit in Fig. 12.8. We observe that although all 72 parameters are floating and could have large correlations, because many are restricted to distinct subsamples of the data, the correlation matrix has significant block diagonal structure. We are effectively performing several smaller fits simultaneously. In the limit of control sample purity, the final fit would cleanly factorize into separate fits to each sample to determine its behavior for extrapolation/subtraction from the signal region. From the correlation matrix, however, we do directly observe several significant correlations far

*We use the MINOS [61] analysis package to extract the true error contours.

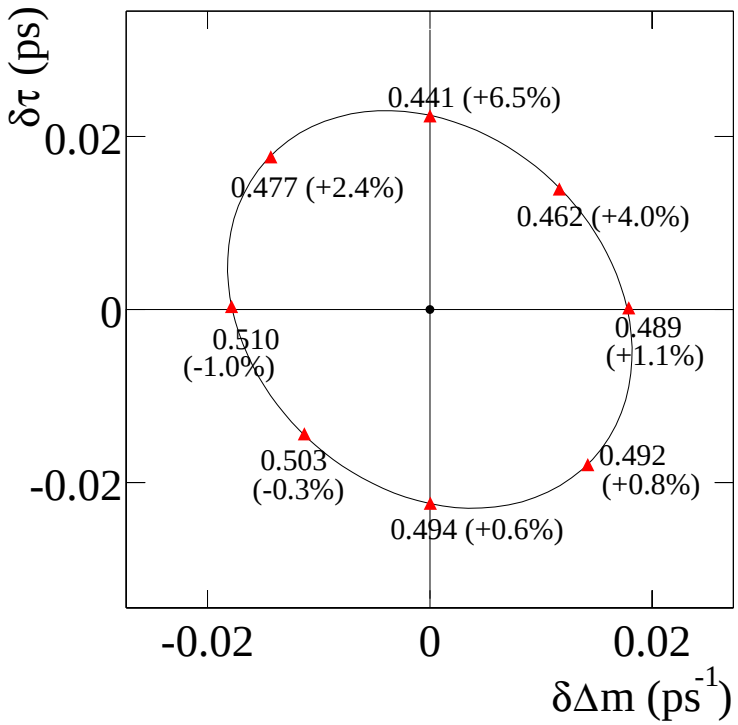


Figure 12.7: The log-likelihood difference between points on one-sigma contours and the minimum. The percentages shown in parenthesis are how much the error should increase in that particular direction if a parabolic curve is assumed.

from the diagonal; these are properly taken into account statistically by the full fit, which is the true power of this technique.

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}, Δm_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. The positive statistical error on τ_{B^0} is increased by 6% to accommodate this observation.

Δt Projections For Backgrounds

The fit result can also be projected onto the background control samples. In Fig. 12.9 the Δt projections for a 99.5% pure sample of combinatoric events is shown, along with the corresponding mixing asymmetry in Fig. 12.10. (This sample was formed by requiring the combinatoric background per-event probability to be greater than 90%, selecting about $33,800$ events.) Similarly, we select a subsample of the Data that is 61% pure in fake lepton backgrounds and project the model onto Δt, as shown in Fig. 12.11 and Fig. 12.12. (This sample was formed by requiring the fake lepton per-event probability to be greater than 40%, selecting about $10,400$ events.) In each case, we observe that the model describes the data well.

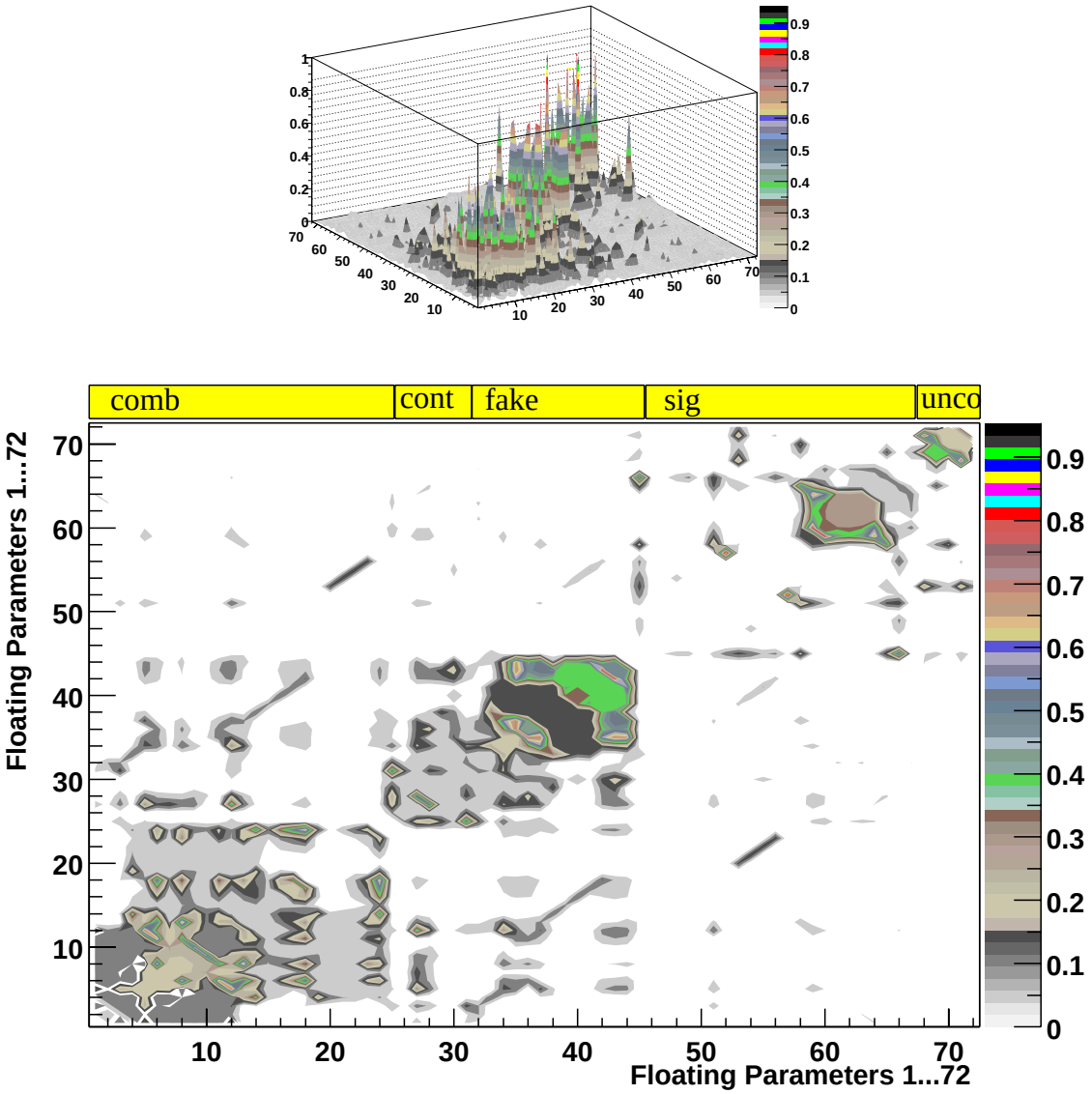


Figure 12.8: Visual representation of the correlation matrix between all 72 free parameters in the final Δt fit. The diagonal elements have been suppressed. The upper plot illustrates the size of the correlations while the lower shows the overall density (same histograms, just different ways to draw them). The band of yellow across the top shows the range of Δt parameters grouped by model type, *i.e.,* signal or combinatoric, continuum peaking, fake lepton peaking, and uncorrelated lepton peaking backgrounds.

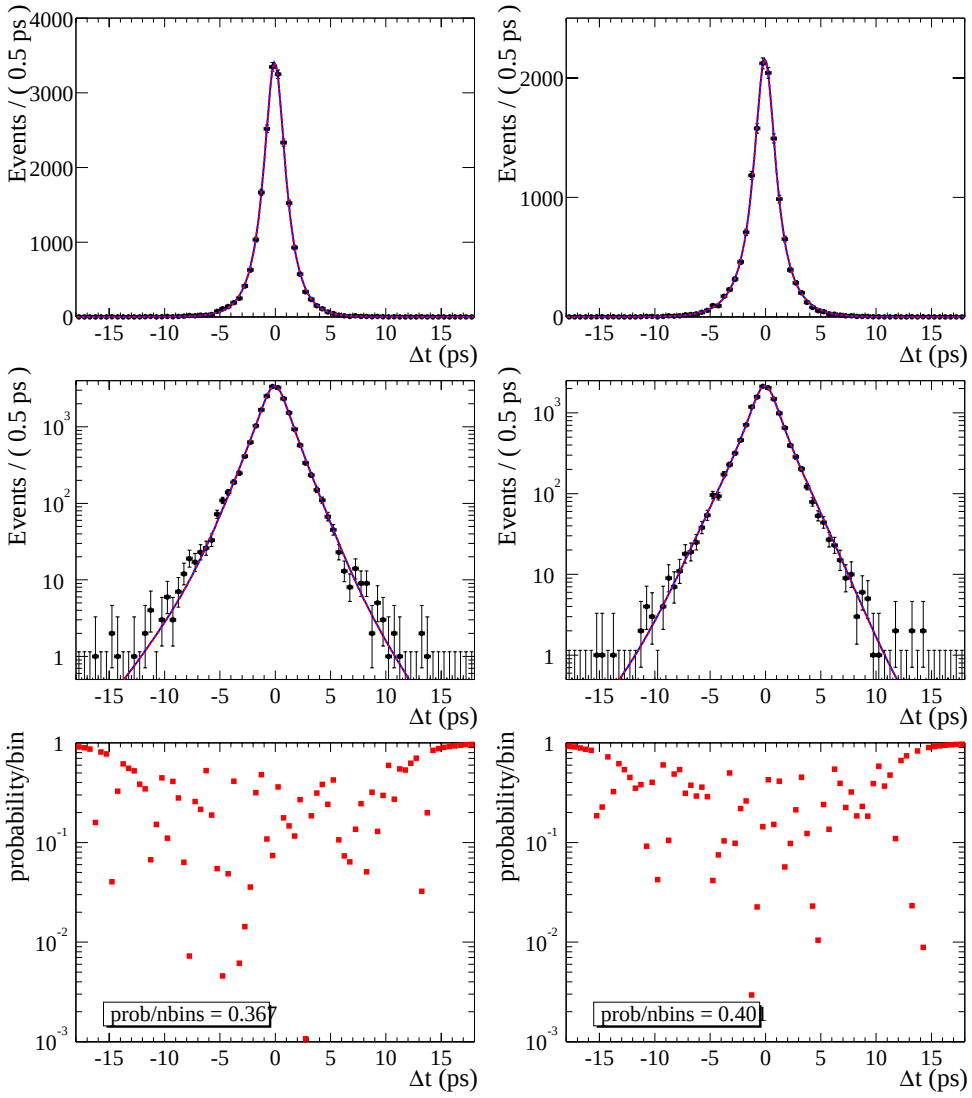


Figure 12.9: The Δt distribution of unmixed and mixed events in a 99.5% pure combinatoric subsample of Data, and the projection of the model. The left hand plots are for unmixed events, the right for mixed events; the middle row is simply a $\log y$ plot. The bottom row is the Poisson probability of 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 $n > \mu$.

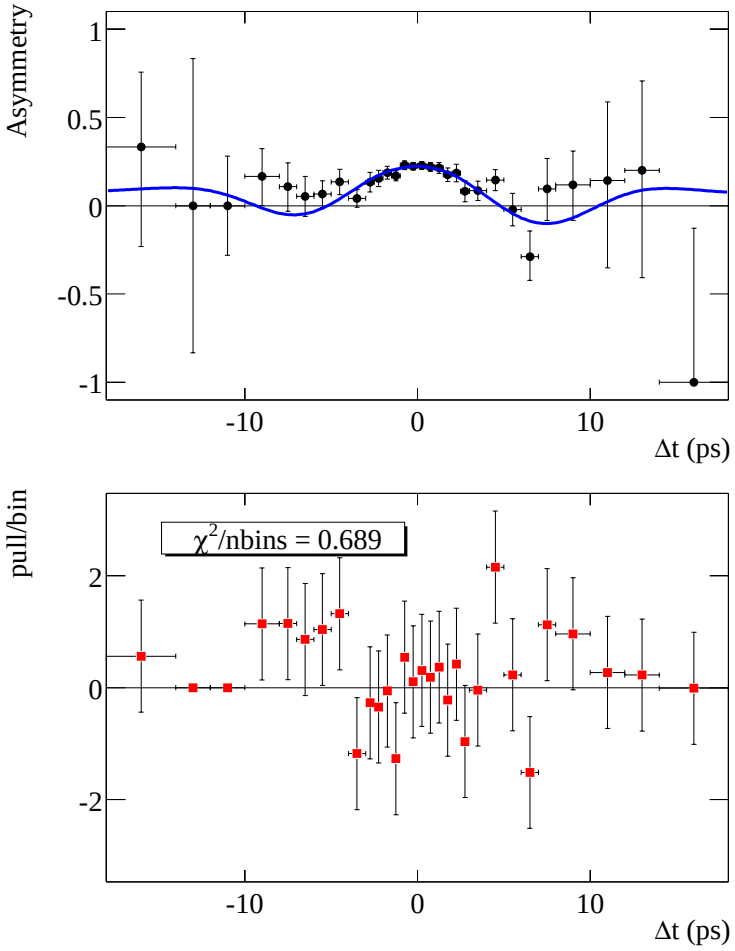


Figure 12.10: The asymmetry plot for the 99% pure combinatoric sample of Data, and the projection of the fit result from Data. The lower plot shows the bin-by-bin difference normalized by the error estimated from Data on each bin.

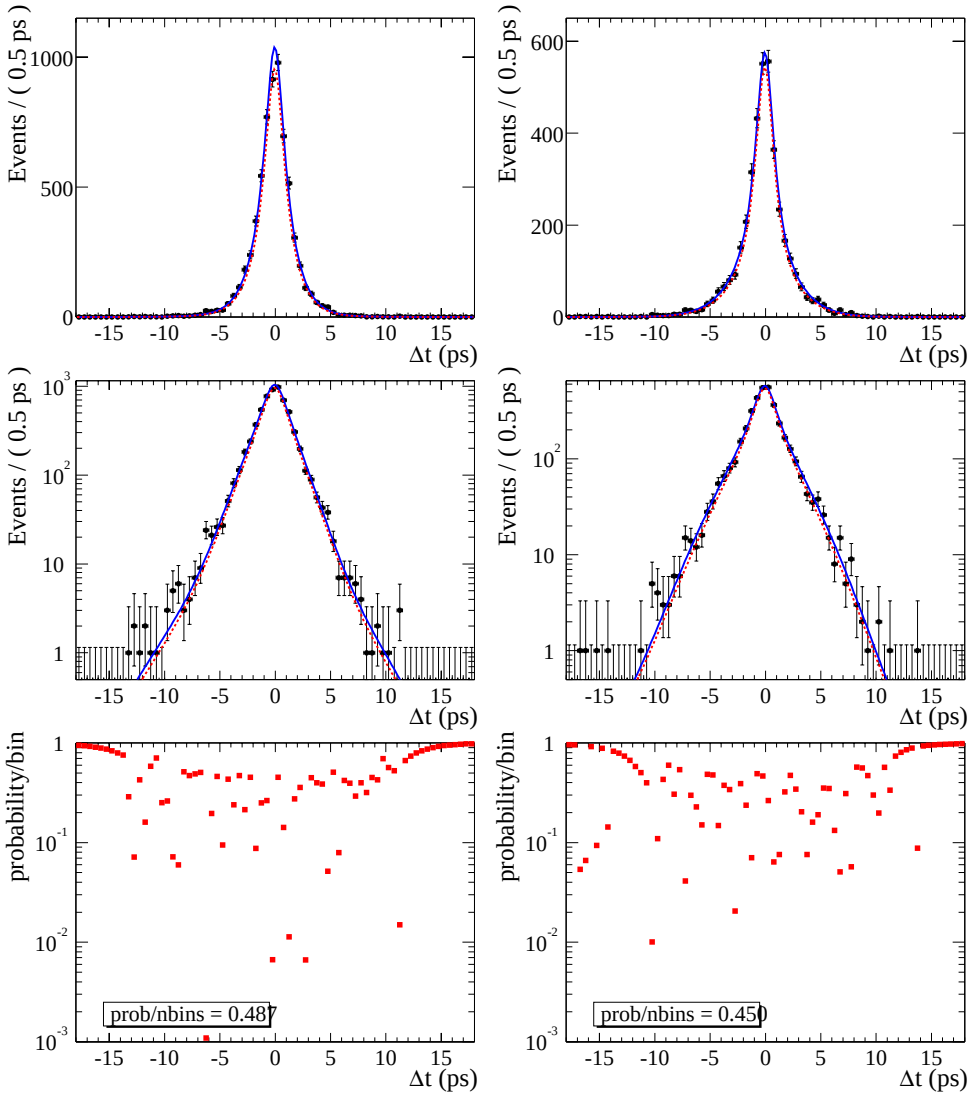


Figure 12.11: The Δt distribution of unmixed and mixed events in a 60% pure fake lepton subsample of Data and the projection of the model. The left hand plots are for unmixed events, the right for mixed events; the middle row is simply a $\log y$ plot. The bottom row is the Poisson probability of observing 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 $n > \mu$. The smooth (blue) curve is the projection of the full model; the dashed (red) curve is the projection of the background component of the full model.

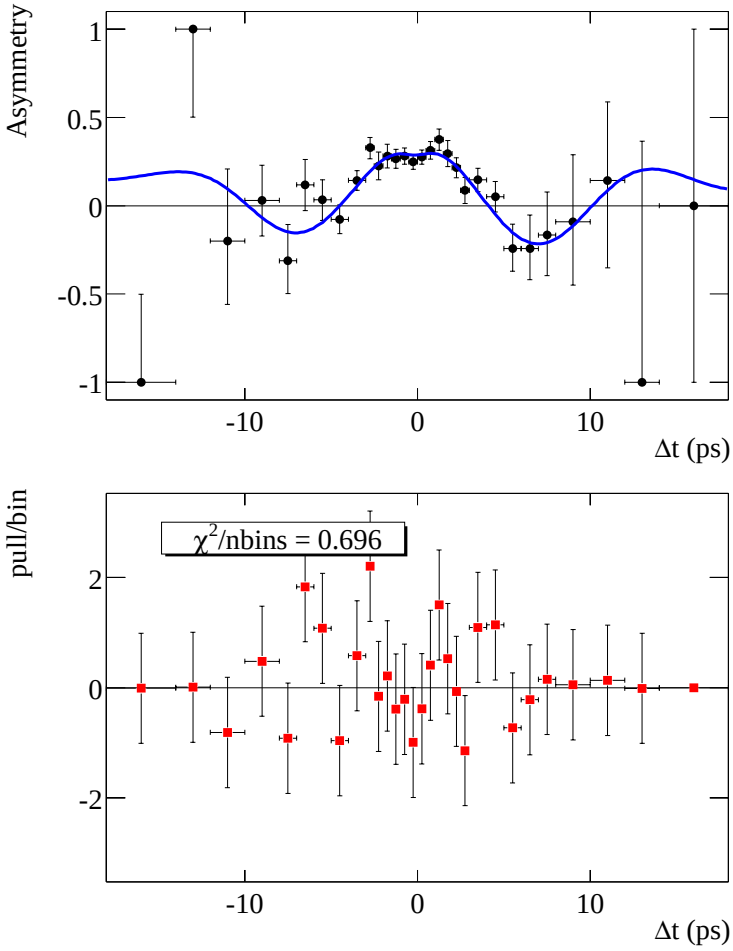


Figure 12.12: The asymmetry plot for the 61% pure fake lepton sample of Data, and the projection of the model fit result from Data. The lower plot shows the bin-by-bin difference normalized by the error estimated from Data on each bin.

Chapter 13

Validation and Consistency Checks

We give the results of performing cross-checks on data, including fitting to different subsamples of the data and fitting with variations to the standard procedure. All tests show the nominal result to be robust and well-behaved. We also describe several tests of the fitting procedure that were performed with both fast parameterized Monte Carlo simulations and full detector simulations.

A small 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. We apply this correction (approximately equivalent to the statistical uncertainty on τ_{B^0} and Δm_d in data) to the final fit results in Data and apply a systematic uncertainty equivalent to the statistical uncertainty on the Monte Carlo fit result.

13.1 Consistency of Subsamples

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 $D^*\ell$ 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. Again, we do not find statistically significant changes in the values of τ_{B^0} or Δm_d. For ease of comparison, tables of results in this section show differences as compared to he baseline measurement from the full Data set. Where possible, statistical errors have been computed to correctly take into account overlaps between samples.

13.1.1 Flavor Subsamples

The sample was split into two subsamples, according to the reconstructed flavor of the signal B^0, $\overline{B}^0$ and by tag side B^0, $\overline{B}^0$ as a consistency check. The results are shown in Table 13.1 for Data. We re-computed the set of per-event (signal and background) probabilities for the two subsamples to maximize our sensitivity to differences in reconstruction between B^0 and $\overline{B}^0$. Because we depend heavily on mixing time structure to separate the charged from neutral B candidates in the sample, these flavor-separated subsample fits have diminished precision. The fit results from the Rec $\overline{B}^0$ subsample did not provide a positive-definite covariance matrix (in part b/c the preferred outlier fraction is so small), but we believe the error estimates to be approximately correct. Additionally, as mentioned earlier, our ability to precisely extract the charged B fraction f_{B^+} is limited. For the Tag $\overline{B}^0$ fit result, we were forced to fix f_{B^+} to be constant and equal to the nominal fit result to guarantee convergence. The errors account for this.

The differences in mistag rates for the Rec B^0, Rec $\overline{B}^0$ categories are consistent with the Δ mistag rates measured in the standard fit. The differences in the Tag B^0, Tag $\overline{B}^0$ samples come from the fact that most events are unmixed, thus there is a strong correlation between an event having a Tag B^0 and a Rec $\overline{B}^0$. The correlation

Full Δt fits (GExp+G) to Data				
	$B_{\rm rec}$		$B_{\rm tag}$	
	B^0	$\overline{B}^0$	B^0	$\overline{B}^0$
$\delta(\Delta m_d)$ (ps^{-1})	0.026 ± 0.018	-0.041 ± 0.018	-0.045 ± 0.010	0.001 ± 0.015
$\delta(\tau_{B^0})$ (ps)	-0.047 ± 0.025	0.041 ± 0.023	-0.015 ± 0.023	0.019 ± 0.025
$f^{\rm out}$	0.005 ± 0.002	0.001 ± 0.019	0.003 ± 0.002	0.001 ± 0.003
$\delta(\omega_{\rm LT})$	-0.011 ± 0.017	0.024 ± 0.012	-0.010 ± 0.011	0.028 ± 0.013
$\delta(\omega_{\rm KT}^{\rm offset})$	0.024 ± 0.031	0.001 ± 0.011	-0.001 ± 0.029	0.006 ± 0.029
$\delta(m_{\rm KT})$	-0.049 ± 0.046	0.043 ± 0.053	0.004 ± 0.043	-0.007 ± 0.042
$\delta(\omega_{\rm N1})$	-0.003 ± 0.023	0.014 ± 0.016	-0.015 ± 0.019	0.017 ± 0.021
$\delta(\omega_{\rm N2})$	-0.017 ± 0.022	0.027 ± 0.016	0.022 ± 0.020	-0.013 ± 0.015
$\delta(\omega_{\rm N3})$	-0.004 ± 0.012	0.011 ± 0.016	0.023 ± 0.011	-0.020 ± 0.012
Others Parameters Suppressed				

Table 13.1: Fitting for τ_{B^0}, mistag rate and Δm_d broken down by reconstructed B^0, $\overline{B}^0$ and tag side B^0, Bzb. The fits use the full data sample and a GExp+G resolution function in the full model. The values listed are the change in central value with respect to the nominal fit. The errors are the difference in quadrature between the subsample and full sample fit results, with the $\Delta\omega_{\rm i}$ parameters fixed to the nominal results, and therefore take into account the correlations due to the common events in both samples.

isn't perfect, so the difference in mistag rates isn't as large. The values of the two physics parameters appear to be consistent across the subsamples.

13.1.2 D^0 Mode Subsamples

The fill fit to Data was performed separately for each D^0 decay mode group. Each `Dmode` category has roughly the same number of peak δm events, although the $K\pi$ mode is the cleanest due to the much reduced combinatoric backgrounds in the three-track final state signal mode. In Table 13.2 we compare the final fit values for each subsample of the Data as grouped by `Dmode`, and find no significant variation between subsamples.

	Full Δt fits (GExp+G) to Data		
Par	$D^0 \to K\pi$	$D^0 \to K\pi\pi\pi$	$D^0 \to K\pi\pi^0/K^0_S\pi\pi$
N_{events}	14523	33621	19760
$\delta(\Delta m_d)$	0.034 ± 0.024	-0.037 ± 0.026	0.019 ± 0.025
$\delta(\tau_{B^0})$	-0.035 ± 0.027	0.017 ± 0.039	0.016 ± 0.032
$\delta(f^{\rm out})$	0.001 ± 0.002	0.000 ± 0.003	-0.002 ± 0.002
$\delta(f_{B^+})$	0.043 ± 0.035	-0.022 ± 0.048	-0.025 ± 0.036

Table 13.2: Difference in fit results between MasterModel fits to the D^0 subsamples and the baseline fit result with Data. Errors listed are the difference in quadrature. We observe no statistically significant difference between the combined D^0 fit to all Data and the separate subsample fits.

13.1.3 Δt-selected Subsamples

Δt selection window

The Data sample was selected applying the criterion $|\Delta t| \leq 18$ ps. We expect the fractional population of outliers to increase at large Δt, so we can test our sensitivity to them by narrowing the Δt selection window. The number of events lost by narrowing the window is small. There are a total of 16 events lost when moving from $|\Delta t|$ <18 ps to $|\Delta t|$ <14 ps in the data sample, and 114 events for the very tight cut on $|\Delta t| < 10$ ps. Table 13.3 shows the effect of using a smaller $|\Delta t|$ range for fits to data.

In these fits, we fix the outlier parameters (bias, width and fraction) to the values determined from the fit to the full ± 18 ps Δt range. Note that the relative fraction of outliers is defined over an infinite Δt range so it makes sense to hold it constant while changing the Δt range for the fit. We are testing the sensitivity of the physics results to the events at large Δt, not our ability to extract outlier parameters when the range of Δt is reduced.

The mistag rates and resolution parameters are insensitive to the Δt selection criterion. The values of τ_{B^0} and Δm_d change very little compared to the statistical errors. The changes are consistent with zero.

Full fits to Data, fixed outliers		
$\|\Delta t\|$ cut (ps)	10	14
$\delta(\Delta m_d)$ (ps^{-1})	-0.0102 ± 0.0188	0.0016 ± 0.0020
$\delta(\tau_{B^0})$ (ps)	0.0381 ± 0.0467	-0.0023 ± 0.0047

Table 13.3: The effect of changing the Δt cut on τ_{B^0}, and Δm_d. The fits use the GExp+G resolution model on the full data sample, and since the physics parameters are blinded, we report the shift with respect to nominal fit result. The parameters of the signal outlier model (bias, width and fraction) are fixed in these fits.

Varying $\sigma_{\Delta t}$ Criterion

The value of the $\sigma_{\Delta t}$ cut was chosen because of the fact that resolution bias parameters scaled linearly with $\sigma_{\Delta t}$ up to a value of 1.8 ps for both resolution models (G+G+G and GExp+G). The events with a $\sigma_{\Delta t}$ larger than 1.8 ps represent less than $\approx$1.5% of the total events in our sample, plus they have the largest uncertainty in the measurement of Δt, so removing these events has a minimal impact on the statistical uncertainty of our measurements. Table 13.4 shows the change in final fit results for the tighter $\sigma_{\Delta t}$ cut of 1.4 ps. Recall that the outlier shape is fixed, although we float the overall fraction. As we select harder on $\sigma_{\Delta t}$, the overall population of outliers decreases, although the physics results are relatively robust. We also performed a full fit to Data with $\sigma_{\Delta t} < 1.4$ ps in which we fixed the outlier fraction of the signal model ($f^{\rm out}$) to the value from a fit to Data with $\sigma_{\Delta t} < 1.8$ ps. In that case, the shift in the fitted value of τ_{B^0} is about twice that shown in Table 13.4, but is still not statistically significant.

13.1.4 Tagging Category Subsamples

The different flavor tagging algorithms have different efficiencies and purities, and are susceptible to different sources of error. To explore the possibility that they might be inconsistent, we alter the signal portion of the model to allow for a separate Δm_d, τ_{B^0}, and Δ-mistag rate for each tagging category. The results are consistent across the five tagging categories, as shown in Table 13.5 for the full fit to Data. The values

Full Δt fits (GExp+G) to Data	
$\sigma_{\Delta t}$ cut (ps)	1.4
$\delta(\Delta m_d)$ (ps^{-1})	0.0012 ± 0.0069
$\delta(\tau_{B^0})$ (ps)	-0.0046 ± 0.0082
$\delta(f_{B^+})$	-0.004 ± 0.007
$\delta(f^{\rm out})$	-0.0013 ± 0.0003

Table 13.4: The effect on τ_{B^0} and Δm_d of changing the $\sigma_{\Delta t}$ cut from the nominal 1.8 ps to 1.4 ps. The fits use the GExp+G resolution model for Data. Mistag parameters are consistent across each fit, but suppressed for clarity. The significant change in outlier population is plausible since the wayward outliers are likely removed by the tighter $\sigma_{\Delta t}$ cut, and thus the outlier population decreases. The central values reported are the change with respect to the baseline value, and the errors are the quadrature difference of statistical errors from the fits.

of τ_{B^0} and Δm_d for each tagging category are consistent with the baseline results. (Comparing quadrature difference of statistical errors is not quite correct due to the different character of this fit.)

13.1.5 Summary

Figures 13.1–13.2 summarize the results of this set of cross-checks on the Data sample.

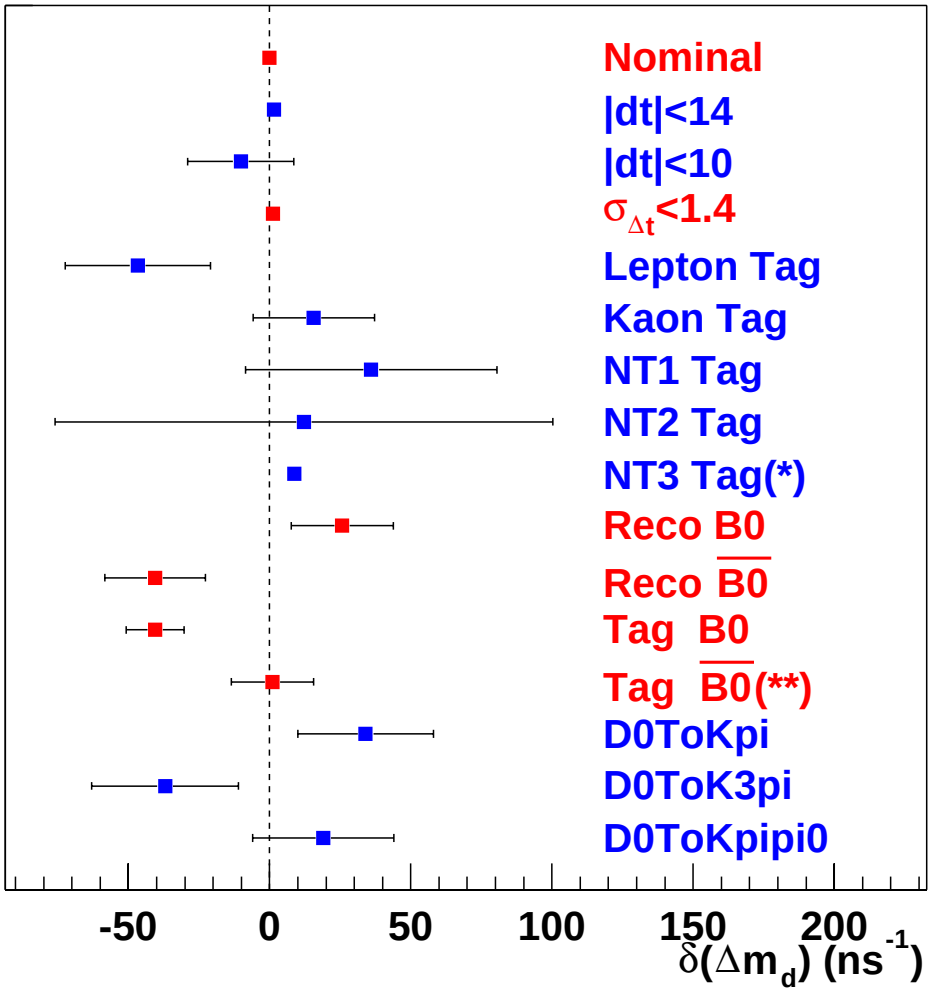


Figure 13.1: These plots show the Δm_d differences from the standard fit for various fits to the Data sample. The units of the plots are ns^{-1}. The errors shown on the plots are the sample and subsample errors subtracted in quadrature. The lines marked with an (*) indicate extremely large errors, which are suppressed. The line marked with (**) represents a fit that only converged when f_{B^+} was fixed to the nominal value.

	Full Δt fit (GExp+G) to Data				
Par	LTag	KTag	NT1	NT2	NT3
$\delta(\Delta m_d)$	-0.047 ± 0.026	0.016 ± 0.022	0.036 ± 0.045	0.012 ± 0.088	0.009 ± 0.183
$\delta(\tau_{B^0})$	0.011 ± 0.048	0.007 ± 0.032	-0.003 ± 0.059	0.023 ± 0.050	-0.066 ± 0.036
$\delta(\Delta\omega)$	-0.003 ± 0.004	0.000 ± 0.003	-0.002 ± 0.005	0.001 ± 0.008	0.000 ± 0.003
$\delta(\omega^{\text{offset}})$	0.007 ± 0.004	-0.005 ± 0.003	-0.008 ± 0.007	-0.001 ± 0.006	-0.001 ± 0.003
$\delta(m_{\text{KT}})$		0.000 ± 0.005			
$\delta(f^1)$	0.125 ± 0.077	0.113 ± 0.073	0.100 ± 0.097	0.100 ± 0.076	0.066 ± 0.064
$\delta(\kappa)$			-0.027 ± 0.079		
$\delta(s^1)$			0.009 ± 0.012		
$\delta(f^{\text{out}})$			-0.002 ± 0.001		

Table 13.5: Master Model fits to the full Dataset using separate Δm_d and τ_{B^0} parameters for each tagging category. The table shows the difference in the fit results with respect to the baseline result which uses the combined statistical power of all tagging categories to measure the physics parameters. The fit is done with the GExp+G resolution model. The units for Δm_d and τ_{B^0} are ps^{-1} and ps. The errors shown are the errors from the fit. The lower section of the table shows parameters which are already split on tagging category and are common for the whole sample; there we report the shift in central values as well.

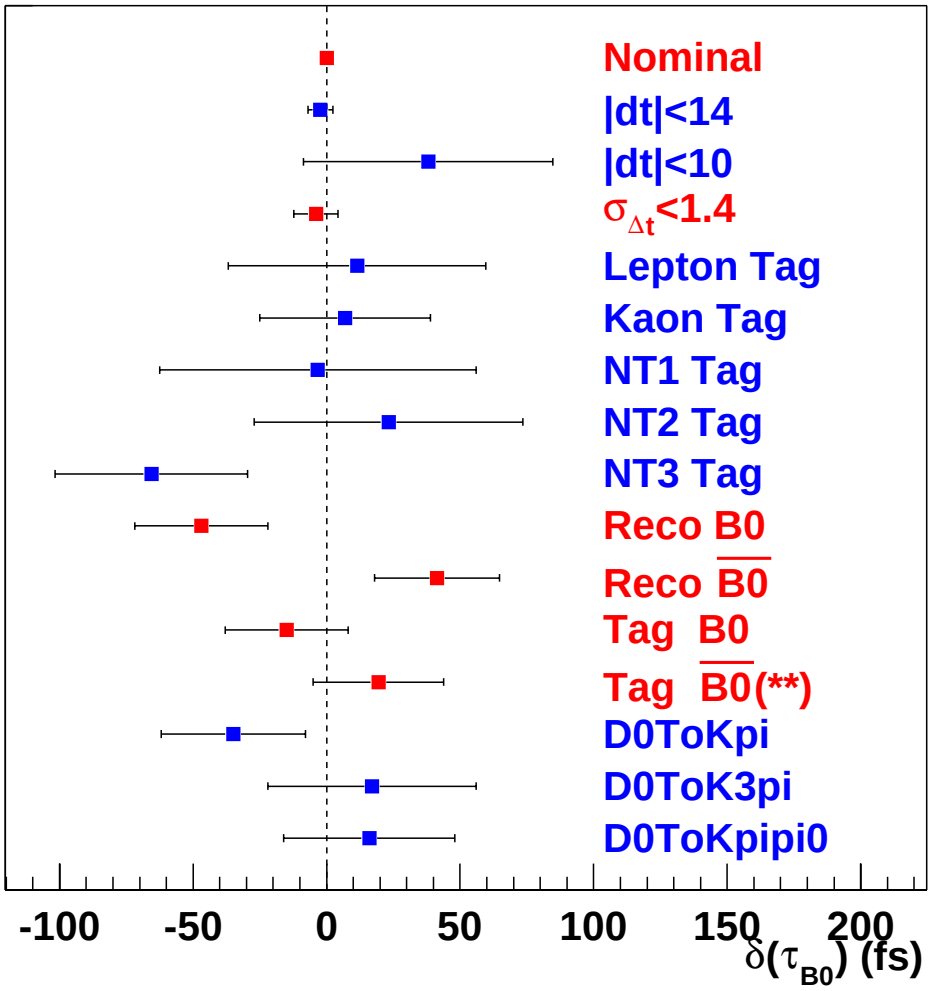


Figure 13.2: These plots show the τ_{B^0} differences from the standard fit for various fits to the Data sample. The units of the plots are fs. The errors shown on the plots are the sample and subsample errors subtracted in quadrature. The line marked with (**) represents a fit that only converged when f_{B^+} was fixed to the nominal value.

13.2 Sensitivity Checks

13.2.1 Upper Limit on $m(D^*) - m(D^0)$

The parameters for combinatoric background are determined primarily by events in the δm sideband. The events in the sideband are similar to the events in the peak region as discussed in Chapter 11.3.1. We test the effect of this assumption in Data by varying the upper limit and repeating the fitting procedure. The default fit uses $\delta m < 165$ MeV. We change this to 162.5, 160.0, 157.5, and 155.0 MeV. The lifetime changes by $+0.6$, -1.2, $+1.0$, $+0.4$ fs, and Δm_d by $+0.7$, $+0.5$, -0.5, -0.2×10^{-3} ps^{-1}, respectively. There is no systematic effect due to the choice of the sideband upper limit.

13.2.2 NT3 Tagging

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 for this 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, which is called the "NoTag" category in other analyses. 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 of NT3 is 0. This is a non-trivial assumption – we are, in effect, assuming a value of χ_d for the NT3 tagged events when we do this; *i.e.,* we fix the mistag probability at 50% rather than fitting for it.

We fix the NT3 mistag rates ($\omega_{\rm NT3}$) 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. Therefore, in this fit, we are not totally fixing NT3 tagging information, but very close.

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 larger (in quadrature); the change in Δm_d 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 actually slightly beneficial. (Clearly, including them in a lifetime-only fit is helpful.)

13.2.3 Δm_d From Shape or Counting Information

In this analysis, we use both Δt-shape information and flavor-tagging information simultaneously to extract Δm_d, and τ_{B^0}. That is, we fit both the mixed and unmixed event sample simultaneously for a common value of Δm_d, using the knowledge that χ_d (the fraction of mixed events in the sample) depends explicitly on Δm_d and τ_{B^0}. An important consistency check is to consider what value of Δm_d (and what statistical error) we might measure using only the time information or only the flavor tagging information (*i.e.,* mixing status) of each event. (By "flavor information," we refer to the relative fraction of the observed number of mixed and unmixed events; *i.e.,* a "time-integrated" mixing measurement.) We use the full signal Monte Carlo sample and the nominal signal model in this study to establish the highest precision for our conclusions.

Δm_d From Δt Only

To estimate Δm_d using only Δt-shape information (and the mixing status of the events, but not their relative proportions), we take advantage of one of the design features of **RooFit** – the lack of a strong distinction between parameters and dependents of a pdf at construction time. The **RooFit** pdf only becomes useful when it is assigned to a dataset, which in turn it uses to distinguish which parameters are dependent variables (i.e. "input" variables, having a value for each event) and which are parameters (to be varied as part of the maximum likelihood fitting procedure). Our standard signal Δt model is associated with datasets containing both Δt and mixing status (tag) information. We eliminate the mixing information from the dataset as a dependent variable in the fit, and then can fit for Δm_d (and even τ_{B^0} and the resolution model if we choose) using only the Δt information. To achieve an

answer most comparable with the the flavor-only result, we fix all parameters in the Δt-only fit to be constant except for Δm_d. For signal Monte Carlo, we then extract $\Delta m_d^{(shape-only)} = 0.47627 \pm 0.00487$ ps^{-1}, which can be compared with the full fit result of $\Delta m_d = 0.47257 \pm 0.00357$ (when only floating Δm_d).

Δm_d From Flavor Only

To estimate Δm_d using only flavor information, we can naïvely calculate it using χ_d as measured in signal Monte Carlo which amount to simply counting mixed and unmixed events. This yields $\chi_d = 0.174 \rightarrow \Delta m_d = 0.4705$ using the measured τ_{B^0} and in the absence of mistag rates. An alternative technique is to return the full signal fit, but fix the resolution function to a very, very wide gaussian (width > 18 ps) so that all Δt information is significantly smeared, and re-fit for Δm_d only. In this case, using the mistag rates obtained from the full fit, we obtain $\Delta m_d^{(flavor-only)} = 0.47361 \pm 0.00533$ ps^{-1}.

We note that the two results are consistent with the full fit, and both achieve a comparable statistical error, indicating that both shape and flavor information are important in obtaining the most precise value of Δm_d from the fit. (Note, for instance, the statistical error on the simultaneous fit is the quadrature sum of the inverses of the independent errors, i.e. $1/\sigma^2_{tot} = 1/\sigma^2_{shape-only} + 1/\sigma^2_{flavor-only}$.)

Δt-Only Fits to Mixed and Unmixed Events

We also consider fitting for both τ_{B^0} and Δm_d using the Δt-shape only method for mixed or unmixed events separately. In order to optimize convergence of the Δt-shape-only models, we fixed the mistag rates at the values obtained from the full fit and set them constant. Table 13.6 illustrates the fitted central values and their statistical errors for τ_{B^0} and Δm_d for this set of Δt-shape-only fits. We note, in particular, that the measured value of τ_{B^0} is more precise in the mixed event sample despite the smaller number of events primarily due to the increased "temporal separation" between oscillation and decay time dependence.

	Δt-shape only fits to signal MC	
Sample	Δm_d [ps^{-1}]	τ_{B^0} [ps]
Unmixed evts	0.47653 ± 0.00711	1.5508 ± 0.0160
Mixed evts	0.47217 ± 0.00903	1.5522 ± 0.0121
All evts	0.47685 ± 0.00520	1.5568 ± 0.0091

Table 13.6: Fitted values from a full signal Δt-shape fit to different samples of signal MC. The mistag rates were held constant in these fits.

We can conclude from the central values that the mixed and unmixed event subsamples yield consistent Δt-shape-only estimates of Δm_d and τ_{B^0} at better than the 0.5σ level.

We also perform a check where we fix all parameters in the fit to those values obtained from the full Δt fit to the full signal MC sample, except for Δm_d. We then extract the two values, $\Delta m_d^{(shape-only)}(mixed) = 0.47867 \pm 0.00595$ and $\Delta m_d^{(shape-only)}(unmixed) = 0.47147 \pm 0.00847$ ps^{-1}, which are consistent at the 1.5σ level, and when combined with a weighted average yield a value of $\Delta m_d^{(shape-only)}(est) = 0.47444 \pm 0.00486$ ps^{-1}. This estimate has nearly the same error and central value as the corresponding simultaneous fit.

13.3 Comparison with Simulation

13.3.1 Signal MC

We perform fits to the signal Monte Carlo sample that includes full detector simulation to check for event selection bias and to validate our procedure. We take advantage of the Monte Carlo truth information to quantitatively evaluate our performance.

To check whether the selection criteria introduce any bias in the lifetime or mixing frequency, we fit the signal physics model to the true distribution of Δt, using true tagging information, for a large sample of signal Monte Carlo events that pass all selection criteria. We obtain the results shown in Table 13.7 which are unbiased

with respect to the generated values of $\tau_{B^0} = 1.548\,\text{ps}$ (significance $= -0.67\sigma$) and $\Delta m_d = 0.472\ \text{ps}^{-1}$ (significance $= -1\sigma$).

Signal Δt fit to Sig MC truth	
τ_{B^0} (ps)	Δm_d (ps^{-1})
1.5446 ± 0.0051	0.4708 ± 0.0013

Table 13.7: Results for τ_{B^0} and Δm_d as fit to the true Δt and tagging information from signal MC.

We also compare the measured mistag rates in the signal Monte Carlo sample with those obtained by comparing the reconstructed mixing status with the true mixing status, as shown in Table 13.8. No significant bias is observed.

Comparison of Measured vs True Mistags in Sig MC					
Source	ω_{LT} (%)	ω_{KT} (%)	ω_{N1} (%)	ω_{N2} (%)	ω_{N3} (%)
Truth-counting	6.3	15.7	19.2	33.8	46.5
Signal Δt Fit	5.6 ± 0.3	15.8 ± 0.9	18.4 ± 0.6	33.9 ± 0.5	46.2 ± 0.4

Table 13.8: Comparison of mistag rates as calculated from the number of wrongly tagged events according to the MC truth versus those extracted from the signal model fit to signal Monte Carlo. (We have used the average value of $\sigma_{\Delta t}$ to extract the average mistag rate from the fit result for ω_{KT}.)

We also fit the measured Δt distribution, using measured tagging information, with the complete signal Δt model described in Chapter 11.2.1, alternately fixing τ_{B^0} or Δm_d, and then floating them both. We find no statistically significant bias in the values of τ_{B^0} or Δm_d extracted in these fits. Table 13.9 summarizes these measured values from the signal Monte Carlo sample.

13.3.2 Generic MC

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

Signal Δt (GExp+G) fits to signal MC			
	$\delta(fix\tau_{B^0})$	$\delta(fix\Delta m_d)$	fit τ_{B^0}, Δm_d
Δm_d (ps^{-1})	0.0007 ± 0.00034	0.472	0.4715 ± 0.0043
τ_{B^0} (ps)	1.548	-0.0004 ± 0.0006	1.5523 ± 0.0091
Mistag Rates			
ω_{LT}	$-0.00016 \pm 1.0e-4$	$-9.4e-05 \pm 3.7e-5$	0.0561 ± 0.0035
$\omega_{\mathrm{KT}}^{\mathrm{offset}}$	$-0.00021 \pm 1.9e-4$	$-4.2e-05 \pm 2.1e-4$	0.0672 ± 0.0084
m_{KT}	$8.0e-05 \pm 2.7e-4$	$-8.0e-05 \pm 3.1e-4$	0.1275 ± 0.0121
ω_{N1}	$-0.00015 \pm 1.8e-4$	$-9.0e-05 \pm 1.3e-4$	0.1840 ± 0.0061
ω_{N2}	$-6.0e-05 \pm 1.6e-4$	$-5.0e-05 \pm 1.4e-4$	0.3391 ± 0.0054
ω_{N3}	$0 \pm 1.2e-4$	$-1.0e-05 \pm 1.1e-4$	0.4617 ± 0.0040
Δ Mistag Rates			
$\Delta\omega_{\mathrm{LT}}$	$-9.9e-06 \pm 6.8e-5$	$3.5e-06 \pm 9.0e-5$	0.0044 ± 0.0058
$\Delta\omega_{\mathrm{KT}}$	$-4.0e-06 \pm 3.0e-5$	$-4.2e-05 \pm 3.0e-5$	-0.0140 ± 0.0045
$\Delta\omega_{\mathrm{N1}}$	$5.0e-06 \pm 4.4e-5$	$-6.0e-06 \pm 6.3e-5$	0.0120 ± 0.0098
$\Delta\omega_{\mathrm{N2}}$	$-4.6e-06 \pm 1.0e-7$	$1.6e-06 \pm 1.0e-7$	-0.0415 ± 0.0083
$\Delta\omega_{\mathrm{N3}}$	$2.0e-07 \pm 1.0e-7$	$1.0e-07 \pm 1.0e-7$	-0.0742 ± 0.0061
Resolution Pars			
κ	0.0053 ± 0.0023	0.0008 ± 0.0017	1.1598 ± 0.0680
s^1	0.0058 ± 0.0025	0.0011 ± 0.0030	1.0017 ± 0.0216
f^1_{LT}	0.00037 ± 0.0024	-0.00031 ± 0.0010	0.8412 ± 0.0230
f^1_{KT}	0.0012 ± 0.0029	$9.0e-05 \pm 0.0011$	0.6864 ± 0.0219
f^1_{N1}	$7e-05 \pm 0.0030$	-0.00026 ± 0.0012	0.8532 ± 0.0289
f^1_{N2}	0.00085 ± 0.0029	0.00018 ± 0.0010	0.7397 ± 0.0248
f^1_{N3}	0.00093 ± 0.0026	0.00026 ± 0.0009	0.7188 ± 0.0202
f^{out}	0.00031 ± 0.0001	$2.1e-05 \pm 7.7e-5$	0.0044 ± 0.0015
b^{out}	-	-	-1.68
s^{out}	-	-	5.44

Table 13.9: The effect of fixing τ_{B^0} and Δm_d to the generator values compared to the baseline result. These fits use all correctly reconstructed signal MC events. All fits use the GExp+G resolution model; the right most column is the "standard" fit to which all other fits are generally compared. The outlier width is fixed to a value of 5.44 ps for all fits, and the bias to -1.68 ps. The errors are statistical, and represent the difference in quadrature between the nominal result w/fixed τ_{B^0} or Δm_d and the result from the corresponding fixed generator value fit.

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.

The generic Monte Carlo sample is a mixture of 14.5 fb^{-1} of $B^0\overline{B}^0$, 16.4 fb^{-1} of B^+B^- and 7.6 fb^{-1} of $c\overline{c}$ events. Because the event selection efficiency for Monte Carlo events is higher than that for data, we have more events (72677) selected from generic MC than from data (67904). Since we do not have off-resonance Monte Carlo, we artificially divide the $c\overline{c}$ sample into two samples; one has seven times as many events as the other, and the smaller sample is marked as off-resonance events.

For this validation exercise, we perform a complete fit to generic Monte Carlo. The ratio of B^+ to B^0 lifetimes is fixed at their generated value, $r = 1.069$, and the ratios of B^+ to B^0 mistag fractions are fixed at the values from MC truth counting.

We conduct a study similar to that described in Chapter 12.1.1 to determine the best values for the signal outlier model shape parameters, $b^{\rm out}$ and $s^{\rm out}$. We cannot expect the population of outliers in Generic MC to be identical to those in true Data, so we undertake the full "scan" of fixed outlier shape parameters as before. As suggested in Chapter 11.2.2, we appear to be unlucky with Generic MC in that the preferred value of the outlier fraction is close to zero and slightly negative.

The likelihood surface for Generic MC in the $b^{\rm out} - s^{\rm out}$ space is much different than that for Data. All of the 33 different fits used for the scan in outlier shape space resulted in an essentially zero outlier fraction, and a large f_{B^+}. (See Fig. 13.3.) We observe that since the outlier fraction is essentially zero, the minimum likelihood value (and fit result) is mostly insensitive to the details of the outlier shape. It is only at the extreme values of the outlier shape parameters ($b^{\rm out} = -8, s^{\rm out} = 10$) ps that the likelihood starts changing dramatically – this is where the outlier fraction is becoming increasingly negative, about -0.3%.

We choose the fixed values for the Generic MC full fit model for two reasons:

- The signal MC fits using just the signal Δt model preferred outlier shape parameters ($b^{\rm out} = -1.68 \pm 0.99$, $s^{\rm out} = 5.45 \pm 0.86$) ps.

- The weight of the (Δm_d,τ_{B^0}) distribution due to variations in the outlier parameters is centered near $(b^{\rm out}, s^{\rm out}) = (-2, 6)$ and $(0, 8)$ ps.

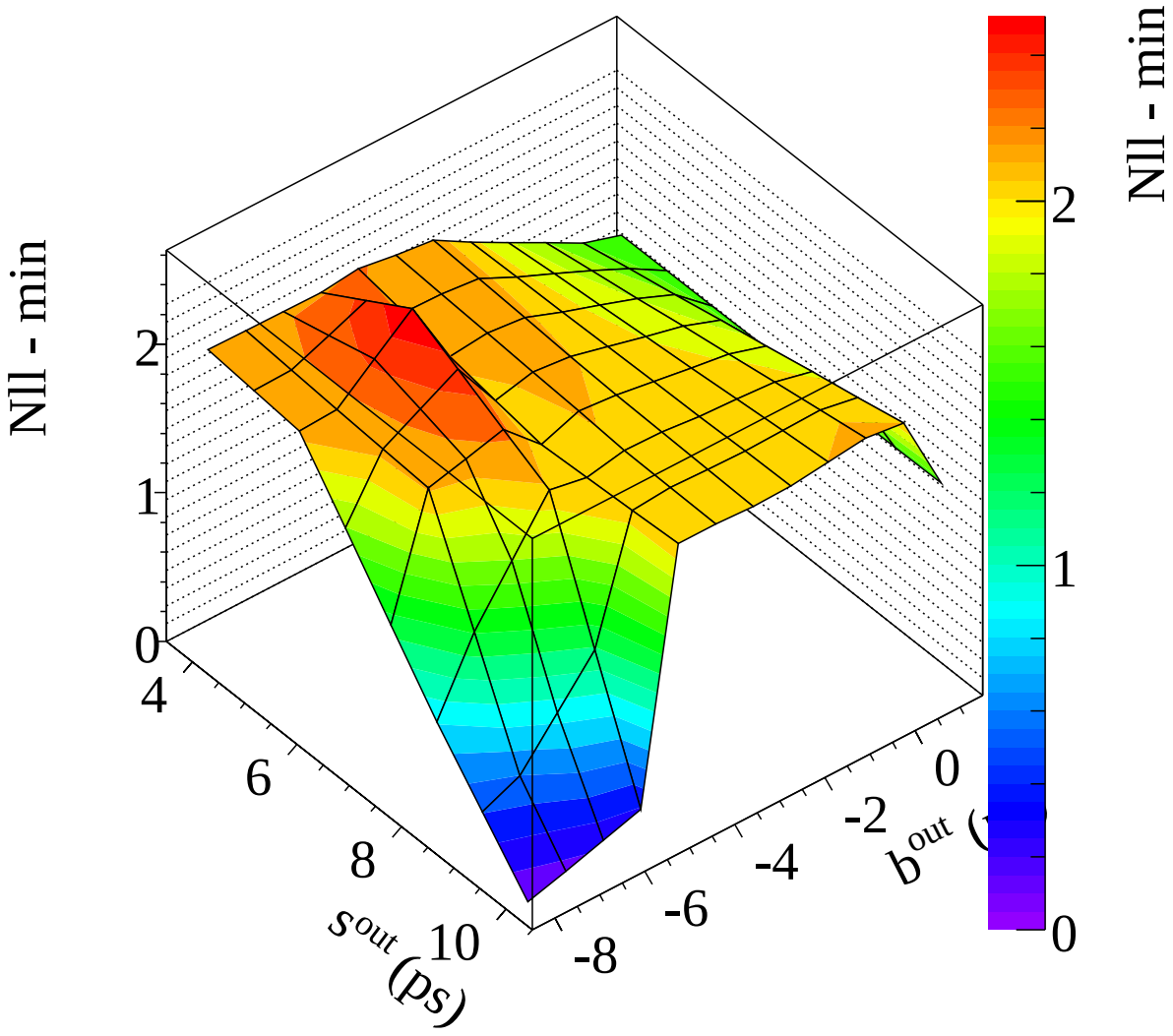


Figure 13.3: The negative-ln-likelihood surface in the space of bias and outlier width of the outlier signal resolution model, as fitted in Generic MC. The vertical scale has been offset so that minimum is at 0. The surface is constructed from 33 fits to Generic MC with different fixed values of the signal outlier parameters. This surface is essentially flat until the edges where the extreme values of the model make it unstable, *i.e.* the fitted outlier fraction becomes significantly less than 0. We do not observe a physical, preferred minimum.

We prefer to use the set of fixed integer values closest to the fit results obtained from the high statistics signal MC sample, so we choose, for Generic MC, to fix the signal outlier shape parameters to be $(b^{\rm out}, s^{\rm out}) = (-2, 6)$. (We note in passing that the other preferred value (8,0) is what other time dependent analyses have assumed for their outlier models.)

Table 13.10 compares the fit results for Generic MC to the true values from the event generator, and to the best-fit results from the signal fit to the truth-matched Generic MC sample as appropriate.

Full Fit to Generic MC, Comparing Signal Model Pars					
$\delta(par)$	value	$\delta(par)$	value	$\delta(par)$	value
Δm_d (ps^{-1})	0.014 ± 0.015	f_{B^+}	0.043 ± 0.028	s^1	0.026 ± 0.064
τ (ps)	0.032 ± 0.024	$\omega_{\rm LT}$	-0.015 ± 0.011	κ	0.041 ± 0.19
-	-	$\omega_{\rm KT}^{\rm offset}$	-0.062 ± 0.027	$f^1{}_{\rm LT}$	0.040 ± 0.055
-	-	$m_{\rm KT}$	0.081 ± 0.039	$f^1_{\rm KT}$	-0.023 ± 0.060
-	-	$\omega_{\rm N1}$	0.041 ± 0.018	$f^1{}_{\rm N1}$	-0.067 ± 0.092
-	-	$\omega_{\rm N2}$	0.041 ± 0.015	$f^1{}_{\rm N2}$	-0.057 ± 0.073
-	-	$\omega_{\rm N3}$	-0.029 ± 0.011	$f^1{}_{\rm N3}$	0.011 ± 0.054
-	-	$\Delta\omega_{\rm LT}$	-0.015 ± 0.018	$f^{\rm out}$	0.0024 ± 0.0022
-	-	$\Delta\omega_{\rm KT}$	-0.020 ± 0.014	$b^{\rm out}$	0.32
-	-	$\Delta\omega_{\rm N1}$	0.041 ± 0.029	$s^{\rm out}$	0.55
-	-	$\Delta\omega_{\rm N2}$	0.041 ± 0.024	-	-
-	-	$\Delta\omega_{\rm N3}$	-0.029 ± 0.018	-	-

Table 13.10: Full fit to generic MC — comparison of signal model results to generator truth (for Δm_d, τ_{B^0}, and f_{B^+}) and truth-matched generic MC results for everything else. (Recall that$s^{\rm out}$ and $b^{\rm out}$ for the signal resolution model are fixed to constant values.) Errors listed are the error on the Gen MC fit result, the appropriate σ for measuring degree of difference.

These fit results bear some discussion, particularly the signal parameters. Several observations are in order:

- The observed outlier fraction is low by, perhaps, about 1σ.
- The measured Δm_d value is 1σ high, the measured τ_{B^0} value is 1.4σ too high, each with respect to the true values from the generator level. They are 1.5σ and

0.58σ too high (respectively) compared to the values from the truth-matched sample of Generic $B\overline{B}$ and B^+.

- The outlier fraction is still consistent with zero, as originally observed in the fits to only only truth-matched candidates from Generic MC.

- The charged B fraction, f_{B^+}, is significantly higher, by about 1.6σ with respect to the true value, or 1.65σ with respect to the result from the truth-matched fit.

- The observed slope for the linear relation between $m_{\rm KT}$ and $\sigma_{\Delta t}$ is significantly higher, by 2.1σ. Additionally, all the mistag rates seem perturbed by at least 1σ, and this is an underestimate because the truth-match signal events are a subset of the full Generic MC data sample.

The kaon mistag rate is modeled as a linear function of $\sigma_{\Delta t}$ in the signal Δt model. The slope from the fit to signal only in generic MC 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 the similar effect in background events, *i.e.,* we assume the kaon mistag rate in the background Δt model is flat. We can confirm that the kaon mistag rate in background events does 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, such that fixing this slope at 12% does not change τ_{B^0} and Δm_d by more than 0.001. Therefore, we can safely ignore the slope for backgrounds.

The large value of f_{B^+} from the fit to Generic MC is surprising, in particular since the result from the full fit should be highly correlated with the result from the fit to only the truth-matched signal events (See Chapter 11.2.3). We suspect that the our Generic MC sample has an unusually small population of outliers in the large Δt tails which degrades our ability to fit for their overall fraction. This inability, in turn, weakens the signal resolution model which convolutes the charged and neutral B time structures together.

To investigate this claim, we observe the correlation coefficients in the final MasterModel Generic MC result shown in Table 13.12. We see that an upward fluctuation in f_{B^+} by 1.6σ would induce a downward fluctuation in $f^{\rm out}$ of about 0.5σ, and cause a positive change in Δm_d of about 0.8σ, The low $f^{\rm out}$ could induce an increase in τ_{B^0} of about 0.4σ. These shifts are not quite sufficient to explain the parameters we observe in the full fit result, however. We do, see, though, that if we fix f_{B^+} to lower values, we can recover performance, as illustrated in Table 13.11.

We have conducted several other studies to understand the source of this bias, but no smoking gun was found. In particular, we used Generic MC truth information to fit with peaking backgrounds or without combinatorics, but the results are hard to interpret and are still consistent with a biased result. With a Generic Monte Carlo sample of this limited size (where the statistical error of the fit result is as large as the possible bias), it is very difficult to isolate cause and effect.

Full fit results to Generic MC w/fixed f_{B^+}						
f_{B^+}	$\delta(f_{B^+})$	Δm_d	$\delta(\Delta m_d)$	τ_{B^0}	$\delta(\tau_{B^0})$	$f^{\rm out}$
0.0474	-0.0047	0.4718	-0.0002	1.5718	0.0238	-0.00008
0.0521	0.0000	0.4732	0.0012	1.5728	0.0248	-0.0003
0.0600	0.0079	0.4756	0.0036	1.5744	0.0264	-0.0006
0.0700	0.0179	0.4786	0.0066	1.5764	0.0284	-0.0011
0.0951	0.0430	0.4856	0.0136	1.5806	0.0326	-0.0019

Table 13.11: Selected results from full MasterModel fits to Generic MC with fixed f_{B^+} to several different values. We directly observe the strong correlations between these four parameters. The columns labelled $\delta()$ reflect the change with respect to the value from MC truth. Recall that the statistical error on Δm_d from these fits is around 0.015 ps^{-1} and about 0.024 ps for τ_{B^0}.

13.3.3 Correction Based on Generic Monte Carlo Sample

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. In order to increase sensitivity,

Corr. coeff's from full fit to Gen MC			
pars	value	pars	value
$\rho(\Delta m_d, \tau_{B^0})$	-0.162	$\rho(m_{\rm KT}, \omega_{\rm KT}^{\rm offset})$	-0.944
$\rho(f^{\rm out}, \Delta m_d)$	-0.169	$\rho(f_{B^+}, \Delta m_d)$	0.552
$\rho(f^{\rm out}, \tau_{B^0})$	-0.402	$\rho(f_{B^+}, \tau_{B^0})$	0.162
$\rho(f^{\rm out}, m_{\rm KT})$	-0.018	$\rho(f_{B^+}, m_{\rm KT})$	0.086
$\rho(f^{\rm out}, \omega_{\rm KT}^{\rm offset})$	0.004	$\rho(f_{B^+}, \omega_{\rm KT}^{\rm offset})$	-0.060
$\rho(f^{\rm out}, f_{B^+})$	-0.333		

Table 13.12: Correlation coefficients from baseline MasterModel fit to Generic MC. (f_{B^+} is correlated with all the mistag rates at only the few % level.)

we include the filtered generic Monte Carlo.* Although the generic plus filtered MC sample has a slightly different background composition from the generic MC, the dominant background (combinatoric) is similar. More over, the biases are consistent between fits to generic MC sample only and fits to the generic plus filtered MC sample. The results are shown in Table 13.13.

The fit to signal events with signal model has been shown to be unbiased with large statistics of signal Monte Carlo sample (Table 13.9). Therefore we test the bias of the full model based on the difference between the fit to signal events with with signal model and the full fit to signal and background events.

The bias is +0.022 ps for τ_{B^0} and +0.020 ps^{-1} for Δm_d. We apply a correction for these biases. For the systematic error due to the statistical uncertainty on this correction, we conservatively apply the full statistical error of the fit result to the generic plus filtered MC sample, *i.e.,* $\pm$0.018 ps for τ_{B^0} and $\pm$0.012 ps^{-1} for Δm_d.

13.3.4 Toy MC

Although there is no definitive test for goodness-of-fit, one can measure how likely the Data is given the full model and fit parameters. We take the final pdf with the

*The "filtered" generic MC is an equivalently sized sample of Generic MC events that have passed a requirement that they have at least one lepton with momentum greater than 1.0 GeV in the center-of-mass system at the truth level. Studies elsewhere [47] have shown this sample to be unbiased for this purpose.

Fits to Generic Monte Carlo Sample					
Type	τ_{B^0}	change(10^{-3})	Δm_d	change(10^{-3})	change wrt
generator	1.548	–	0.472	–	-
truth fit	1.538 ± 0.011	-10 ± 11	0.469 ± 0.005	-3 ± 5	generator
sig model	1.566 ± 0.023	$+28 \pm 20$	0.462 ± 0.013	-7 ± 12	truth fit
full model	1.580 ± 0.024	$+14 \pm 6$	0.486 ± 0.015	$+24 \pm 6$	sig model
Fits to Generic plus Filtered Monte Carlo Sample					
Type	τ_{B^0}	change(10^{-3})	Δm_d	change(10^{-3})	change wrt
truth fit	1.556 ± 0.009	$+8 \pm 9$	0.470 ± 0.004	-2 ± 4	generator
sig model	1.554 ± 0.016	-2 ± 13	0.459 ± 0.011	-12 ± 10	truth fit
full model	1.576 ± 0.018	$+22 \pm 9$	0.479 ± 0.012	$+20 \pm 5$	sig model

Table 13.13: Fit results and biases for generic Monte Carlo sample and for generic plus filtered Monte Carlo sample. The errors on changes are calculated as the difference in quadrature between two errors. This method of estimating error on the change may be a gross overestimate.

fitted parameters and generate dozens of data-sized samples from it, using the Monte Carlo technique. By fitting each sample with the original pdf, we expect to obtain unbiased estimates of the parameters in the case of a consistent description, *i.e.,* we should get out what we put in, which is only the pattern of correlation specified in the model. We can examine the distributions for each fitted parameter to check for bias and accurate error estimation (that is, we expect the RMS of the distribution parameter estimates to be commensurate with the reported statistical error). Finally, we compare the minimized log-likelihood value obtains from the fit to Data with the distribution of those values from the set of Monte Carlo experiments. In the event that the model is as likely to describe the generated dataset as the real Data, the likelihood value from Data should lie within the distribution of those from the Monte Carlo experiments.

We conduct a test of our fitting procedure using this fast parameterized Monte Carlo simulations, where 87 experiments are generated with signal and background control sample sizes and compositions corresponding to that obtained from the full likelihood fit to data. The mistag rates and Δt distributions are generated according to the model used in the likelihood fit. The full fit is then performed on each of these

experiments. We find no statistically significant bias in the average values of τ_{B^0} and Δm_d for the 87 fits. The RMS spread in the distribution of results is consistent with the mean statistical error from the fits and the statistical error on the results in data, for both τ_{B^0} and Δm_d. We find that 17 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. (The distribution of log-likelihood is drawn in Fig. 13.4.)

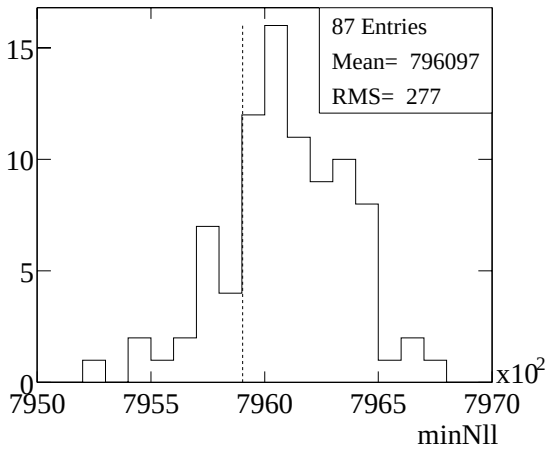


Figure 13.4: The distribution of minimixed negative log-likelihood values from 87 of the fast parameterized Monte Carlo studies described in the text. The dashed line indicates the value found in the Data.

Chapter 14

Systematic Studies

We estimate systematic uncertainties on the parameters τ_{B^0} and Δm_d with studies performed on both data and Monte Carlo samples. We consider each of the key assumptions or conditions under which we perform the analysis and vary it to observe the change in the measured values. Effects considered include:

- *Uncertainty in the estimate of the decay-time difference due to assumptions about the z-scale of the detector, knowledge of the $\Upsilon(4S)$ boost in the lab-frame, internal alignment of the SVT's sensors, and effect of the beamspot constraint in the $B_{\rm rec}$ and $B_{\rm tag}$ vertex fits.*
- *Assumptions about properties of the B^+ meson, such as the lifetime ratio with respect to the B^0, and the assumed ratio of B^+ to B^0 flavor misidentification probabilities.*
- *The accuracy of the background fractions which we use to weight each events as signal or specific background. These probabilities are determined from a separate fit to the δm spectrum and therefore are subject to uncertainty.*
- *Sensitivity to the choice of shape for Δt model of the largest source of background, combinatorics.*
- *Assumptions about the shape and parameterization of the signal resolution function, including the outlier term.*
- *Uncertainty in the size of the correction we apply to the final results because of the bias observed when applying the analysis to Generic Monte Carlo.*

We find an overall systematic uncertainty (assuming the sources are uncorrelated) of $\pm 0.013 ps^{-1}$ on Δm_d and $0.022\,\mathrm{ps}$ for τ_{B^0}.

14.1 Motivation

We now develop an estimate for our uncertainty on the measurements due to systematic effects in our technique. Systematic effects are not statistical by nature; in fact, some define them as sources of uncertainty which do not yield immediately to more data. In this analysis, for instance, we might imagine a systematic error in our procedure due to an error in understanding the z-scale of the inner tracking detector, the SVT. To address systematic error sources, one typically constructs a recipe to vary some condition under which the measurement was made within a range of values consistent with the uncertainty on precise knowledge of that condition. The measurement is then re-performed, and the change in the final result is used as an indication of the degree of sensitivity to the original condition.

In this study, we conduct systematic studies on both the Data and the Monte Carlo event samples, often using the full model incorporating both Δt and δm. We then assume that each source of uncertainty is uncorrelated with the others, and sum them in quadrature to form the final estimate. We neglect, at this point, correlated changes between τ_{B^0} and Δm_d. For instance, the beamspot position is varied, both τ_{B^0} and Δm_d are re-extracted, and the shift in each is computed by comparing with the nominal central value. There is no attempt made to disentangle whether or not the change in τ_{B^0} "cancels" out any of the potential change in Δm_d. This appears to be a safe assumption, as most of the systematic uncertainties are small enough compared to the overall error that a correction due to statistical correlation would still be negligible.

14.2 Uncertainty in Decay-Time Difference

The calculation of a decay-time difference Δt for each event assumes a nominal detector z-scale, PEP-II boost, vertical beam-spot position, and SVT internal alignment. We vary each of these assumptions and assign the variation in the fitted parameters as a corresponding systematic uncertainty.

z Scale

The z scale systematic error is common to many analyses which measure the time difference between B^0 decays. The difference in the z position is used to determine Δt for an event, so a bias in the z scale will directly lead to a bias in the measured Δt. A conservative estimate of the z scale uncertainty is $\pm 0.4\%$ [63], as measured elsewhere by reconstructing electron-neutron that are scattered into the detector volume because of interactions in the beampipe walls. These measurements determine the uncertainty in the z-scale at the radius of the beampipe, and we conservatively enlarge them by a factor of 2 to account for possible differences at the interaction point (where the Δz measurement is made). Using this recipe, the uncertainties are:

- $\delta(\Delta m_d) \leq \pm 0.002$ ps^{-1}
- $\delta(\tau_{B^0}) = \pm 0.006$ ps.

PEP-II Boost Uncertainty

Since Δt is directly proportional to the measured average PEP-II boost, the errors on τ_{B^0} and Δm_d are also directly related to the uncertainty on the boost. At *BABAR*, the $\Upsilon(4S)$ boost in the z-direction is known to relative uncertainty of 0.1%, based on an analysis of PEP-II knowledge of the beam energies. [40] We use this 0.1 % prescription for the systematic uncertainty, *i.e.,*

- $\delta(\Delta m_d) = \pm 0.0005$ ps^{-1}
- $\delta(\tau_{B^0}) = \pm 0.0015$ ps.

Beamspot Position

The beamspot 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 assumed beamspot position by using the entire signal Monte Carlo and varying the beamspot position in y direction, where it is most important. The nominal position

of the beamspot is determined on a detector run-by-run basis using reconstructed Bhabha and dimuon events.

We use beamspot information returned from the **eventInfo** → **beamSpotBFlight()** method, which incorporates a 25 μm width due to the flight distance of the B candidate. This 25 μm is combined with the roughly 10 μm intrinsic width of the beamspot in the y-direction to form an effective spread of 30 μm. For this systematic study we examine two scenarios: one where we systematically shift the beamspot position (via the **BtaLoadBeamSpot::offsetY** parameter), and the other where we smear the position randomly according to a Gaussian of fixed width (via the **BtaLoadBeamSpot::errorYPos** tcl parameter). Tables 14.1-14.2 describe the effect we observe in the final parameter values after performing the full fit (floating the resolution functions) on the signal sample. For these results, we use the GExp resolution model exclusively. (Note: For this study to be generally valid, we assume that detector effects and event selection are symmetric in the y-direction.)

We note that in particular, the lifetime τ_{B^0} is very robust with respect to movements of the beamspot position. The resolution function model adapts easily even to the most significant smearing and offset of the beamspot position. The mixing frequency, Δm_d appears more sensitive, and we note that the measured dilution parameters change slightly as we systematically disturb the beamspot. To investigate this further, we perform another series of signal Δt fits to the different Monte Carlo samples where we use Monte Carlo truth information to determine flavor tagging status. This increases our statistical precision by nearly a factor of two , and allows us to write another (abbreviated) table, Table 14.3.

By studying the z-residuals of each vertex (reco and tag) separately, we also find that this uncertainty in the beamspot constraint affects each vertex about equally. From figures similar to Fig. 6.1 for each of the transformed-beamspot datasets, we write into Table 14.4, which records the effect of the beamspot transformation on the z-residual distribution's width and bias.

Based on these observations, and knowing that (a) the difference in estimated beamspot position between different algorithm is than 10 μm and (b) that average resolution of the beamspot is about 30 μm, we estimate the systematic error due to

Change in fit results to signal MC w/ beamspot variations					
		Beamspot Variation [μm]			
Parameter	nominal err	shift 10	shift 30	shift 80	shift/smear 20
Δm_d	± 0.00435	-0.00054	-0.00129	-0.00202	0.00008
τ_{B^0}	± 0.00878	-0.0042	-0.0001	0.0012	-0.0024
mistag rates	All values very consistent				
κ	± 0.0847	-0.004	-0.0125	-0.0853	0.0307
$f^1_{\rm KT}$	± 0.0304	-0.00371	-0.0106	-0.0214	0.0149
$f^1_{\rm LT}$	± 0.0278	0.00203	0.00343	0.00606	0.0167
$f^1_{\rm N1}$	± 0.034	0.00443	0.00566	0.0005	0.0224
$f^1_{\rm N2}$	± 0.0312	0.00372	-0.00967	-0.0162	0.0169
$f^1_{\rm N3}$	± 0.0271	-0.00274	-0.00909	-0.0177	0.00023
$b^{\rm out}$	± 0.699	-0.124	-0.476	-0.821	-0.269
$f^{\rm out}$	± 0.00318	-0.000225	-0.00101	0.000076	-0.000683
s^1	± 0.0227	0.0159	0.019	0.0997	0.0203

Table 14.1: Difference in final parameter values from the full signal fit to signal Monte Carlo samples for different variations applied to the beamspot position, as compared to a fit to the nominal sample. Here we shift the beamspot position systematically before applying reco- and tag-side vertexing. (The first column of numbers lists the statistical errors on the fitted parameters for the $20\,\mathrm{fb}^{-1}$ nominal signal Monte Carlo sample.)

uncertainties in the beamspot position to be

- $\delta(\tau_{B^0}) = \pm 0.005$ ps
- $\delta(\Delta m_d) = \pm 0.001$ ps^{-1}.

SVT alignment

This analysis relies on the precision vertexing possible with the combined resolution of the SVT and DCH. However, the technique is most sensitive to internal misalignments of the SVT, i.e. to the relative positions of the wafers and strips which report individual hits per track. To estimate this effect, we study the fit results using two different alignment scenarios with signal Monte Carlo, and then consider the consistency of two large subsamples of the Run-1 Data using different SVT alignment

Change in fit results to signal MC w/ beamspot variations				
	Beamspot Variation [μm]			
Parameter	smear 20	smear 40	smear 80	shift/smear 80
Δm_d	0.00053	−0.00022	−0.00295	−0.00577
τ_{B^0}	−0.005	−0.0036	−0.0029	0.0024
mistag rates	All values consistent.			
$f^1_{\rm KT}$	0.00685	0.0104	−0.0126	−0.0271
$f^1_{\rm LT}$	0.00078	0.00036	−0.0197	−0.0108
$f^1_{\rm N1}$	0.0158	0.00839	0.0124	−0.00667
$f^1_{\rm N2}$	−0.00742	0.0107	−0.0265	−0.0328
$f^1_{\rm N3}$	0.00065	0.00945	−0.0107	−0.0189
$b^{\rm out}$	0.383	0.303	0.0398	−0.461
$f^{\rm out}$	−0.000712	−0.000468	0.000463	0.00187
s^1	0.0248	0.0314	0.0947	0.146

Table 14.2: Difference in final parameter values from the full signal fit to the signal Monte Carlo samples for different variations applied to the beamspot position. Here we smear the beamspot position randomly by sampling from a gaussian of the indicated width, before applying reco- and tag-side vertexing.

	Fits to signal MC using truth-tagging					
Par	Nominal Value	shift 10	shift 80	smear 20	smear 80	shft/smr 80
Δm_d	0.472 ± 0.00216	0.00022	−0.00106	0.00008	−0.00026	−0.00262
τ_{B^0}	1.54 ± 0.00686	−0.0017	0.0045	−0.0022	−0.0012	0.0053
κ	0.928 ± 0.0668	−0.0266	−0.0482	−0.00771	−0.00666	−0.0341
s^1	1.02 ± 0.0182	0.0117	0.085	0.0177	0.0874	0.14
$b^{\rm out}$	-1.48 ± 0.357	−0.151	−0.405	0.192	−0.0836	−0.52

Table 14.3: Fit results for select parameters from signal Δt fits to Monte Carlo samples with a shifted and/or smeared beamspot constraint. As described in the text, Monte Carlo truth information was used to determine the flavor status of each candidates, thereby improving the statistical precision on the other physics parameters. The first column of numbers shows the fit results to the nominal signal Monte Carlo sample and their statistical error, the other columns show the change in central value.

sets “D” and “E”.

Characteristics of z-vtx residuals from signal MC					
Variation	Size	z_{reco} Bias	z_{reco} RMS	z_{tag} Bias	z_{tag} RMS
nominal	0	0.05 ± 0.3	77.5 ± 0.2	34.4 ± 0.5	156.1 ± 0.4
shift	10	0.07	77.4	34.6	156.8
shift	80	1.92	85.5	36.7	162.2
smear	20	0.09	77.9	34.1	156.8
smear	80	0.03	85.0	34.8	161.6

Table 14.4: Shape parameters for tag- and reco- vertex z-residual distributions of correctly reconstructed signal Monte Carlo under several different beamspot transformation scenarios. [All units are μm.] As expected, systematic offsets of the beamspot increase the bias and RMS of the z-residual for each vertex separately, while smearing only degrades the RMS of the residuals.

Estimate of bias using signal MC

We form a signal Monte Carlo cocktail of 20 fb^{-1} by combining SP3 simulated events (in luminosity proportion) for the four different D^0 decay modes for the three months of largest luminosity, Jul2000, Aug2000, and Oct2000. The cocktail created with this recipe has an approximate composition as shown in Table 14.5.

D^0 Mode	Evts / 20 fb^{-1}
$D^0 \to K^-\pi$	52k
$D^0 \to K^-\pi\pi\pi$	102k
$D^0 \to K^-\pi\pi^0$	188k
$D^0 \to K^0_S\pi^+\pi^-$	36k

Table 14.5: Numbers of events per D^0 decay mode in the signal Monte Carlo cocktail with luminosity weight totalling 20 fb^{-1}.

An SVT alignment scenario consists of a set of transformations that are applied to the set of individual wafers and modules. The alignment scenario can be used to describe the difference between one state of SVT internal alignment and another. So-called "misalignment" occurs when the particle hits and tracks are reconstructed using incorrect knowledge of the relative positions of the sensors. "Perfect" alignment is therefore the default case, where tracks are reconstructed from hits with completely

accurate knowledge of their relative positions.

The SP3 Monte Carlo default scenario uses so-called "perfect" alignment, i.e. it uses the same alignment scenario for event generation and event reconstruction. In the real detector, however, we do not have the luxury of perfect knowledge of the sensors' locations when the event was "generated," or rather, detected. Through a variety of sophisticated fitting algorithms using dimuon events, we derive an *alignment scenario* for the real detector which describes the "true" positions of the sensors with respect to the nominal positions. During Run-1 and Run-2 data-taking, the SVT internal alignment algorithm was substantially improved, leading to several different procedures for constructing alignment scenarios and some data was processed with each of these alignment scenarios, or "sets."

To estimate the effect of the systematic uncertainties inherent in the estimated positions of the SVT wafers in the real detector, we "misalign" the signal Monte Carlo cocktail (described above) to match the alignment scenarios used in the real data processing and reconstruction. We then look for biases in the parameter values after the full signal fit to the "misaligned" signal Monte Carlo. The key point is that the data sample is reconstructed with imperfect knowledge of the SVT's internal alignment; to mimic the effect of that uncertainty, we perturb the SVT in signal Monte Carlo and the reconstruct the data.

To achieve this in practice, we follow a recipe suggested by the Tracking Group. We assume that alignment set "L" is perfect for Run-1 data, which we then associate with the default perfect alignment of SP3 Monte Carlo. We then assume the alignment sets used in Release-8 processing (28% set "D" and 56% set "E") are additively related to this near-perfect "L" alignment. We derive the alignment transformations "diffDL" and "diffEL," which we claim are reasonable descriptions of the changes in internal SVT alignment (with respect to nominal) necessary to emulate performance under alignment sets "D" and "E". (*i.e.,* We degrade the internal alignment with respect to the alignment used for event generation to match the scenario used for reconstruction. The degradation is chosen to be consistent with the two different misalignments used for data-taking/processing.) We then apply these alignment transformations to the raw Monte Carlo data, and re-run the entire event processing

chain from Bear (reconstruction) to ascii file production (input to the signal fits). We report here on a study of three different datasets, each corresponding to reconstruction, event selection, and signal Δt fitting under a different alignment scenario for the 20 fb^{-1} signal Monte Carlo cocktail:

- Default (nominal SP3),
- diffEL,
- diffDL.

In the full signal fit, we use Monte Carlo truth information to determine the mixing status of each candidate. This improves the statistics used in the study, and focuses our attention on the changes in Δt due to the degraded alignment scenarios. (This also eliminates changes in the fitted physics parameters due to correlations between reconstructed flavor and vertexing. All three datasets were fitted successfully with the full signal model, and yield consistent, less precise conclusions.)

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 14.6 summarizes the change in physics and resolution parameters in the two (mis)alignment scenarios diffDL and diffEL with respect to the default case.

We observe that the central value of Δm_d decreases in both misalignment scenarios, and that τ_{B^0} increases. In the case of Δm_d, the change is nearly comparable to the statistical precision of the nominal fit, indicating that misalignment is probably an important systematic uncertainty. (Note: The statistical error on Δm_d increases from ± 0.00516 in the perfect alignment case to ± 0.00530 in the misalignment scenario. This might suggest that the misalignment itself contributed ± 0.001 to the error on Δm_d, while the resolution model absorbed other effects.)

Based on these results, we assign the following systematic uncertainty to our measurements of Δm_d and τ_{B^0}:

- $\delta(\tau_{B^0}) = -0.0056$ ps

	Effect of misalignment on signal Δt fits to MC		
Parameter	Nominal Value	Δ diffDL	Δ diffEL
Δm_d	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

Table 14.6: 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 different SVT alignment scenarios: the nominal scenario typical of SP3, misalignment scenario diffDL, and misalignment scenario diffEL.

- $\delta(\Delta m_d) = -0.0030$ ps^{-1}

With the understanding that the alignment transformations diffDL and diffEL distort the SVT alignment from nominal (perfect) to imperfect, we recognize that our estimate of Δm_d (τ_{B^0}) from Run-1 data (using the imperfect alignment scenarios setD and setE) may be biased low (high).

Consistency in Data between different alignment scenarios

For this study, we separate the Run-1 Data sample into three subsamples according to the SVT alignment scenario used in event processing and reconstruction: set "D" (29%), set "E" (53%), and "other" (18%). (We perform the grouping by matching event run numbers and conditions information.) We fit the full model to each of the "D" and "E" subsamples, independently, including the background δm analysis for evaluation of background composition/fractions. Although the statistical power of the test is poor, we find results as shown in Table 14.7. The fitted central values for the key parameters are statistically consistent between these two subsets of the Data.

	Full Δt fits to Data	
Parameter	Set "D"	Set "E"
$\delta(\Delta m_d)$	-0.0202 ± 0.0449	-0.0001 ± 0.0308
$\delta(\tau_{B^0})$	-0.0342 ± 0.0434	-0.0083 ± 0.0227
$\delta(f^{\rm out})$	0.0034 ± 0.0035	-0.000088 ± 0.0024
$\delta(\kappa)$	-0.254 ± 0.442	-0.133 ± 0.304

Table 14.7: Results of the full fit to subsamples of the Data, as grouped by SVT alignment set (here, sets "D" and "E"). We show the difference in central values as compared to the baseline fit, but quote the statistical error on the subsample fit result, as we wish to compare the results from Sets D and E with each other.

14.3 Fixed $B^{\pm}$ Properties

The model of the charged B background assumes fixed B^+/B^0 ratios for the mistag rates and lifetimes. We vary the mistag ratio by the uncertainty determined from separate fits to hadronic events. We vary the lifetime ratio by the statistical uncertainty on the world average. [5] The resulting change in the fitted physics parameters is taken as the systematic uncertainty.

In the MasterModel fit to Data, the B^+ to B^0 lifetime ratio is fixed. We vary this ratio up and down by one sigma according to PDG2002 lifetime ratio [5] and repeat the full fit. The mixing and lifetime results are shown in Fig. 14.8. We assign half of the observed difference as the systematic error.

The B^+ to B^0 mistag ratios for all five tagging categories are also fixed. We vary all five mistag ratios jointly by $\pm$ 1σ according to Table 11.3 and repeat the full fit. The results are shown in Table 14.9. Again we assign half of the difference as the systematic due to fixing mistag ratios.

We conclude that the uncertainties on the two physics parameters due to the assumed properties of the $B^{\pm}$ mesons are:

- $\delta(\Delta m_d) = 0.0003$ ps^{-1}
- $\delta(\tau_{B^0}) = 0.0019$ ps

Full Δt fits to Data			
Parameter	$(\tau_{B^+}/\tau_{B^0}) + 1\sigma$	$(\tau_{B^+}/\tau_{B^0}) - 1\sigma$	diff./2
$\delta(\Delta m_d)$ (ps^{-1})	-0.00002	0.0006	0.0003
$\delta(\tau_{B^0})$ (ps)	-0.0016	0.0021	0.0019
$\delta(f_{B^+})$	-0.0025	0.0029	

Table 14.8: Fitted results from 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.

Full Δt fits to Data			
	$(\omega_{B^+}/\omega_{B^0}) + 1\sigma$	$(\omega_{B^+}/\omega_{B^0}) - 1\sigma$	diff./2
$\delta(\Delta m_d)$ (ps^{-1})	-0.00007	-0.00023	0.00008
$\delta(\tau_{B^0})$ (ps)	-0.0002	0.0004	0.0003
$\delta(f_{B^+})$	-0.0020	-0.0026	

Table 14.9: Fitted results from Data for the two physics parameters for different values of the fixed B^+ to B^0 mistag rate ratio.

14.4 Background Fractions

The modeling of the background contributions to the sample determines the probability we assign to each event for it to be signal, and the Δt distribution we expect for background events. We estimate the uncertainty due to the signal probability calculations by repeating the full fit using an ensemble of different signal and background parameters for the $m(D^*) - m(D^0)$ distributions, varied randomly according to the measured statistical uncertainties and correlations between the parameters. We assign the spread in each of the resulting fitted physics parameter as the systematic uncertainty.

The per-event probabilities use two key ingredients: fit results that describe the peak and sideband lineshapes for δm in 360 sub-categories of the data, and background fractions which essentially count the fraction of peak and sideband events in each of the sub-categories (using the lineshape fit results). The fits and subsequent computations that produce this set of values (which are held fixed in the full Δt fit

because we expect Δt and δm to be independent variables) introduce many correlations between them (due to the simultaneous nature of the result). To understand the effect of the statistical error on the calculations needed for per-event probabilities, we therefore perturb this set of values jointly and randomly, taking into account the individual errors on each result as well as the correlation between any two results (via the covariance matrix). We then have a "new" set of values from which to compute per-event probabilities in the full Δt fit.

We examine the distribution of fitted τ_{B^0} and Δm_d for 111 different perturbations of the parameters used to compute the per-event probabilities. In order to optimize our use of computing resources, (and possibly be conservative on our estimate of the error), we fix the other 70 parameters in the full fit to the baseline values. These 2-parameter (τ_{B^0} and Δm_d) run quickly, and allows us to gather sufficient statistics for a conclusion. We also ran full Δt fits to the Data for 20 different random perturbations to verify that the 2-parameter fit results were not an inappropriate estimate. Our results are shown in Fig. 14.1 and Tab. 14.10. When floating only the two physics parameters, we achieve a statistical error from the fit of ± 0.012 ps^{-1} on Δm_d and ± 0.017 ps on τ_{B^0}.

Full fits to Data, floating only Δm_d, τ_{B^0}		
	$\delta(\Delta m_d)$ (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

Table 14.10: 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.

In this study, we performed two separate perturbations to affect the per-event probability calculations (varying the δm shape fit values, and varying the yields fit results used to compute the background fractions). We also treated each perturbation separately, and found that for the 111 fits floating only the 2-physics parameters, the dominant source of spread in the fitted values was due to the perturbation of the background fractions. For instance, the RMS of the distribution of the fitted Δm_d

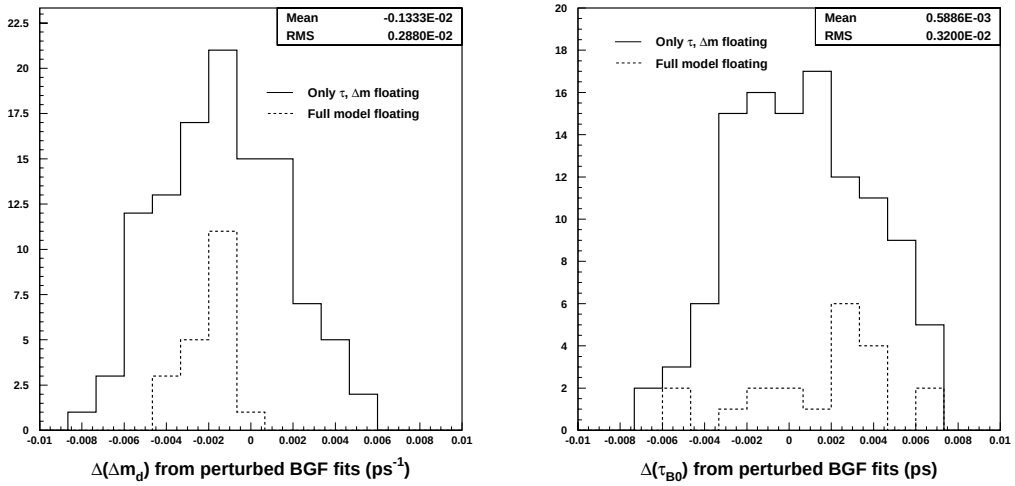


Figure 14.1: Histograms of the spread in fitted results (with respect to the baseline blinded value) from Data using 111 different statistical variations of the input δm parameters and floating only τ_{B^0} and Δm_d in the final Δt fits. The superimposed shaded histograms show the distribution of fit results for the 20 fits where the entire model was floating in the fit.

values from just the shape-fit perturbation was small, about 0.00055 ps^{-1} whereas the RMS of the Δm_d values fitted using just the perturbed background fractions was 0.0033 ps^{-1}. (Likewise for τ_{B^0}, the numbers were 0.0015 ps and 0.0034 ps, respectively.) This suggests that the statistical uncertainty on the δm shape fits is negligible compared to the statistical uncertainty on the background fractions due to statistical fluctuations of the Data populations in each of the 360 categories.

An additional cross-check used each of 20 different sets of perturbed parameters in a full fit to Data. Although the precision is poor, the RMS of the distributions of Δm_d and τ_{B^0} values are consistent with the 111 two-parameter fits, and the bias with respect to the baseline fit is very small. The fitted Δm_d values from the 20 different fits using perturbed background parameters yield an average shift (wrt the nominal fit) of -0.002 with an rms of 0.0012 ps^{-1}. Likewise for τ_{B^0}, the distribution of fitted values have an average of 0.0016 and an RMS of 0.0033 ps. (The error on these RMS values is roughly 20%.) This increases our confidence in assigning the

systematic errors as listed at the end of this section.

Finally, we examine the correlation between τ_{B^0} and Δm_d as a function of the background fraction perturbations. Fig. 14.2 shows the empirical lack of correlation between τ_{B^0} and Δm_d from the 111 different perturbed background fraction results. Using the 20 full fits, however, we can examine the measured correlation coefficient between the two physics parameters. Unfortunately, only 10 of the fits yielded a good measure of the error matrix. From these fits, we find that the average correlation between τ_{B^0} and Δm_d is -0.213 ± 0.003 with an RMS of 0.011.

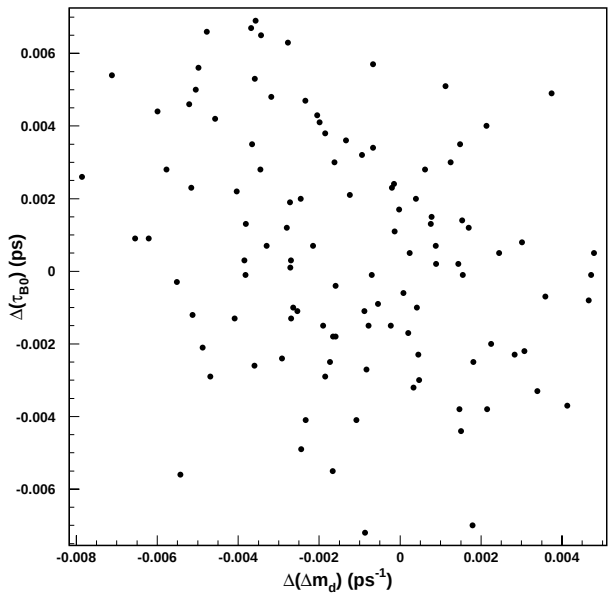


Figure 14.2: Scatter plot of pairs of (τ_{B^0},Δm_d) values (wrt baseline fit) from the 111 2-parameter floating Master fits to the Data, each using a different set of perturbed background fractions in the fit. We observe no induced correlation on τ_{B^0} and Δm_d due to varying the δm fit results and background fractions.

Based on these results, and recognizing that ∼100 tests give us a fractional precision of order 7% on the rms of the distribution, we assign the following systematic errors:

- $\delta(\Delta m_d) = \pm 0.0029$ ps^{-1}

- $\delta(\tau_{B^0}) = \pm 0.0032$ ps.

14.5 Background Models

We estimate the systematic uncertainty due to the assumed background Δt distributions as the shift in the fitted parameters when we replace the model for the largest background (due to combinatoric events) with a pure lifetime model. An important feature of the Δt model we've selected is its effort to accommodate the mixing structure of the combinatoric background. As a suitable variation, we consider modelling the combinatoric background Δt and flavor distribution 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 larger than the default fit value, 795903.8. The charged B fraction, f_{B^+}, 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, we use the size of this shift as an estimate of our systematic uncertainty:

- $\delta(\Delta m_d) = \pm 0.0012$ ps^{-1}
- $\delta(\tau_{B^0}) = \pm 0.0063$ ps.

14.6 Signal Resolution Models

The final category of systematic uncertainties is due to assumptions about the resolution model for signal events. We have largely avoided assumptions by floating many parameters to describe the resolution simultaneously with the parameters of interest. However, two sources of systematic uncertainty remain: the shape of the outlier contribution, which cannot be determined from data alone, and the assumed parameterization of the resolution for non-outlier events.

Outlier Resolution Model

Based on the analysis described in Section 12.1.1, the shape of the outlier model for signal (the Gaussian's width and bias) was fixed. We study the sensitivity to this outlier shape by repeating the full fit with an ensemble of different shapes described by the 36 different pairs of outlier shape parameters. We assign half the full spread of lifetime and mixing fit results as the uncertainty due to the choice of model. The resulting systematic errors are:

- $\delta(\Delta m_d) = \pm 0.0010$ ps^{-1}
- $\delta(\tau_{B^0}) = \pm 0.0054$ ps.

Core Resolution Model

We estimate the uncertainty due to the assumed resolution parameterization by repeating the full fit with a triple-Gaussian resolution model (see Eq.6.16) and assigning the shift in the fitted values as the uncertainty. This test isolates the sensitivity to the shape of the core resolution model, because the outlier gaussian shape is identical, *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 GExp fit, the lifetime shifts by +0.0034 ps and Δm_d shifts by -0.0009 ps^{-1}. The minimized negative log-likelihood is 795904.4, (for reference, the value for the default fit is 795903.8.). We assign the full scale of the shift as the systematic uncertainty due to the core resolution model:

- $\delta(\Delta m_d) = \pm 0.0034$ ps^{-1}
- $\delta(\tau_{B^0}) = \pm 0.0009$ ps.

14.7 Selection and Fit Bias

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 modelling) with Monte Carlo events.

We check for event selection bias by fitting to the truth Δt distributions of signal Monte Carlo sample using the physics-only pdf. There is a slightly negative bias for both mixing and lifetime, as mentioned in Chapter 13.3. We also expect a small positive bias in lifetime due to ignoring the resolution dependence on true Δt, as discussed in Chapter 6.3. However, the result from the fit to signal Monte Carlo events with the default signal Δt model (as shown in Table 13.9) does not show any significant deviation from the generated values. Hence we do not assign any systematic errors for this source.

However, the full fit to generic Monte Carlo does show a bias, also as discussed in Section 13.3. We assign the statistical errors of a full fit to the Monte Carlo samples including background to estimate this systematic uncertainty:

- $\delta(\Delta m_d) = \pm 0.0123$ ps^{-1}
- $\delta(\tau_{B^0}) = \pm 0.0178$ ps.

14.8 Summary

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 modelling) with Monte Carlo events. We assign the statistical errors of a full fit to Monte Carlo samples including background to estimate this systematic uncertainty.

The total systematic uncertainty on τ_{B^0} is 0.022 ps and on Δm_d is 0.013ps^{-1}.

Table 14.11: Summary of systematic uncertainties on the two physics parameters, τ_{B^0} and Δm_d.

Source	$\delta(\Delta m_d)$ (ps^{-1})	$\delta(\tau_{B^0})$ (ps)
Selection and fit bias	0.0123	0.0178
z scale	0.0020	0.0060
PEP-II boost	0.0005	0.0015
Beamspot position	0.0010	0.0050
SVT alignment	0.0030	0.0056
Background / signal prob.	0.0029	0.0032
Background Δt models	0.0012	0.0063
Fixed B^+/B^0 lifetime ratio	0.0003	0.0019
Fixed B^+/B^0 mistag ratio	0.0001	0.0003
Fixed signal outlier shape	0.0010	0.0054
Signal resolution model	0.0009	0.0034
Total systematic error	0.013	0.022

Chapter 15

Discussion and Outlook

The central value and precision of our measurement is comparable to other recent measurements from both the BABAR and Belle experiments at the asymmetric B-factories. The technique has been successful enough to merit extension to larger datasets where the analytical technique can be improved. Despite some challenges, it is worth considering extending the definition of the signal model to parameterize additional features of $B^0\overline{B}^0$ time evolution, such as the difference in lifetimes of the mass eigenstata, $\Delta\Gamma_d$.

15.1 Significance of This Result

This study uses about 20 fb^{-1} of *BABAR* data, and achieves measurements of τ_{B^0} with almost 2% overall precision and Δm_d with better than 5% overall precision. Other experiments and analyses have performed equivalently, and these results compare well, both in terms of central value and overall uncertainty, as summarized in Fig. 15.1.

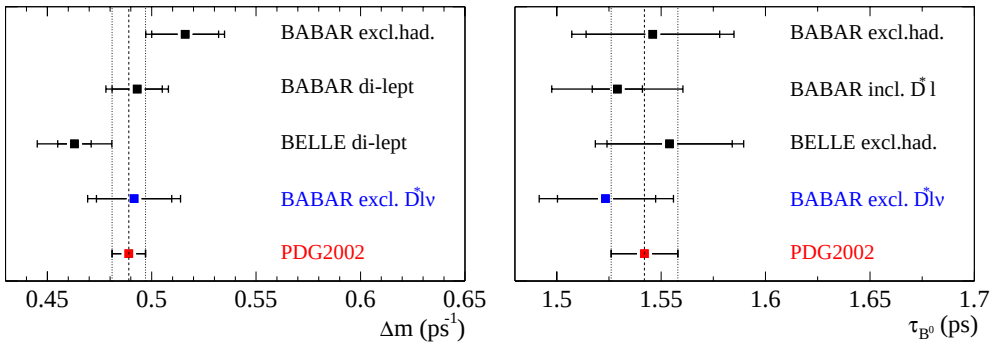


Figure 15.1: Comparison of this measurement with the world average and other measurements made at *BABAR* and Belle. The measurements described in this note are not included in the world average. Left: mixing frequency: *BABAR* exclusive hadronic measurement [64], *BABAR* dilepton measurement [65], Belle dilepton measurement [66], and the world average (PDG2002 [5]). Right: lifetime: *BABAR* exclusive hadronic measurement [50], *BABAR* inclusive D^* ℓ measurement [67], Belle exclusive hadronic measurement [68], and world average (PDG2002 [5]).

One obvious conclusion from this analysis is that it was critical to perform it blind. The correction applied to the fit result to account for the bias observed in Monte Carlo data moved the central value on the order of one statistical σ. If the true central value had been known when the correction was being considered, it would have been much more difficult to make the decision, as we could have compared the corrected value with the world average, rather than letting our analysis speak for itself.

The OPAL result [37] using a time-dependent analysis of semileptonic B-decays also reports the statistical correlation between τ_{B^0} and Δm_d, and they observe -0.14. Despite the difference in details of the analysis, the general mechanisms causing the

correlation are in common (namely, detector resolution), and our measured correlation coefficient of -0.22 is comparable. We note, in passing, that the correlation between τ_{B^0} and Δm_d does depend on the charged B composition of the sample. (See Ref. [69] for a discussion of this effect.)

15.2 Extracting Additional Performance

This study has pioneered a complex technique and has been successful in several different areas. In terms of physics results, we have measured the lifetime and mixing frequency of the neutral B meson with high precision. In terms of analysis strategies, this measurement is the first of its kind in that we also report the statistical correlation due to experimental technique between the two independent physics observables. And finally, this analysis has raised the bar in terms of maximizing sensitivity of the Data by incorporating a simultaneous analysis of signal and background to reduce systematic uncertainties.

Yet there is still room for improvement. Because of the strength of this technique, it is appropriate to extend it for larger Data samples and to address additional issues. The prospect for an impressive new B-factory Data sample is good, as suggested in Fig. 15.2. The current $100\,\mathrm{fb}^{-1}$ available is exciting enough to merit extension of this analysis, but a projected $1000\,\mathrm{fb}^{-1}$ is breath-taking. In the following list, we discuss areas where future work can be directed to strengthen this analysis.

- First and foremost, the apparent bias observed in the fit results when analyzing Monte Carlo data with simulated backgrounds needs to be characterized and understood. A necessary condition for this resolution is a significantly larger sample of trusted Monte Carlo data in order to establish the statistical significance of the potential bias. If the bias proves to be significant, areas in which to direct attention might be the outlier fraction of the signal resolution model (observed to be negative in the current generic Monte Carlo sample), fit results with only e or μ subsamples, and studies restricted to samples of signal and one class of background events at a time.

- Perhaps on the heels of such an undertaking, there are some unanswered questions surrounding the final yields of D^*e and $D^*\mu$ in the Data (and Monte Carlo) samples. Although we measure the selection efficiencies to differ by more than 20%, the signal yields in the δm peak regions are comparable. This could be due to differences in reconstruction/selection effects with the high momentum cut on the lepton; for instance, electrons might experience greater energy losses (material interactions) than muons.

- The ability to measure the charged B fraction in the sample, f_{B^+}, can be improved by including the spectrum of another variable in the analysis – $\cos\theta_{B,D^*\ell}$. By parameterizing the signal probability with this variable as well, sensitivity to f_{B^+} could be increased.

- *BABAR* has introduced improved flavor-tagging algorithms [70], and the improvements in tagging power are worthwhile. It may be necessary to re-examine the dependence of the mistag rates on $\sigma_{\Delta t}$ since the new tagging algorithms categorize physics processes differently.

- In the current study, we have assumed (a) that f_{B^+} does not depend on any event characteristics, and (b) that all candidates from B^+ parents behave identically as those from B^0 parents, except in the signal region where we separate the charged B sample by "physics" Δt structure. Current studies show that, for instance, splitting f_{B^+} by tagging category yields statistically consistent values. We also assume that charged B candidates have the same signal resolution model as the neutral B candidates. With access to an even larger sample size of Monte Carlo, these assumptions should be checked, and then tested in Data.

- The background control sample identified in Section 8.3 for uncorrelated backgrounds is relatively impure. To improve the Data-driven characterization of uncorrelated-lepton peaking background events, a cleaner sample can be selected. It has been suggested by C. LeClerc [71] and others that a more powerful control sample might be formed by requiring instead that $\cos\theta_{D^*,\ell} > 0$ and $\cos\theta_{B,D^*\ell} > 1.1$. Preliminary studies suggest that this proposal might better

match the Δt structure of uncorrelated lepton events in the signal region.

- A additional benefit of using an updated Data sample is the advantage of more advanced reconstruction algorithms available in new releases of the software. Specifically, the internal alignment of SVT has been more carefully parameterized and measured, and we expect reprocessed Data and new Data to be better aligned, *i.e.,* that source of systematic error could decrease. Bug fixes and improvements to the track finding algorithms are also available which could increase the signal selection efficiencies.

- As explained in Section 7.4, we observe correlations between the per-event Δt error ($\sigma_{\Delta t}$) and the mistag fraction, largest in the `KTag` category. We do not allow for this affect in our models for background Δt structure, nor for any of the other tagging categories. The linear correlation model could be extended to relax both of these assumptions *i.e.,* allowing for a dependence of mistag rate on $\sigma_{\Delta t}$ for all tagging categories and for all background Δt models.

- We have ignored some of the issues of multiple $D^{*-}\ell^+\nu_\ell$ candidates per event. In reality, we should expect at most only two $B_{\rm rec}$ in the event, corresponding to the semileptonic decays of both B mesons. The current implementation selects only one $B_{D^*\ell}$ candidate per D^0 decay mode, determined by the candidate with a D^0 mass closest to the nominal value. We therefore allow multiple candidates per event to selected * provided that they differ in D^0 mode, lepton identification, or in angular correlation of the $D^* - \ell$ system. For instance, about 3.5% of the events from Data have identical timestamps, but are distinct in at least one of the aforementioned three characteristics, and (of course) have different Δt and δm values. It would be appropriate to explore this more fully in the future. The difficulty is in the technique for determining the "best" candidate, as we select both signal and background candidates from an event, and it is therefore non-trivial to determine which is the "better" candidate to retain. †

*As distinguished by event timestamps.

†On an unrelated note, the rate of overlapping event selection with other time-dependent analyses per was estimated to be less than a few percent [72].

- As was illustrated in Section 12.1.1, the current Dataset is insufficient to completely characterize the outlier component of the signal resolution model. One could imagine developing strategies to completely eliminate outliers. An initial step could involve examining the handful of outliers in Monte Carlo (for instance, those eliminated when tightening the $|\Delta t|$ selection window) with the event display and/or the new *mini* event store.

- Finally, the calculation of per-event probabilities relies on the *BABAR* particle identification algorithms and characterizations developed by the Particle Identification (PID) Group [73]. Studies suggest that the results are not strongly dependent upon the absolute efficiencies of the lepton identification algorithms, but we do note that the efficiencies used in this analysis were developed for slightly different selection criteria [‡] With the statistical power of additional Data and Monte Carlo, it would be appropriate to regenerate the so-called PID tables, tailored to the $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ needs.

While this may seem like a depressingly long list of future work, it is, on the contrary, an exciting opportunity to move forward and harvest more Data. It is a tribute to our success that we have the freedom to use new Data to relax even more assumptions and let the Data help us determine the true parameters and their relationships.

Last but not least, the **RooFit** technology [60] was used in building, maintaining, and fitting the multi-dimensional probability density functions. This analysis suite is incredibly powerful, but due to the level of sophistication required by the analysis, the computational overhead became intense. For future work with even more Data and even more complex models, one may need to re-evaluate or re-optimize this technology for meaningful progress.

[‡]That is, our lepton selection criteria requires the charged tracks to pass the GoodTracksTight selection, while the particle identification tables only require GoodTracksLoose.

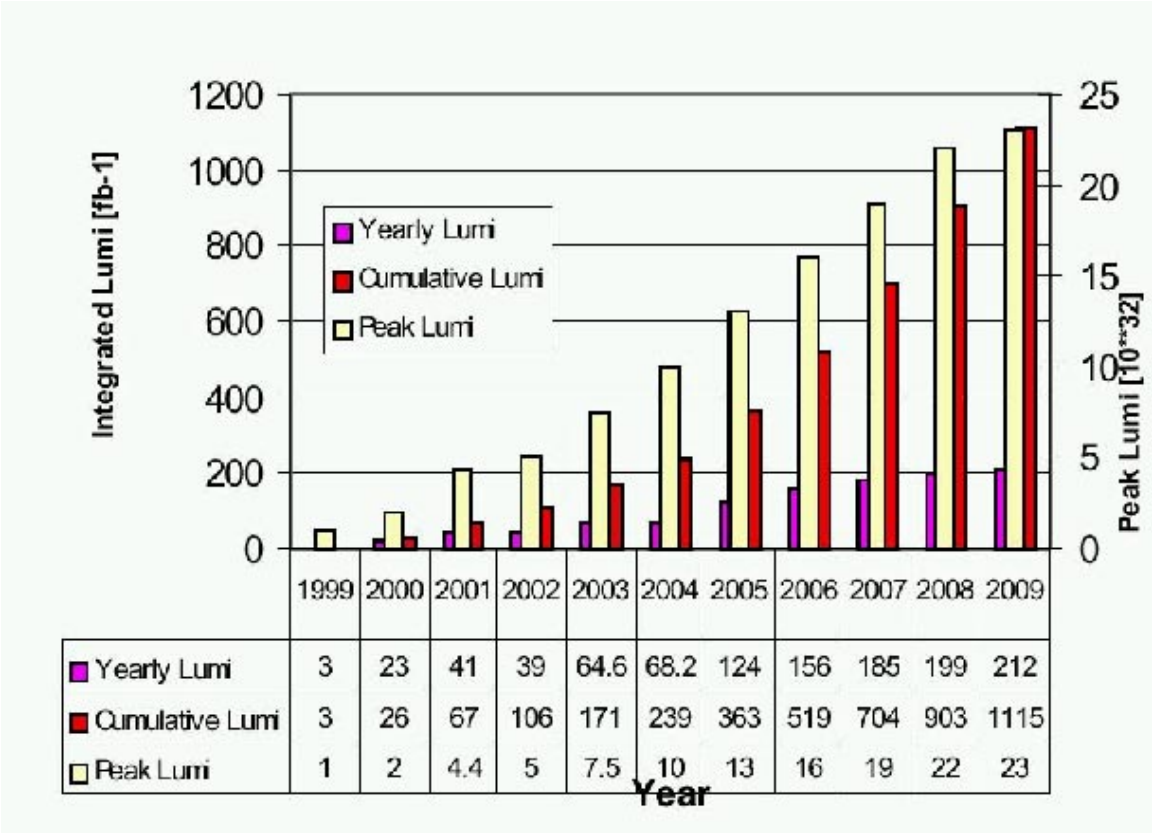


Figure 15.2: One model for the accumulation of integrated luminosity by *BABAR* for the next few years.

15.3 Extracting Additional Physics

15.3.1 General Comments

Because of the high precision offered by the large $B\overline{B}$ data samples at the asymmetric B factories, it is natural to consider the reach for new types of physics as well as the precision measurements of traditional parameters. Given the framework and definitions provided in Chapter 3, one avenue of exploration is the discrete symmetries of $B^0\overline{B}^0$ time evolution. Consider the cartoon in Fig. 15.3. In neutral B meson oscillations, we can gain access to the probabilities for

- $B^0 \to B^0$,

- $\overline{B}^0 \to \overline{B}^0$,
- $B^0 \to \overline{B}^0$,
- $\overline{B}^0 \to B^0$.

For example, if B mixing obeyed the T symmetry, we would expect the rate for $B^0 \to \overline{B}^0$ to equal the rate of $\overline{B}^0 \to B^0$. Alternatively, CPT symmetry would require that $prob(\overline{B}^0 \to \overline{B}^0) = prob(B^0 \to B^0)$, and so on. In our discussion of Chapter 3, we required CPT symmetry, which allowed us to take the elements M_{11} and M_{22} to be equal, and similarly for Γ_{11} and Γ_{22}. If, as the Standard Model suggests, CPT is a good symmetry but CP and T are separately violated, we would expect M_{12} and Γ_{12} to have identical phases. To determine which of these scenarios Nature has provided is the challenge to the experimentalist.

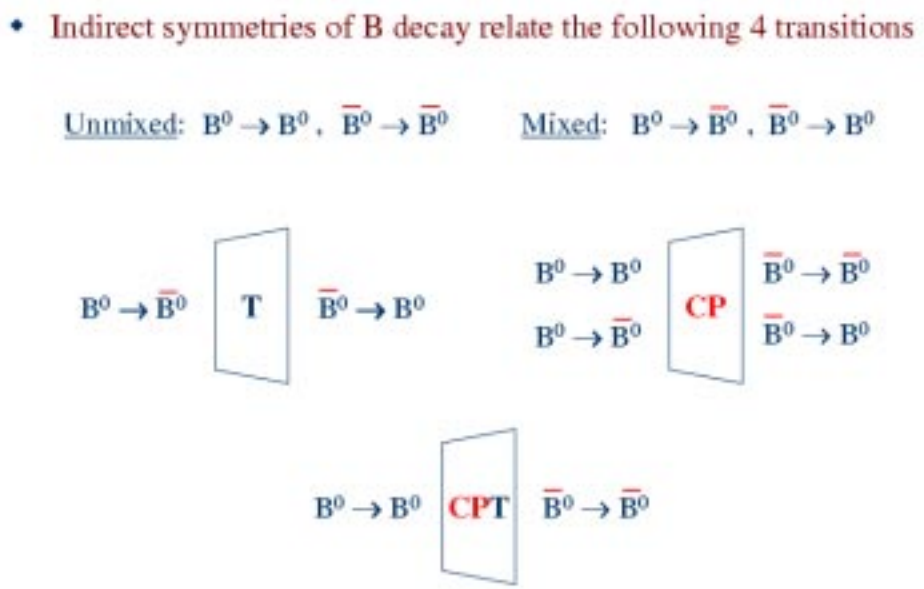


Figure 15.3: Artist's rendition of the possible relations between $B^0\overline{B}^0$ time evolution and combinations of the discrete symmetries C, P, and T.

Several such measurements are already underway at *BABAR*. A comparison of the time-integrated rates for inclusive $B^0\overline{B}^0 \to B^0B^0 \to \ell^+\ell^+$ and $B^0\overline{B}^0 \to \overline{B}^0\overline{B}^0 \to \ell^-\ell^-$

at *BABAR* directly examines the probabilities for $B^0 \to \overline{B}^0$ and $\overline{B}^0 \to B^0$. [74] By measuring the time-dependent asymmetry of like-sign lepton pairs

$$\begin{aligned} A_{T/CP}^{meas}(\Delta t) &= \frac{N(\ell^+\ \ell^+, \Delta t) - N(\ell^-\ \ell^-, \Delta t)}{N(\ell^+\ \ell^+, \Delta t) + N(\ell^-\ \ell^-, \Delta t)} \\ &\approx \frac{1 - |q/p|^4}{1 - |q/p|^4} \end{aligned} \tag{15.1}$$

using a large sample of inclusive dileptons, the study concluded that $|q/p| = 0.998 \pm 0.006(stat) \pm 0.007(syst)$, very consistent with unity as expected if *CPT* is conserved. Additionally, work is in progress to measure $\Delta\Gamma$ and test CPT/T violation using a large variety of B decay modes. [75] (See Section 15.3.2 for additional remarks.)

This is just one example of work that has yet to be done. With the precision environment offered by the B-factories and the expected large samples of Data, [§] searches for these asymmetries could become a very rewarding industry.

15.3.2 A Specific Case: $\Delta\Gamma$

The discrete symmetry violations discussed above are all interesting, and they probe the Standard Model description. However, we should recall that the Standard Model prescribes a non-zero value for $\Delta\Gamma_d$, and recent theoretical estimates place the ratio $\Delta\Gamma/\Gamma \sim 0.3\%$ [22]. In the natural progression of bigger and better, we might first expect to measure $\Delta\Gamma$ and then add in searches for additional sources of *CP* violation, for instance. In fact, as we saw in Section 3.1.2, $\Delta\Gamma$ appears as a hyperbolic trigonometric term in the model for B time-evolution, and will therefore need to be included in future higher precision fits for parameters such as τ_{B^0}, Δm_d, and $\sin 2\beta$. We will discuss some properties of $\Delta\Gamma$ in the B_d system, and then make some observations about techniques to measure it.

[§]Keep in mind, though, that the $b\overline{b}$ production cross-section at the upcoming hadron machines is significantly larger, and those Data samples could be even more gigantic.

Comments on $\Delta\Gamma$

From our discussion in Sections 2.1.5 and 3.3, we recall that in the limit of zero *CP* violation in the Standard Model, $\Delta m_d = 2\ |M_{12}|$ and $\Delta\Gamma_d = 2\ |\Gamma_{12}|$. As mentioned before, $\Delta\Gamma$ is the sum of contributions from real intermediate states to which both B^0 and $\overline{B}^0$ can couple, states such as $u\overline{u}d\overline{d}$ ($\pi^+\ \pi^-$), $c\overline{c}d\overline{d}$ ($D^{*+}\ D^{*-}$), and $c\overline{u}d\overline{d}$ ($D^{*\pm}\ \pi^{\pm}$). And, as we stated earlier, these couplings are heavily CKM-suppressed, of order λ^6. Δm_d is related to virtual transitions in the box diagram, which, due to heavy top-exchange, is dominated by $|V_{td}V_{tb}|^2$ which is also of order λ^6. (See Fig. 15.4 for the two different "aspects" of the box diagram to which Δm_d and $\Delta\Gamma_d$ are sensitive.) As $\Delta\Gamma_d$ does not access the intermediate states involving the top quark ¶, we expect that

$$\frac{\Gamma_{12}}{M_{12}} \sim \mathcal{O}(\frac{m_b^2}{m_t^2})\ . \tag{15.2}$$

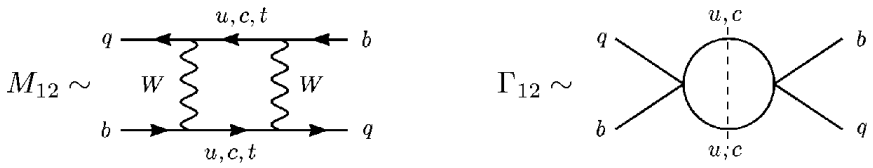


Figure 15.4: The B mixing diagrams [22] that contribute to Δm_d (left) and $\Delta\Gamma$ (right).

Borrowing heavily from the analysis by Dighe *et al.* in Ref. [22], an effective theory below the m_W scale would suggest

$$\Gamma_{12} = \frac{1}{2m_{B_d}}\langle\overline{B}_d|i\ \mathcal{I}m\ \int d^4x\mathcal{T}(\mathcal{H}_eff^{\Delta B=1}(x)\mathcal{H}_{eff}^{\Delta B=1}(0))|B_d\rangle\ , \tag{15.3}$$

where the effective Hamiltonian $\mathcal{H}_{eff}^{\Delta=1}$ can be written as the sum of short-distance physics contributions (Wilson coefficients, CKM elements, and four-quark operators) from u and c interactions. We see here that while M_{12} and Γ_{12} both produce $\Delta B = 2$ transitions, because $\Delta\Gamma$ is sensitive to long distance physics (real intermediate

¶Note that Ref. [26] disagrees with this comment on pg. 619; this has been corrected in the updated version.

states), the hadronic matrix element is really the product of two $\Delta B = 1$ operators evaluated at different points. However, because the bottom quark mass offers another short-distance scale, we return to short-distance physics and expand in terms of local $\Delta B = 2$ operators scaled by $1/m_b$. Jumping to the conclusion, we find that Γ_{12} can be written as [26]

$$\Gamma_{12} = \frac{G_F^2 m_b^2 \eta_b' m_{B_d} B_{B_d} f_{B_d}^2}{8\pi} \left[(V_{td}^* V_{tb})^2 + V_{td}^* V_{tb} V_{cd}^* V_{cb} \mathcal{O}(\frac{m_c^2}{m_b^2}) + (V_{cd}^* V_{cb})^2 \mathcal{O}(\frac{m_c^4}{m_b^4}) \right] \tag{15.4}$$

where we only wish to observe the presence of the CKM matrix elements. If we approximate the sum as $(V_{td}^* V_{tb} + V_{cd}^* V_{cb})^2$ we can use the unitarity condition to replace it with $(V_{ub}^* V_{ud})$ and identify the phase of Γ_{12} as that of $-\eta$ from the Wolfenstein parameterization of the CKM matrix (see Section 2.1.3). Recall now that M_{12} is controlled by $V_{td}^* V_{tb}$ which conveniently has the same phase. This suggests that

$$\phi_{M_{12}} - \phi_{\Gamma_{12}} \sim 0 + \mathcal{O}(\frac{m_c^2}{m_b^2}) \; . \tag{15.5}$$

Finally, the best estimate for $\Delta\Gamma/\Gamma$ in the Standard Model is $\sim +(0.3^{+1.2}_{-1.6})\%$ where Dighe *et al.*have included $1/m_b$ contributions and some of the next-to-leading order QCD corrections. [‖]

Modified Time Dependence

Before discussing some experimental techniques, let's recall the modified time dependence when we include the natural presence of $\Delta\Gamma$. Effectively, we add one more parameter, replacing the 1 with $\cosh((\Delta\Gamma/2)\ t)$:

$$\begin{aligned} |\langle B^0 | B^0_{phys}(t)\rangle|^2 &\sim \frac{e^{-\Gamma|t|}}{2} \left[\cosh(\frac{\Delta\Gamma_d}{2} t) + \cos(\Delta m_d t) \right] \\ |\langle \overline{B}^0 | B^0_{phys}(t)\rangle|^2 &\sim \frac{e^{-\Gamma|t|}}{2} |\frac{q}{p}|^2 \left[\cosh(\frac{\Delta\Gamma_d}{2} t) - \cos(\Delta m_d\ t) \right] \; . \end{aligned} \tag{15.6}$$

[‖] We have used the sign convention of Dighe *et al.*, and therefore the Standard Model prediction is that the $\Gamma_L > \Gamma_H$.

Note that in the absence of *CP* violation in mixing, the factor $|\frac{q}{p}|$ is unity. We illustrate the change to the decay time distributions in Fig. 15.5. A non-zero $\Delta\Gamma$ effectively reduces the mixing amplitude with increasing $|\Delta t|$. Because cosh is an even function, and cosh differs from unity for arguments near zero only by the square of the argument (*i.e.,* the Taylor expansion near 0), sensitivity to $\Delta\Gamma$ in the explicit time dependence is reduced to $(\Delta\Gamma/\Gamma)^2$.

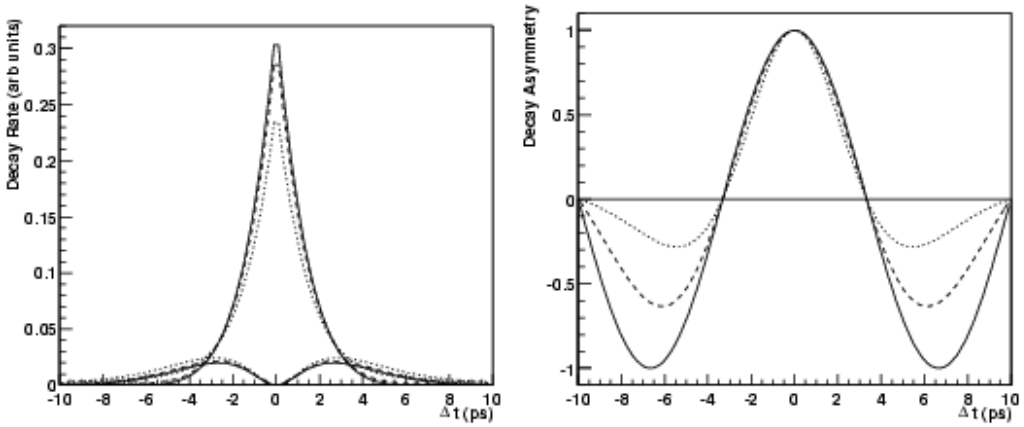


Figure 15.5: Plots of the mixed and unmixed decay time distributions (left) and the time-dependent mixing asymmetry (right), assuming perfect tagging and resolution. The three curves correspond to three different values of $\Delta\Gamma/\Gamma$: 0 (solid), 0.5 (dashed), and 1 (dotted).

For so-called *CP* final states $f_\pm$, we can write the tree-level decay amplitudes as

$$\begin{aligned} A_+ &= \langle f_+|B^0\rangle = +e^{-2i\beta} \ , \\ A_- &= \langle f_-|B^0\rangle = -e^{-2i\beta} \ , \end{aligned} \tag{15.7}$$

where f_+ is the *CP*-even final state. Substituting these and their conjugates into

Eq. 3.27 and neglecting "direct" *CP*-violation, we arrive at

$$\begin{aligned}\Gamma(B_d(t) \to f_\pm) &\propto e^{-\Gamma t}\left[\cosh(\frac{\Delta\Gamma}{2}t) \mp \cos(2\beta)\sinh(\frac{\Delta\Gamma}{2}t) \pm \sin(2\beta)\sin(\Delta m_d t)\right] \\ &\propto e^{-\Gamma_L t}(1 \pm \cos(2\beta)) + e^{-\Gamma_H t}(1 \mp \cos(2\beta)) + \text{oscillating terms}\ .\end{aligned} \tag{15.8}$$

For a final state such as the semileptonic decay $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$, we have (as before) $\overline{A}_{D^{*-}\ell^+\nu_\ell} = 0$ and then

$$\begin{aligned}\Gamma(B_d(t) \to f_\pm) &\propto e^{-\Gamma t}\left[\cosh(\frac{\Delta\Gamma}{2}t) + \cos(\Delta m_d t)\right] \\ &\propto e^{-\Gamma_L t} + e^{-\Gamma_H t} + \text{oscillating terms}\ .\end{aligned} \tag{15.9}$$

Comments on Measuring $\Delta\Gamma_d$

At present, there are no good experimental measurements, or even upper limits on ratio $\Delta\Gamma_d/\Gamma$. ** Because the B^0-$\overline{B}^0$ mixing phase relative to decay ($\sin 2\beta$) is large for charmonium final states, the *CP* eigenstates are measurably different from the physical states (with widths Γ_H and Γ_L). That is, decays to *CP* states involves both lifetimes, as do the flavor-tagging decays (such as the semileptonic modes). There is no final state to which the B^0 decay involves one of the decay widths (unlike $K^0_L \to \pi^+\pi^-$). As such, the measurement of $\Delta\Gamma$ is difficult, but we press forward anyway.

We consider two general strategies for measuring $\Delta\Gamma$, the so-called "untagged" and "tagged" measurements, by which we refer to flavor-identification of the decaying B_d meson. Assuming balanced acceptances and efficiencies for B^0 and $\overline{B}^0$, measurements without flavor-tagging "integrate out" the dependence on the mixing status, eliminating sensitivity to Δm_d. Conventional wisdom claims that Δt measurements are only sensitive quadratically to $\Delta\Gamma/\Gamma$. One can observe this explicitly by considering

**CLEO has observed that their time-integrated χ_d measurement can be converted into a limit for $\Delta\Gamma$ using the Δm_d measurement from other experiments, but the technique is troubled by the assumption of $\Delta\Gamma_d \equiv 0$ in order to measure Δm_d in the first place.

the calculation

$$\left(\frac{1}{\sigma_\alpha}\right)^2 = N \int \frac{1}{f}\left(\frac{\partial f}{\partial \alpha}\right)^2 \tag{15.10}$$

for the time dependence in Eq.15.6 and finding that $\sigma_{\Delta\Gamma/\Gamma} \sim N^{-1/4}$. [76] Dighe *et al.*propose a general argument by showing that all moments of the non-oscillating part of the decay time distribution are linear in $(\Delta\Gamma/\Gamma)^2$ for any analysis using a single final state. However, as shown below, there are methods to get around this, namely by using more than one mode.

At the basic level, non-zero $\Delta\Gamma$ implies that the lifetime distributions for B decays are actually the sum of two slightly different negative exponentials. By considering two final states that have different contributions from B_H and B_L we can extract two different "lifetime" measurements and combine them to find $\Delta\Gamma$. [††] In the general case for untagged measurements (integrating out the mixing dependence), the non-oscillating part of the time dependence reduces to

$$f(t) = \frac{1}{2}\left[(1+b)e^{-\Gamma_L t} + (1-b)e^{\Gamma_H t}\right] \ , \tag{15.11}$$

where we have effectively pulled the overall lifetime exponential through each of the terms. For two different final states f_1, f_2 we can write:

$$\frac{\tau_1}{\tau_2} = 1 + \frac{b_2 - b_1}{2}\frac{\Delta\Gamma}{\Gamma} + \mathcal{O}(\frac{\Delta\Gamma}{\Gamma}^2) \ . \tag{15.12}$$

We can immediately apply this result to the previous situations of decays to charmonium CP eigenstates and semileptonic decays, identifying b as

$$\begin{aligned} b_{SL} &= 0 \\ b_{CP+} &= +\cos(2\beta) \\ b_{CP-} &= -\cos(2\beta) \ . \end{aligned} \tag{15.13}$$

The experimental technique then is to measure the partial lifetimes and use their ratio to extract $\Delta\Gamma$. For the best precision, we might use the semileptonic final state

[††] Actually, this is a technique often used in the D^0 system.

and the charmonium $CP-$ final state because of the higher reconstruction efficiencies:

$$\frac{\tau_{SL}}{\tau_{CP+}} \approx 1 - \frac{\cos(2\beta)}{2}\frac{\Delta\Gamma}{\Gamma} , \tag{15.14}$$

assuming we know the CKM angle 2β to good accuracy. There are also techniques that involve using only one final state, such as $B^0 \to J/\psi K^*$, where a transversity analysis can be used to separate the even- and odd- CP contributions. Another popular notion is to construct the untagged time-dependent asymmetry between the decays $B^0 \to J/\psi K^0_S$ and $B^0 \to J/\psi K^0_L$, which (again) requires good knowledge of 2β.

Up to this point, we have not used the tagging information in the time-dependent analysis. As T. Kittleman suggests in Ref. [77], we can gain sensitivity if we include the oscillation terms (*i.e.,* those depending on the flavor of B). His analysis is summarized in Fig. 15.6, and suggests that the statistical error on $\Delta\Gamma$ can decrease substantially with the inclusion of the additional tagging information. *

The full time-dependent analysis appears to be the best option for measuring $\Delta\Gamma$, and by including both CP final states and flavor-tagging final states (that is, so-called *self-tagging* final states) for $B_{\rm rec}$, we gain linear sensitivity to $\Delta\Gamma$ (CP events) and high statistics for the resolution function(s) (flavor/mixing events). † Such an analysis would need to simultaneously extract $\sin2\beta$, because at the precision required for measuring $\Delta\Gamma$, corrections to the measurement of $\sin2\beta$ also arise:

$$\begin{aligned} A_{CP} &= \sin(\Delta m_d t)\sin2\beta \quad \text{ if } \Delta\Gamma = 0 , \\ A_{CP} &= \frac{e^{-\Gamma t}\sin(\Delta m_d t)\sin2\beta}{\cos^2\beta e^{-\Gamma_L t} + \sin^2\beta e^{-\Gamma_H t}} \quad \text{ otherwise.} \end{aligned} \tag{15.15}$$

The fractional error introduced on A_{CP} is approximately $-(\cos(2\beta)\Delta\Gamma\ t)$ so that for $t \sim 1/\Gamma$ the fractional error on the $\sin2\beta$ extraction is about 0.5%. Ambitiously,

*This is not surprising, as the τ_{B^0}-Δm_d analysis described in this paper relies on the same principle: adding "good information" leads to enhanced sensitivity.

†One might imagine fitting just to a semileptonic sample, thereby removing the CP information and recovering a large sample. Unfortunately, this again raises the feared quadratic sensitivity dragon.

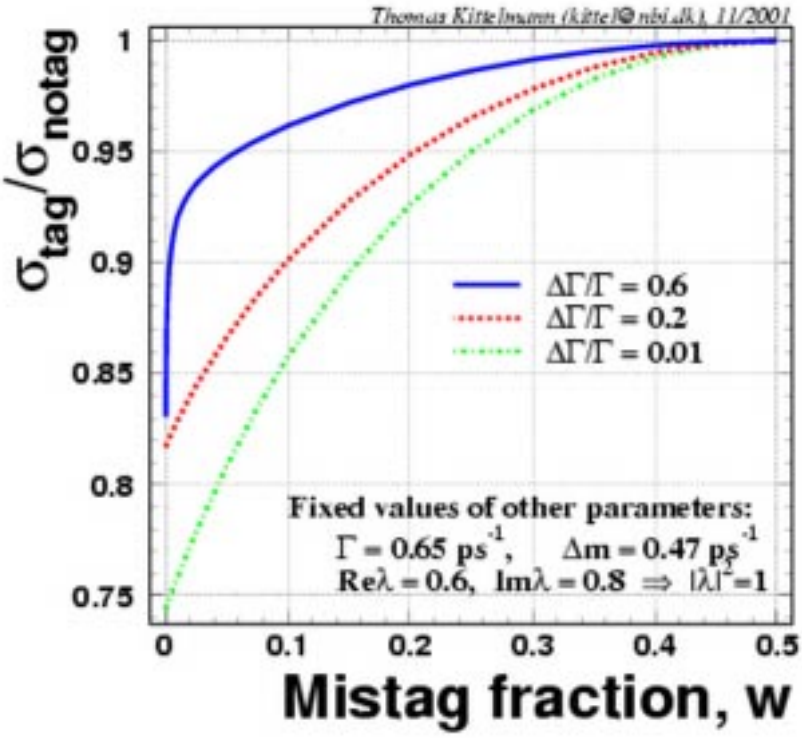


Figure 15.6: Estimate of the ratio of statistical error on $\Delta\Gamma$ when ignoring flavor-tagging information over including it in the fit. The three curves correspond to different values assumed for $\Delta\Gamma$. (Figure courtesy Ref. [77].)

one could then use the full fit to the decay time difference distributions of both *CP* and flavor events to extract all four parameters: $\sin 2\beta$, τ_{B^0}, Δm_d, and $\Delta\Gamma$. There is another wrinkle, however, in the plan. Recall that we chose to neglect *CP* violation in mixing, by assuming that $|q| = |p|$. It might become necessary to parameterize and simultaneously measure $|q|/|p|$ for both $B^0\overline{B}^0$ and $K^0\overline{K}^0$ mixing as well since they are expected to be of the same order of magnitude.

Finally, from Eq. 3.27 (Section 3.1.3), we recall the interference term in the model for the correlated time dependence of the two B mesons. While we typically expect this term to be small, it too involves $\Delta\Gamma$. A proper analysis to extract these parameters needs to accommodate the interference term as well, significantly complicating the situation. For instance, one should also consider the effects of so-called doubly

CKM-suppressed decays on the B_{tag} side, as they will contribute an interference term that is not parameterizable with only uniform mistag rates. [24] A naive guess might have been that such DCSD decays only contribute to an effective mistag rate, since they will necessarily have "wrong-sign" decay products, but the interference term can play an important role. Appropriate modelling of these experimental effects is at the forefront of modern investigation.

Despite these complications, interest in measuring $\Delta\Gamma$, perhaps necessarily in combination with new parameters sensitive to *CP* violation (*i.e.,* q and p) remains strong, and we can look forward to exciting results from the B-factories on this matter.

Chapter 16

Closure

On the advice of the review committee, and with the full powers and privilege vested within us by the *BABAR* Collaboration, we unblinded the fit results on 05 July, 2002. On the Run-1 Data sample using Release 8 processing, we measure the following properties of B^0 time evolution with a sample of approximately 14,000 exclusively reconstructed $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ signal events,

- $\Delta m_d = 0.492 \pm 0.018 \pm 0.013$ ps^{-1},
- $\tau_{B^0} = 1.523\ ^{+0.024}_{-0.023} \pm 0.022$ ps
- statistical correlation $\rho(\Delta m_d, \tau_{B^0}) = -0.22$.

We compare our measurement with the world average: [78]

- $\Delta m_d = 0.503 \pm 0.006$ ps^{-1}
- $\tau_{B^0} = 1.540 \pm 0.014$ ps,

and find that our results are very consistent (Δm_d and τ_{B^0} are both within 0.5σ). The world averages for both lifetime and mixing are dominated by recent measurements at the B-Factories (*BABAR* and Belle).

This result was announced at the 31^{st} International Conference on High Energy Physics (ICHEP02) in Amsterdam, The Netherlands on July 25, 2002. [79] Updated

world-averages of B^0 lifetime and mixing were calculated by the *LEP B Working Groups* and incorporated this result. The technology used to extract the world average is advanced, but at the time of July 2002, was not sufficient to accept the correlated measurement of τ_{B^0} and Δm_d obtained here. For instance, the Working Group adjusts the traditional measurements of Δm_d to accommodate more recent (and higher precision) measures of τ_{B^0}. Using our measured statistical correlation coefficient, we converted our simultaneous measurement to an equivalent "independent measurement" of τ_{B^0} and Δm_d from this sample. Our reformulated contribution was incorporated into the new world averages announced at the ICHEP02 conference. See Figs. 16.1-16.3 for details on the updates.

This study has focused on maximizing the sensitivity to the physics under study (lifetime and mixing) by using few assumptions, by selecting additional background control samples, and by employing an unbinned maximum likelihood fit with Data-driven parameterizations wherever possible to simultaneously extract signal and background parameters. By bootstrapping the final fit with preliminary studies using enriched subsamples, we were able to develop not only an accurate understanding but also sets of starting values for the final analysis. The full fit took into account the correlations between all 72 Δt parameters and transformed a traditionally systematics-limited analysis into a statistics-limited one; with more data, the performance can only improve. The general form of this analysis also lays important groundwork for even more challenging measurements of the future.

We look forward to new analyses which expand upon this effort, and are proud to have helped pioneer this technique. As we seek out the secrets of the Universe with sophisticated analyses such as these, we can only benefit from being as inclusive and objective as possible.

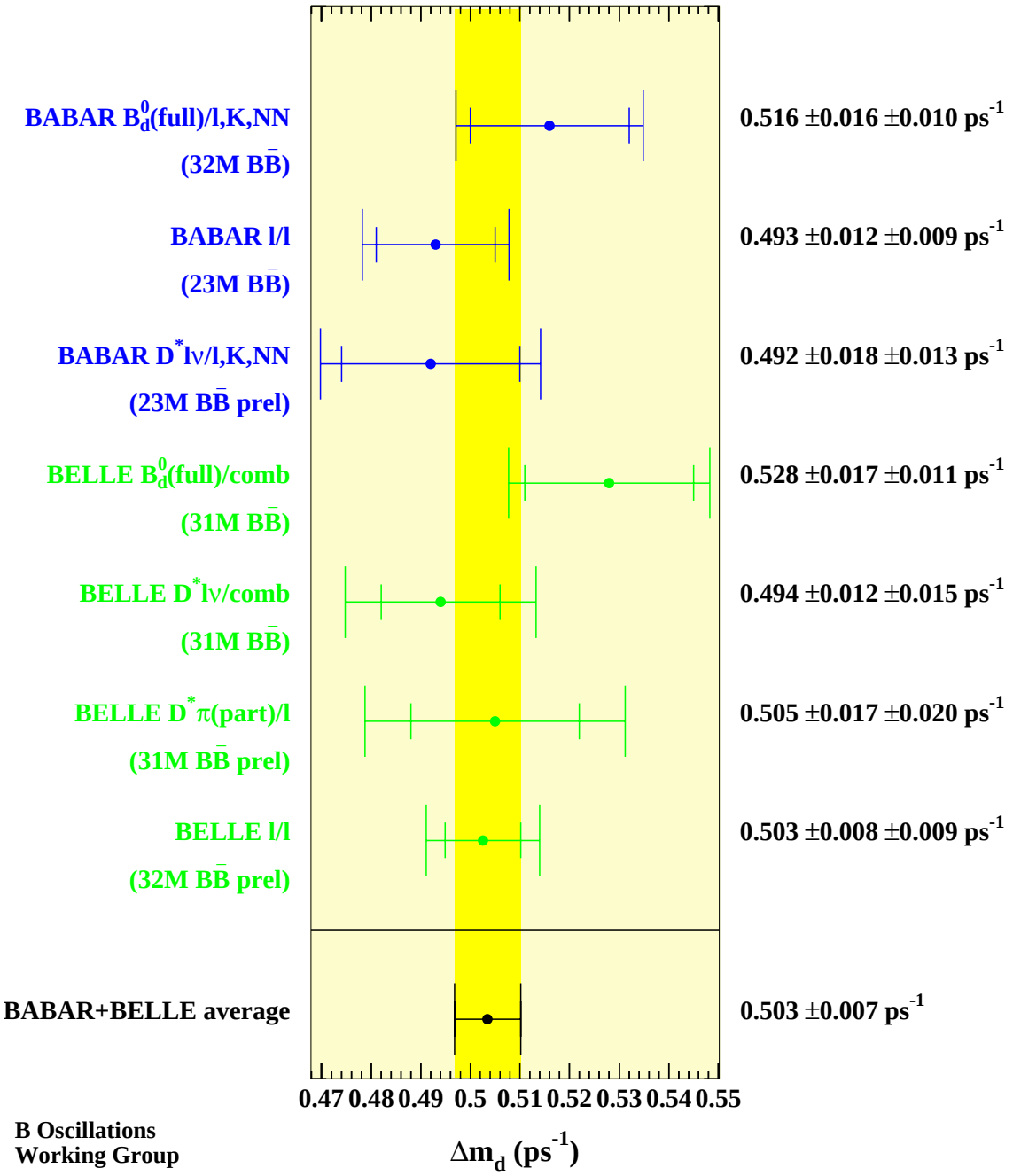


Figure 16.1: The updated world average for the mixing frequency of B^0-$\overline{B}^0$ oscillations as of July 2002, showing only the measurements from the asymmetric B-factories. Note that this result includes the measurement outlined in this study.

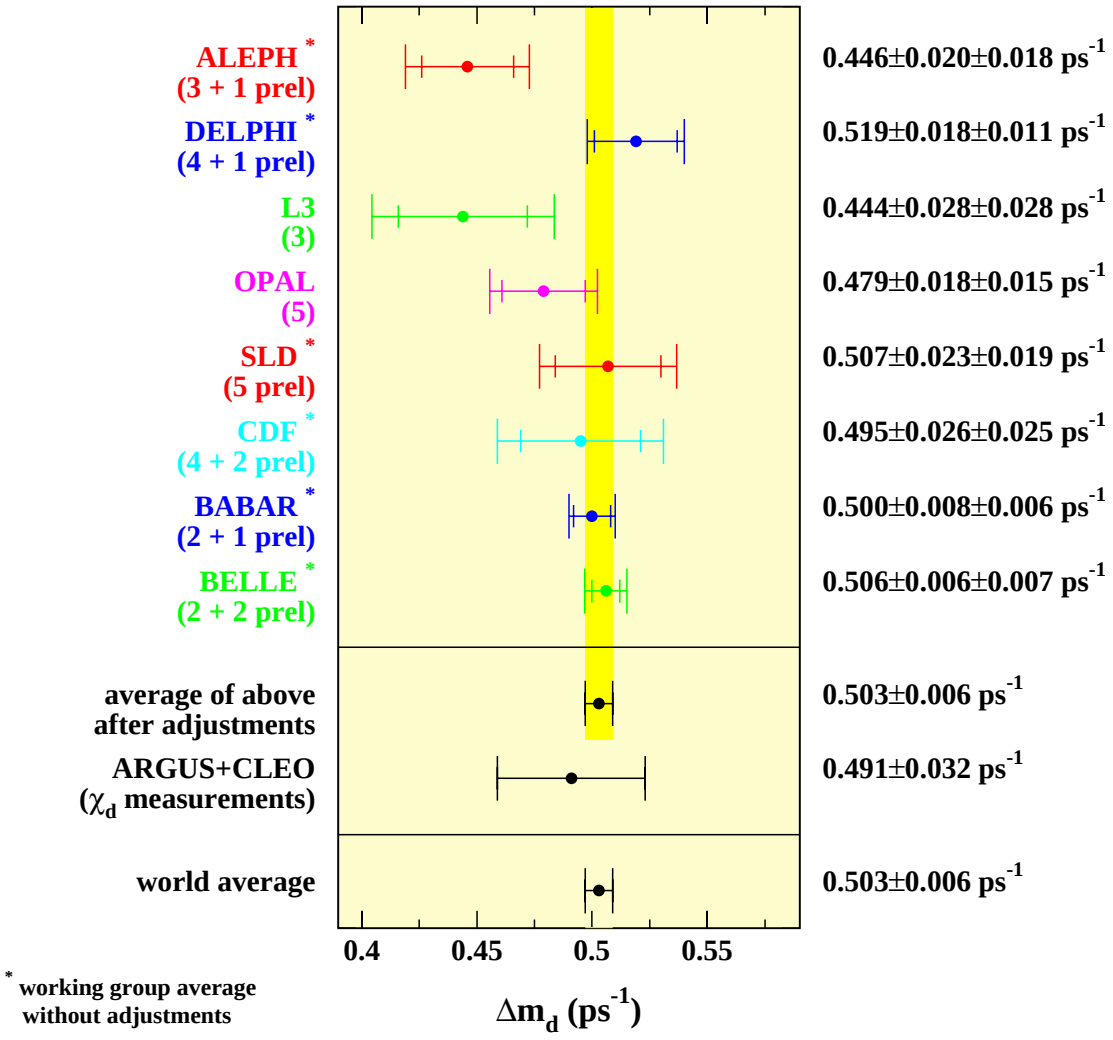


Figure 16.2: The new world average for the B^0-$\overline{B}^0$ mixing frequency Δm_d as of July 2002, showing the averaged contributions from each experiment. This average includes our result.

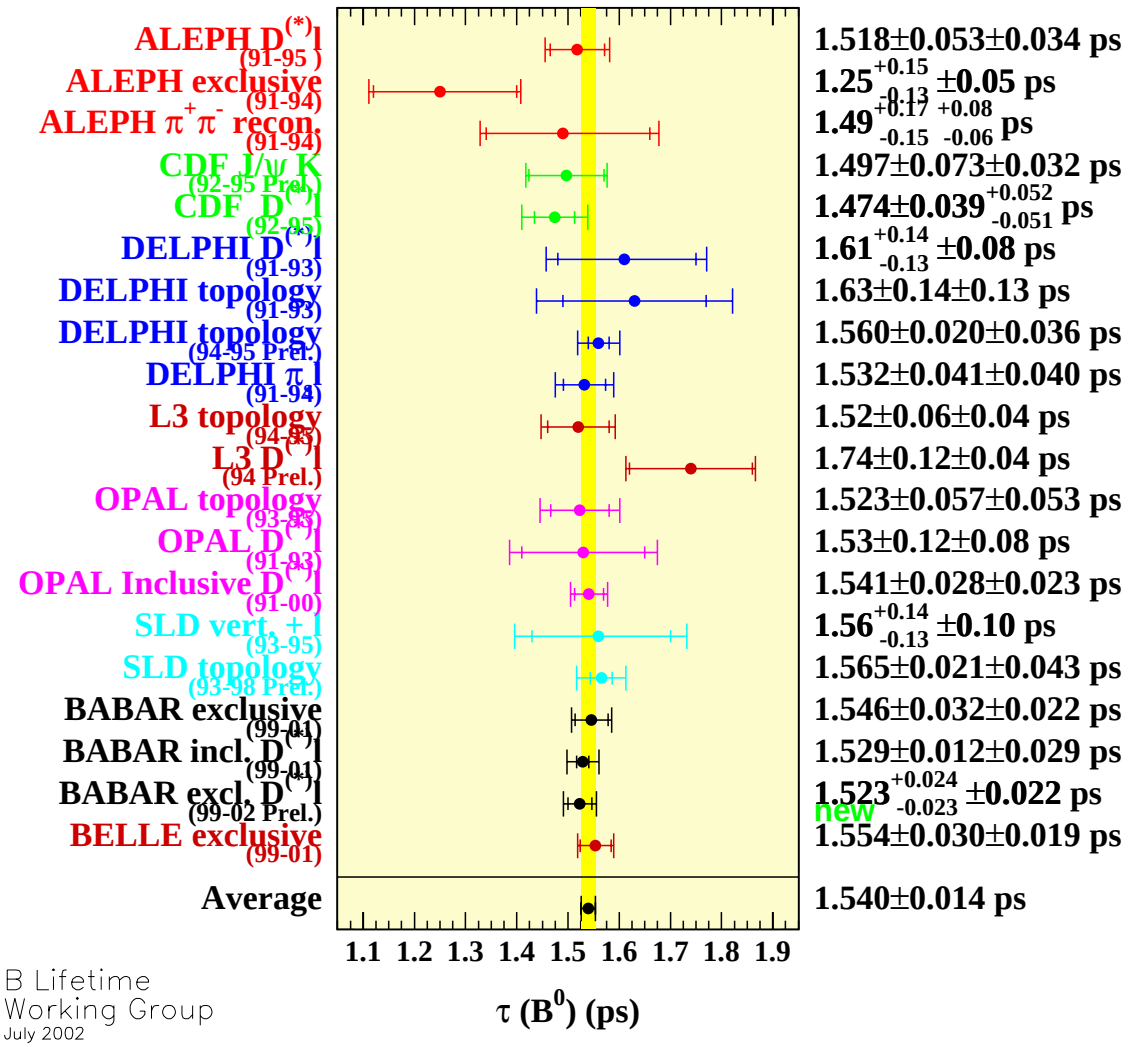


Figure 16.3: The world average for τ_{B^0} as of July 2002, including our result.

Chapter 17

Afterword

I entered graduate school for selfish reasons – because I enjoyed doing physics. In particular, I found the crafty and clever ways that experimentalists teased out the secrets of Nature to be very inspiring. But I didn't choose graduate school to secure a job in research physics or to acquire a set of technical credentials to advance my career plans. I wasn't even so much as interested in the skills I would take *beyond* graduate school as I was in the process of learning and developing them *in* graduate school. I think this attitude has given me an important sense of freedom, allowing me to fully commit to the graduate school process without being anxious about the future.

That journey is now complete. *BABAR* has taught me a lot, my advisor even more, and 6 years of living the most of all. I'm grateful, ready, and open to whatever comes next.

Chapter 18

Appendices

I include here additional materials that might be relevant or interesting to the dedicated reader.

Appendix A

Additional Plots of Event Sample

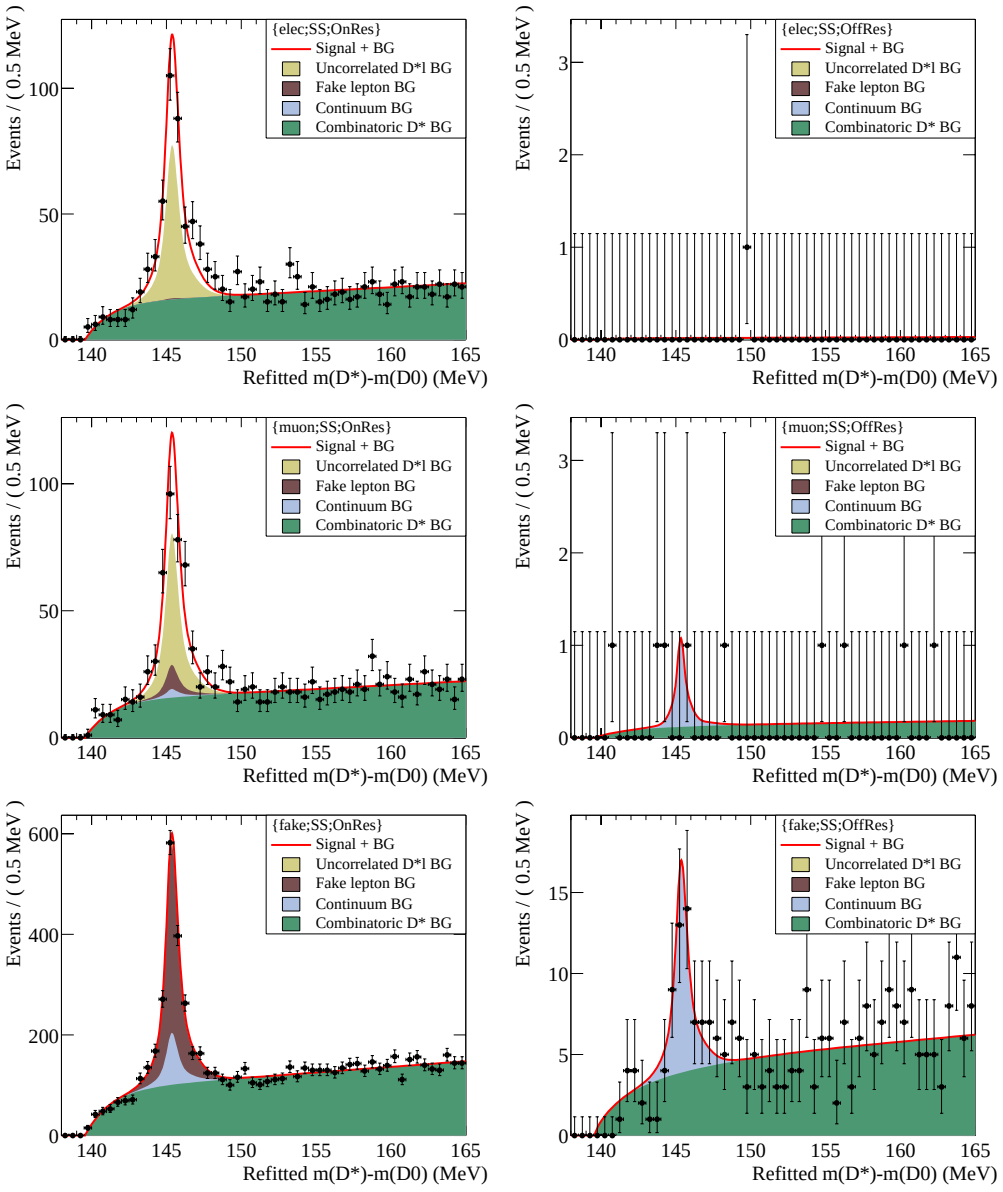


Figure A.1: Uncorrelated $D^*\ell$, fake ℓ, continuum D^*X, and combinatoric D^* contributions for same-side samples. From top to bottom: electron, muon and fake control samples; left column: on-resonance and right column: off-resonance.

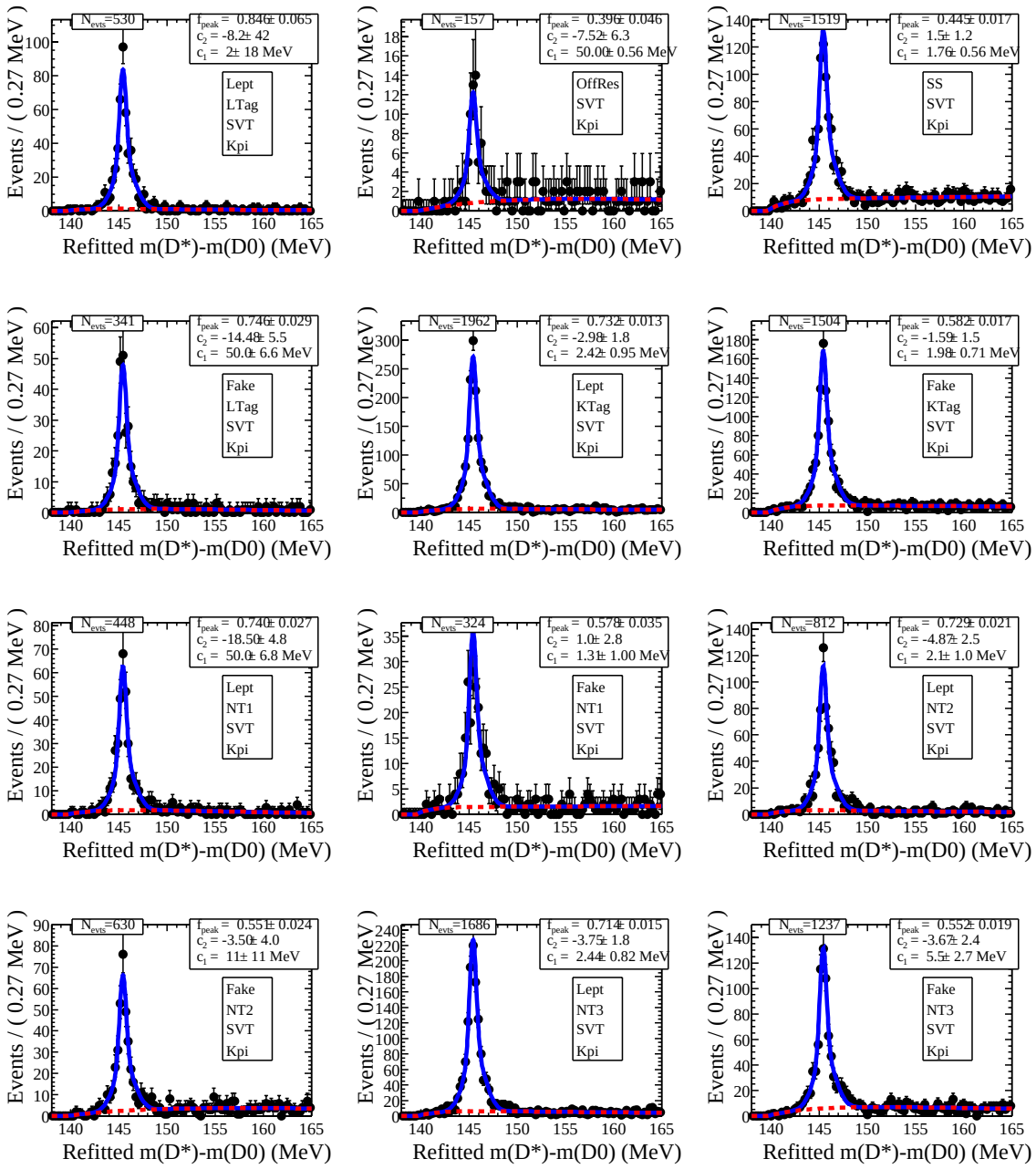


Figure A.2: Projection of the fit result on Data for the δm combinatoric background fits described in the text. The 12 different plots show the distinct subsamples of the (SVT x $D^0 \to K\pi$) sample from Data. The dashed curve shows the fitted contribution from combinatorics.

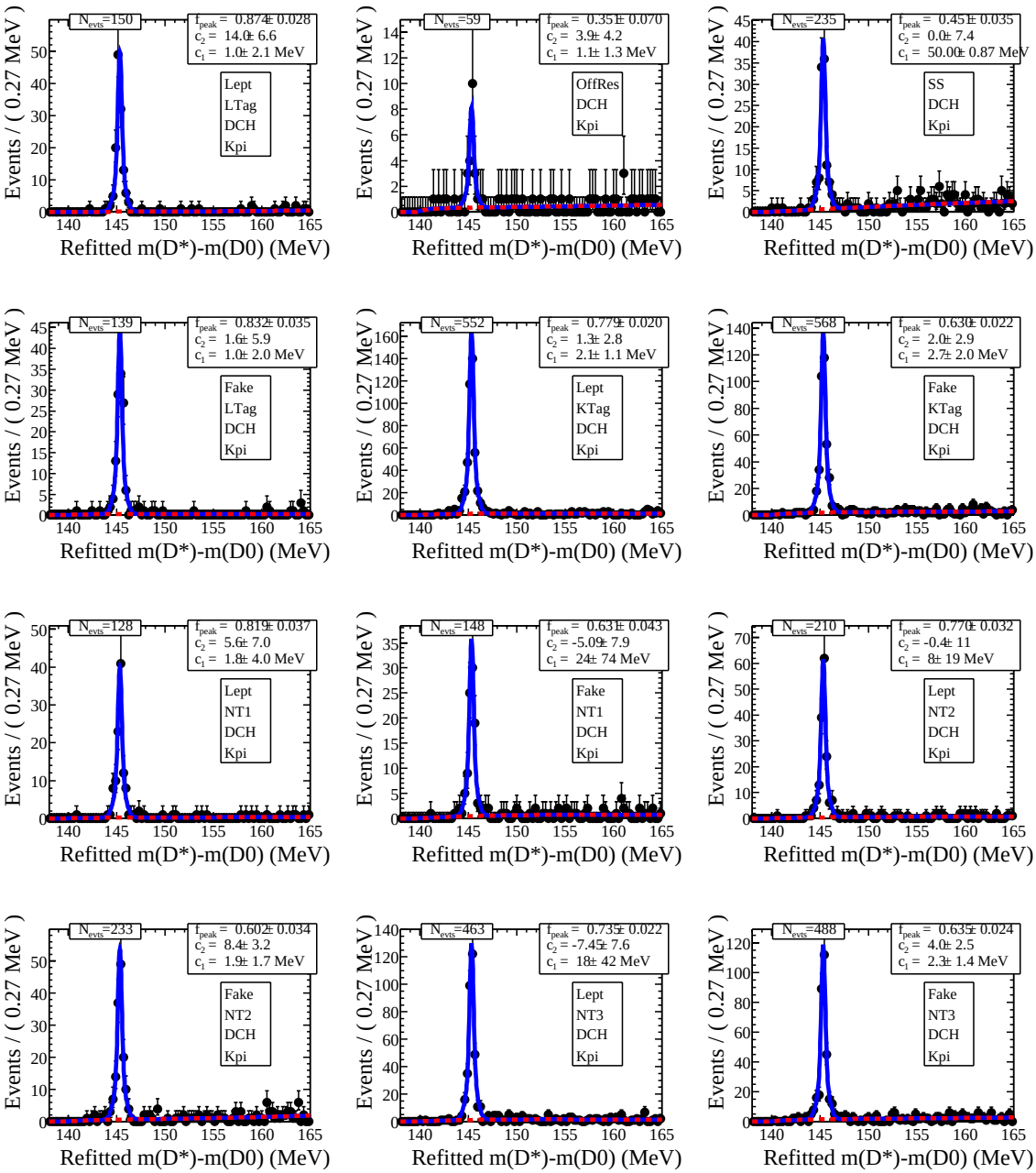


Figure A.3: Projection of the fit result on Data for the δm combinatoric background fits described in the text. The 12 different plots show the distinct subsamples of the (DCH x $D^0 \to K\pi$) sample from Data. The dashed curve shows the fitted contribution from combinatorics.

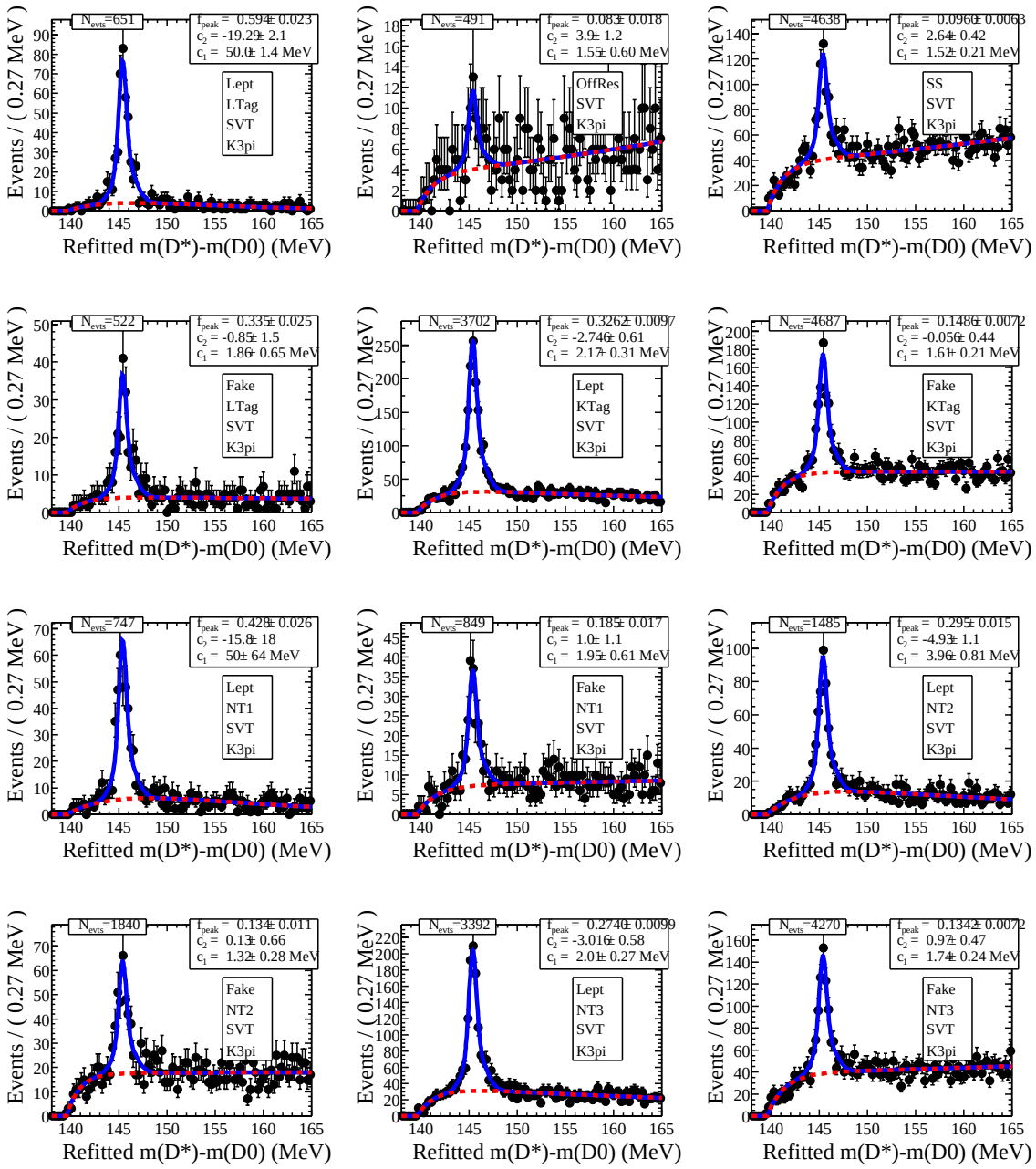


Figure A.4: Projection of the fit result on Data for the δm combinatoric background fits described in the text. The 12 different plots show the distinct subsamples of the (SVT x ($D^0 \to K\pi\pi\pi$ + $D^0 \to K^0_S\pi\pi$) sample from Data. The dashed curve shows the fitted contribution from combinatorics.

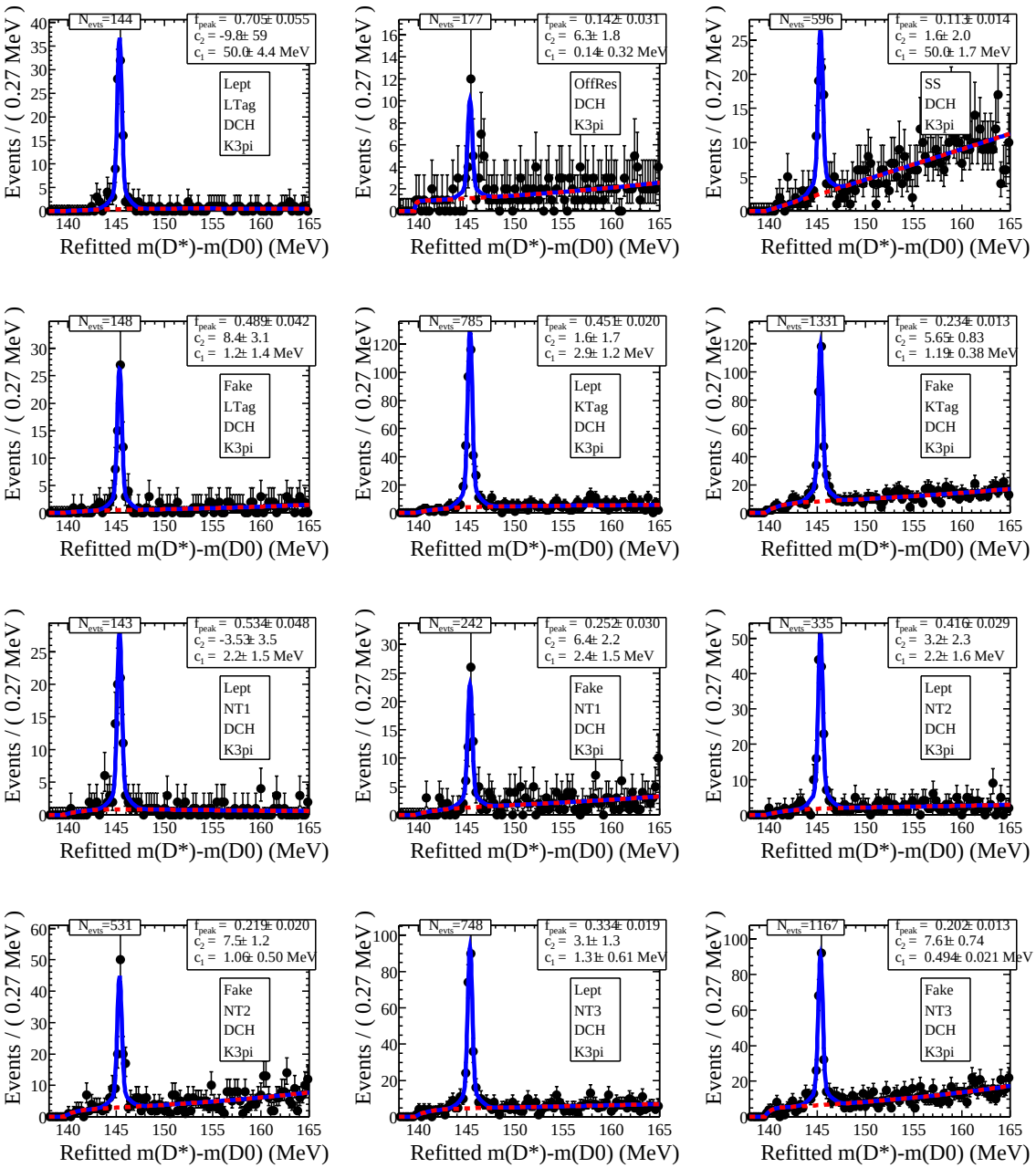


Figure A.5: Projection of the fit result on Data for the δm combinatoric background fits described in the text. The 12 different plots show the distinct subsamples of the (DCH x ($D^0 \to K\pi\pi\pi + D^0 \to K^0_S\pi\pi$) sample from Data. The dashed curve shows the fitted contribution from combinatorics.

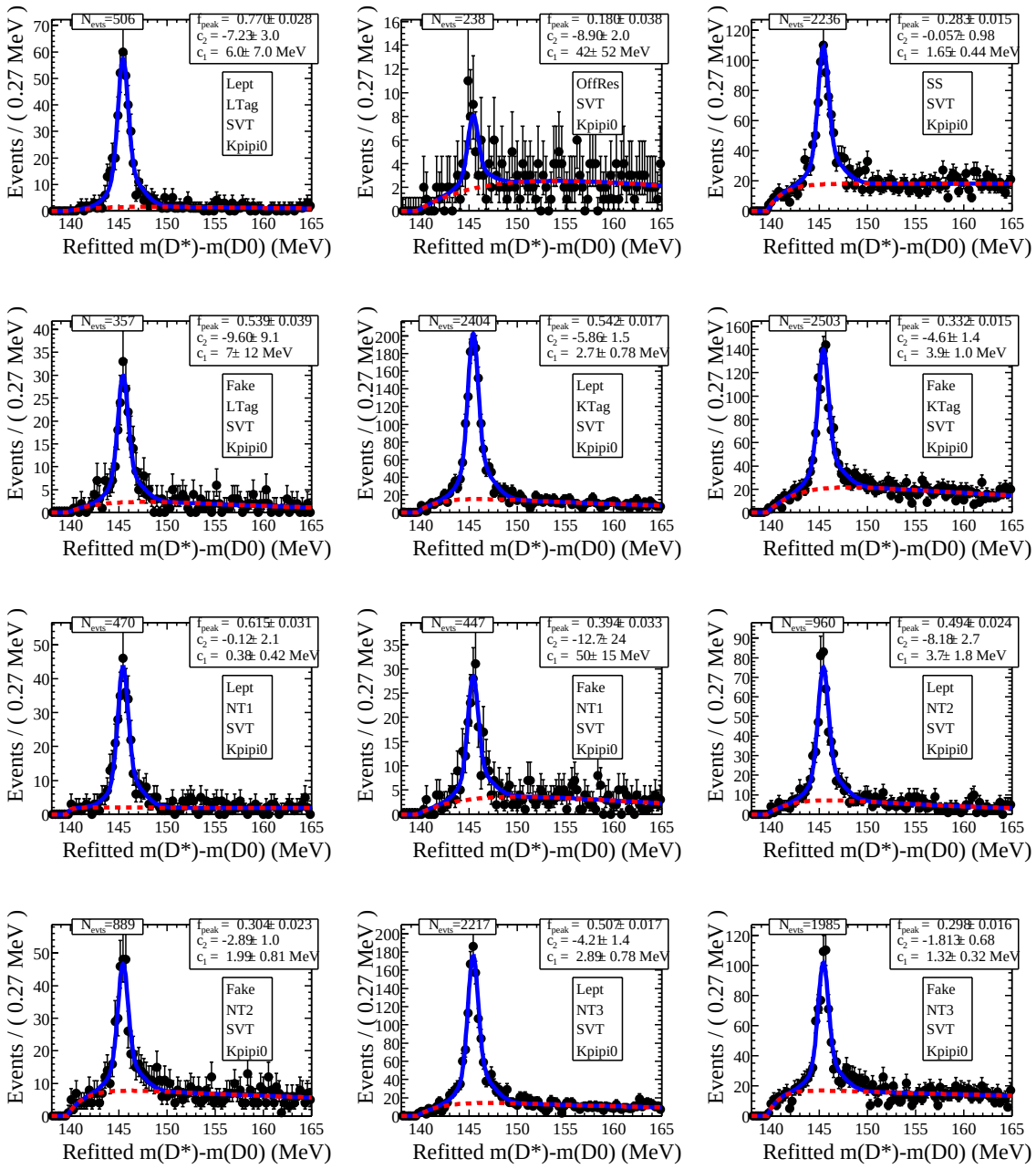


Figure A.6: Projection of the fit result on Data for the δm combinatoric background fits described in the text. The 12 different plots show the distinct subsamples of the (SVT x $D^0 \to K\pi\pi^0$) sample from Data. The dashed curve shows the fitted contribution from combinatorics.

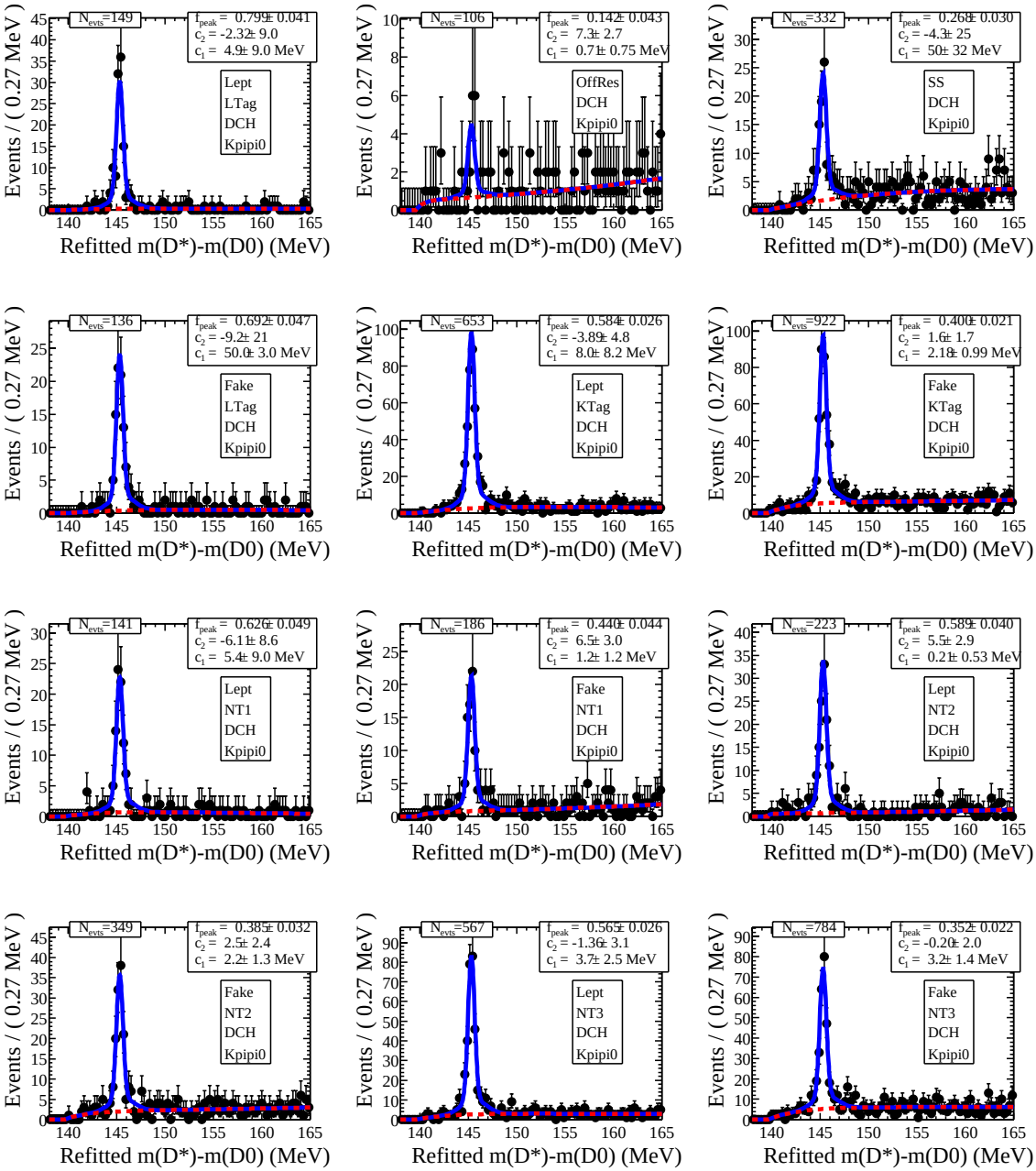


Figure A.7: Projection of the fit result on Data for the δm combinatoric background fits described in the text. The 12 different plots show the distinct subsamples of the (DCH x $D^0 \to K\pi\pi^0$) sample from Data. The dashed curve shows the fitted contribution from combinatorics.

Appendix B

Event Selection Criteria

For the sake of completeness, we list the long, long list of specific event selection criteria here. Refer to Chapter 5.5 for more details on the track lists, and Refs. [31] and [41] for implementation specifics.

- $|\cos\theta^*_{\rm thrust}| < 0.85$ where $\theta^*_{\rm thrust}$ is the angle between the thrust of the $D^*\ell$ candidate and that of the rest of the event.
- The charged tracks of the D^* and ℓ candidates have $r_{\rm doca} < 1.0$ cm and $|z_{\rm poca}| < 3$ cm where *poca* is the point-of-closest approach to the origin.
- The lepton candidate is selected from the GoodTracksTight list.
- The lepton momentum in the CM frame satisfies $p^*_\ell > 1.2$ GeV.
- For e and μ samples, the lepton candidate passes the veryTight e or μ selector. For fake samples, the lepton candidate fails both e and μ Loose selectors.
- π^0 is reconstructed from two photons with raw invariant mass within ± 15.75 MeV of the PDG π^0 mass. The π^0 is fit with a mass constraint and has χ^2 probability greater than 1%.
- K^0_S is reconstructed from a pair of charged π's (from ChargedTracks list) with raw invariant mass within ± 15 MeV of the PDG K^0_S mass. The vertexing χ^2 probability is greater than 1%.

- For the charged daughters of D^0 candidates, π's are selected from GoodTracksVeryLoose; the K is selected from GoodTracksVeryLoose list for $K\pi$ mode and GoodTracksLoose list for $K\pi\pi\pi$ and $K\pi\pi^0$ modes.
- Raw D^0 mass is within ± 17 MeV for $K\pi$, $K\pi\pi\pi$ and $K^0_S\pi\pi$ and ± 34 MeV for $K\pi\pi^0$ of the PDG D^0 mass. D^0 is vertexed with mass contraint and the χ^2 probability is greater than 0.1%.
- For $K\pi\pi^0$ and $K^0_S\pi\pi$ modes, the Dalitz probability density* is greater than 0.1.
- Charged kaon passes Tight criterion for $K\pi\pi\pi$ and $K\pi\pi^0$, or notPion for $K\pi$ mode using KaonSMSSelector.
- Soft $\pi^\pm$ is selected from GoodTracksVeryLoose list.
- Soft $\pi^\pm$ satisfies $p_{\pi_{\rm soft}} < 450$ MeV, and $p^t_{\pi_{\rm soft}} > 50$ MeV.
- Momentum of the D^* in the CM frame satisfies $0.5 \text{ GeV} < p^*_{D^*} < 2.5$ GeV.
- $D^* - D^0$ mass difference δm before $D^*\ell$ refitting (but using mass-constrained D^0) is less than 165 MeV.
- $D^*\ell$ vertex is fitted with beamspot contraint. The χ^2 probability is required to be greater than 1%.(Later, it is refitted with the multiply-constrained $(D^*, \ell, beamspot)$ fit which must converge as well.)
- For the opposite-side sample, the $D^*\ell$ candidate satisfies $\cos\theta^*_{D^*\ell} < 0$ and $|\cos\theta^*_{B-D^*\ell}| < 1.1$. For the same-side sample, it satisfies $\cos\theta^*_{D^*\ell} > 0$ and $|\cos\theta *_{B-D^*(-\ell)}| < 1.1$, where $-\ell$ represents the lepton four-momentum after it is flipped by 180°.
- Fit for $z_{\rm tag}$ using $B_{\rm tag}$ and beamspot constraint converges.
- $|\Delta t| < 18$ ps.
- $\sigma(\Delta t) < 1.8$ ps.
- The number of tracks used in the tag-side vertexing is greater than 1.

*We calculate the decay amplitude squared based on measurements of amplitudes and phases by E687[57] and the four-momenta of the D^0 decay products in our data, assuming perfect resolutions. The maximum is normalized to unity.

Appendix C

Comments on the *BABAR* Event Store

"The Data" is a broad term. The *BABAR* detector provides a complex set of time-dependent information, and it was only afters years of careful thought that the final design emerged. In order to maximize the utility of the *BABAR* dataset, information is organized into several gross domains under the *BABAR* Database Management System, composed of the Event Store, the Conditions database, and the Online databases (the Ambient and Configuration databases). The Configuration database is managed by the online group, and is implemented in Oracle to track and record which detector calibrations and configurations are used when data is recorded. The Event Store and the Conditions databases are the most important for physics analysis.

A *database* is a mechanism by which data elements and relationships (often called associations) between elements are preserved for index/retrieval. *BABAR* uses the commercial technology *Objectivity* for many of its database needs. Objectivity is one of the first commercially available implementations of a true object-oriented database that offers real-time control and large-scale systems support. In the *BABAR* implementation, an Objectivity database is more specific: it represents a unit of stored database information (usually just a file) with a given size, currently around 20 GB. A *federation* is an Objectivity reserved word defining a set of related database files and objects that can be accessed by a single application. Each federation consists

of possibly multiple databases distributed across multiple file servers, together with their indices and dictionaries. At *BABAR* there are several different federations in use for the Event Store, to prevent the several different central read/write activities from colliding. It is also nice to physically separate some groups of data, for instance, the simulated from the real detector events.

The Event Store

This document focuses on the Event Store part of the Data. Event data arrives from the detector, and if it passes the Level-3 Trigger selection criteria, it is written to a file on disk, called an XTC file because of the *tagged container* data structuring. The L3 trigger information is saved for each event as a set of trigger bits, the most important of which are the L3OutDch and L3OutEmc bits. (The presence of either of these bits guarantees an with a so-called physics trigger.) Simulated data is produced by BBSIM and SIMAPP and saved in XDR files. The L3-output XTC files are the ones processed by OPR, served primarily by BBR-SRV02. (OP)Reconstruction runs in the form of ELF for detector data (XTC) and BEAR for simulated data (XDR). The "reconstruction" executable does double-duty on the data: it extracts and computes the raw data format using digis (aka units of digitization), and also calculates (many) new quantities for each event. Events are subjected to a set of "filters", usually called BGFILTER, and their pass/fail state for each filter component is recorded in the event. The results of these conversions and computations are stored in the Objectivity databases. Each Run's set of events is stored as an Objectivity collection.

Hierarchy of Objects

In *BABAR* millions of useful events are recorded each year, but not all event are useful to everyt physics analysis. If the analyst can make a selection decision based on, say, only a dozen event summary variables instead of having to load and consider a few dozen tracks in the event, he or she can make significant performance gains. Many experiments, including *BABAR*, have elected to store several "versions" of the data, each with more reduced and more summarized information content. This allows the

analyst to breeze through the high-level features of each event and quickly determine if the event is useful. When a useful event is uncovered, the analyst requests the more detailed information on that event, performs more precise calculations to determine if the event meets stricter criteria, and so on and so on. There is some redundancy in this duplication, since the summary information stored in the high-level rep of the data is computed from quantities stored in the most-detailed representation. For most physics quantities, however, retrieval of the stored high-level quantity is faster than recomputation involving the retrieval of yet many more more detailed quantities. In *BABAR*, the words RAW, RECO, MINI, MICRO and NANO are used (respectively) to describe increasingly summarized levels of information about each event. (Tagbits are, for instance, stored in the NANO level of the Event Store.) We refer to these as levels or stages of information about an event in the Store. The nano and micro levels of information are well described in Ref. [80].

How does one store the different levels of detail about each event? *BABAR* uses the tools of Objectivity (Objy) to manage these different layers of information in a minimally redundant way. Each event is represented with an event header which serves as a portal: The event header object holds references to child objects, each corresponding to an additional stage of processing. Each child stage object contains a stage header and a summary object. *BABAR* uses each stage of object to store a level of information about the event. For instance, once the reco stage object is attached to the event header, another processing pass (often in OPR) computes the global R2 value for the event and adds the quantity to the nano stage (the event summary object). Objectivity calls the summary object for each stage the tag object because it provides fast query capabilities. For historical reasons, the nano level of information is also known generally as the event Tag, the summary object for the entire event (unrelated to B-tagging!). The power of the Objectivity implementation lies in the organization of these levels (stages) of information; typically, a database file will contain only one stage of an event's processing. Event access and retrieval per processing stage is therefore streamlined.

Kanga

Due to initial concerns about the throughput, accessibility, and distribution of the Objectivity Event Store, an alternative technique for compact data access was proposed and implemented. Kanga uses a ROOT-based technology to store an event's micro and nano information (only) independent of Objectivity. Each Kanga file is equivalent to an Objectivity collection, and in fact, Kanga files are produced from the Objy Event Store by a separate production task. Kanga files offer efficient access and portability but do NOT contain the full event information. In March 2000, Kanga file production became centrally and uniformly managed. Kanga files are produced for:

- skimmed events from good runs processed by the central skim production
- isPhysicsEvents from good runs processed by the central skim production
- SP3 Objy collections

Skims

For additional performance, *skims* have been developed. Known in some other experiments as "strips" or "substrips" of the data, a skim contains a reduced set of events that have been selected according to some physics criteria. Looping over a smaller set of "interesting" events will always be faster than examining every event in the Store. Each group of analysts working on similar physics processes (i.e. within an Analysis Working Group) defines their own private set of event selection criteria that are loose and efficient. The central production team produces the skimmed data sets which the individual analysts can process with their more detailed selection more quickly. A well-planned skim selection criteria may actually only involve a single bit in the event's nano layer of detail, i.e. a tagbit. This greatly simplifies the selection process and additionally preserves "always and forever" (in the event itself) which skim criteria are satisfied by the event. For some skimming procedures, the exectuable will add a "tagbit" to the event's nano information to indicate its selection decision.

It is important then for the central skim production team to run the same set of executables over all events so that an event looks "the same" in every skim.

The Objectivity implementation of the skim is again designed to minimize redundancy: collections in an Objy skim only contain pointers to the selected events' headers in the Event Store. An Objy skim doesn't contain any new information, it is just a list of pointers to events in the Store that passed certain (typically nano based) selection criteria. Files in a Kanga skim, however, are actually standalone files holding a smaller number of events with their own private nano and micro quantities.

Conceptually a *stream* and skim are equivalent : a subset of events that are peeled off for a purpose, typically because they appear to contain a particular physics feature. In these terms, a stream is just a big skim. Historically, the two words have meant different things at different times, especially when comparing Kanga and Objy implementations. Typically, a stream has referred to a collection of events that are packaged in a manner to facilitate portability to remote sites. A skim is typically defined by a set of physics processes or selection criteria.

Appendix D

Comments on Parameter Splitting

To build the case for parameter "splitting," consider the following exercise. Suppose we have a class of events that have Gaussian distributions in some variable x. Additionally, half of the events have an attribute value of "Narrow" for some characteristic, and the other have the value "Wide" for the same characteristic. If we were to draw their collective distribution in x, we might arrive at Fig. D.1 which we fit with a single, common Gaussian function.

Guessing that the fit is rather poor in terms of estimating the width of the Gaussian, we might then separate the events completely by their Color category, "splitting" the Gaussian's mean and width on the values "Wide" and "Narrow". We now perform another fit, but it is now two separate, disjoint fits. We discover that the the widths σ of the two Gaussians are indeed statistically distinct, but the means μ look similar, as shown in Fig. D.2.

Finally, we elect to "share" the mean between "Narrow" and "Wide" because there will be twice as many events used to determine the common mean parameter, greatly improving our statistical power. We are "splitting" the width, σ, between the two subtypes of events, but "sharing" the mean parameter. The simultaneous fit to the x distribution now looks as in Fig. D.3, and we observe the improved statistical error on μ. Unveiling the source distributions in Fig. D.4, we observe that this final simultaneous fit splitting across categories is indeed the correct way to represent the joint distribution in the variable x.

This exercise is trivially simple, but it outlines one of the strategies we use to determine if parameters should be "shared" or "split" across categories. Note that we do not need to split all parameters in the same way. We perform exercises in this vein in Data using the full set of 360 possible subdivisions within each control sample.

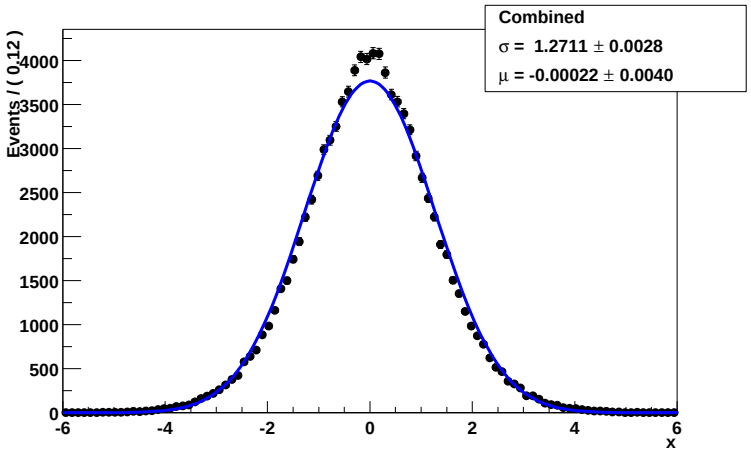


Figure D.1: The summed distribution in x for "Narrow" and "Wide" events and the the projection of a single Gaussian fit.

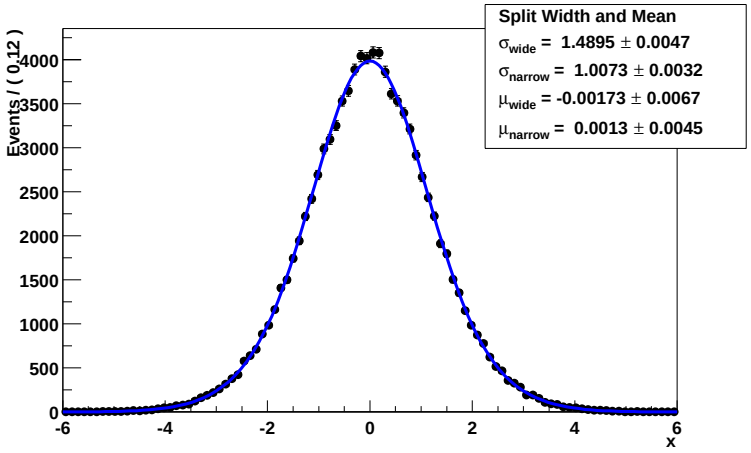


Figure D.2: The summed distribution in x for "Narrow" and "Wide" events and the projection of the two, distinct Gaussian fits.

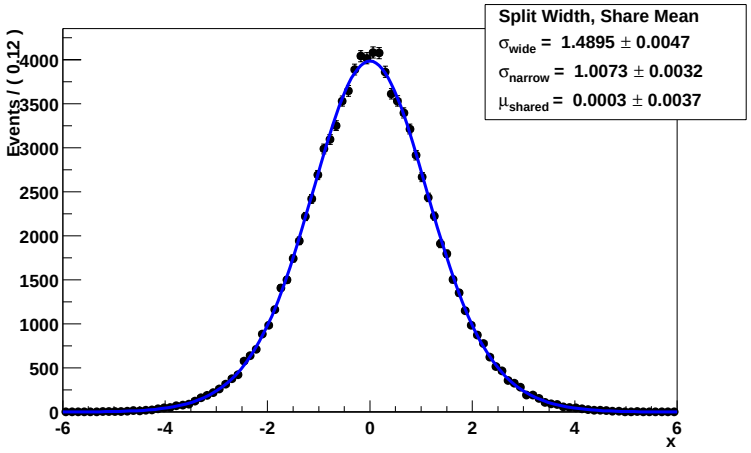


Figure D.3: The summed distribution in x for "Narrow" and "Wide" events and the projection of a fit result using a simultaneous pdf that shares and splits different parameters. In particular, the width σ is distinct for "Narrow" and "Wide" events, and the mean μ is common.

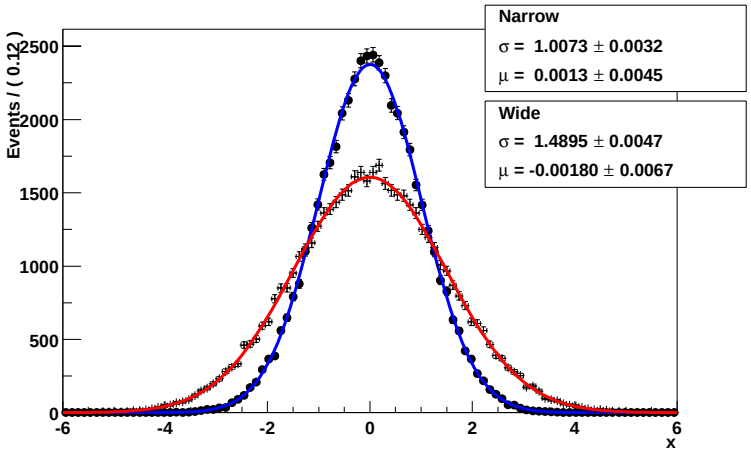


Figure D.4: The distributions in x for "Narrow" and "Wide" events, drawn separately and in different colors.

Appendix E

SVT Radiation Monitoring and Protection

This Appendix brings together several different resources to define the *BABAR* Silicon Vertex Tracker's radiation monitoring and protection system, *SVTRAD* for which I was primarily responsible. It is impossible to distill the full spectrum of documentation into one brief set of comments. We therefore choose to briefly describe the system andinvite the reader to review some hand-picked references.

Reference	Brief Description
Ref. [81]	Homepage for complete online documentation
Ref. [82]	Proceedings contributed to DPF2000 Conference
Ref. [83]	Draft of future plans and upgrades
Ref. [84]	Overview of the *BABAR* Radiation Protection System

Table E.1: Brief list of top references for the SVTRAD system.

E.1 Motivation

The demand for high luminosity at B-factories makes accelerator-induced backgrounds an important issue. The need for strong final focusing, high beam currents,

and tightly spaced bunch trains creates potential background sources, such as synchrotron radiation, and lost particles from beam-gas interactions or the scraping of beam-tails. The *BABAR* detector at the PEP-II B-factory is no exception. These backgrounds can have a significant influence on the detector and its physics goals by potentially causing: long-term damage to the detector hardware and its electronics through integrated radiation exposure; performance reduction of the detector due to hardware or bandwidth limitations; complications to the extraction of physics signals due to spurious background tracks, photons, and triggers.

Radiation monitoring

Experience at accelerator-based experiments elsewhere has indicated that successful control of radiation backgrounds relied strongly upon precision measurements and characterization of background sources and dependencies: efficient remediation of backgrounds requires an understanding of their origin. Based on detailed experiments and testing before final assembly, many of the *BABAR* subsystems assigned a lifetime radiation budget to their sensing hardware and front end electronics. To ensure the careful rationing of each sub-detector's lifetime, detailed cumulative radiation monitoring is necessary. The design of a system that could accurately report/record instantaneous and long-term radiation dose rates for the *BABAR* detector was critical.

Radiation protection

An additional challenge at *BABAR* is the severity of machine-induced backgrounds due in part to the high very beam currents of the PEP-II storage ring. An accident in the machine could easily dump MRads of ionizing energy into the detector, damaging it severely. A radiation monitoring system therefore needed an active "feedback" component, a protection system that could automatically and independently abort the beams in the machine if radiation levels within the detector became dangerously large.

The decision to abort the beams because of a high risk of permanent radiation damage is not easy – it must be weighed against the experiment's needs for integrated

luminosity and the operational efficiency of the detector. A hardware protection system needs to be extremely flexible, then, in order to meet the changing demands of the detector and accelerator as each evolves and matures. In addition, a rational radiation protection system requires an expert monitoring system that can forewarn, diagnose, and evaluate radiation accidents.

Radiation and the SVT

The unprecedented beam currents and luminosity at the PEP-II accelerator make machine backgrounds a significant challenge to the *BABAR* experiment. Despite significant use of radiation hard technology, the Silicon Vertex Tracker (SVT) is the most radiation vulnerable sub-detector because of its proximity to the beamline. The SVT has established a rigorous program of radiation monitoring and protection (SVTRAD) to ensure the proposed lifespan, including the ability to automatically abort the beams in PEP-II when radiation exceeds programmable thresholds

The innermost subsystem of the *BABAR* detector is the Silicon Vertex Tracker (SVT), consisting of 5 concentric layers of double-sided silicon microstrip detectors. The inner two layers are within 3 cm of the beampipe, and hence receive some of the most intense radiation within the entire detector. The proximity to high radiation, combined with the inherently vulnerable Si nature of the detector modules and the novel front-end CMOS electronics placed on the silicon wafers themselves, makes radiation safety a top priority for the SVT.

The SVT is vulnerable to radiation damage by several different mechanisms. Both the front-end readout electronics and the actual silicon wafers themselves are susceptible to gross radiation damage. (For the silicon wafers, the pedestal leakage current increases with damage and the CMOS readout electronics suffer a degradation in noise performance.) Radiation damage to the SVT sensors and readout chips are increased surface leakage current, increased interstrip capacitance, and increased noise from the input transistor. These effects reduce the signal-to-noise ratio. The SVT has been designed as a rad-hard device, such that test beam studies and extrapolation to the full size detector give the SVT a minimum 2.4 MRad radiation dose budget over its

lifetime.

Additionally, there was strong evidence that the detector modules could be vulnerable to large radiation transients which can short-circuit an effective capacitor near the p-stop implants. It is estimated that an acute dose exposure is damaging at levels of 1 Rad deposited in less than 1 second. The SVT is required to meet its resolution performance specifications up to an integrated dose that is 10 times the expected nominal dose, or 2.4 MRad over 10 years for the first-layer sensors and electronics.

Because of the sensitivity of the SVT to radiation damage, and based on preliminary experiences with the diversity of backgrounds at PEP-II during the commissioning phase, it was necessary to design and implement a system specifically dedicated to the SVT. The final solution described here fulfills both the radiation monitoring and protection needs.

E.2 Specifications

The systems solution (SVTRAD) to the radiation monitoring and protection needs for the SVT evolved through several years of prototyping and detailed studies during the construction and commissioning phase of PEP-II. SVTRAD needed to use a set of sensors which are compact and easy to distribute within the SVT volume. Using these sensors and cutsom electronics, the systems needed to provide radiation monitoring and protection, while conforming to *BABAR* and PEP-II standards of interface and control. This section details these specifications.

Mechanical constraints

Early simulations accurately predicted a complicated pattern of radiation exposure around the central beampipe near the interaction point. The scaling of the background levels is more complex than a naive $1/r^2$ flux fall-off because of beamline component shadowing and complex upstream and local scattering centers. This flux consists primarily of secondaries from showers initiated by scattered beam particles.

To obtain the most accurate (spatially segmented) monitoring of radiation exposure, it was necessary that the radiation sensors be placed as close as possible to the SVT's silicon detector modules.

The narrow mechanical envelope allocated for the sensors was the most stringent constraint. The sensing volume was restricted to two annular regions of 19.7 mm in z and 3.6 mm in r, at $z = +12.1$ cm and -8.5 cm.

Performance requirements

The primary requirement of this system is to protect SVT components from avoidable radiation damage by early detection of increasing radiation levels. Provision of real-time monitoring of slowly-varying (few second) radiation levels in the interaction region for accelerator tuning and diagnostics is also necessary. The ability to integrate measured radiation dose rates to form estimates of the total dose received by the SVT and its front-end electronics is critical as well. The radiation monitoring and protection system needs to be operational at all times when beams are present in the accelerator.

A standard unit for measuring absorbed radiation dose, which we use almost exclusively here, is the Rad; we additionally use mR/s for measuring instantaneous radiation dose rates. The design radiation dose budget of the SVT is 2.4 MRad over 10 years. A naive model for the accumulation of this dose yields an average dose rate of 24 mR/s. We therefore require that the radiation monitoring system be sensitive to changes in dose rate on the order of 5 mR/s. Dose rate samples should be provided at the rate of a few Hz, and should have an accuracy of at least 10%. Systematic uncertainties (such as temperature dependence and radiation damage) need to be controlled carefully to preserve the overall 10% uncertainty on radiation dose integrals over the expected 10 year lifetime.

We require a minimum number of false-positive trips from the radiation protection system. The shortest timescale for a response to avoid damage to the SVT is 100 *mu*s, being equivalent to a few tens of ring turns, or a few times the minimum response time of the beam abort kicker magnets. We therefore require that the system

respond to acute radiation doses within 100 mus. The protection system must also prevent chronic exposure at rates higher than the budget. These two requirements are combined by defining a damaging exposure as a minimum dose absorbed at a minimum dose rate. The minimum dose integrated above threshold which results in a trip should be programmable between 100 mRad and 10 Rad, and the chronic dose rate threshold should be usable between 10 mR/s and 1000 mR/s with an absolute error better than 10%. The implementation of these thresholds must also be immune to temperature variations induced by normal operations of the SVT and/or effects induced by beam heating.

The SVTRAD system is only one component within the general *BABAR* Radiation Protection System (RPS), and as such, it must also not unnecessarily interfere with other systems' performance. A sketch of the full *BABAR* RPS is shown in Fig. E.1.

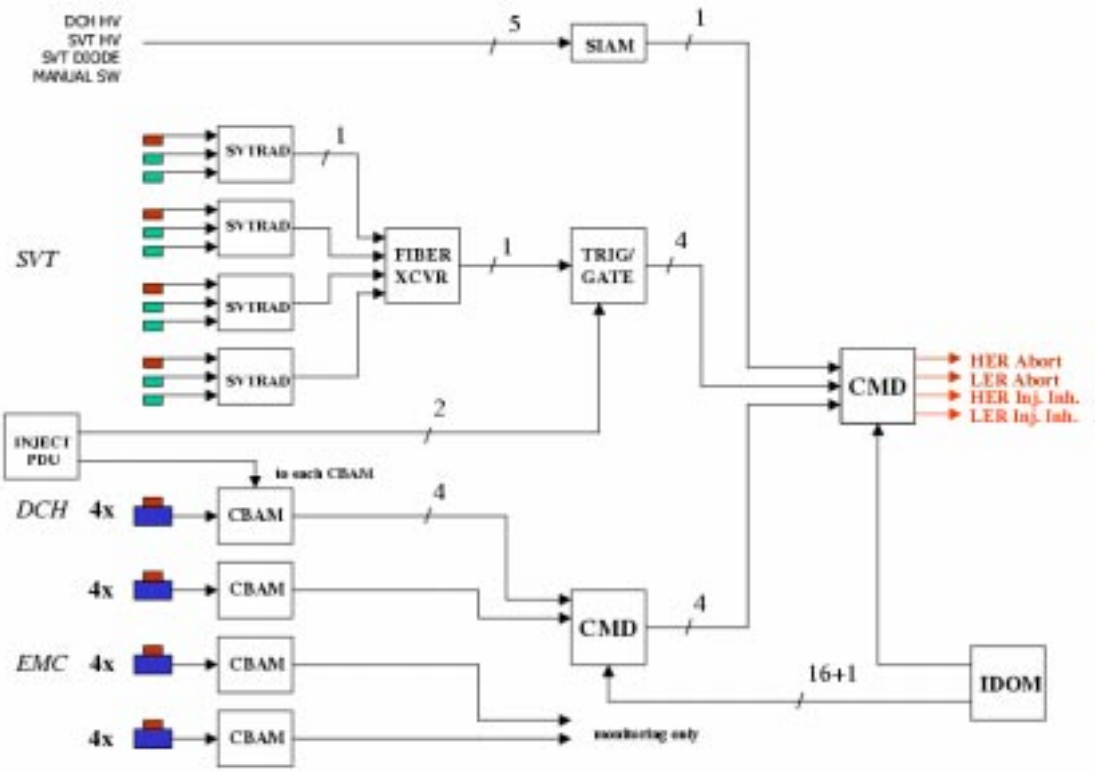


Figure E.1: A schematic of the *BABAR* Radiation Protection System.

E.3 Sensors

Based on the mechanical constraints and the need for radiation monitors with sensivitiy and response comparable to silicon in the SVT, large area PIN photodiodes were selected as the optimal sensors.

The SVTRAD system uses 12 Hamamatsu PIN photodiodes (S3590-08) and 24 YSI thermistors (55036). The reverse-biased diode behaves like a current source whose output varies with the instantaneous dose rate, the integrated dose received, and the temperature. The accuracy for measuring dose rate (= total current - leakage current) depends initially on the current-measuring accuracy of the readout, and then eventually (after sufficient radiation damagce) on the accuracy to which the diode temperature can be monitored to correct for leakage current variations. The accuracy for measuring integrated dose is limited by how well a diode's annealing history has been tracked, and the uncertainty in the damage coefficient. These issues are discussed more completely in the next sections.

PIN photodiodes

The sensor of choice is a large-area Hamamatsu S3590-08 PIN photodiode. The diode consists of a 300 μm thick N-type silicon substrate (intrinsic layer) with a thin P-layer at the active side with the active area 1 cm $\times$ 1 cm and an N^+-layer at the rear surface. The capacitance of this diode is about 50 pF. For our detectors, the diodes are reverse-biased at 50V (the full depletion voltage); the leakage current for an undamaged PIN diode is typically 1–2 nA at room temperature. Because of the relatively thick intrinsic layer, the fully depleted bias is relatively low, and since almost the whole intrisic layer is depleted, the active region is almost the same size as the whole material.

Because the diode relies on the details of the Fermi-Dirac distribution function for majority and minor carrier population, the response of the diode is highly temperature sensitive. Theoretically, one can show that the leakage current of the reverse biased

diode should behave as

$$I(T) = I(T_0) \times (T/T_0)^2 e^{\frac{-E}{k}(\frac{1}{T}-\frac{1}{T_0})}, \tag{E.1}$$

where E is the bandgap energy of silicon, or about 1.2 eV, and k is the Boltzmann constant. With a Taylor expansion in the limit of small temperature changes expected during normal operation, this temperature dependence can be parameterized by a simpler exponential model,

$$I(T) = I(T_0) \times e^{\alpha(T-T_0)}, \tag{E.2}$$

with the temperature coefficient, α typically 10%/°C.

Signal extraction

Signal current is formed by the passage of charged particles through the diode's depletion region which deposit energy, creating electron-hole pairs along the particle trajectory. (The number of the pairs produced is proportional to the energy lost by the particle.) Under the applied electric field, the electron-hole pairs flow out of the bulk and contribute a signal portion to the total leakage current. Taking into account the bandgap energy in silicon, each mRad of energy absorbed by the diode yields about 0.2 pC of charge.

To extract radiation dose rates one needs not only to accurately measure the total leakage current, but also to have a good model for the pedestal.

The integrated radiation effects in semiconductors are complicated. With significant exposure to radiation, the diode becomes damaged and the leakage current increases. The basic radiation damage mechanisms are:

1. Displacement damage: Incident radiation displaces silicon atoms from their lattice sites. The resulting defects alter the electronic characteristics of the crystal.

2. Ionziation damage: Energy absorbed by electronic ionization in insulator layers, predominantly SiO_2, librates charge carriers that diffuse or drift to other locations, usually at pre-existing defects, where they are trapped, leading to unintended concentrations of charge and, as a consequence, parasitic fields.

Ionization effects depend primarily on the energy deposited in the material, independant of the energy and the type of the incident particles, and displacement damage depends on the non-ionizing energy loss, namely, the energy transferred to the lattice atoms, which depends on the mass and energy of the incident particles. The ionization damage is considered to be more transient with relaxation time of the order of days, which is highly dependent on the temperature, while the displacement damage is a long-term effect. These mechanisms are not fully understood The effect of damages on our device is the increase of the leakage current. The relation between the increase of leakage current and the radiation dose depends on the types of the radiation, and is assumed to be linear over a wide range. The proportionality factor is of the order of 1 nA/cm^2krad according to the previous studies using proton and neutron sources.

- distributing bias voltage to each photodiode sensor
- receiving, filtering, accurately digitizing and simple time averaging of each diode's leakahe current
- forming and buffering a continuous data stream for acquisition via the BaBar standard protocol for slow-control communication
- interpreting real-time diode leakage currents as radiation signals on-board to provide an interlock alarm to the accelerator control system when radiation levels exceed user-defined thresholds

Thermistors

Because of the leakage current's significant temperature dependence, dedicated temperature sensors were needed for each radiation sensor. Although thermal gradients within the first layer of the SVT were expected to be small due to a chilled

water cooling system, two negative temperature coefficient thermistors were mounted adjacent to each photodiode.

E.4 Design Choices

Many of the design choices were guided by *BABAR*-wide system standards and PEP-II interface considerations. Several key principles were emphasized in the design phase of this project.

1. Systems flexibility and expandability. Many configuration options were left as parameters, and conscious effort was made to keep the hardware rigorous and failsafe and yet not value-specific.

2. Modularity. The separate importance of the radiation monitoring and protection tasks and their relative decoupling suggested that interdependencies be minimized between them. Each subsystem was designed to operate with and without remote support, such that a breakdown in communications would not compromise the hardware protection, nor cause a configuration ambiguity when communications were restored.

3. Synchronization and transparency. The tiered structure of the task description (hardware, local firmware, remote communication protocols, and high-level user interfaces) called for deliberate and careful bookkeeping. By instantiating certain tasks of the hardware as high level objects, organization of error detection and configuration became simpler and more intuitive.

Monitoring

An obvious choice for precise monitoring of the diode leakage currents is high efficiency charge-integrating ADCs. Relevant time scales for average values to be sent to scrolling displays were judged to be a few Hz. This was in contrast with the expected time development of dangerous radiation accidents, which could be anywhere from a single beam turn around the ring (7.4 μs) to several seconds for a feedback

instability. The dynamic range of 250 kHz - 0.5 Hz was too broad to span, so the upper limit on ADC digitization was set around 20 kHz. To accommodate this range of time-resolved ADC samples, a two-stream model of data transport was devised.

A front-end high-speed ADC with programmable sampling capabilities was necessary, combined with a bank of accumulating registers implemented in a Field Programmable Gate Array (FPGA) and a set of deep memories to permit asynchronous polling of the data from a remote logging client. Individual ADC samples (internally averaged and/or filtered) are directed to the accumulator. Each sample updates the accumulator's contents, which is echoed onto a 5000 sample-deep circular buffer. This very deep buffer forms the "fast" datastream, and after start-up equilibration, carries at least 2500 samples of recent history. The accumulator rolls over every N samples, copying its final sum value to the "slow" datastream buffer. (The rate of samples in the "slow" buffer is chosen to match the remote polling rate.) At the user's request, the ADCs are frozen, and the "fast" circular buffer can be read-out at the standard "slow" readout rate.

Analog low-pass filtering with a 5 ms time-constant was selected to form a primary guard against high frequency noise. Optimal powerline noise (60 Hz) rejection was achieved in the "slow" datastream by requiring the samples' averaging time to be a multiple of the fundamental 60 Hz. The circular buffer was implemented with a pair of back-to-back deep FIFOs such that when one became full, the other was cleared, and the sample stream transferred to it.

Protection

Solving the radiation protection problem with PIN diodes is intrinsically difficult, especially in the high-dose/long-lifetime environment of *BABAR*. The design chosen must allow precise setting of thresholds as low as 10 mR/s (10 nA) when diode leakage current pedestals range from 2 nA to 1000 nA over their lifetime of radiation exposure. The accuracy of the threshold depends almost entirely upon one's knowledge of the diode pedestal current. Because of the large temperature coefficient (approximately 10% / deg C), any meaningful radiation trip threshold also needs to allow for front-end

temperature compensation. (For instance, the SVT DAQ system changes the diodes' ambient temperature by as much as 0.5 deg C, which for an uncorrected diode signal with pedestal current of 100 nA amounts to an excess signal of 5 nA or 25 mR/s.) As radiation damage increments the diode pedestal current, the radiation protection circuitry must also be capable of efficient re-calibration to adjust for damage and re-compute the radiation trip thresholds. Finally, the radiation protection circuitry should preserve the input diode signal for transmission to the monitoring system for logging, computation, and display. (We have allowed for a degradation in signal performance by a factor of 10 after radiation protection processing.)

Several analog schemes based on comparators with DAC-generated reference levels were prototyped, but they typically failed due to poor noise immunity, thermal drift of components, and a gross inability to effectively distinguish between events of damaging dose integral and transient events with a high peak but negligible width over the full range of time scales. A scheme of competing frequencies into a "leaky-bucket counter" was designed, which additionally offered straightforward remote configuration and control. This solution best instantiates the goals of minimizing radiation exposure of the SVT while providing a programmable level of "forgiveness" for the machine's radiation backgrounds.

Commercially available ICs exist which can convert an input current to a frequency pulse train in a reliable, 0% deadtime fashion. The leaky-bucket uses these frequency inputs to implement a digital counter of programmable depth. The three inputs are proportional to:

- total diode leakage current
- voltage-biased thermistor current
- programmable up/down clock rate

When the bucket overflows, the circuit trips, registering a radiation alarm. The three primary inputs form a net count rate for the bucket; when the count rate is greater than zero, the bucket will inevitably overflow and a trip is imminent. Of course, if the net count rate is negative, the bucket will remain empty. The acute dose

threshold is the delta-function extreme on the diode current input; in the limit of high radiation, the time until trip is only governed by the depth of the counter, and hence is proportional to the total dose integrated until the trip. The chronic dose threshold takes over in the limit of many clock cycles from each of the inputs. The diode current input increases with radiation and temperature, filling the bucket. The thermistor input forms a negative temperature coefficient signal, which counteracts the changes in the diode input due to thermal variations. The third input, a programmed clock signal, sets the effective chronic dose threshold. To reach a trip condition, the diode input signal must remain greater than the clock and thermistor combination long enough to fill the bucket. The depth of the bucket is now the hysteretic threshold, putting a cap on the dose integrated until trip. The time until trip is inversely proportional to the input radiation dose rate, with a "forgiveness" equal to the depth of the leaky-bucket.

The successful detection of dangerous radiation conditions and the accurate generation of a radiation trip signal is not the end of the story. The SVTRAD system must interface with the PEP-II Machine Control System, to gain access to beam abort and injection inhibit inputs. The SVTRAD radiation trip signals are routed to the accelerator control system via fiber optics to eliminate impedance and level-shifting issues. At the machine interface, the fiber light levels are translated to locally-referenced TTL and compared to individual injection triggers derived from PEP-II. Because injection losses can easily dominate the radiation backgrounds in the interaction region and because the cost of aborting stored beams in both rings is so high, the radiation trip signals from SVTRAD are gated against individual injection triggers. If the radiation trip is coincident with an injection pulse, the trip is steered toward the injection kicker magnet and latches an injection inhibit. (Timing coincidence is independent for each ring, and hence each ring can be separately inhibited.) If another radiation trip occurs before the injection inhibit is manually reset, the stored beam is deemed the guilty party and the stored beams are aborted.

Remote protocols and controls

The SVTRAD system was designed to comply with *BABAR* interface standards, namely through the use of EPICS and CAN-bus. Controller Area Network (CAN) is a bus-protocol with high bandwidth (up to 25 kbits/s) and reliable arbitration. The SVTRAD module operated independently of remote clients, retained configuration in EEPROM across power-cycles, and could also host an elementary RS232 serial-port sessions. A polling process living in EPICS requested status and slow ADC data from the SVTRAD module at a constant rate.

E.5 Implementation

The SVTRAD system is the combination of several different pieces of hardware, an array of sensors, and a dynamic set of EPICS-based applications running on single-board computers and UNIX clients. We describe here the details these each in turn.

Overview

The front-end hardware of the SVTRAD system consists of printed circuit board sporting three programmable memories, a microcontroller, and two communications ports. Each SVTRAD module services three separate diodes, distributes bias voltage and a commonly referenced grounding scheme. Each diode is supported by two thermistors in close thermal proximity; these thermistors are also connected to the SVTRAD module.

The SVTRAD module is a custom designed 6U x 13 3/8" x 8 H PCB fabricated specifically for insertion in a double depth VME crate, although there is no dynamic link through the VME crate's backplane. (Only a system reference for the grounding scheme is provided on the backplane connector.)

The microcontroller is the heart of centralized operations of the board, although both the monitoring and protection tasks (once configured, either at power-up or by specific commands from the mcu) can run without support from the mcu. The radiation monitoring task is handled by a dedicated FPGA, as is the radiation protection

tasks' leaky-bucket counters. Each of the module's three channels uses a hardware configuration jumper to determine if the diode current will be used exclusively for monitoring. Any diode participating in the interlock/protection circuit is fed back into the monitoring ADCs via a signal-reconstruction circuit. A combination of the three diodes' radiation trip states are used to form the entire module's radiation alarm, which is passed out on a dedicated fiber optic as a negative going 4 MHz pulse. Fig. E.2 depicts the separation and collaboration between the two tasks of monitoring and protection.

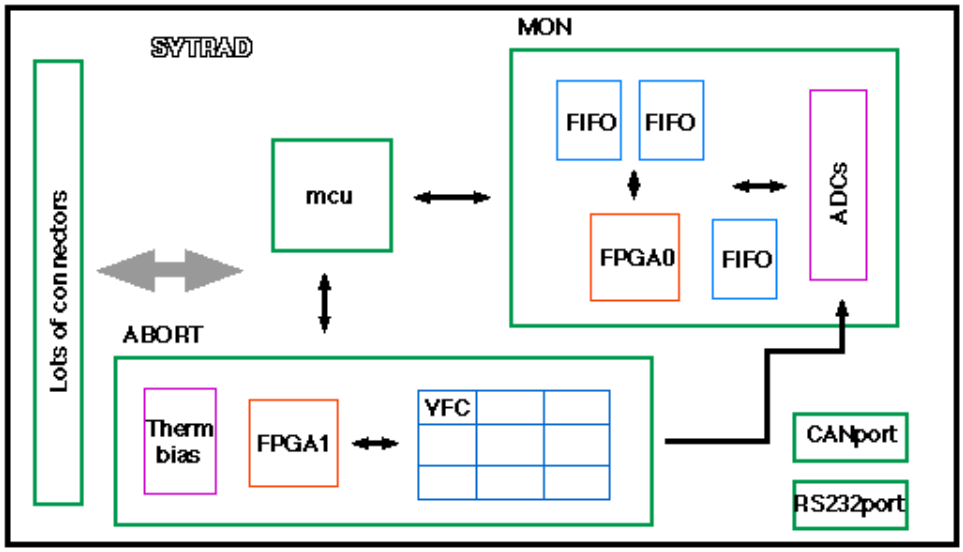


Figure E.2: Toy block diagram of the SVTRAD module.

Radiation monitoring system

When a particular diode enters the monitoring stream, the input level is digitized using a precision Burr-Brown DDC101-U ADC. Most of the ADC conversion options (rate, oversampling, averaging, etc.) are programmable and are set equally for all three channels. These ADCs are operated continuously (with less than 1% deadtime), but have programmable parameters that include number of integrations per sample and number of bits transmitted after conversion (up to 20). The digital output

from the ADCs is transferred to Xilinx FPGA on a dedicated databus, where it is deserialized and stored in an accumulating register. The averaging cycle of the accumulator can be user defined, as can the number of bits from the ADCs (up to 20) that are used. The accumulator stores its averaged values in a "slow" data stream that fills a FIFO; the mcu begins a read sequence whenever the FIFO reports that it is not empty. There is also a "fast" data stream stores all of the ADC samples in a circular buffer that can store at least 100 ms of data at the maximum sampling rate of 15.6 kHz. The slow data stream divides down the data rate by averaging over N (remotely configurable) samples. The circular buffer is normally not read out during continuous sampling, but can be accessed at the receipt of a remote command. The slow data stream is normally read out in real time during continuous sampling and transmitted via CAN bus through the mcu.

During normal operations, the slow data was acquired at 0.47 Hz, with the ADC integration time 0.250 ms. The ratio of fast/slow data rates was 1024. With this configuration, stable input currents were resolved to better than 20 pA rms, with equivalent accuracy. The SVTRAD modules were polled at 2 Hz so that nearly every response message contained ADC data as well as status information.

The fast history buffer readout was configured with a post-mortem trigger managed by a polling task in EPICS. When the SVTRAD module reported a radiation trip to EPICS (via the routine status message), EPICS responded with a latency of less than 500 ms and issued a series of commands to the module's microcontroller that stopped the ADC digitization and transferred the contents of the fast history buffer to disk at 1600 bits/second over CAN. A manual fast history buffer request was implemented as well. The 480 Hz fast history sampling rate was more than sufficient to accurately reconstruct the 60 Hz fundamental and its harmonics in the noise spectrum. The combination of the fast history buffer and a simple FFT analysis package formed a valuable diagnostic tool to verify the integrity of the module's grounding and shielding scheme.

Radiation protection system

The radiation protection system is based on the Analog Devices AD652-KP Synchronous Voltage-to-Frequency (VFC) converter. The frequency output is fanned to a Frequency-to-Voltage (FVC) converter (to be sent back into the radiation monitoring stream) and passed directly into the Lucent Orca FPGA as a digital pulse train of TTL levels. The converters are so-called synchronous, and therefore have a programmable collected charge to pulse rate ratio, based on a supplied clock speed. (The slower the driving clock speed, the more charge a given output pulse will correspond to.) The VFC operates according to:

$$f_{out} = \frac{I_{in}}{1.0mA} \times f_{clk} \tag{E.3}$$

Thermistors near each diode are voltage biased with -5.0 V, passed through a transimpedance stage, and then sent to a corresponding bank of VFCs. The FPGA uses the thermistor frequency pulse trains for crude corrections to pedestal leakage current due to temperature changes. The diodes have a temperature coefficient of 10% / degC, the thermistors -5% / deg C, but because of the biasing scheme, both frequency pulse trains from the VFC are positively correlated with temperature.

The Orca programmable logic device fulfills several functions. Included on each channel of processing is a watchdog timer that monitors the nominal leakage current of each interlocked diode, as well as the current from each included thermistor. Failure to reset the timer results in an interlock alarm; this protects against bias supply failure and breaks in cabling upstream of the SVTRAD board.

The leaky bucket counter within the AbortBrain FPGA pits a programmable clock against the frequency pulse train from the thermistor and the diode (via the VFCs). The counter has three inputs then: an INCREMENT input from the diode VFC, a DECREMENT input from the thermistor VFC, and a programmable INCREMENT/DECREMENT input from the internal timer. Thus, the net frequency into the counter is:

$$f_{net} = \pm f_{timer} + f_{diode} - f_{temp} \tag{E.4}$$

An interlock condition is achieved when the counter reaches its maximum value. (i.e. Whenever f_{net} is positive for sufficient time to overflow the counter.) Note that the total dose required to exceed the threshold depends on the rate at which it is delivered: below a certain rate, the module never trips, but if this threshold is exceeded for long enough, the unit will trip after absorbing a maximum dose that decreases with increasing dose rate. Thus, the user has control over two independent trip parameters that correspond to the instantaneous radiation dose threshold (acute) and average dose rate threshold (chronic).

Upon power up, the Orca FPGA enters a configuration state, which by default sets the radiation alarm. The on-board microcontroller can only address the when it is in the configuration state, during which it can initiate calibration procedures or set channel parameters. Because the uses a programmable depth counter, programmable rate timer, and provides the clock pulses to each VFC, the unit can be calibrated to subtract off pedestal leakage current of the diode, and to be immune to temperatures changes in the diode leakage current.

In configuration mode, the user can initiate a calibration sequence (for diode leakage or thermistor current) which counts the VFC pulses for each channel during a precisely defined amount of time (at present, 16.7 seconds). The final value in the counter (for each channel) then reflects the pedestal current for that device. The user can then select the best set of parameters so that the net count-up input is nearly temperature independent, so that the average state of the counter represents only radiation dose. The user defines the two radiation thresholds (instantaneous and cumulative dose rate) by constraining the relationship between the depth of the counter, the VFC clock signals, and rate (and direction) of the timer. For a pictorial overview of this design, see this image.

Knowledge of the SVT's susceptibility to p-stop shorts came late in the SVTRAD design phase, and it is only a tribute to the system's flexibility that we have been able to include hard protection against the sub-millisecond type of radiation damage. In order to meet this compromise most satisfactorily, an additional timescale of chronic dose threshold was necessary. By requiring the acute dose threshold to be near 100 mRad (100 mR / ms is the SVT's best estimate of the scale of p-stop short formation,

the module's fastest response time is 500 usec), the hysteresis on the chronic dose threshold is small. As successful commissioning of the collider ramped up, the chronic dose threshold was raised to accommodate higher ambient background levels due to stored beam backgrounds. A 100 mRad tolerance threshold became unforgivable. To increase the hysteresis, another tier in the timescale hierarchy of thresholds was added. By artificially raising the chronic dose threshold to, say, 300 mR/s, the front-end circuits sensitivity to short-lived 150-200 mR background fluctuations was reduced. A higher level administrative control was implemented in EPICs whereby a chronic dose rate of 45 mR/s was enforced on the timescale of 5 minutes. When 5 consecutive minutes passed with the average radiation dose rate over the long-term chronic dose rate threshold, an EPICS process tripped the SVTRAD module, forcing a software abort of the hardware interlock system. This "softAbort" system was very effective in decreasing sensitivity to transient ∼few hundred mR background fluctuations while continuing to enforce the SVT's daily radiation budget.

The acute response time of the front-end radiation protection hardware scales linearly with the value of the acute dose threshold such that

$$t_{trip} = \frac{1ms}{1Rad} \tag{E.5}$$

but plateaus around 450 μs. The injection trigger gating window was set to 2.0 ms to allow for some jitter and non-infinitely tall acute dose events.

Control and monitoring

The SVTRAD modules are essentially autonomous in their protection task, aside from updating the configuration and/or calibration. To meet the requirements of dynamic monitoring of radiation levels, a remote client service was implemented using a multi-level custom EPICS package. A poll-response protocol was used on the CAN bus, with EPICS polling regularly for status and recent data and the module pushing messages in response. Within EPICs the CAN messages are stripped down to status contents and ADC data. The ADC samples are converted to measurements of

integrated charge, converted to radiation dose rate by pedestal subtraction and temperature correction, farmed to various display clients, and simultaneously archived to local disk.

The conversion of measured diode leakage current to radiation dose rate is a non-trivial matter, and one of the strengths of the EPICS implementation of the SVTRAD system. The thermistors not used in the radiation protection circuit are handled by another dedicated hardware module, HTEMP. The HTEMP module provides current biased measurements of the thermistor resistances which are then converted to temperature through the manufacturer's reported R-T conversion table. The temperature is instantiated as an EPICS record which then feeds the temperature correction machinery. The entire dose rate conversion package uses an event processing state machine; in the "run' state, incoming ADC samples are assumed to contain a pedestal and signal contribution, and the last know good pedestal is converted to the instantaneous temperature reading, subtracted from the instantaneous diode current sample, converted to dose rate and pushed out. When the state machine is in the "calibration" state, incoming ADC samples are assumed to contained only diode pedestal current. These samples and their temperature are accumulated sequentially to form an increasingly more precise estimate of the diode's true pedestal leakage current. At 0.47 Hz sampling, 10-15 points (take over the course of 20-30 seconds) were entirely sufficient to determine the diode pedestals to better than 0.25 nA precision, and converge the diode dose rates to 0 mR/s with an rms resolution of 1 mR/s.

Because of the rate of radiation exposure by PEP-II during high luminosity running (often 1-3 kRad / day), the diode pedestal currents increment several nA per day. In order to re-measure the pedestals and stay ahead of drift, an automatic diode dose rate calibration system was implemented. Taking advantage of the widespread EPICS support in *BABAR*, an automatic EPICS task was created that monitored the state of the accelerator. when sufficient conditions (typically beam currents and associated warning indicators) indicated that PEP-II was in a stable, no beam state, the task automatically denied PEP-II injection, changed the event processing state machine to "calibration" for 35 seconds, and then cleared and exited.

Additional radiation protection algorithms are implemented in the online software

as well. For instance, operator alarms can be triggered when the integrated radiation dose for a sensor exceeds a daily, or hourly maximum. These "administrative" alarms provide valuable feedback to the accelerator team, and advises their policy in short-term planning. Recently, a new algorithm has been included that monitors radiation dose absorbed during any particular PEP-II injection situation and, based on comparisons to established limits on acceptable radiation exposure during injection, can either inhibit the rate of injection, or abort the beams. Because radiation policy has a direct impact on experiment up-time (i.e. aborting the beams costs integrated luminosity), sophisticated algorithms have been necessary to carefully optimize the balance between radiation exposure, risk, and the need for continuous up-time. The EPICS control system has generously supported these efforts.

The combination of EPICS and the host UNIX platform posed an excellent environment in which to develop a streamlined task for calibrating the front-end protection hardware and computing the optimal parameters for a given set of radiation thresholds. A native $C++$ program, a PERL script, and a well-documented EPICS GUI screen were all that were needed to reduce the challenging calibration to a simple button-click. Fig. E.3 shows a sample of the user-interface panels implemented under EPICS.

E.6 Performance

A diode used in the protection circuit is typically referred to as "abort diode" and those used in only the monitoring system are termed "monitoring diodes." At present, the four diodes in the horizontal plane have been configured as abort diodes, and the other 8 TOP and BTM diodes are used only for radiation monitoring.

The SVTRAD system has met or exceeded most design expectations, and has successfully protected the SVT from premature aging and catastrophic damage. It has proven flexible and robust in the face of changing requirements and greater sensitivity to radiation damage that hoped. As indicated in the left-hand portion of Fig. E.4, the system has aborted the beams in PEP-II more than 10,000 times since inception in order to protect the SVT and maintain an integrated dose that is less than the original

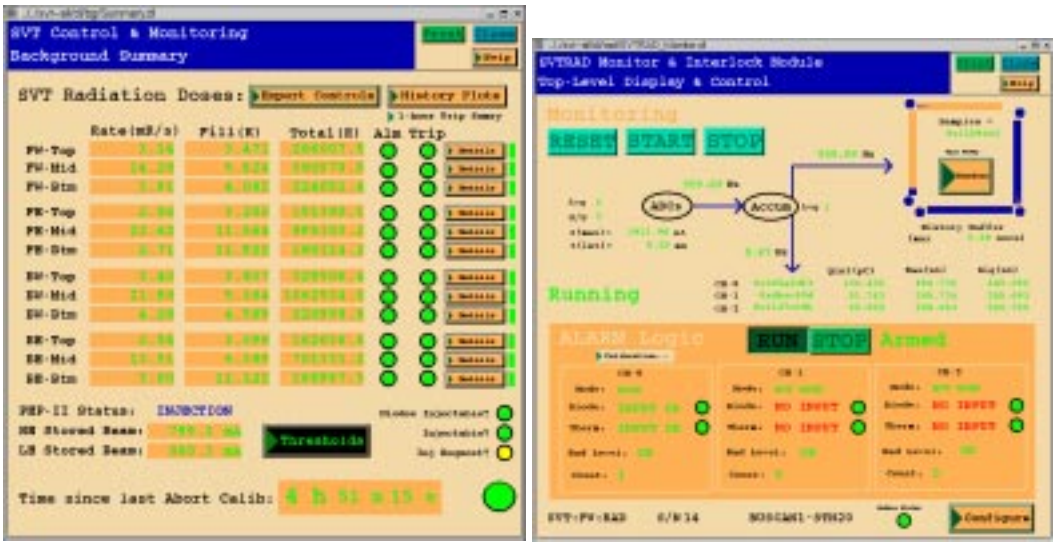


Figure E.3: Sample EPICS screenshots. The left hand screen shows summary information for each diode, including dose rates, integrated dose, alarm status, and trip alarms, and links to more detailed expert controls. The right hand screen shows the more primitive EPICS proxy for the SVTRAD module, whose ADC samples and configuration information are used to form the radiation dose rates in the summary panel on the left.

budget determined in 1999 (solid black line in the right-hand portion of Fig. E.4). The instantaneous dose rates delivered by the system have served as valuable tools in characterizing and remediating backgrounds in the interaction region, and are often used by the accelerator operators to tune the collision parameters to maximize luminosity by minimize radiation damage to *BABAR*.

The redundancy of sensors in the system has proved important, as there is ample evidence that the different locations of the diodes experience different patterns of radiation backgrounds, often from different sources. Typically, diodes in the horizontal measure dose rates 5-8 times larger than those out of the mid-plane, as expected due to beam losses in the machine's bean-plane. A catastrophic failure mode for one diode ("BW:MID") has also been observed, although it is suspected that this is due to mechanical or electrical failure).

The single largest difficulty with the system has been accurate modeling of the

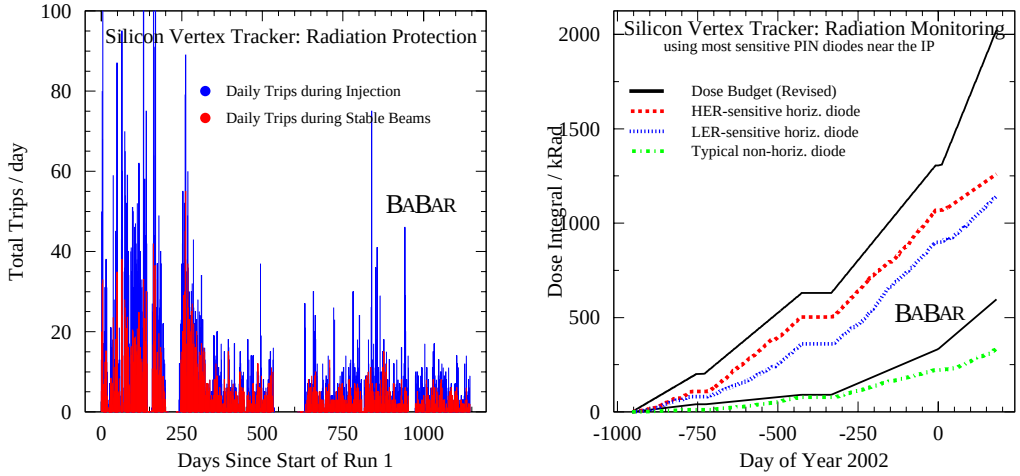


Figure E.4: History of daily trip rate and integrated dose absorbed by the most exposed sensors.

pedestal current for each diode. Calibration of the dose rates requires precise knowledge of the pedestal current in order to perform the sensitive subtraction. Competing effects (resulting from the accumulation and time-dependent annealing of radiation damage) make absolute, precise knowledge of the diode's pedestal current challenging. With radiation exposure exceeding 1 kRad per day in many locations, the pedestal current can be change by as much as 5 nA, which, if unaccounted, would appear as an excess dose rate of 25 mR/s, or a radiation trip threshold reduced by 25 mR/s.

As the requirements of radiation protection for the silicon have evolved, so has the system. Bear in mind that this description is already dated.

Appendix F

Background Remediation Group

The Background Remediation Group was formed in the fall of 1998 under Guy Wormser with a charge to identify, characterize, and remediate the unprecedented high levels of accelerator-related backgrounds at the Interaction Region. The *BABAR* Detector is sensitive to, and vulnerable to, radiation damage, and the Background Remediation Group was created to protect it.

As a member of this team, I participated in both on- and offline analysis tasks. By conducting experiments with the accelerator, such as variations of circulating beam currents, fill patterns, vacuum pump rates, deviations from nominal orbits, and final focus optics, we determined the primary sources of background radiation in the IR-2 region. With the assistance of a detailed simulation and analysis of subdetector response, we developed specific recommendations for the remediation of backgrounds, such as the installations of collimators and "super-low" vacuum pipe sections.

The primary set of conclusions from our work are excerpted below from Ref. [85].

Operationally, the acceptable level of backgrounds at an unprecedented machine like PEP-II is determined primarily by the radiation hardness of the sub-detectors (SVT, EMC) and by the tolerable drift chamber current. The Level-1 (L1) trigger rate and the occupancy in other detectors (DIRC) also constitute occasional limitations. Careful measurement, analysis and simulation of the background sources and of their impact, has led to a detailed understanding and an effective remediation of these effects.

The primary causes of backgrounds in PEP-II are, in order of increasing importance:

- *Synchrotron radiation (SR) generated in the bending magnets and final focussing quadrupoles in the incoming high- and low-energy ring (HER and LER) beam lines.*

Careful layout of the interaction-region area and a conservative SR masking scheme have proven very effective.

- *Two-beam backgrounds from three sources: enhanced beam-gas interactions in the HER, due to low-energy ring IP synchrotron radiation impinging onto the incoming HER beam pipe; photons and low-energy e+- from radiative-Bhabha scattering hitting nearby vacuum components; and tails generated by the beam-beam interaction and/or by the electron-cloud-induced blowup of the low-energy beam.*

- *The interaction of beam particles with residual gas around the rings (beam-gas), which constitutes the primary source of radiation damage, and that with the largest impact on operational efficiency.*

While instantaneous background conditions do vary because of IP orbit drifts, and of the sensitivity of beam tails to small tune adjustments, reproducible patterns have emerged.

The HER beam-gas contribution is typically dominant: the combination of a 40m long straight section, almost devoid of magnetic bending upstream of the final doublet, and of the magnetic beam-separation scheme, conspire to direct abundant bremsstrahlung-induced electromagnetic debris into the IP vacuum pipe. The same process occurs in the LER, but to a lesser extent because of a shorter drift section and lower primary energy. Most BaBar subdetectors, therefore, exhibit occupancy peaks at ϕ=0 and 180 degrees, reflecting the fact that the separation dipoles bend energy-degraded particles in opposite directions. Such local beam-gas interactions dominate the SVT instantaneous dose rates, the total drift chamber current, and the L1 trigger rate. Limited vertexing of L1 pass-through events identifies the beampipe wall and several aperture restrictions within 100cm of the IP as the primary impact points of lost particles. Maintaining a low pressure in the region from 4m to 60m upstream of the IP in the incoming HER beam line, is vital to minimize this background. Scrubbing, which has reduced the HER-averaged dynamic pressure below 50% of its design value, also proved effective.

Both the DIRC and the drift chamber appear sensitive to beam-gas Coulomb scatters around the entire LER. In addition, the DIRC proved particularly vulnerable to tails generated by beam-beam or electron-cloud induced blowup of the low-energy beam. Even though partially eliminated by a set of betatron collimators in the last arc, such tails tend to scrape near the highest-beta point of the final LER doublet, located inside the SOB. This results in photomultiplier counting rates sometimes exceeding 200 kHz; the problem has been alleviated by local lead shielding, and additional collimation will be installed during the fall shutdown.

Whereas trigger-rate and occupancy considerations define acceptable dynamic running conditions, it is the total integrated radiation dose that determines the lifetime of the subdetectors. Despite a significant investment in radiation-hard technology, the innermost layers of the SVT silicon and front-end electronics remain the most susceptible to radiation damage. The accumulated dose has been maintained below budget, through a strict program of hardware interlocks, administrative controls, and real-time monitoring. To date, the most irradiated portion of the SVT has been exposed to approximately 475 kRad, 20 to 30% of which is contributed by injection periods.

The Background Remediation Group transformed into the initial core team of the *BABAR*-PEP-II Liaisons in Summer 1999, the set of personnel who were qualified and authorized to represent *BABAR* in the SLAC Main Control Center on a 24-hour, 7-days a week basis (during data-taking). Liaison shifts began in May, 1999 and were a critical element in the highly efficient detector-accelerator relationship.

For more details about the Background Remediation Group's work, please see Refs. [86] and [87].

Appendix G

High Luminosity Backgrounds Task Force

The High Luminosity Backgrounds Task Force was commissioned in June 2000 under Witold Kozanecki to estimate the *BABAR* background levels that could be expected as the B-factory luminosity increased by an order of magnitude over several years. The particulars of the charge included developing estimates of *BABAR* background levels at luminosities up to 3×10^{34}, estimating the impact of any potential improivements, and identifying any areas of additional improvement.

In order to identify and characterize the various components of the background, * a number of machine physics experiments were carried out in June and July 2000. These consisted primarily in measuring the dependence of background levels in each *BABAR* subdetector on beam currents and on luminosity. In addition, highly valuable data, particularly on so-called vacuum scrubbing rates, were collected parasitically during physics running or extracted from the *BABAR* and PEP-II history buffers. These experimental studies were complemented by existing and new simulations of beam-gas scattering around the PEP-II ring, and of the propagation of beam-gas induced shower in the *BABAR* detector. Finally, a small number of collimation and vacuum pump experiments were also carried out.

The task force completed its mission in October 2000, and the findings were

*Borrowing from the introduction of the published report...

published in Ref. [88]. [†] The overarching conclusion, at the qualitative level, was that extrapolations to higher luminosities indicated no extreme performance limitations, although concerns were noted about the trigger rates and the SVT absorbed radiation dose.

[†]We urge the interested reader to pursue this document not only for its conclusions, but because it serves as an excellent reference for the methods and procedures required for background studies.

Appendix H

Beam Abort Reduction Task Force

The Beam Abort Reduction Task Force was formed in mid-2000 in order to specifically address the issue of overall PEP-II uptime. At that point, *beam aborts* were occurring frequently (a dozen per day) enough to impact the integrated luminosity deliverable to the *BABAR* experiment. Beam aborts were being caused by PEP-II elements (bad orbit alarms, RF cavity trips, key magnet failures) and by the *BABAR* Radiation Protection Systems (SVTRAD, DCH and EMC diodes). Additionally, so-called *trapped events* were studied in detail; in these situations, background levels at *BABAR* would increase by 3 time nominal, and then stay constantly high for several minutes or until the PEP-II operators brought the beams out of collision or aborted the beams.

The task force implemented a beam abort monitoring system which recorded the source and type of each beam abort so that attention could be focused on the most significant sources. The *BABAR* SVTRAD system and the PEP-II RF system were identified as the leading causes. Optimizing the SVTRAD thresholds, in part with new knowledge of the SVT's sensitivity to "fast" radiation damage, significantly reduced the number of detector beam aborts. Tuning the control software and operating parameters of the RF cavities assisted in reducing that source of beam aborts as well. The final activity of the task force was the design and implementation of a dedicated high-speed data acquisition system that monitored two dozen different sensors near the interaction region with sampling rates more than 10 kHz. When

triggered by a beam abort, the system archived precision data from the sensors for up to 60 seconds before the beam abort. Analysis of this data provided key insights into the mechanism of the trapped-events and supported the hypothesis of heavy-element micron-sized dust particles becoming entangled in the electron beam's orbit and causing the elevated backgrounds.

Appendix I

Physics Impact of Beam Backgrounds Task Force

The Physics Impact of Beam Backgrounds Task Force was convened in January 2001 under Steve Robertson to evaluate the effect on *BABAR* data of current and possible future background running conditions. A secondary goal was to investigate the possibility of supporting a long-standing investigation into the effects of machine-related backgrounds on *BABAR* physics.

The task force used simulated performance data as well as samples of so-called *background triggers* taken from real detector running (*i.e.,* events recorded without physics triggers present) to investigate the effect of beam backgrounds on issues such as π^0 mass resolution, SVT vertex resolution and efficiency, physics candidate reconstruction efficiency, and generic track-finding. Ultimately, conclusions were limited. The ability to disentangle beam background and detector noise contributions was poor, and the random trigger background samples from Data quickly became dated as background conditions in the machine changed. For instance, there was no toolset to easily associate short spans of data with background conditions information from which to build a self-consistent "high" vs "low" background comparison. However, the clear presence of machine backgrounds in the physics data sample was established.

The task force's analyses are more completely described in Ref. [89].

Bibliography

[1] H. Albrecht *et al.*, (ARGUS Collaboration), "Observation Of B0 - anti-B0 Mixing," Phys. Lett. **B192**, 245 (1987).

[2] Q. R. Ahmad *et al.*, (SNO Collaboration), "Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory," Phys. Rev. Lett. **89**, 011301 (2002), nucl-ex/0204008.

[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. Theor. Phys. **49**, 652–657 (1973).

[5] K. Hagiwara *et al.*, (Particle Data Group Collaboration), "Review of Particle Physics," Physical Review D **66**, 010001+ (2002).

[6] S. W. Herb *et al.*, "Observation of a Dimuon Reonsance at 9.5-GeV in 400-GeV Proton - Nucllus Collisions," Phys. Rev. Lett. **39**, 252–255 (1977).

[7] M. Herrero, "The Standard Model," June 1998, hep-ph/9812242, Presented at 10th NATO ASI on Techniques and Concepts of High-Energy Physics, St. Croix, U.S. Virgin Islands.

[8] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Addison-Wesley (1995). Reading, USA. 842 p.

[9] D. Griffiths, *Introduction to Elementary Particles*, John Wiley and Sons, Inc. (1987).

[10] W. Pauli, "The connection between spin and statistics," Phys. Rev. **58**, 716–722 (1940).

[11] "The Particle Adventure," (2002), `http://particleadventure.org/particleadventure/`, Educational website.

[12] B. Aubert *et al.*, `(BABAR Collaboration)`, "A study of time dependent CP-violating asymmetries and flavor oscillations in neutral B decays at the Upsilon(4S)," (2002), `hep-ex/0201020`.

[13] J. Albert, *Measurement of CP Violation in $b \to c\bar{c}d$ Exclusive Decays at the BABAR Experiment*, PhD thesis, Princeton University (2002).

[14] 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).

[15] A. Angelopoulos *et al.*, `(CPLEAR Collaboration)`, "First Direct Observation of Time-Reversal Non-Invariance in the Neutral Kaon System," Phys. Lett. **B444**, 43–51 (1998).

[16] S. Weinberg, "A Model Of Leptons," Phys. Rev. Lett. **19**, 1264–1266 (1967).

[17] V. D. Barger and R. J. Phillips, *Collider Physics*, Addison-Wesley (1987).

[18] D. H. Perkins, *Introduction to High Energy Physics*, Addison-Wesley (1987).

[19] L. Wolfenstein, "Parameterization of the Kobayashi-Maskawa Matrix," Phys. Rev. Lett. **51**, 1945 (1983).

[20] I. I. Y. Bigi, "Lifetimes of Heavy-flavour Hadrons: Whence and Whither?" Nuovo Cim. **109A**, 713–726 (1996), `hep-ph/9507364`.

[21] P. J. Oddone, "A Conceptual Design For a High Luminosity Linear Collider B Anti-B Factory," Proceedings, UCLA Workshop page 423 (1987).

[22] A. S. Dighe, T. Hurth, C. S. Kim, and T. Yoshikawa, "Measurement of the Lifetime Difference of B_d mesons: Possible and Worthwhile?" Nucl. Phys. **B624**, 377–404 (2002), `hep-ph/0109088`.

[23] H. R. Quinn *et al.*, "The *BABAR* Physics Book," Technical Report (1998). SLAC-R-504.

[24] D. Kirkby and R. Cahn, `(BABAR Collaboration)`, "A General Model for Neutral B Decay Time Distributions," BaBar Analysis Document **188** (2002).

[25] T. O. Meyer, `(CLEO Collaboration)`, "CLEO Measurements of the CKM Elements V_{ub} and V_{cb}," (2001), Presented at the 2001 Lake Louise Winter Institute, Lake Louise, Alberta, Canada.

[26] D. Groom *et al.*, `(Particle Data Group Collaboration)`, "Review of Particle Physics," Eur. Phys. Jour. **C15**, 1+ (2000).

[27] M. Neubert, "A fresh look at the B semileptonic branching ratio and beauty lifetimes," (1996), `hep-ph/9605256`.

[28] L. Lellouch and C. J. D. Lin, `(UKQCD Collaboration)`, "Standard model matrix elements for neutral B meson mixing and associated decay constants," Phys. Rev. **D64**, 094501 (2001), `hep-ph/0011086`, Revised Feb 2002.

[29] T. Inami and C. S. Lim, "Effects Of Superheavy Quarks and Leptons In Low-Energy Weak Processes K(L) to mu anti-mu, K+ to pi+ neutrino anti-neutrino and K0 — anti-K0," Prog. Theor. Phys. **65**, 297 (1981).

[30] W. McDonough, "Private communication," (2001).

[31] D. Kirkby *et al.*, `(BABAR Collaboration)`, "Selection of D^* ℓ Events for B Lifetime and Mixing Analyses," BaBar Analysis Document **034** (2001).

[32] D. R. Perticone, `(CLEO Collaboration)`, "Measurements of B0 anti-B0 Mixing and the Lifetime Ratio tau(B+)/tau(B0) with the CLEO II Detector," (1992), Prepared for Particles & Fields 92: 7th Meeting of the Division of Particles Fields of the APS (DPF 92), Batavia, Illinois, 10-14 Nov 1992.

[33] E. Fernandez *et al.*, (MAC Collaboration), "Lifetime of Particles Containing b Quarks," Phys. Rev. Lett. **51**, 1022 (1983).

[34] N. Lockyer *et al.*, (Mark II Collaboration), "Meausrement of the Lifetime of Bottom Hadrons," Phys. Rev. Lett. **51**, 1316 (1983).

[35] C. Albajar *et al.*, (UA1 Collaboration), "Search For B^0 $\overline{B}^0$ Oscillations at the CERN proton-antiproton Collider," Phys. Lett. **B186**, 247–254 (1991), [Erratum: *ibid* B197, 565 (1987)].

[36] B. H. Behrens *et al.*, (CLEO Collaboration), "Precise measurement of B0 anti-B0 mixing parameters at the Upsilon(4S)," Phys. Lett. **B490**, 36–44 (2000), hep-ex/0005013.

[37] G. Abbiendi *et al.*, (OPAL Collaboration), "Measurement of the B^0 Lifetime and Oscillation Frequency Using $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}$ Decays," Phys. Lett. **B493**, 266–280 (2000), hep-ex/0010013.

[38] "PEP-II: An Asymmetric B Factory, Conceptual Design Report," Technical Report (June 1993). SLAC-418, LBL-5379.

[39] M. Sullivan, "Private communication," (1998).

[40] B. Aubert *et al.*, (BABAR Collaboration), "The BaBar Detector," Nucl. Instrum. Meth. **A479**, 1–116 (2002), hep-ex/0105044.

[41] P. Burchat, C.-H. Cheng, D. Kirkby, and T. I. Meyer, (BABAR Collaboration), "Semileptonic B Sample for Lifetime and Mixing Analyses," BaBar Analysis Document **147** (2002).

[42] P. Billoir, "Track Fitting with Multiple Scattering: A New Method," Nucl. Instr. Meth. **A225**, 352 (1984).

[43] U. Langenegger, (BABAR Collaboration), "Cut-based Electron Identification," BaBar Analysis Document **90** (2001).

[44] F. Fabozzi and L. Lista, (BABAR Collaboration), "Muon Identification in the BaBar Experiment," BaBar Analysis Document **60** (2000).

[45] T. I. Meyer, "Comparing $D^*\ell$ Vertexing Algorithms," November 2001, http://www.slac.stanford.edu/~meyertim/analysis/Dstarlnu/recoVtxFitting.html, Presentation at *BABAR* AWG meeting.

[46] F. Martinez-Vidal *et al.*, (BABAR Collaboration), "The *BABAR* Vertexing," BaBar Analysis Document **102** (2001).

[47] C.-H. Cheng, C. LeClerc, D. Kirkby, and T. I. Meyer, (BABAR Collaboration), "Measurement of the B^0 Lifetime and Oscillation Frequency Using Exclusively Reconstructed Semileptonic Decays," BaBar Analysis Document **137** (2002).

[48] P. Burchar *et al.*, (BABAR Collaboration), "Supporting Document For the Run 1 sin2β Analysis," BaBar Analysis Document **115** (2001).

[49] S. Prell *et al.*, (BABAR Collaboration), "Measurement of $B^0\overline{B}^0$ Mixing With Fully Reconstructed Hadronic Decays," BaBar Analysis Document **125** (2001).

[50] B. Aubert *et al.*, (BABAR Collaboration), "Measurement of the B0 and B+ meson lifetimes with fully reconstructed hadronic final states," Phys. Rev. Lett. **87**, 201803 (2001), hep-ex/0107019.

[51] D. Kirkby, "Private communication," (2000).

[52] J. Beringer *et al.*, (BABAR Collaboration), "B Tagging in *BABAR*: Status for sin2beta Winter Conference Results," BaBar Analysis Document **119** (2002).

[53] 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).

[54] C.-H. Cheng, *A Measurement of the Lifetime and Mixing Frequency of Neutral B Mesons with Semileptonic Decays With the BABAR Detector*, PhD thesis, Stanford University (2002).

[55] C.-H. Cheng and P. Burchat, (BABAR Collaboration), "Correlation Between Mistag of Kaon Tag and Calculated Error on Δt," BaBar Analysis Document **349** (2002).

[56] C. Touramanis *et al.*, (BABAR Collaboration), "Reconstruction of π^0 in *BABAR*," BaBar Analysis Document **020** (2000).

[57] P. L. Frabetti *et al.*, (E687 Collaboration), "Analysis of three D $\rightarrow$ K pi pi Dalitz plots," Phys. Lett. **B331**, 217–226 (1994).

[58] T. O. Meyer, (CLEO Collaboration), "Dalitz Analysis of $D^0 \rightarrow K^0_S\pi^+\pi^-$," (2002), Presented at the 2002 April Meeting of the American Physical Society, Alburquerque, New Mexico.

[59] G. Hamel de Monchenault *et al.*, (BABAR Collaboration), "Recommendations from the Blind Analysis Group (BAG)," BaBar Analysis Document **91** (2000).

[60] W. Verkerke and D. Kirkby, "RooFit Toolkit for Data Modelling," (2002), http://www.slac.stanford.edu/BFROOT/www/Computing/Offline/ROOT/RooFit/.

[61] F. James *et al.*, "MINUIT: Function Minimization and Error Analysis," (1998), http://wwwinfo.cern.ch/asdoc/minuit/, CERN Program Library Long Writeup D506.

[62] R. Brun *et al.*, "ROOT: An Object-Oriented Data Analysis Framework," March 2002, http://root.cern.ch/root/, User's Guide v3.02b.

[63] W. Dunwoodie *et al.*, "Radial and Longitudinal Length Scale From Beampipe Structure," December 2000, http://www.slac.stanford.edu/~wmd/beampipe/dec00.talk, Presented at the Dec. 2000 *BABAR* Collaboration Meeting.

[64] B. Aubert *et al.*, (BABAR Collaboration), "Measurement of B0 - anti-B0 flavor oscillations in hadronic B0 decays," Phys. Rev. Lett. **88**, 221802 (2002), hep-ex/0112044.

[65] B. Aubert *et al.*, (BABAR Collaboration), "Measurement of the B0 - anti-B0 oscillation frequency with inclusive dilepton events," Phys. Rev. Lett. **88**, 221803 (2002), hep-ex/0112045.

[66] 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.

[67] B. Aubert *et al.*, (BABAR Collaboration), "Measurement of the B0 lifetime with partially reconstructed $B0 \to D* -l + \nu_l$ decays," Phys. Rev. Lett. **89**, 011802 (2002), hep-ex/0202005.

[68] 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.

[69] M. A. Mazzoni, "Systematics Associated with the B Lifetimes," February 2002, http://www.slac.stanford.edu/BFROOT/www/Physics/Analysis/AWG/BBMixingHadr/Meetings/14Feb2002/Feb14_02.pdf, Presentation at *BABAR* Lifetime & Mixing AWG Meeting.

[70] J. Beringer, (BABAR Collaboration), "BTagger - A Multivariate Tagging Algorithm with Categories Based on the Physics of the $B_{\rm tag}$ Decay," BaBar Analysis Document **317** (2002).

[71] C. LeClerc *et al.*, "Uncorrelated Lepton Study," October 2001, http://www.slac.stanford.edu/~cleclerc/plots/flip_lepton_sample.html, Presentation at *BABAR* Lifetime & Mixing AWG meeting.

[72] D. Kirkby, "Sample Overlaps for Lifetime & Mixing Measurements," October 2000, http://www.slac.stanford.edu/BFROOT/www/Physics/Analysis/AWG/BBMixingHadr/Meetings/26Oct2000/SampleOverlaps.html, Presentation at *BABAR* AWG meeting.

[73] A. Roodman *et al.*, (BABAR Collaboration), "*BABAR* Particle Identification Group," http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/pid.html.

[74] B. Aubert *et al.*, (BABAR Collaboration), "Search for T and CP violation in B0 - anti-B0 mixing with inclusive dilepton events," Phys. Rev. Lett. **88**, 231801 (2002), hep-ex/0202041.

[75] F. Martinez-Vidal and F. Sandrelli, (BABAR Collaboration), "Measurement of dGamma/Gamma and Study of CPT and T Violation Using CP and Flavor Eigenstates," BaBar Analysis Document **436** (2002).

[76] R. Cahn, (BABAR Collaboration), "Some Things About $\Delta\Gamma$," (2001), http://www.slac.stanford.edu/~cahn/internal/delta_Gamma.ps.

[77] T. Kittleman, "Measuring $\Delta\Gamma$ with *CP* Events," December 2001, http://www.slac.stanford.edu/BFROOT/www/Physics/Analysis/AWG/BBMixingHadr/Workshop/session4/kittel.pdf, Presentation at *BABAR* Lifetime & Mixing Workshop.

[78] (LEP B Working Groups Collaboration) http://lepbosc.web.cern.ch/LEPBOSC/combined_results/amsterdam_2002/.

[79] B. Aubert *et al.*, (BABAR Collaboration), "Simultaneous Measurement of B^0 Meson Lifetime and Mixing Frequency with $B^0 \to D^{*-}\ell^+\overline{\nu}_\ell$ Decays," (2002), hep-ex/0207071, Contributed to the 31^{st} International Conference on High Energy Physics, Amsterdam, The Netherlands.

[80] C. Roat, "Nano/Micro Database Summary," (2001), http://www.slac.stanford.edu/BFROOT/www/doc/workbook/nanomicro/v8.8/.

[81] T. I. Meyer, (BABAR Collaboration), "Radiation Monitoring and Protection in the BaBar Silicon Vertex Tracker : The SVTRAD System," (2001), http://www.slac.stanford.edu/BFROOT/www/Detector/SVT/Operations/SVTRAD/index.html.

[82] T. I. Meyer, (Babar Collaboration), "PIN Photodiodes for Radiation Monitoring and Protection in the *BABAR* Silicon Vertex Tracker," Int. J. Mod. Phys. **A16S1C**, 1084–1086 (2001), hep-ex/0010058, Presented at the DPF2000 Conference, Columbus, Ohio.

[83] T. I. Meyer *et al.*, (BABAR Collaboration), "SVT Radiation Protection: Upgrade Requirements & Scenarios," BaBar Analysis Document **339** (2001).

[84] T. I. Meyer, (BABAR Collaboration), "The *BABAR* Radiation Protection System," (2001), http://www.slac.stanford.edu/BFROOT/www/Detector/SVT/Operations/SVTRAD/babar_radiation_protection.html.

[85] U. Wienands *et al.*, "PEP-II: An Asymmetric Energy B-factory at SLAC," in preparation.

[86] G. Wormser, (BABAR Collaboration), "*BABAR* Background Remediation Home Page," (2000), http://www.slac.stanford.edu/BFROOT/www/Detector/Backgrounds/BackgroundsOldPage.html.

[87] T. I. Meyer and T. Schietinger, (BABAR Collaboration), "The *BABAR*-PEP-II Liaison Shift Manual," (2001), http://www.slac.stanford.edu/BFROOT/www/Detector/Backgrounds/liaisons/manual/index.html.

[88] W. Kozanecki *et al.*, (BABAR Collaboration), "Report of the High Luminosity Backgrounds Task Force," BaBar Note **522** (2000).

[89] T. I. Meyer *et al.*, (BABAR Collaboration), "Report of the Physics Impact of Beam Backgrounds Task Force," BaBar Analysis Document **246** (2001).

Index