$B^0\overline{B}^0$ Mixing and CPT Violation with the Belle Experiment

Christos Leonidopoulos

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE

BY THE DEPARTMENT OF PHYSICS

NOVEMBER 2000

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

Daniel R. Marlow
(Principal Adviser)

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

Eric J. Prebys

Approved for the Princeton University Graduate School:

Dean of the Graduate School

Abstract

We have measured the mixing parameter Δm and searched for CPT violation in the B^0 mixing from the time evolution of dilepton events on the $\Upsilon(4S)$ resonance. The study was performed on a sample corresponding to an integrated luminosity of 5.1 fb^{-1} of data collected with the Belle detector between January and July of 2000.

If we invoke the CPT symmetry we measure

$$\Delta m = 0.456 \pm 0.008 \text{ (stat)} \ \pm 0.018 \text{ (sys) ps}^{-1}$$

This is the first direct measurement of Δm from a time-dependent analysis on the $\Upsilon(4S)$ resonance.

If we fit simultaneously for Δm and the CPT-violating parameter $\cos\theta$ we measure

$$\Delta m = 0.456 \pm 0.009 \text{ (stat)} \ \pm 0.019 \text{ (sys) ps}^{-1}$$

$$\text{Re}(\cos\theta) = 0.00 \pm 0.21 \text{ (stat)} \ \pm 0.30 \text{ (sys)}$$

$$\text{Im}(\cos\theta) = 0.019 \pm 0.031 \text{ (stat)} \ \pm 0.061 \text{ (sys)}$$

The $\cos\theta$ measurement corresponds to the following mass and lifetime differences between B^0 and $\bar{B}^0$:

$$\frac{|m_{B^0} - m_{\bar{B}^0}|}{m_{B^0}} < 4.1 \times 10^{-14}, \text{ at 95\% C.L.}$$

$$\frac{\Gamma_{B^0} - \Gamma_{\bar{B}^0}}{\Gamma_{B^0}} = [-2.9 \pm 4.7 \text{ (stat)} \ \pm 9.2 \text{ (sys)}]\%$$

The results are consistent with CPT conservation.

This is the first measurement on a search for a mass difference between B^0 and $\bar{B}^0$.

Acknowledgments

It has been a long and stimulating journey, and a very valuable personal experience.

I feel fortunate to have worked with and for the Belle collaboration. The completion of this thesis was made possible through the tremendous efforts of many people that I have never met in person. That saves an awful lot of space here.

I would like to thank my advisor, Dan Marlow, for his patience and the discrete guidance; the faculty of the Physics Department of Princeton University for finding the time to discuss my questions; Boris Kayser for helping me see the magic of the B phenomenology; Abe-san, Kay Kinoshita and Sakai-san for the fruitful conversations we had at KEK; and Hanagaki-san for showing me how to become a "san".

Finally, I would like to thank Antonis, Eva and Keir, for being there.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Since the discovery of the $\Upsilon(4S)$ resonance in 1977 [1], considerable effort has been devoted to the study of B mesons. Even before that discovery, physicists had realized that the existence of the 3rd generation b-quark in the B mesons could open the door to large CP violating effects [2]. It was quickly realized that B pairs could be obtained in large numbers through resonant production in e^+e^- collisions via the reaction $e^+e^- \to \Upsilon(4S) \to B\overline{B}$. With a generous source of B pairs available, the field has since become very active, and progress has been steadily achieved.

Even though many measurements from very successful experiments, such as ARGUS at DORIS and CLEO at CESR, have been carried out, until recently, two obstacles have limited progress. First, the most interesting CP channels have small branching ratios ($\sim O(10^{-4})$), and therefore a very high luminosity e^+e^- machine (10^{33} - 10^{34} cm^{-2}s^{-1}) is needed. Second, because B pairs are born in an antisymmetric state, most of the time-integrated CP asymmetries (otherwise, expected to be large) vanish. Since $\Upsilon(4S)$s have barely enough energy to produce B pairs, B mesons at the $\Upsilon(4S)$ rest frame do not typically travel measurable distances. Therefore, their time evolution during their lifetime of $\sim$ 1.6 ps cannot be observed. This difficulty was circumvented by P. Oddone's ingenious idea [3]: at a collider with unequal e^+ and e^- energies (also called an asymmetric e^+e^- collider), the boosted

B mesons will travel distances large enough to measure, and therefore the time dependent asymmetries could be extracted from their decay vertex reconstruction. It is only today that we are technologically advanced enough to build high luminosity asymmetric e^+e^- colliders and have finally overcome these two obstacles. We belong to the fortunate first generation of scientists, researchers and scholars to have access to such rich samples of B mesons with the time evolution information.

The high energy physics community has studied effects related to the mixing of neutral kaons for four decades. The phenomenology of neutral B mesons is similar to that for kaons but there are substantial differences between the two systems that could result in interesting, yet thus far unexplored, physics. The key points that are the foundation of mixing in the B system are:

- The B lifetime is much larger than what one would naively expect from examining other heavy mesons, such as the D mesons. Assuming a m^{-5} dependence for the lifetime[1], we would expect $\tau_B/\tau_D = (m_D/m_B)^5 \sim 5 \times 10^{-3}$. B and D mesons, however, have comparable lifetimes, on the order of one ps. This is a direct consequence of the CKM matrix suppression. The b-quark, being much lighter than the t-quark, has no other choice than to decay to quarks from another generation, typically c-quarks. The small value of $|V_{cb}| \sim 0.04$ pushes the lifetimes for the B mesons (charged and neutral) to relatively large values. Since b-quarks are much heavier than the quarks from the first two generations, when a B decays, it has a large selection of decay modes to choose from. So, the picture for the two eigenstates is the following: many decay channels and small branching fractions. The existence of many (Cabibbo suppressed) modes to which both B^0 and $\bar{B}^0$ can decay, leads to a very small lifetime difference for B_H and B_L, on the order of 1%, or less [4]. This is very different from the kaon system, where the phase-space for $K_S \to 2\pi$ and $K_L \to 3\pi$ differ significantly, producing a dramatic lifetime difference ($\tau_{K_L}/\tau_{K_S} \sim 500$).

[1] This is guessed by the muon lifetime formula: $\tau_\mu = 192\pi^3/G_F^2\, m_\mu^5$.

- According to Standard Model predictions, large CP violating effects are expected in the neutral B meson system. Although these are somewhat beyond the scope of this dissertation, a brief comment is in order. With oscillation rates comparable to their decay rates, there are points in the time evolution of neutral B mesons where B^0 and $\overline{B}^0$ contribute equally to the overall amplitude for decay to CP eigenstates. Indeed such interference effects are thought to be the most promising way to study CP violation since poorly understood strong-interaction effects do not enter. To take it one step further, CP violation in the B system *depends* on mixing. B mixing is therefore a crucial element in the study of CP violation. In view of its important role, it is essential to acquire a firm theoretical and experimental understanding of neutral-B mixing.

  The approximate equality between the oscillation period in the neutral B system and the B^0 lifetime is a "coincidence": it just so happens that the small value of V_{cb} and the heaviness of the t-quark conspire to produce a mixing parameter $x_d \equiv \Delta m/\Gamma$ of order unity.

- In the B mesons the width difference of the physical states is expected to be very small, when compared to the mass difference:

$$\Delta m \gg \Delta\Gamma$$

  It is, therefore, mainly Δm that controls the oscillation. Δm is a tiny mass difference, thirteen orders of magnitude smaller than the B mass. This interferometry makes the neutral B system very sensitive compared to, e.g. the charged B mesons, and therefore miniscule effects can, in principle, be detected. This is the reason that people believe that the B^0-$\overline{B}^0$ mixing is not just the place to look for large CP violating effects, but also a powerful probe that will allow us to look for physics beyond the Standard Model.

In this analysis, we focus on tests of CPT symmetry. A straightforward violation of the CPT conservation theorem would be mass or lifetime differences between a particle and its antiparticle: here B^0 and $\bar{B}^0$. Δm sets the scale for the sensitivity in the measurement of a hypothetical mass difference between the two particles: $\sim O(10^{-14}\ \mathrm{GeV}/c^2)$. On the other hand, since $\Delta m \sim \Gamma$, the sensitivity on the B^0-$\bar{B}^0$ fractional lifetime difference measurement is not as large.

There are several theoretical models that suggest the breaking of CPT symmetry [6]. In some cases, CPT violation arises from terms in the Lagrangian whose coupling strength is quark dependent. Therefore, it is important to make separate tests for CPT invariance in neutral meson systems with different quark content (B, K).

In this dissertation we limit ourselves to the phenomenology[2] and parameterization of a CPT-violating signal (The basic equations are given in Ch. 2; more details can be found in Appendix, Sec. A.1). The phenomenology used here does not *explain* a hypothetical CPT violation signal. Assuming that CPT is not a good symmetry, we use the phenomenology to parameterize the deviation from the Standard Model, so that we can measure it. Ch. 3 gives an overview of the asymmetric "B-factory" KEKB, the Belle detector and the software used. The complete analysis is described in Ch. 4. We discuss the findings in Ch. 5.

[2]For a suggestive list of phenomenological studies on ways of testing a CPT violation hypothesis see Ref. [7] - [9].

Chapter 2

Physics and Formalism: Overview

Neutral B mesons resemble neutral K mesons in that they share the property of having mass eigenstates that differ from their flavor eigenstates. Since we understand the production and decay processes of the B mesons in terms of quark content (flavor eigenstates) but the time evolution in terms of mass and lifetime (mass eigenstates), we need the transformation that leads from one language to the other. The phenomenology that has been developed predicts the mixing between the two flavor states and parameterizes a possible CPT violating signal. Here we only state the essential results: details can be found in Appendix (Section A.1).

2.1 Mixing in the B system

The phenomenon of neutral particle-antiparticle oscillation is very general. Just like kaons, D mesons, or even neutrinos, B^0 and $\overline{B}^0$ mesons evolve with time and mix into each other. In the Standard Model, the mixing is described with a weak interaction involving two W bosons. If we write down the terms that contribute to this process[1]

$$\sum_{i=u,c,t} \frac{V_{id} V_{ib}^*}{(q^2 - m_i^2)} \simeq \sum_{i=u,c,t} \frac{V_{id} V_{ib}^*}{q^2} \left[1 + \frac{m_i^2}{q^2} + \cdots\right] \tag{2.1}$$

[1]The approximation corresponds to the dominant (logarithmically diverging) part of the interaction.

we see that the first term vanishes due to the unitarity condition

$$\sum_i V_{id} V_{ib}^* = 0 \tag{2.2}$$

and therefore the interaction is proportional to the square of the mass of the quark involved. In the limit that all quarks are the same, there is no mixing:

$$\sum_{i=u,c,t} \frac{V_{id} V_{ib}^*}{(q^2 - m_i^2)} \longrightarrow \frac{1}{(q^2 - m^2)} \sum_{i=u,c,t} V_{id} V_{ib}^* = 0 \tag{2.3}$$

Since

$$m_t^2 \gg m_c^2 \gg m_u^2$$

the t-diagram is the main contribution to the mixing (Fig. 2.1). The calculation of the Feynman diagrams gives for the mixing parameter x_d [3]:

$$x_d \equiv \frac{\Delta m}{\Gamma} \simeq \frac{G_F^2 F_B^2 m_B m_W^2}{\Gamma} \cdot \frac{\eta_t E(x_t) |V_{tb}^* V_{td}|^2 B_B}{6\pi^2} \tag{2.4}$$

where G_F is the Fermi constant, m_B, m_W and m_t are the B^0, W and t-quark masses, $x_t \equiv m_t^2/m_W^2$, F_B is the decay constant, B_B the bag parameter, V_{ij} the CKM matrix elements and $\eta_t \sim 0.55$ the QCD correction. The function $E(x_t)$ is given by

$$E(x_t) = \frac{4x_t - 11x_t^2 + x_t^3}{4(1-x_t)^2} - \frac{3x_t^3 \ln x_t}{2(1-x_t)^3} \tag{2.5}$$

The mixing parameter $x_d = 0.723 \pm 0.035$ [5] is much larger than what physicists initially expected. Theoretical prejudice that the t-quark should be relatively light led to early "discoveries" of a ~ 40 GeV/c^2 top quark[2], which forced theorists to think of a B system without oscillations, which originally led experimentalists into "measuring" a B mixing consistent with zero. The large mass of the t-quark is the main reason for the large oscillation rate, comparable to the B decay width. ARGUS soon did see the the mixing, LEP showed that the top was heavy and CDF eventually found it.

[2]The mass of the t-quark is 173.8±5.2 GeV/c^2 [5].

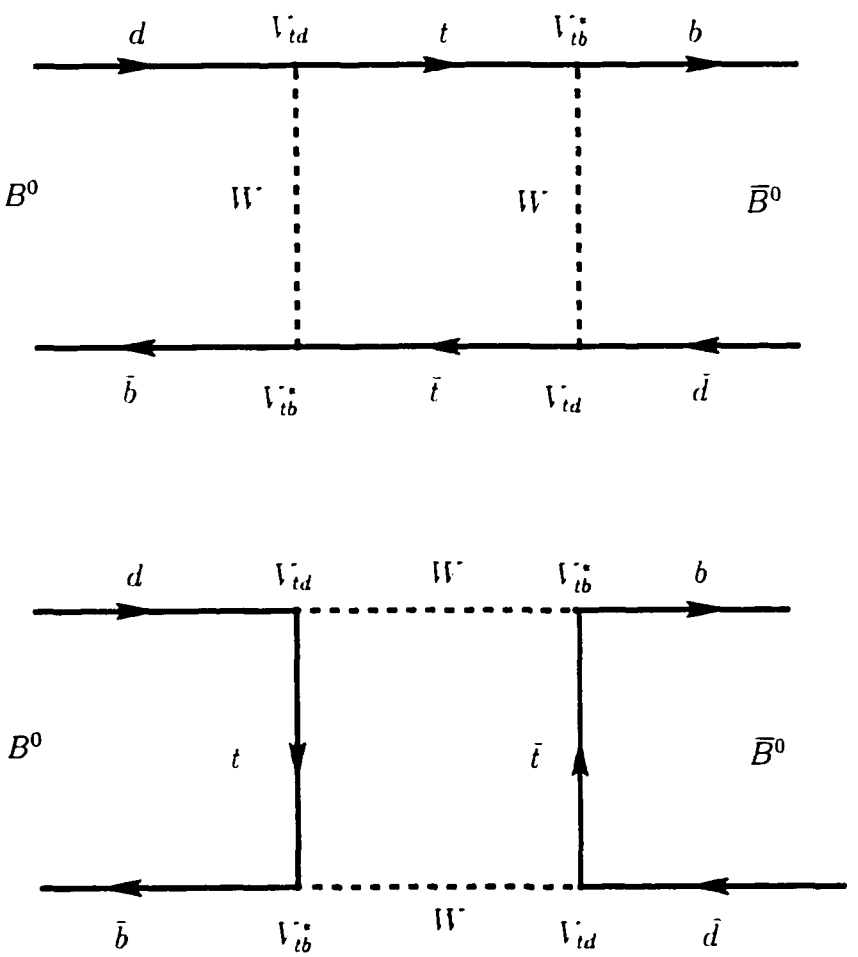

Figure 2.1: Feynman diagrams for the B^0-$\overline{B}^0$ mixing: the t-quark is the main contribution to the oscillation.

The Δm measurement is considered to be one of particular value because it leads to the determination of the $|V_{td}|$ CKM matrix element. Currently the uncertainty in the estimation of $|V_{td}|$ is limited by the Δm ($\sim 4\%$) and m_t (3%) measurements, but is especially hampered by the theoretical uncertainty in the calculation of $F_B\sqrt{B_B}$ (15-20%). As mentioned in Ch. 1, the B^0-$\bar{B}^0$ mixing is an essential ingredient of the CP violation process. The B^0 decay rates are comparable to the oscillation rate, so there are both B^0 and $\bar{B}^0$ contributions to CP final states. A precise Δm measurement is therefore necessary for the determination of the size of CP violation in the B meson system.

Apart from the QCD uncertainties, B-mixing is thought to be well understood in the context of the Standard Model, since the Feynman diagrams shown in Fig. 2.1 can be reliably calculated. Thus if we see a deviation from what is expected, it is likely brought about by "new physics," not included in the Standard Model. In this dissertation we consider CPT violation as a possible signature of such a deviation. The next section describes how such a CPT-violating signal can be parameterized.

2.2 The CPT violation parameter $\cos\theta$

The most general effective Hamiltonian that governs the B mixing can be described with 4 complex parameters. In the $B^0 - \bar{B}^0$ basis, the representation of the Hamiltonian is [7, 8, 9]

$$H = \begin{pmatrix} E\cos\theta - iD & E\sin\theta e^{-i\phi} \\ E\sin\theta e^{i\phi} & -E\cos\theta - iD \end{pmatrix} \tag{2.6}$$

The parameterization of CPT violation arises from the remark that CPT invariance requires that the diagonal elements should be equal

$$\langle B^0|H|B^0\rangle = \langle\bar{B}^0|H|\bar{B}^0\rangle$$

or $\cos\theta = 0$[3]. The parameter ϕ is related to CP violation in the $B^0 \leftrightarrow \bar{B}^0$ mixing: CP invariance would require $|\langle\bar{B}^0|H|B^0\rangle| = |\langle B^0|H|\bar{B}^0\rangle|$, or $\mathrm{Im}\phi = 0$.

A CPT-violating $\cos\theta \neq 0$ corresponds to differences in the mass and the lifetime of a particle and those of its antiparticle: B^0 and $\bar{B}^0$. By using the good approximation[4] [4]

$$\frac{\Delta\Gamma}{\Gamma} \simeq 0 \tag{2.7}$$

we obtain

$$\mathrm{Re}(\cos\theta) \simeq \frac{\Delta m_0}{\Delta m}, \qquad \mathrm{Im}(\cos\theta) \simeq \frac{-\Delta\Gamma_0}{2\,\Delta m} \tag{2.8}$$

or

$$\frac{m_{B^0} - m_{\bar{B}^0}}{m_{B^0}} \simeq \frac{\Delta m}{m_{B^0}} \times \mathrm{Re}(\cos\theta) \quad \text{and} \quad \frac{\Gamma_{B^0} - \Gamma_{\bar{B}^0}}{\Gamma_{B^0}} \simeq -\frac{2\,\Delta m}{\Gamma_{B^0}} \times \mathrm{Im}(\cos\theta) \tag{2.9}$$

where $\Delta m_0 \equiv m_{B^0} - m_{\bar{B}^0}$, $\Delta\Gamma_0 \equiv \Gamma_{B^0} - \Gamma_{\bar{B}^0}$ the mass and width differences between B^0 and $\bar{B}^0$, and $\Delta m \equiv m_H - m_L$ the mass difference of the eigenvalues of the Hamiltonian

$$\lambda_{H,L} = \pm E - iD \equiv m_{H,L} - \frac{i}{2}\Gamma_{H,L} \tag{2.10}$$

with $m_{H,L}$ and $\Gamma_{H,L}$ defined to be real.

Using the world average values for $\Delta m = 3.1 \times 10^{-13}$ GeV/c^2, $m_{B^0} = 5.28$ GeV/c^2 and $\Gamma_{B^0} = 4.1 \times 10^{-13}$ GeV/$\hbar$ [5] we obtain

$$\frac{m_{B^0} - m_{\bar{B}^0}}{m_{B^0}} \sim 5.8 \times 10^{-14} \times \mathrm{Re}(\cos\theta) \quad \text{and} \quad \frac{\Gamma_{B^0} - \Gamma_{\bar{B}^0}}{\Gamma_{B^0}} \sim -1.5 \times \mathrm{Im}(\cos\theta) \tag{2.11}$$

So, a non-zero $\mathrm{Re}(\cos\theta)$ corresponds to a mass difference between B^0 and $\bar{B}^0$, whereas a non-zero $\mathrm{Im}(\cos\theta)$ corresponds to a lifetime difference between the two flavor states. $\mathrm{Im}(\cos\theta)$ is practically equivalent to $|\Gamma_{B^0} - \Gamma_{\bar{B}^0}|/\Gamma_{B^0}$, however, $\mathrm{Re}(\cos\theta)$ probes mass differences fourteen orders of magnitude smaller than the B mass.

[3]This is a short proof: If $CPT \equiv \zeta$, then $(\zeta|\alpha\rangle, \zeta|\beta\rangle) = \langle\beta|\alpha\rangle$. By setting $\alpha \equiv HB^0$ and $\beta \equiv B^0$, we have $\langle B^0|H|B^0\rangle = (\zeta H|B^0\rangle, \zeta|B^0\rangle)$. If CPT is a good symmetry then $\zeta H \zeta^{-1} = H^\dagger$, or $\langle B^0|H|B^0\rangle = (H^\dagger\zeta|B^0\rangle, \zeta|B^0\rangle) = \langle\bar{B}^0|H|\bar{B}^0\rangle$.

[4]The effect of a non-zero $\Delta\Gamma$ is discussed later in the analysis (Sec. 4.13).

2.3 Decay rate of $B^0\overline{B}^0$ pairs: Dileptons

An initially ($t = 0$) pure flavor eigenstate ($|B^0\rangle$ or $|\overline{B}^0\rangle$) will evolve in time. Due to mixing, at a time $t > 0$, the two B meson states $|B^0(t)\rangle$ and $|\overline{B}^0(t)\rangle$ will be linear combinations of $|B^0\rangle$ and $|\overline{B}^0\rangle$ components.

The standard approach for the study of mixing is the use of a mode where both B mesons decay to a *flavor-specific* state, in other words, a state that tags the flavor of the B meson (b or $\bar{b}$). This is a final state that cannot be accessed from an unmixed B. For example, in the leading order, a B^0 can give a ℓ^+X^- state, but an unmixed $\overline{B}^0$ cannot, according to the SM prediction:

$$\begin{aligned} A_{\ell^+} &\equiv \langle \ell^+X^-|T|B^0\rangle \neq 0, \quad \bar{A}_{\ell^-} \equiv \langle \ell^-X^+|T|\overline{B}^0\rangle \neq 0 \\ A_{\ell^-} &\equiv \langle \ell^-X^+|T|B^0\rangle = 0, \quad \bar{A}_{\ell^+} \equiv \langle \ell^+X^-|T|\overline{B}^0\rangle = 0 \end{aligned} \tag{2.12}$$

In this analysis we use exclusively leptons to tag the flavor of the mesons. This mode is referred in the bibliography as "dileptons": events where both B mesons are flavor-tagged by leptons. With the term "dileptons" it is implied that both of the B mesons decay via a semileptonic process.

$$B^0(t),\ \overline{B}^0(t) \longrightarrow \ell^\pm X^\mp$$

We are actually interested only in *primary* lepton pairs.

A primary lepton is one that is generated through a $b \to c$ quark transition[5] (Fig. 2.2.a). However, a cascade decay can potentionally produce a *secondary* lepton of opposite ($c \to s$, Fig. 2.2.a), or same ($\bar{c} \to \bar{s}$, Fig. 2.2.b) sign. In this analysis, all secondary leptons are considered to be background.

From Eqs. (2.12) it follows that $B^0(t)$ and $\overline{B}^0(t)$ can only give a (primary) ℓ^+ through the $|B^0\rangle$ component and a (primary) ℓ^- through the $|\overline{B}^0\rangle$ component. The sign of the leptons is used to tag the meson as a $|B^0\rangle$ or a $|\overline{B}^0\rangle$. $B^0 \leftrightarrow \overline{B}^0$ mixing is the only way we can get two primary leptons of same sign (B^0B^0 or $\overline{B}^0\overline{B}^0$).

[5]Text and figures refer to a $\overline{B}^0$ decay. All conjugate modes are implied.

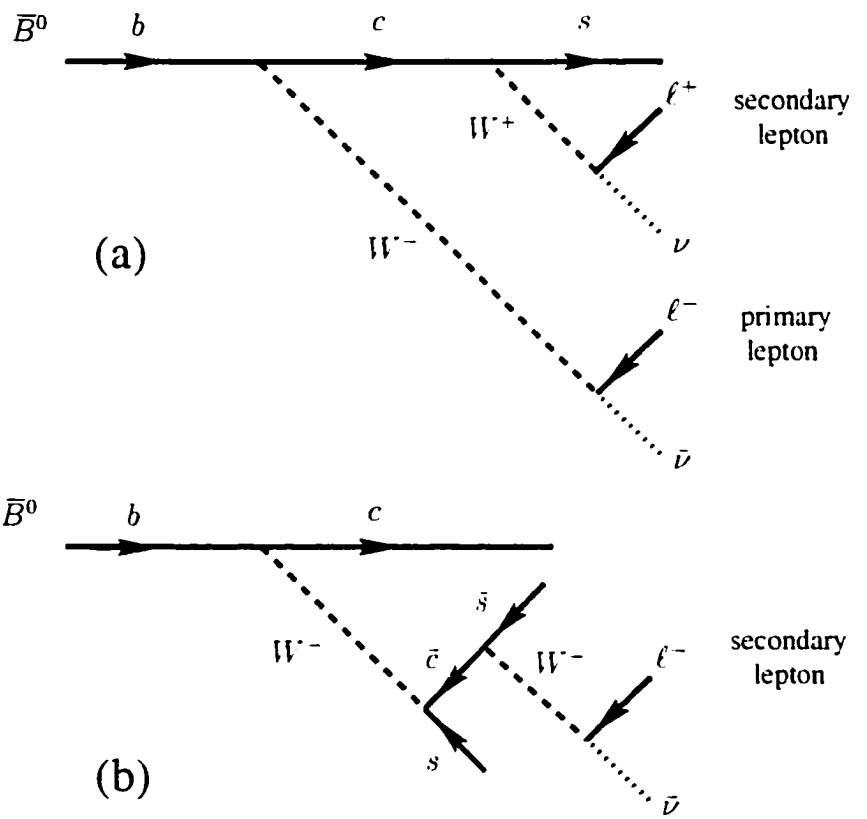

Figure 2.2: **Tagging the B flavor.** (a) The b quark of the B meson decays to a c quark and a W^- boson, which can give a primary lepton. The $c \to s$ cascade decay can give a secondary lepton of opposite charge. (b) It is also possible for a $\bar{c} \to \bar{s}$ cascade decay to give a secondary lepton with charge equal to that of a primary lepton, even though there is no primary lepton in this case. Secondary leptons have typically smaller energies/momenta than primary leptons.

One expects that a primary lepton will typically have higher momentum than a secondary one, since the former comes with the "hard" W^- boson of the primary $b \to c$ vertex, whereas the latter originates from a much "softer" W^- of a cascade decay. Therefore, the standard technique is to set some momentum threshold for the tagging lepton in the center-of-mass, thus eliminating a large fraction of the secondary leptons.

Note that although $X^{\pm}$ of the final states in Eqs. (2.12) could be any hadronic state, in practice ℓ can only be an e or μ. We avoid at this level the reconstruction needed for the τ leptons, which decay inside the detector. Because of the heaviness of the τ mass, there is a substantial phase space suppression in $b \longrightarrow c\,\tau\,\nu$ decays:

$$\mathcal{B}(B \longrightarrow \tau\,\nu\,X) = (2.2 \pm 0.5)\% \quad [40]$$

whereas

$$\begin{aligned} \mathcal{B}(B \longrightarrow e\,\nu\,X) &= (10.4 \pm 0.3)\% \\ \mathcal{B}(B \longrightarrow \mu\,\nu\,X) &= (10.3 \pm 0.5)\% \end{aligned} \quad [5]$$

So, we end up with a primary electron or muon decay $\sim 21\%$ of the time. When we combine the primary lepton branching fractions for both B mesons, we see that primary dilepton pair events occur $\sim 4\%$ of the time. A high efficiency and low fake rate lepton ID system is needed.

2.4 Proper time distributions

The most general (and complex) case is presented first, which describes the B mixing without assuming CP or CPT symmetry (Sec. 2.4.1). The standard expressions in the bibliography of the mixing phenomenology are a simplified version, and are derived from the general case assuming CPT invariance (Sec. 2.4.2).

2.4.1 If CPT is *not* a good symmetry

Same-Sign Dileptons (SS)

The proper time distributions for same-sign lepton pairs are

$$P(\ell^+\ell^+, \Delta t) \sim \frac{|A_{\ell^+}|^4}{4\Gamma} |\sin\theta e^{-i\phi}|^2 e^{-\Gamma|\Delta t|} [1 - \cos(\Delta m\, \Delta t)]$$
$$P(\ell^-\ell^-, \Delta t) \sim \frac{|\bar{A}_{\ell^-}|^4}{4\Gamma} |\sin\theta e^{+i\phi}|^2 e^{-\Gamma|\Delta t|} [1 - \cos(\Delta m\, \Delta t)] \tag{2.13}$$

where Γ is the mean width of the mass eigenstates: $\Gamma \equiv (\Gamma_1 + \Gamma_2)/2$, and Δm their mass difference.

One notices that there is no way to time order the experimentally indistinguishable same sign leptons. Therefore, it only makes sense to talk about $|\Delta t|$ and treat the (symmetric) distributions of Eqs. (2.13) as a function of Δt expanding from 0 to ∞.

To get the total number of same-sign dileptons we integrate:

$$N^{++} \equiv N(\ell^+\ell^+) = 2\int_0^\infty P(\ell^+\ell^+, \Delta t)\, d(\Delta t) \sim \frac{|A_{\ell^+}|^4}{2\Gamma^2} |\sin\theta e^{-i\phi}|^2 \frac{x_d^2}{1+x_d^2} \tag{2.14}$$

$$N^{--} \equiv N(\ell^-\ell^-) = 2\int_0^\infty P(\ell^-\ell^-, \Delta t)\, d(\Delta t) \sim \frac{|\bar{A}_{\ell^-}|^4}{2\Gamma^2} |\sin\theta e^{i\phi}|^2 \frac{x_d^2}{1+x_d^2} \tag{2.15}$$

where

$$x_d \equiv \frac{\Delta m}{\Gamma} = 0.723 \pm 0.035 \quad [5]$$

is the mixing parameter for the $B^0\bar{B}^0$ pair.

Opposite-Sign Dileptons (OS)

The proper time distribution for opposite-sign lepton pairs is

$$P(\ell^+\ell^-, \Delta t) \sim \frac{|A_{\ell^+}\bar{A}_{\ell^-}|^2}{2\Gamma} e^{-\Gamma|\Delta t|} \Big[(1 - |\cos\theta|^2)\cos(\Delta m\, \Delta t) - 2\,\mathrm{Im}(\cos\theta)\sin(\Delta m\, \Delta t) + (1 + |\cos\theta|^2)\Big] \tag{2.16}$$

Here Δt is always defined as $t_{\ell^+} - t_{\ell^-}$.

The total number of opposite-sign dilepton events from neutral B pairs is

$$\begin{aligned} N^{+-} \equiv N(\ell^+\ell^-) &= \int_{-\infty}^{\infty} P(\ell^+\ell^-, \Delta t)\, d(\Delta t) \\ &\sim \frac{|A_{\ell^+}\bar{A}_{\ell^-}|^2}{\Gamma^2} \frac{(2 + x_d^2 + x_d^2 |\cos\theta|^2)}{1 + x_d^2} \end{aligned} \tag{2.17}$$

In Eq. (2.17), N^{+-} refers to final products independent of order, corresponding to both $\ell^+\ell^-$ and $\ell^-\ell^+$ cases.

Combining Eqs. (2.14), (2.15) and (2.17), we obtain the fraction of the same-sign dileptons in the total sample from neutral B pairs

$$\chi_d \equiv \frac{N^{++} + N^{--}}{N^{++} + N^{--} + N^{+-}} \tag{2.18}$$

For direct CP invariance

$$|A_{\ell^+}|^2 = |\bar{A}_{\ell^-}|^2 \tag{2.19}$$

so

$$\chi_d = \frac{|\sin\theta|^2(|e^{-i\phi}|^2 + |e^{i\phi}|^2)\, x_d^2}{|\sin\theta|^2(|e^{-i\phi}|^2 + |e^{i\phi}|^2)\, x_d^2 + 2(2 + x_d^2 + x_d^2 |\cos\theta|^2)} \tag{2.20}$$

At this point we will assume a small $\mathrm{Im}(\phi)$[6] and approximate

$$|e^{\pm i\phi}|^2 \simeq 1 \mp 2\mathrm{Im}(\phi) \tag{2.21}$$

So

$$\chi_d \equiv \chi_d(x_d, \theta) = \frac{|\sin\theta|^2\, x_d^2}{|\sin\theta|^2\, x_d^2 + (2 + x_d^2 + x_d^2 |\cos\theta|^2)} \tag{2.22}$$

[6]This is related to the expectation that CP violation should be small in the B mixing; see Ref [3].

2.4.2 If CPT *is* a good symmetry

Same-Sign Dileptons (SS)

The proper time distributions for same-sign lepton pairs are

$$P(\ell^+\ell^+, \Delta t) \sim \frac{|A_{\ell^+}|^4}{4\Gamma}|e^{-i\phi}|^2 e^{-\Gamma|\Delta t|}[1 - \cos(\Delta m\, \Delta t)]$$
$$P(\ell^-\ell^-, \Delta t) \sim \frac{|\bar{A}_{\ell^-}|^4}{4\Gamma}|e^{+i\phi}|^2 e^{-\Gamma|\Delta t|}[1 - \cos(\Delta m\, \Delta t)] \tag{2.23}$$

The integrated numbers are

$$N^{++} \sim \frac{|A_{\ell^+}|^4}{2\Gamma^2}|e^{-i\phi}|^2 \frac{x_d^2}{1+x_d^2} \tag{2.24}$$

$$N^{--} \sim \frac{|\bar{A}_{\ell^-}|^4}{2\Gamma^2}|e^{i\phi}|^2 \frac{x_d^2}{1+x_d^2} \tag{2.25}$$

Opposite-Sign Dileptons (OS)

The proper time distribution for opposite-sign lepton pairs is

$$P(\ell^+\ell^-, \Delta t) \sim \frac{|A_{\ell^+}\bar{A}_{\ell^-}|^2}{2\Gamma} e^{-\Gamma|\Delta t|}\Big[1 + \cos(\Delta m\, \Delta t)\Big] \tag{2.26}$$

The total number of opposite-sign dilepton events from neutral B pairs is

$$N^{+-} \sim \frac{|A_{\ell^+}\bar{A}_{\ell^-}|^2}{\Gamma^2}\frac{(2+x_d^2)}{1+x_d^2} \tag{2.27}$$

and the expression for χ_d becomes the standard one

$$\chi_d \equiv \chi_d(x_d) = \frac{x_d^2}{2(1+x_d^2)} \tag{2.28}$$

The above equations define the parameters of the problem: the dilepton proper time distributions are simple, and contain oscillation terms that should be easy to detect, at least in principle. The mass difference of the B eigenstates defines the scale of the microscopic world that we are probing: $O(10^{-14}$ GeV). Any deviation

from the SM equations (Sec. 2.4.2), comparable in size to Δm (or Γ) should be apparent. Δm is a very small mass difference, thirteen orders of magnitude smaller than the B mass. Therefore, the B mixing is not just the mechanism that gives birth to large CP violation, but also a great tool for testing the CPT symmetry.

2.5 Asymmetric e^+e^- colliders and proper time determination

Since the B lifetime is ~1.6 ps, a time scale that cannot be measured directly, the decay vertex determination is the only method to draw conclusions about the time evolution of a B before its decay. The effect of an asymmetric collider on the vertex measurement will be examined.

The $\Upsilon(4S)$ mass (10.58 GeV/c^2) is just above the threshold of B pair production (~10.56 GeV). Therefore, the B mesons in the $\Upsilon(4S)$ rest frame are born quite non-relativistic:

$$\gamma = \frac{M_{\Upsilon(4S)}}{2M_{B^0}} \simeq 1.002, \quad \text{or} \quad |\vec{\beta}| \simeq 0.064$$

To estimate the average distance traveled by a B in the $\Upsilon(4S)$ frame we use the same boost: $\ell \sim \gamma\beta c\tau_B \simeq 30\mu$m. So, in a symmetric collider, the B mesons typically travel only a few tens of microns, a distance that is too short to measure.

In an asymmetric collider, however, because of the electron-positron energy asymmetry, the $\Upsilon(4S)$ (and the B pair) is boosted. Since the two B mesons are practically at rest in the $\Upsilon(4S)$ rest frame, in the laboratory they move together and at the same speed until the first of them decays. The distance between the two decay vertices corresponds to the difference between the decay times (in the $\Upsilon(4S)$ frame), which is the quantity appearing in the time distributions.

In Belle, the difference in the electron (8.0 GeV) and positron (3.5 GeV) energies

yields for the boost from the laboratory to the $\Upsilon(4S)$ rest frame

$$|\vec{\beta}| = \frac{|\vec{p_{e^-}} + \vec{p_{e^+}}|}{E_{e^-} + E_{e^+}} \simeq 0.391, \quad \text{or} \quad \beta\gamma \simeq 0.425$$

The average decay time difference for the two B mesons at their center of mass is $\tau_B \sim 1.6$ ps. This time corresponds to decay vertices separated by $\ell \sim \gamma\beta c\tau_B \simeq 200\ \mu$m. As we will see later, the Belle decay vertex resolution ($100 - 110\ \mu$m) allows us to measure distances of this scale.

It was mentioned earlier that the Re($\cos\theta$) measurement corresponds to mass differences fourteen orders of magnitude smaller than the B mass. The Δm interferometry does not give the same sensitivity for the lifetime difference measurement, since $\Delta m \sim \Gamma$. However, an assumed Γ_{B^0}-$\Gamma_{\bar{B}^0}$ x Im($\cos\theta$) difference would introduce antisymmetric terms in the decay time distributions (Eq. (2.16)). Therefore, if there is a difference between the B^0 and $\bar{B}^0$ lifetimes an asymmetric e^+e^- machine is the right place to look for it.

Chapter 3

The Tools: Machine, Detector, Software

The primary goal of a "B factory", such as KEKB or PEP-II, is the study of CP violation. An asymmetric high luminosity e^+e^- collider provides an ideal combination of a clean environment (e^+ and e^- beams), lots of B pairs (high luminosity and machine tuned on $\Upsilon(4S)$ resonance) and sensitivity to time dependent asymmetries (asymmetric machine with the addition of a suitable vertex detector). The Belle experiment operating at the B factory (KEKB) of the High Energy Accelerator Research Organization (KEK) is optimized for B Physics studies. The Belle detector has a good Particle Identification system (PID), high reconstruction efficiency for charged and neutral particles, and makes precise vertex measurements with a silicon vertex detector. In Sec. 3.1 we give a short description of the KEKB machine, and in Sec. 3.2 we give a brief overview of the Belle detector. Sec. 3.3 describes the software used in this study.

3.1 The KEKB accelerator

KEKB is located in Tsukuba, Japan. It has an 8.0 GeV electron beam (or High Energy Ring — HER) and a 3.5 GeV positron beam (LER). The B mesons are produced

in pairs at the $\Upsilon(4S)$ resonance of the e^+e^- annihilation: $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B^0\overline{B}^0$ (or B^+B^-), with $E_{CM} \simeq 10.58$ GeV. The $4S$ is the 4th excited state of the Υ, a $b\bar{b}$ bound state (Fig. 3.1). The first three states ($1S$, $2S$ and $3S$) have much narrower peaks than the $4S$, as they decay through a 3 gluon exchange and are, therefore, suppressed. $\Upsilon(4S)$ has energy just barely above the $B\overline{B}$ production threshold. So, $\Upsilon(4S)$ decays almost instantaneously to a B pair without any additional byproducts, resulting in a broad peak. The misfortune is that the $\Upsilon(4S)$ peak is much shorter than the three lowest lying Υ states, so we have to compromise with a cross section of 1.05 nb [10]. The $\Upsilon(4S)$ sits on a continuum background of a ~ 3.7 nb cross section.

The HER and the LER are two separate rings with a 3 km circumference (Fig. 3.2). They are located in the TRISTAN tunnel. A crossover at the Fuji area reverses the relative position of the two rings so as to give them the same length. The single interaction point order for them to have the same length. The single interaction point (IP) is found in the Tsukuba Hall ("Tsukuba Area").

A unique feature of KEKB is that electrons and positrons do not collide "head-on", but at a small angle (± 11 mrad), in order to reduce parasitic collisions near the IP. To achieve the design luminosity 10^{34} cm^{-2}s^{-1} (corresponding to about 100 million neutral and charged B meson pairs a year), 5000 bunches need to be injected in each ring. That sets the spacing between them to just 60 cm, or 2 ns. As of July 2000, the achieved luminosity is 2×10^{33} cm^{-2}s^{-1}, a factor of five lower than in the full operation mode. The number of bunches is 1150 and the spacing between them 240 cm. Finally, the achieved beam currents are 465 mA for the LER and 420 mA for the HER. The design values are 2600 mA and 1100 mA, respectively. The design specifications of KEKB are listed in Table 3.1 [12].

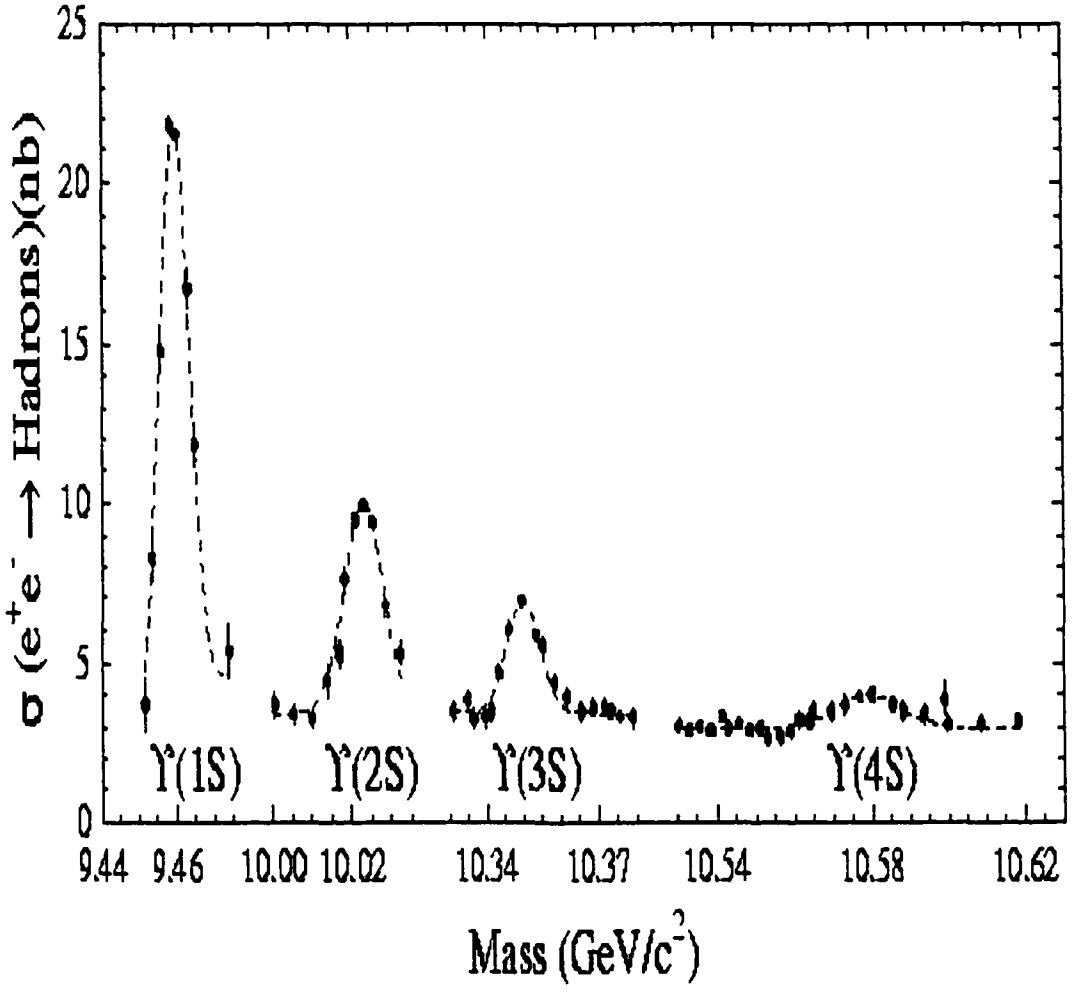

Figure 3.1: The first four Υ excited states. KEKB is operating on the $4S$ resonance. The figure is reproduced from the CLEO www page [11].

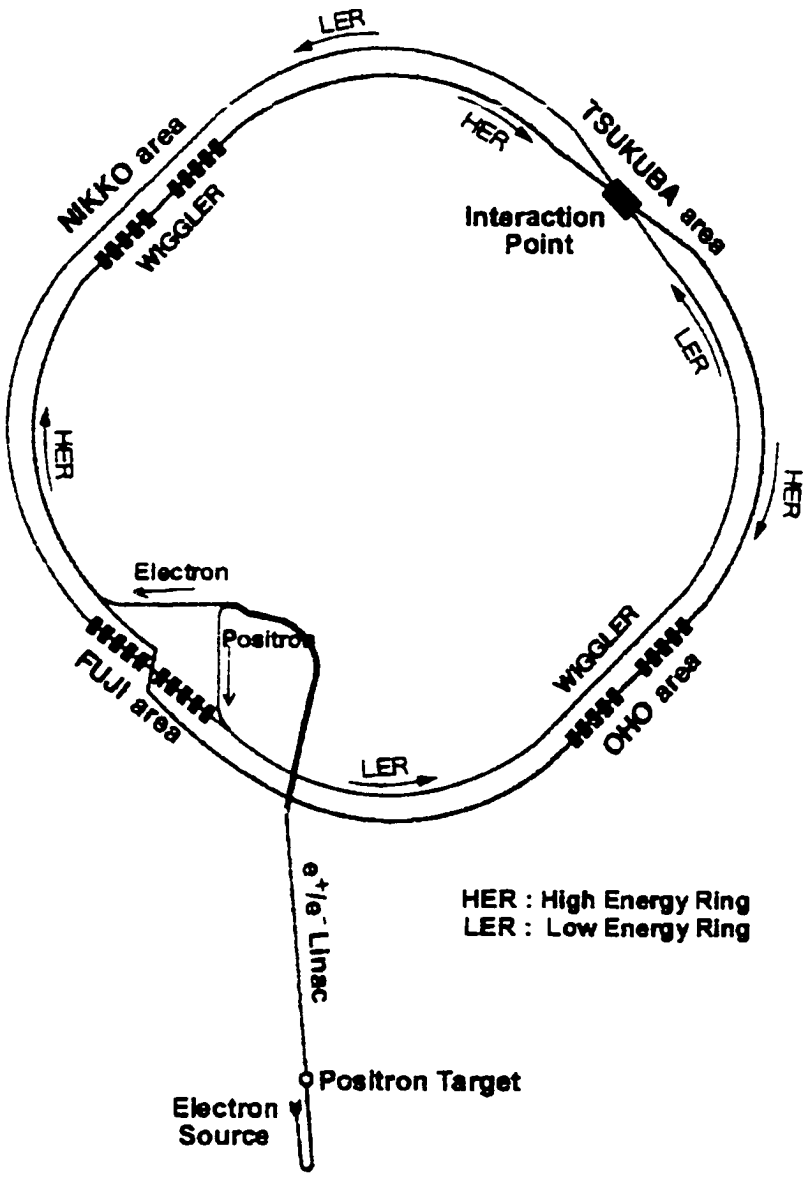

Figure 3.2: The KEKB B factory.

Table 3.1: The KEKB accelerator design parameters.

Parameter	Symbol	LER	HER	Units
Energy	E	3.5	8.0	GeV
Circumference	C	3016.26		m
Luminosity	$\mathcal{L}$	10^{34}		$cm^{-2}s^{-1}$
Crossing angle	θ_x	± 11		mrad
Tune shifts	ξ_x/ξ_y	0.039/0.052		—
Beta function at IP	β_x^*/β_y^*	0.33/0.01		m
Beam current	I	2.6	1.1	A
Natural bunch length	σ_z	0.4		cm
Energy spread	σ_ε	7.1×10^{-4}	6.7×10^{-4}	—
Bunch spacing	s_b	0.59		m
# of particles/bunch	N	3.3×10^{10}	1.4×10^{10}	—
Emittance	$\varepsilon_x/\varepsilon_y$	$1.8 \times 10^{-8}/3.6 \times 10^{-10}$		m
Synchrotron tune	ν_s	0.01-0.02		—
Betatron tune	ν_x/ν_y	45.52/45.08	47.52/43.08	—
Momentum compaction factor	α_p	$1 \times 10^{-4} - 2 \times 10^{-4}$		
Energy loss/turn	U_o	0.81/1.5	3.5	MeV
RF voltage	V_c	$5 - 10$	$10 - 20$	MV
RF frequency	f_{RF}	508.887		MHz
Harmonic number	h	5120		—
Longitudinal damping time	τ_ε	43/23	23	ms
Total beam power	P_b	2.7/4.5	4.0	MW
Radiation power	P_{SR}	2.1/4.0	3.8	MW
HOM power	P_{HOM}	0.57	0.15	MW
Bending radius	ρ	16.3	104.5	m
Length of bending magnet	ℓ_B	0.915	5.86	m

3.2 The Belle detector

The Belle group is an international collaboration of about 250 scientists from more than 50 Institutes and from eleven countries around the world (Australia, China, India, Korea, Japan, Philippines, Poland, Russia, Taiwan, Ukraine, USA). Along with the CP measurement in the "gold-plated" $J/\psi K_S$ decay, there are numerous other decay modes, through whose reconstruction we can make precision measurements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements [2]. By completing the mosaic, in the long run, people hope to test and verify to a large degree the Standard Model (SM) picture for the B system, or reject it. A parallel set of studies is performed on charm, two-photon and tau physics, and of course, the search for physics beyond the SM, as is a large part of this dissertation. The properties that a detector with such a wide range of goals must include are:

- the ability to make precise vertex measurements for the determination of the B decay times,
- a good PID system for the identification of charged particles (e, μ, π, K, p) and a strong detection system for neutral particles (γ, π^0, K_L),
- efficient charged-particle tracking
- high-resolution photon detection,
- high performance in the reconstruction of particles that decay inside the detector and determination of exclusive final states, and
- an efficient trigger and DAQ system.

A picture of the Belle detector is shown in Fig. 3.3.

For the vertexing, the trajectories of the charged particles must be reconstructed. This is accomplished with a combination of a vertex detector (in Belle it is called

Silicon Vertex Detector, or SVD) and a drift chamber (Central Drift Chamber, or CDC, for Belle) surrounding the vertex detector. For measurements of the charged particle momenta, the SVD and CDC are embedded in a 1.5 T magnetic field, which runs parallel to the direction of the z-axis (defined by the electron beam). Tracks are defined by the pattern of the hit wires in the CDC. Refined position values of the tracks with respect to the wires are obtained using drift-time information. The curvature of the tracks in the magnetic field reveals the charges of the particles and their momenta. The SVD provides precise vertex information near the IP.

For the PID, three subsystems are employed: the CDC (which measures the mean energy loss dE/dx), the Aerogel Čerenkov Counter (ACC) and the Time of Flight (TOF) systems. The ACC is a set of Čerenkov radiation modules located outside the CDC volume. The TOF is a system of scintillators providing time information. Besides the PID, it is also used for the trigger.

A pair of calorimeters, the Electromagnetic Calorimeter (ECL) and the Electromagnetic Forward Calorimeter (EFC), is used for the detection and identification of photons and electrons by the amount and the pattern of the energy deposited. The picture is completed by the K_L and muon Detector (KLM), which is used for the identification of muons and K_L's.

The following sections give details on the specifications of the detector components [13]. A cylindrical coordinate system (r, ϕ, θ) with origin at the IP is assumed throughout.

3.2.1 The Silicon Vertex Detector (SVD)

The SVD provides information for the reconstruction of decay vertices close to the IP. As mentioned in Sec. 2.5, precision vertexing is necessary for the B decay time difference extraction. The tracking is done with a combination of hits from the CDC. The SVD improves the particle momentum resolution with the addition of a number of associated hits and by expanding the lever arm over which the curvature

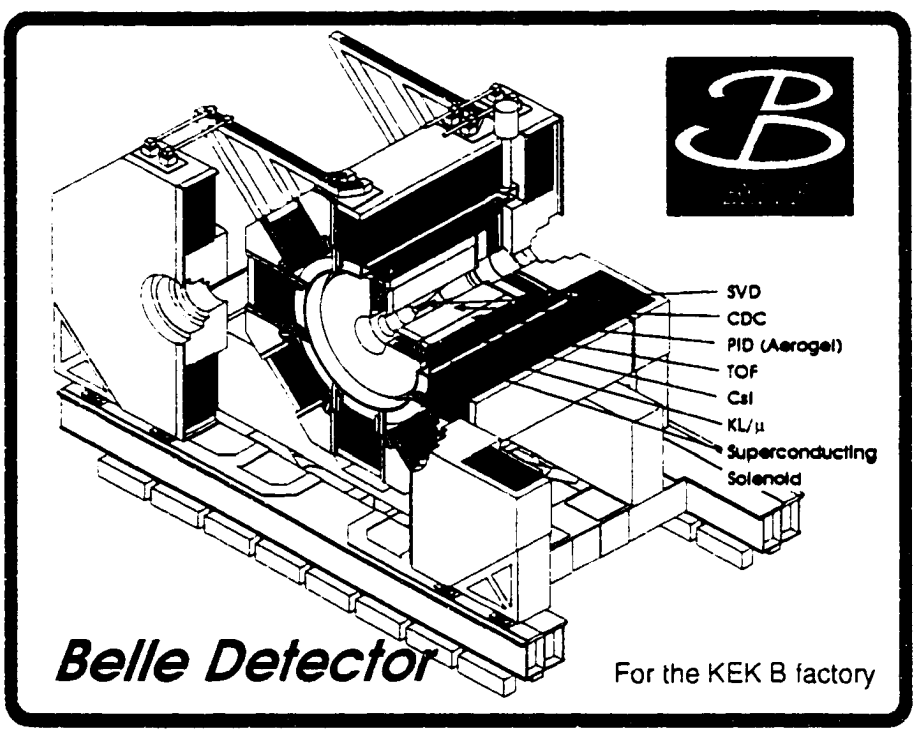

Figure 3.3: The Belle detector.

of the track is calculated.

The SVD consists of three layers of 300-μm-thick double sided silicon sensors, surrounding a double walled beryllium beam pipe of total thickness 1.0 mm and diameter 4.6 cm (Fig. 3.4). The configuration of the sensors is such that there is an overlap between adjacent pairs. This geometry ensures that a trajectory will necessarily pass through at least one sensor from each layer, providing the maximum number of hits. The radial position of the layers are at $r = 3.0$ cm, 4.55 cm and 6.05 cm, which have 8, 10 and 14 sensor ladders in the ϕ direction, respectively (Fig. 3.5). The polar angle coverage is from 20° to 140°, which corresponds to 87% of the 4π solid angle in the center of mass frame. The number of readout channels is 81900. The impact parameter resolution at the interaction point is $(21 + 69/p\beta \sin^{3/2}\theta)$ μm in $r\phi$ and $(41 + 48/p\beta \sin^{5/2}\theta)$ μm in the z direction.

During the first two months of its operation in the summer of 1999, Belle suffered from beam backgrounds (mainly synchrotron radiation) that seriously damaged the SVD. Specifically, 5 keV X-rays from the HER steering magnet close to the IP were hitting the inner layers. As the dose monitor was not sensitive to this particular radiation, the SVD was eventually destroyed and had to be replaced. The runs for Experiments 5 and 7 (Fall 1999 through Summer 2000) were taken with SVD 1.2, with a 200 kRad radiation hardness. At the end of the summer of 2000, SVD 1.2 will be replaced by SVD 1.5 (500 kRad). Finally, the plan for the summer of 2001 is to install an even more tolerant to radiation vertex detector: SVD 2.0, which can withstand up to 2 MRad.

A more detailed description of the SVD and its performance is given in [14].

3.2.2 The Central Drift Chamber (CDC)

The CDC is used for the detection of charged particles. This is achieved with the partial collection of the charge produced when high momentum particles ionize the gas contained in the chamber. The positive and negative ion avalanches very quickly

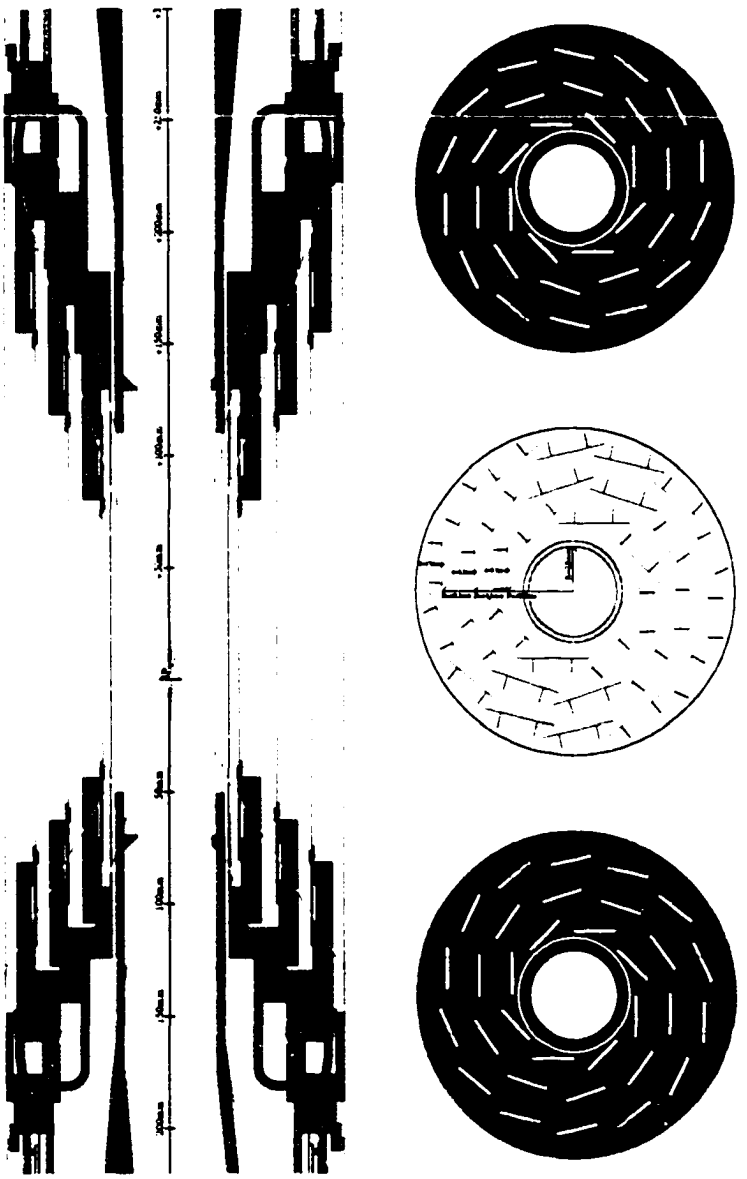

Figure 3.4: The three double-sided silicon sensors of the SVD (profile and cross section).

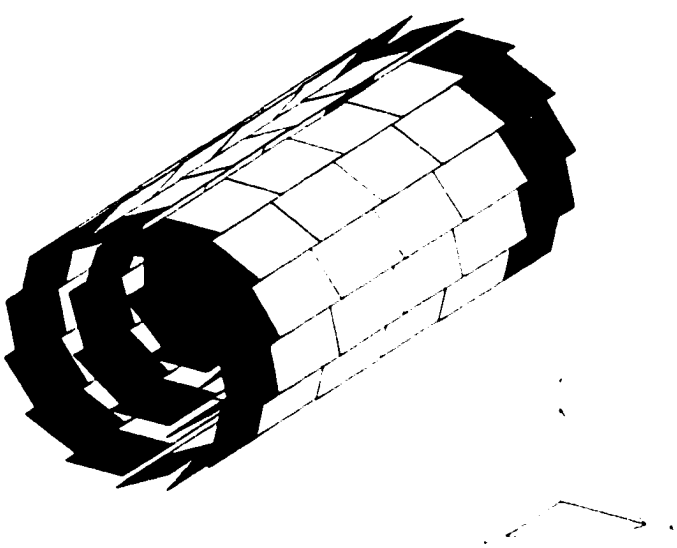

Figure 3.5: The configuration of the SVD ladders.

reach their maximum speed in the gas (typically 50 μm/ns for electrons) and "drift" towards cathode and anode wires, which are included in the same structure.

The tracks of the event are defined by the hits in the drift chamber. The time ordering of the hits depends on radial differences of the particle trajectory from the wires. An algorithm called a Kalman filter combines the CDC with the SVD hits by taking into account the small variations of the magnetic field [15]. The result is three-dimensional reconstructed trajectories.

Besides the tracking information, the CDC hit pattern leads to the determination of the mean ionization deposition dE/dx. dE/dx scales as $\ln(\beta^2\gamma^2)/\beta^2$ for moderately relativistic particles and as $\ln(\beta\gamma)/\beta^2$ for higher energies [5]. From the dE/dx and the measured momentum of the particle a PID hypothesis is derived. The CDC is employed in this analysis both for tracking (vertexing and momentum determination) and for lepton identification.

The CDC consists of 32 axial anode wire layers ($x - y$ measurement), 18 small-angle stereo wire layers (z measurement) and three cathode-strip layers at the inner radius with fine segmentation for high precision z measurement at the entrance point of the track. It occupies the region from $r = 8$ cm to $r = 88$ cm, and covers polar angles between 17° and 150°. This is equivalent to 92% of the 4π solid angle in the $\Upsilon(4S)$ rest frame. There are 8400 readout channels for the anode wires and 1792 for the cathode strips. A low Z gas mixture (50% helium - 50% ethane) is used in order to minimize the multiple Coulomb scattering contribution to the momentum resolution.

Using muon tracks from cosmic rays and from $e^+e^- \longrightarrow \mu^+\mu^-$ interactions, a resolution of $\delta p_t/p_t = (0.36 \oplus 0.28 p_t)\%$ (the transverse momentum p_t in GeV/c) was obtained. The dE/dx resolution for hadron tracks is 6.9%.

More details on CDC can be found in [16].

3.2.3 The Aerogel Čerenkov Counter (ACC)

The ACC system is the backbone of the PID system. It is designed to separate kaons from pions in the 1.5 GeV/c < p < 3.5 GeV/c region. The ACC consists of blocks of silica aerogel in 0.2 mm aluminum boxes. This material is a colloidal form of glass, in solid form, transparent and very light. The choice of aerogel as the radiation material was made because its index of refraction (n) can be controlled to be $1.01 < n < 1.05$, a regime otherwise difficult to reach. Fine-mesh photo-multipliers (FM-PMT) that can operate inside the strong magnetic fields are attached to the aerogel radiator modules for the detection of Čerenkov radiation. This is the light emitted when a particle travels faster than the speed of light in the material, i.e. when $\beta > 1/n$. The index of refraction of the aerogel in the barrel region varies with the polar angle ($n = 1.010 - 1.028$) and has been optimized to match the kinematics of two-body decays from the boosted B mesons (Fig. 3.6). On the other hand, the index of refraction for the endcap ($n = 1.030$) has been tuned for kaon identification, which will ultimately be used for B flavor tagging. Even though in this analysis we are not using kaons for tagging, the ACC information is important for the suppression of fake leptons, mainly from kaons.

There are 960 ACC modules in the barrel region (16 z and 60 ϕ segments) and 228 ACC modules in the endcap in 5 layers (36, 36, 48, 48 and 60 segments going outwards). The number of readout channels is 1560 for the barrel and 228 for the endcap.

Ref. [17] gives more details on the ACC system.

3.2.4 The Trigger/Time of Flight Counter (TOF)

The TOF is part of the PID system, mainly for slow particles (p < 1.2 GeV/c). The TOF modules are mounted on the inner surface of the CsI electromagnetic calorimeter. They measure the elapsed time between a collision at the IP and the

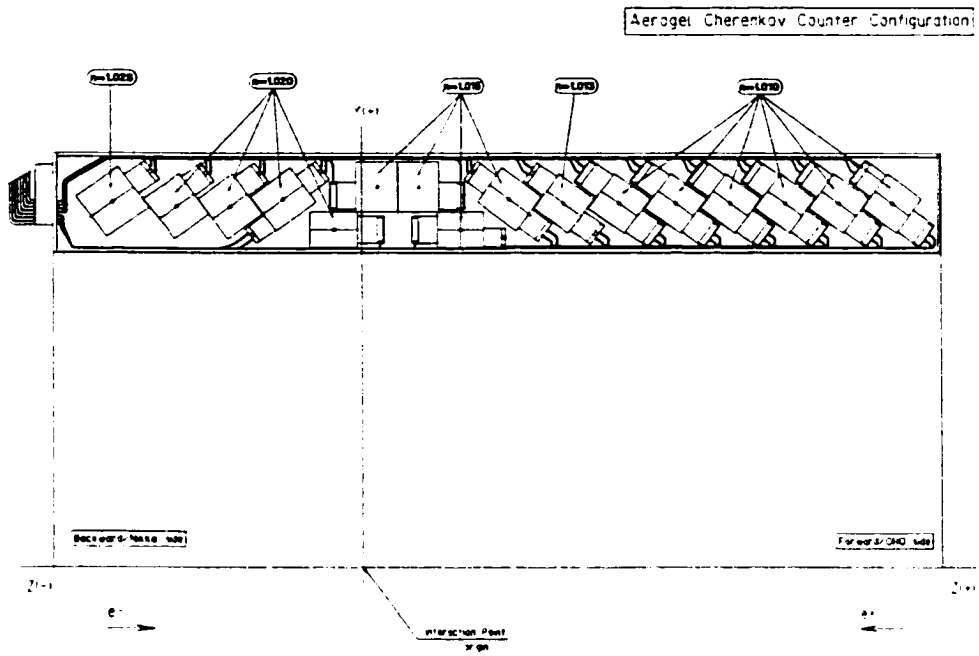

Figure 3.6: **The ACC configuration in the barrel region.** The index of refraction n varies with the polar angle in order to match the kinematics of B decays. The design is based on Monte Carlo studies.

moment at which a particle crosses the TOF layer. With the particle momentum measured by the CDC, this time difference gives an estimate of the particle mass.

There are 128 TOF scintillation ϕ-sector counters. Their dimensions are 4 cm $\times$ 6 cm $\times$ 255 cm, and are located at r = 120 cm. The time resolution achieved is 100 – 120 ps, and the track matching efficiency $\sim$ 90%.

More details on TOF can be found in Ref. [18].

3.2.5 The Electromagnetic Calorimeter (ECL)

The main purpose of the ECL is the detection of the numerous photons from π^0 decays or de-excitation of D^* and K^* states, with high efficiency and good resolution. It is also used for the PID of electrons, making good use of the characteristic electromagnetic showers that they create through the Bremsstrahlung and pair-production mechanisms. In principle, showers in the ECL from photons and electrons are

indistinguishable. However, a charged track in the CDC associated with an electromagnetic shower in the ECL separates electrons from photons. Muons, being much heavier than electrons, lose little energy in the ECL while ionizing the atoms, and therefore leave a very different energy deposition pattern than electrons. Pions, on the other hand, interact strongly with the absorber of the calorimeter and leave a much longer and larger shower than electrons, which fall very quickly below the critical energy. The pattern and the amount of the deposited energy ("calorimetry") can be combined with the momentum of the track from the CDC and used for electron-hadron separation. Good ECL energy resolution is important in order to tell electrons from hadrons and to keep the electron fake rates (mainly from pions) to a low level.

The ECL comprises 8736 blocks of 30.0 cm × 5.5 cm × 5.5 cm CsI (Tl) crystals situated just inside the solenoid and the end yokes. The thickness of 30 cm corresponds to 16.1 radiation lengths. The crystals are placed in the barrel region (6624), and forward (1152) and backward (960) endcaps. The barrel part has 46 θ and 144 ϕ segments, and occupies the region between $r = 125$ cm and $r = 162$ cm. The forward endcap has 13 θ and 48 to 144 ϕ segments and is located at $z = 196$ cm, whereas the backward endcap has 10 θ and 64 to 144 ϕ segments and is located at $z = -102$ cm. The energy resolution is $\sigma_E/E = (1.3 \oplus 0.07/E \oplus 0.8/E^{1/4})\%$, with E given in GeV. The ECL covers the same angular region as the CDC.

Fig. 3.7 summarizes the ECL performance.

A detailed description of ECL can be found in Ref. [20].

3.2.6 The Superconducting Magnet

Belle has a magnetic field of $B = 1.5$ T parallel to the beam axis. Charged particles with momentum vectors forming an angle λ with respect to the field will be forced to move on a helix. The momentum can be measured by the curvature of the helix

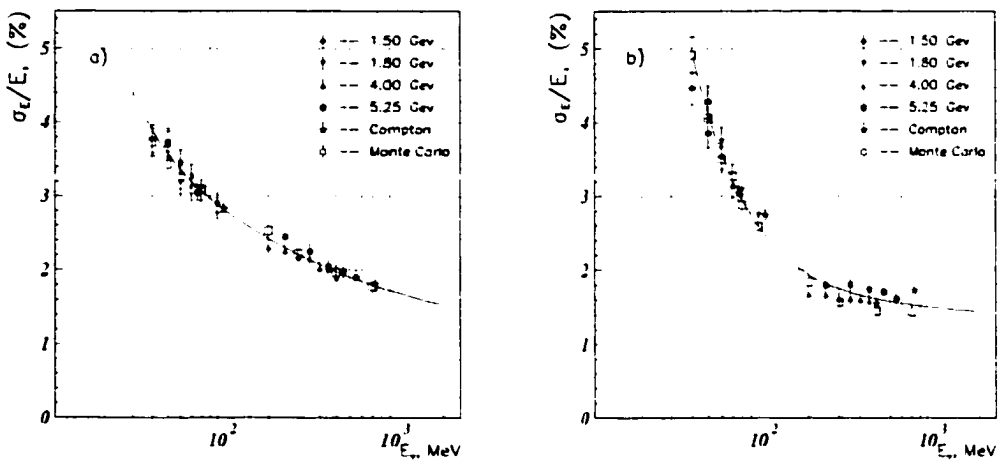

Figure 3.7: **ECL performance.** Energy resolution *vs.* incident photon energy for 3×3 (left) and 5×5 (right) matrices. Results obtained for different values of electron beam energies. Compton edge and Monte Carlo are shown [19].

(defined by the CDC and SVD hits). R:

$$p = \frac{BqR}{\cos\lambda}$$

where $q = \pm1$ is the usual value of the charge of the particle.

The field is generated by a superconducting coil, which consists of a single layer of a niobium-titanium-copper alloy embedded in a high purity aluminum stabilizer. It is wound around the inner surface of an aluminum support cylinder of 3.4 m in diameter and 4.4 m in length. The cooling is provided by liquid helium circulating through a tube on the inner surface of the aluminum cylinder. The iron yoke in the barrel and endcap regions serves as the magnetic flux return path outside the coil.

3.2.7 The K_L and muon Detector (KLM)

The KLM, as its name suggests, is used for the detection of muons and K_L's. It consists of 14 layers of Resistive Plate Counters (RPC). It is the only sub-detector placed outside the coil. The RPC's are sandwiched betwen 4.7 cm thick iron plates

and are filled up with a gas mixture (30% argon, 8% butane and 62% HFC134a). High voltage is applied to the glass plate electrodes. Charged particles going through the RPC's ionize the gas, and the amplified signal is picked up by two sets of strips: x and y in the barrel, and θ and ϕ in the endcaps. The role of the iron is double: to interact with K_L's, giving a small shower that we can detect, and to serve as the return path for the magnetic flux outside the coil. A shower in the KLM without an associated charged track in the CDC is a K_L candidate. An example is shown in Fig. 3.8.

Muons, as heavy leptons, have much cleaner tracks (see Fig. 3.9 [21]) since they lose energy only through ionization. Pions, on the other hand, interact strongly with the iron and scatter randomly usually without managing to penetrate more than a few KLM layers. This is a dramatic signature difference that helps us keep the fake rates from the (numerous) pions at relatively low numbers. A finite number of pions decaying in-flight to muons, however, cannot be avoided.

The picture of a shower associated with strong interactions applies also to kaons, with the difference that pions, just like muons, fire the ACC's, whereas the kaons generally do not.

The KLM polar angle coverage is from 20° to 155°. The number of readout channels is 21856 in the barrel and 16128 in the endcap sections.

Reference [22] gives more details about the KLM sub-detector.

3.2.8 The Electromagnetic Forward Calorimeter

The EFC is an additional electromagnetic calorimeter for very small polar angles: $6.2° < \theta < 11.6°$ and $163.1° < \theta < 171.5°$. It consists of 320 radiation-hard BGO (Bismuth Germanate - $Bi_4Ge_3O_{12}$) crystals, with 12.0 and 10.5 radiation lengths in the forward and the backward regions, respectively. The role of the EFC is the monitoring of the luminosity, and the study of a series of processes, such as Bhabha, $e^+e^- \to e^+e^-\gamma$, $e^+e^- \to e^+e^-\mu^+\mu^-$, or even $e^+e^- \to e^+e^-\gamma\gamma$ for the search of

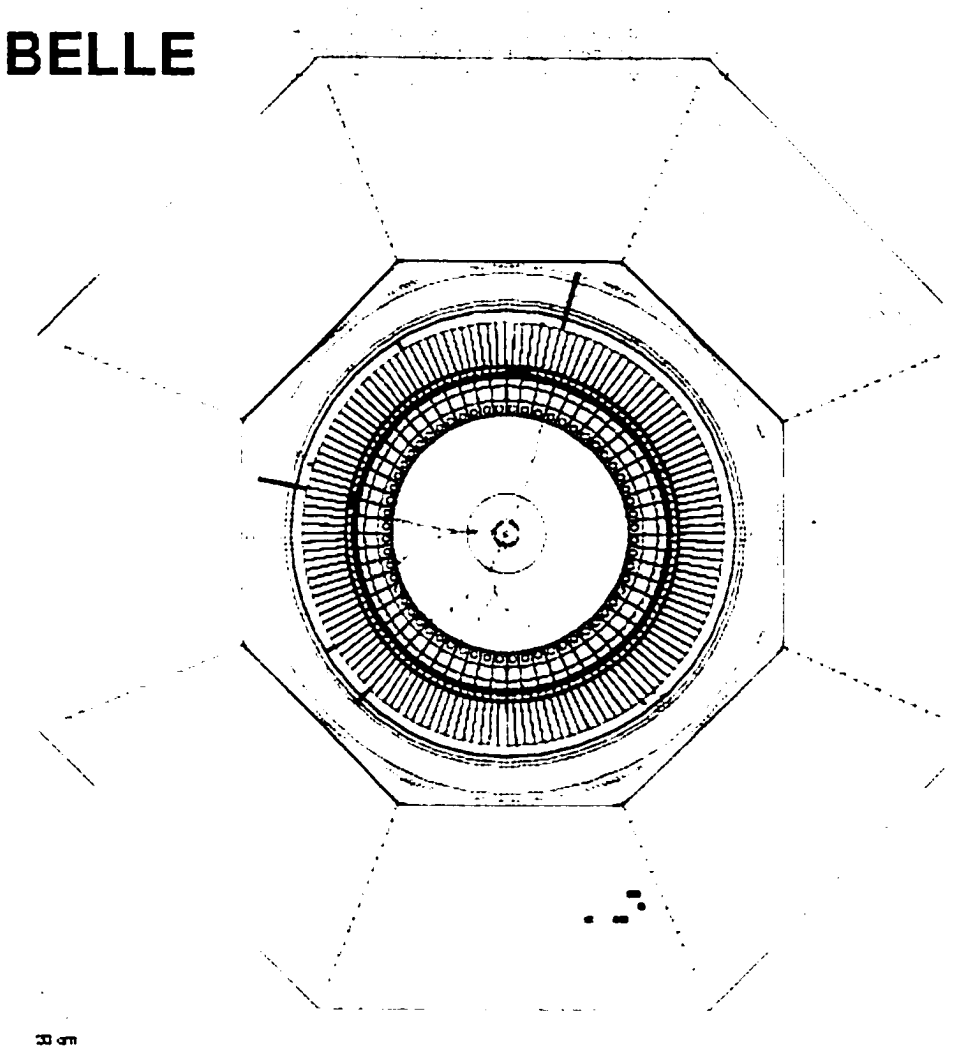

Figure 3.8: **$x-y$ view of an event recorded at Belle.** A small shower is detected at the lower side of the KLM detector, without any associated charged tracks in the CDC. This is a K_L candidate.

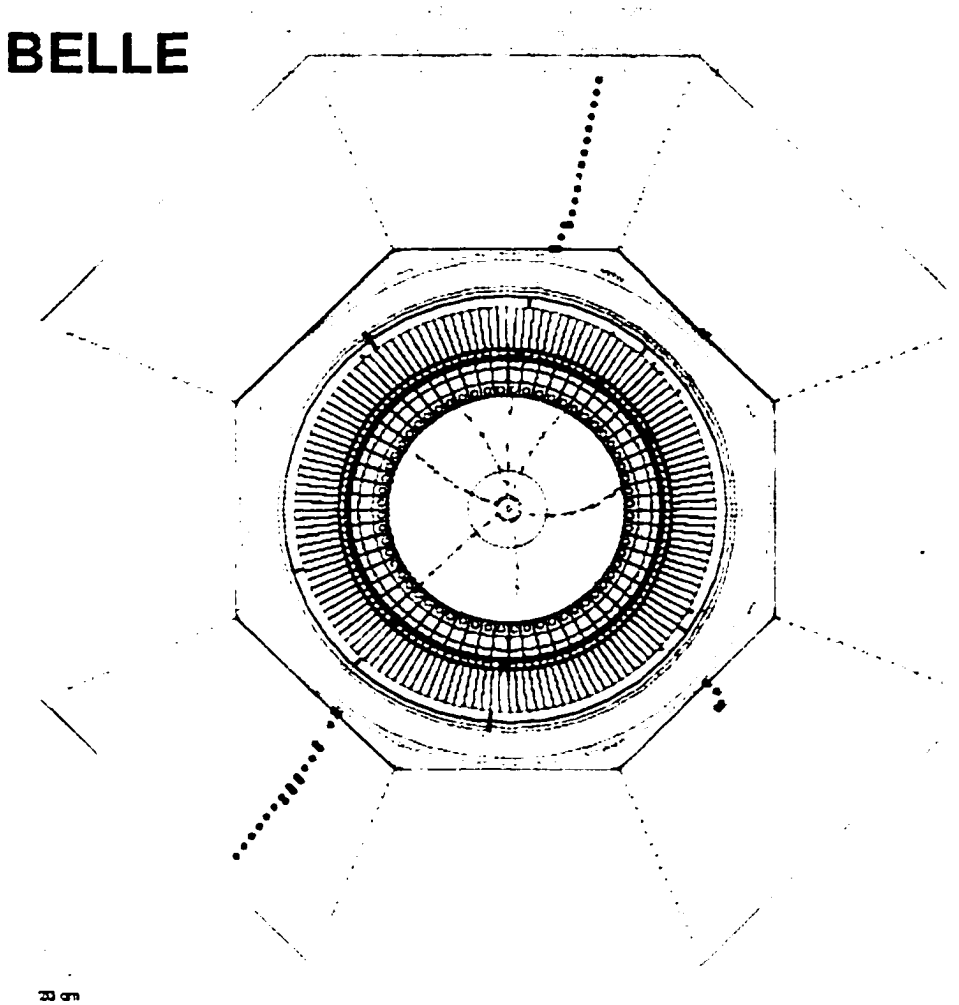

Figure 3.9: **$x - y$ view of another event recorded at Belle.** Clean muon tracks at the top (N-NE) and bottom (SW) diagonal of the KLM detector, which can be traced back to the CDC.

exotic resonances and gluon balls. These reactions have a very large cross section for small angles. By including the EFC, we increase the hermiticity of the detector and the photon detection efficiency, and we hopefully reduce the background in rare B decays channels. The potential problem with the the EFC is that high beam currents give high levels of background.

3.2.9 The Trigger and Data-Acquisition System (DAQ)

The concept of the trigger is that of a fast signal that controls the DAQ. If the event is interesting, then information (the data) needs to be stored. Otherwise, the event is ignored. The decision is based on a pre-designed combination of logic steps, which are implemented electronically.

In Belle, the sub-triggers from the sub-detectors are combined in what is called the Global Decision Logic (GDL), to form the master trigger. A decision for the trigger is made within 2.2 μs after the beam collision at the IP. Using a pipelined trigger system reduces the dead time, a critical factor for the trigger system, especially in full operation modes when the typical time between two beam crossings is only a few ns. The logic diagram of the trigger is given in Fig. 3.10.

If a decision to keep the event is made, the "news" is passed to the DAQ system by the Sequence Control Unit. The data from the sub-detectors has to be digitized in 200 μs. This corresponds to dead time of 10% for a 500 Hz trigger rate[1]. The Event Builder combines the data from the sub-detectors to form full event records, which will in turn be fed to the on-line computer farm. There is one last step before the data is saved on disc: the on-line event reconstruction, which takes place at the computer farm. Fig. 3.11 shows the logic diagram of the DAQ system.

[1]The average trigger rate for the runs in the summer of 2000 is approximately 200 Hz.

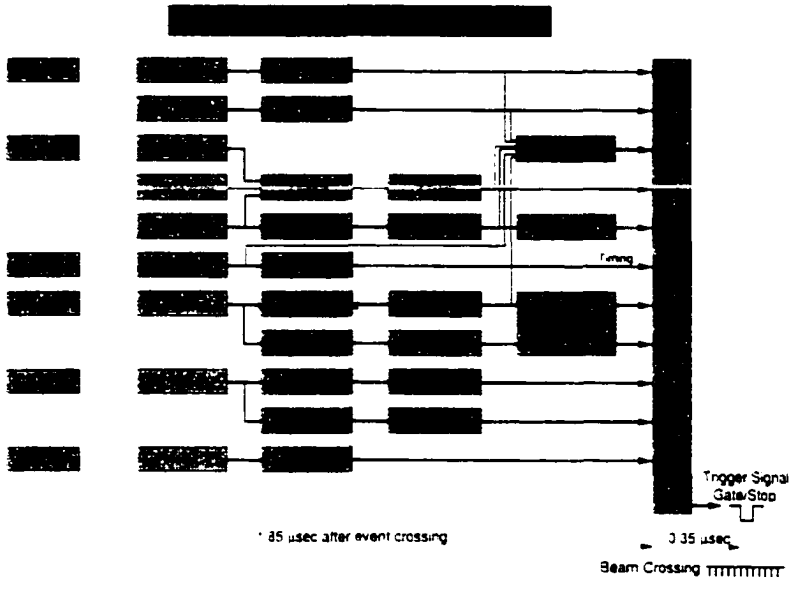

Figure 3.10: Belle detector - Logic diagram of the trigger system.

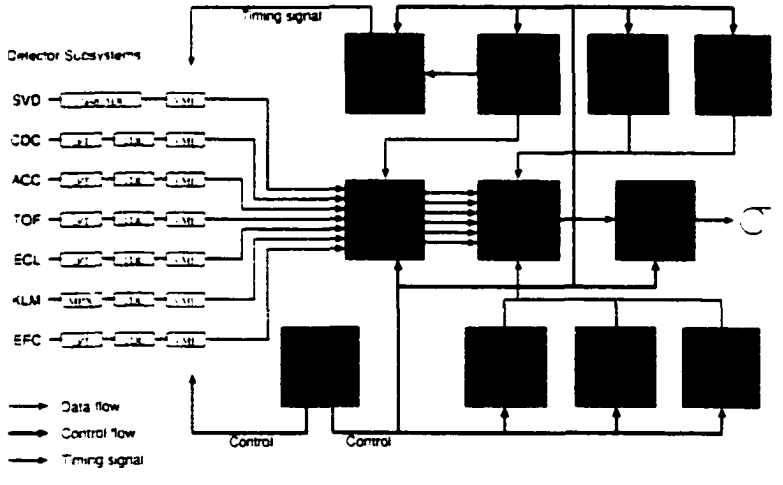

Figure 3.11: Belle detector - Logic diagram of the DAQ system.

3.3 The Software

We give a brief overview of the structure of the analysis environment and the description of Belle's Monte Carlo simulation program.

3.3.1 The Computing Environment

The Monte Carlo simulation used in this dissertation invokes the Belle Analysis and Simulation Framework (BASF) [23]. BASF is the main generic structure for the Belle analysis software and combines different blocks of software, or "modules", to build the analysis program. The user typically provides the specific purpose analysis code as a "plug-in" module and makes use of existing external software that is necessary for the analysis. The convenience that BASF offers is that the user need not worry about the interface between the different modules. There is also a choice of more than one supported computer languages: Fortran, C and C++. Using BASF, the user can

- choose the input and output for his/her reconstruction code.
- use the provided event data management to create objects of his/her own.
- choose between single- and multi-CPU[2] (parallel processing) mode for reading and analyzing events.
- use existing utilities, such as initialization of constants and functions, histogram and N-tuple management, etc.

The philosophy of BASF is to provide a common analysis environment for Monte Carlo and data. This offers the user the advantage of analyzing and comparing real data to the simulation by using the very same software (reconstruction code, utility functions, data format), which provides automatic control over a potentially

[2]The multi-processing mode is developed based on a FPDA package [24].

important source of systematic error. A central event data management called PANTHER [25] is used for the derivation of structures and objects containing the event information. The user has access to a series of tables called MDST tables [26] that contain or combine information from the detector sub-systems on an event-by-event basis, in a user-friendly form. In the case of Monte Carlo studies, additional tables are obviously available to the user (containing information about the true nature of the event, independent of the traces in the detector: the "Truth Tables"). Apart of course from the origin of the events, this is the only important difference between Monte Carlo and real data.

Along with the tables containing hits from the sub-detectors, a set of generic use modules are available. Among other things, they automatically calculate and store several parameters: primary vertices, lists of charged and neutral particles, etc.; and implement more advanced functions, such as PID, vertexing, flavor tagging, etc.

3.3.2 The Monte Carlo simulator

There are problems that can be solved best or only by the Monte Carlo method. The study of a high energy physics detector's behavior belongs to the latter category. The motive for the simulation of the physical geometry of the sub-detectors and their expected responses lies in the complexity of the system: many components, many particles, many interactions, varying conditions. Some of the modes of interest are very rare and leave traces in the detector similar to those from other, maybe less interesting, processes. A Monte Carlo simulation can show us what we should expect and hopefully help us realize ways to increase the detector sensitivity. Another great advantage is that with the current CPU power available, we can simulate large numbers of "interactions" and get projections in the future for an equivalent of many years of data taking.

A Monte Carlo simulation has two main parts: The event generator, which simulates the physics of the collision and the subsequent decays of the daughter particles, and the detector simulation, which parameterizes its behavior and is obviously unique for every experiment.

We use the QQ (and QQ98) event generator program [27], which was developed by the CLEO collaboration [28] for the study of B mesons in the $\Upsilon(4S)$ resonance. The program setup includes a set of parameters that are controlled by the user, such as the electron and positron beam energies, the crossing angle, the decay products ($\Upsilon(4S)$, continuum, Bhabha, muon and tau pairs, etc.) and the number and the kind of allowed interactions. So, even though QQ was designed for a symmetric collider (CESR) and another experiment (CLEO), it has been adjusted to describe the Belle experiment. We can study specific interactions of interest by adjusting a "decay table" that defines the interactions and the final products, and can be modified by the user.

For the event generation of this analysis, a standard decay table has been used, based on early CLEO data on B decays, and modified accordingly, as new measurements have become available (QQ98). For $\Upsilon(4S)$ decays, charged and neutral B pairs are produced via a virtual photon, typically with a $\sin^2\theta$ angular distribution at their center of mass. The mixing parameter used is $x_d = 0.66$ and $x_d = 0.723$ for the generation of two large sets of Monte Carlo events. The continuum generation uses the LUND (JETSET 7.3) program [29], in which the subsequent hadronization process is based on the Lund string fragmentation model [30].

The detector is simulated by a full Monte Carlo (GEANT [31]) program called GSIM. The performance of the sub-detectors in the simulation is based on beam tests, cosmic ray runs, and finally, on- and off-resonance data. Discrepancies in efficiencies, tracking performance, or global positioning are monitored and the software library is periodically updated. A simpler and faster version of GSIM is available, the "fast" simulator, or FSIM [32], which uses parameterized detector responses and

is mainly invoked for debugging or less detailed studies.

The CPT-violating event generator that we use is a modified version of the official QQ software. Details on the developed subroutine can be found in the Appendix (Sec. A.2). Whereas the Monte Carlo samples used in the fit are generated with GSIM, the modified subroutine and the sensitivity of the analysis in detecting CPT-violating signals have been tested on FSIM.

Chapter 4

Analysis

4.1 The data set

The analysis is performed on data from Belle's Experiment 7 (January - July 2000) "on-resonance" runs. The total integrated luminosity is 5.12 fb^{-1}. The runs are listed in Table 4.1.

4.2 Hadronic event selection

First, the hadronic events are selected. The following criteria eliminate events from other sources, such as beam backgrounds, QED, continuum, etc.

- Use only "good" tracks. A good track is defined by

  $P_t > 0.1$ GeV/c, $|\Delta r| < 2.0$ cm, $|\Delta z| < 4.0$ cm,

  with Δr and Δz being the closest distance of a track to the z-axis and the z position of the closest point, respectively.

  For charged tracks a pion mass hypothesis is used.

- Use only "good" clusters. A good cluster is defined by

  $E > 0.1$ GeV

Table 4.1: "On-resonance" runs from Experiment 7 (January - July 2000) used in this analysis.

Dates	Run #	$\int \mathcal{L}\, dt$ (pb^{-1})
Jan 15 - Feb 14	0006-0243	245.5
Feb 14 - Feb 21	0244-0454	262.9
Feb 21 - Feb 22	0455-0475	28.0
Feb 25 - Mar 04	0537-0727	290.7
Mar 10 - Mar 17	0729-0827	179.8
Mar 31 - Apr 08	0858-0982	239.1
Apr 08 - Apr 16	0983-1138	353.5
Apr 16 - Apr 24	1141-1229	358.3
Apr 24 - Apr 25	1231-1248	80.9
Apr 25 - Apr 26	1251-1256	15.6
Apr 26 - Apr 30	1257-1320	169.3
Apr 30 - May 05	1321-1413	362.3
May 05 - May 07	1414-1439	93.6
May 07 - May 09	1441-1486	188.1
May 13 - May 20	1538-1639	243.6
May 20 - May 23	1641-1746	234.4
May 23 - May 27	1752-1825	252.1
May 27 - May 31	1826-1897	250.4
Jun 05 - Jun 11	1913-1996	260.2
Jun 11 - Jun 14	1997-2059	210.7
Jun 14 - Jun 18	2063-2123	234.8
Jun 18 - Jun 20	2124-2156	135.3
Jun 27 - Jul 04	2294-2430	289.8
Jul 04 - Jul 07	2432-2515	140.2
Total		5119.1

- A primary event vertex must have radial (V_r) and z (V_z) components satisfying

  $V_r < 1.5$ cm, $|V_z| < 3.5$ cm

- Number of good tracks ≥ 5

- Total visible energy detected by the ECL and CDC:

$$\sum_{\substack{\text{all charged}\\ \text{tracks}}} p^i + \sum_{\substack{\text{all cluster}\\ \text{energy}}} E^i \geq 50\% \text{ of CM energy}$$

- z-component of total visible energy:

$$\left| \sum_{\substack{\text{all charged}\\ \text{tracks}}} p_z^i \right| + \sum_{\substack{\text{all cluster}\\ \text{energy}}} E_z^i \leq 30\% \text{ of CM energy}$$

- Total cluster energy:

$$2.5\% \text{ of CM energy} \leq \sum_{\substack{\text{all cluster}\\ \text{energy}}} E^i \leq 90.0\% \text{ of CM energy}$$

4.3 Lepton identification

In this analysis, electrons and muons have been used for B flavor tagging. Their PID performance is given separately.

4.3.1 Electrons

For electron[1] ID we use a likelihood function, which returns the probability that the detected particle is an electron. The input parameters of this function are:

- energy to momentum ratio: E/p,

- mean energy loss rate in CDC: dE/dx,

[1]The term "electrons", throughout the analysis, stands for both electrons and positrons.

- information about the cluster energy, position and shape of the electromagnetic shower in ECL,
- matching between the track in the CDC and the cluster in the ECL, and
- TOF and ACC hit information for suppression of hadrons mimicking electrons producing fake leptons.

Optimization of the above parameters yields a $\sim 90\%$ ID efficiency for electrons and a $\sim 0.5\%$ misidentification probability for hadron tracks with $p > 1$ GeV/c.

A more detailed description of the electron likelihood function and its performance is given in Ref. [33].

4.3.2 Muons

For muon ID, the process is the following: we first find the track in the CDC, and extrapolate it to the KLM. We estimate the expected penetration in the KLM layers, given the (measured) momentum of the particle. We construct a variable that examines

- the difference between the expected and the actual penetration, and
- the χ^2 of the distances of the hits from the fitted track.

A muon ID is assigned if the track is not consistent with hadron track patterns. The efficiency is $\sim 85\%$ for muon tracks and the misidentification probability $\sim 2\%$ for particles with $p > 1$ GeV/c.

A more detailed description of the muon ID can be found in Ref. [34].

4.4 Cuts

Once an event is classified as "hadronic" (Sec. 4.2), but before it is tested with the "dilepton candidate" selection cuts, there are still two criteria it must satisfy. These

are:

1. the cutoff on R_2[2] (Fox-Wolfram second moment), mainly for continuum and Bhabha, muon and tau pair rejection. The event is rejected if $R_2 > 0.7$

2. the J/ψ veto: charmonium events and primary dileptons are mutually exclusive. The invariant mass $M_{\rm inv} = M_{\ell^+\ell^-}$ is therefore calculated for all combinations of each identified lepton of the event and all oppositely charged tracks (without PID). The mass window is somewhat looser for the low edge if the lepton is an electron or a positron, to account for radiative photons. The event is rejected if at least one pair is found with

   $|M_{J/\psi} - M_{\mu^+\mu^-}| < 50$ MeV/c^2, or

   $M_{J/\psi} - 150$ MeV/$c^2 < M_{e^+e^-} < M_{J/\psi} + 50$ MeV/c^2

   Monte Carlo studies show that the J/ψ veto discards $\sim 2.0\%$ of B^0B^0 or $\overline{B}^0\overline{B}^0$ and $\sim 3.6\%$ of $B^0\overline{B}^0$ pairs that have generated primary lepton pairs. This cut efficiency difference is taken into account.

From the list of electrons and positrons of the event, the ones that are consistent with pair production are rejected. To this end, the invariant mass of each particle identified as electron or positron and all oppositely charged tracks (without PID) is calculated. The track is rejected if $|M_{e^+e^-}| < 100$ MeV/c^2. Monte Carlo studies show that with the $\gamma \to e^+e^-$ conversion cut we lose $\sim 0.2\%$ of the primary lepton pairs.

All remaining leptons of the event are ordered according to the magnitude of their momentum at the $\Upsilon(4S)$ rest frame. Dilepton event candidates must have at

[2] R_2 is defined as

$$R_2 = \frac{H_2}{H_0}, \quad H_m = \frac{\sum_{i,j} |\vec{p_i}|\,|\vec{p_j}|\,P_m(\cos\theta_{ij})}{(\sum_i E_i)^2}$$

R_2 is a measure of the "jettiness" of the event, i.e. discriminates between $\Upsilon(4S)$ and continuum events. See Ref. [35].

least two particles identified as leptons. The leptons with the highest and the second highest momenta in the CM are selected as "dileptons".

The final set of cuts is applied on the selected dileptons:

3. Lepton momentum threshold at CM (mainly for suppression of fake and secondary leptons):

   $p_\ell^* > 1.1$ GeV/c (applied on both leptons)

4. Angle between lepton momenta directions at the CM (mainly for suppression of fake and secondary leptons):

   $-0.80 < \cos\theta^*(\ell_1, \ell_2) < 0.95$

5. Lepton momentum cutoff at CM (mainly for suppression of continuum):

   $p_\ell^* < 2.3$ GeV/c (applied on both leptons)

6. z-decay vertex differences (for suppression of beam backgrounds):

   $|\Delta z| \equiv |z_{\ell_1} - z_{\ell_2}| < 1850\ \mu$m

7. r and z distances of decay vertex from interaction point[3] (for suppression of beam backgrounds):

   $|\Delta r^{\mathrm{IP}}| < 0.05$ cm
   $|\Delta z^{\mathrm{IP}}| < 2.00$ cm (applied on both leptons)

8. # of hits in SVD (r, ϕ) and SVD (z) layer strips (for better vertex resolution):

   $N_{\mathrm{SVD}(r,\phi)} \geq 1$ and $N_{\mathrm{SVD}(z)} \geq 2$ (applied on both leptons)

9. Lepton polar angle in the laboratory frame (for better vertex resolution):

   $30^\circ < \theta_\ell^{\mathrm{lab}} < 135^\circ$ (applied on both leptons)

[3]The description of the run-by-run interaction point is given in Sec. 4.5.

The above values of the cuts have been optimized based on Monte Carlo studies.

A set of kinematic parameter distributions is plotted for the selected events. The expected distributions for various sources from a Monte Carlo simulation are superimposed on the same plots.

- Fig. 4.1 shows the lepton momenta distributions of the dilepton events in the $\Upsilon(4S)$ rest frame. The leptons from $\Upsilon(4S)$ events are separated into primary,
• Fig. 4.1 shows the lepton momenta distributions of the dilepton events in the
secondary and fakes. The continuum contribution is shown separately.

- Fig. 4.2 shows the distributions of R_2 (Fox-Wolfram second moment), calculated at the laboratory frame (top), and the opening angle between the two leptons at the $\Upsilon(4S)$ rest frame (bottom). The latter plot reveals the expected partial correlation for events with at least one fake or secondary lepton, and continuum. On the other hand, primary lepton pairs show no correlation.

- Fig. 4.3 shows the invariant mass of the dileptons. The J/ψ mass cut is apparent.

Applying the cuts listed above yields samples of 7429 same-sign (SS) and 35646 opposite-sign (OS) dilepton events.

4.5 Vertexing and proper time differences

The proper time difference can be calculated from the z-component of the distance between the decay vertices of the B meson pair, Δz. Using the approximation that the B mesons are at rest in the $\Upsilon(4S)$ frame we have[4]

$$t_{\text{proper}} \simeq t_{\Upsilon(4S)} \tag{4.1}$$

And therefore

$$\Delta z = \gamma(\Delta z_{\text{proper}} + \beta c\, \Delta t_{\text{proper}}) \simeq \beta\gamma c\, \Delta t_{\text{proper}} \implies \Delta t_{\text{proper}} \simeq \frac{\Delta z}{\beta\gamma c} \tag{4.2}$$

[4]We will use this approximation throughout this analysis.

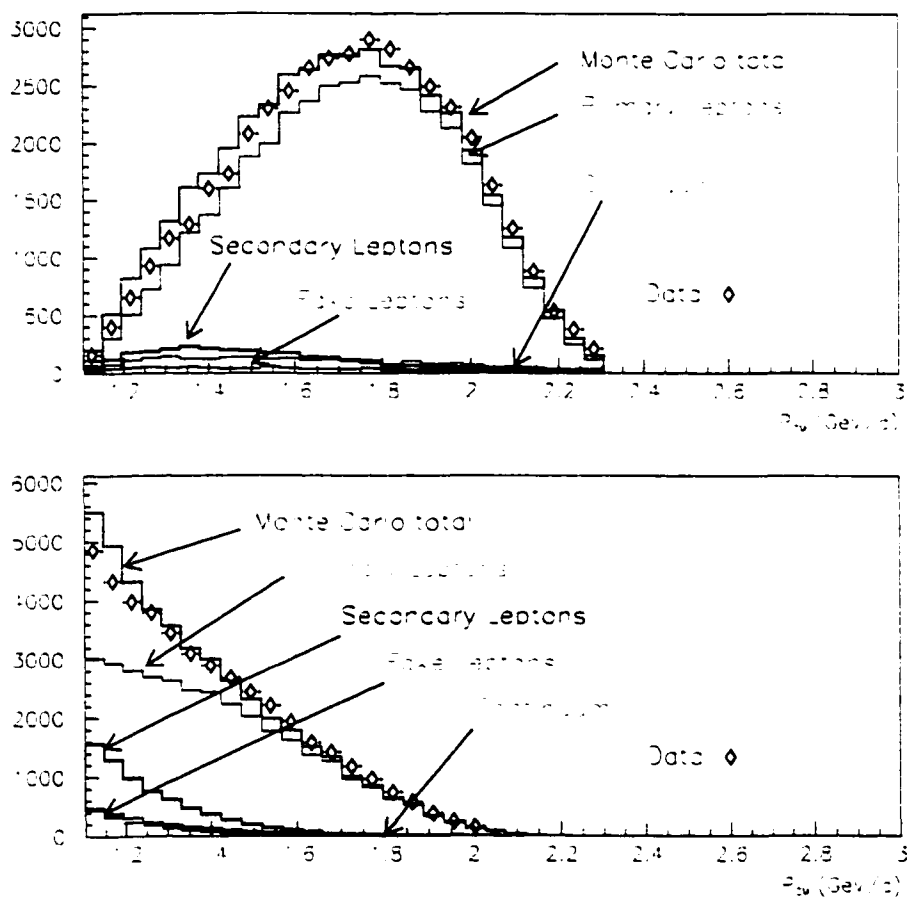

Figure 4.1: Comparison between Monte Carlo and data for the high (top) and the low (bottom) momentum distributions of the selected lepton pairs at the $\Upsilon(4S)$ rest frame.

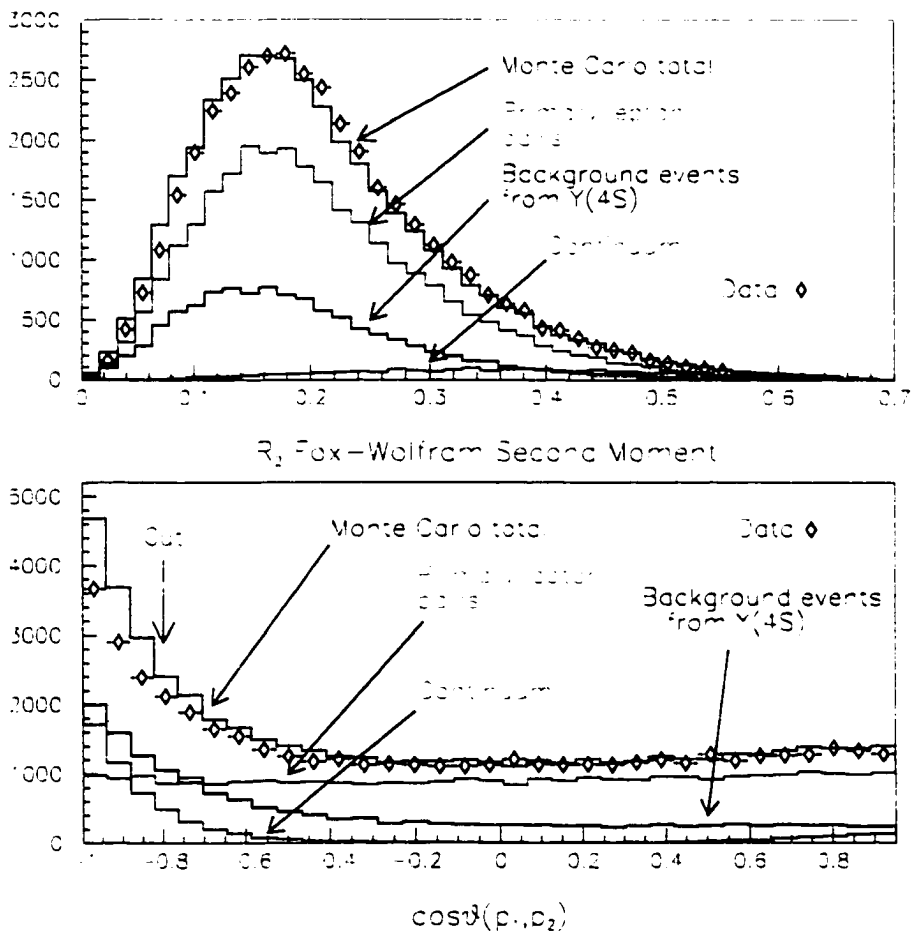

Figure 4.2: Comparison between Monte Carlo and data for R_2 (Fox-Wolfram second moment) (top) and $\cos\theta^*(\ell_1, \ell_2)$ (bottom) distributions for the selected dilepton events. R_2 is calculated in the laboratory frame and $\cos\theta^*(\ell_1, \ell_2)$ is given at the $\Upsilon(4S)$ rest frame. The $\Upsilon(4S)$ background is defined as events with at least one secondary or fake lepton.

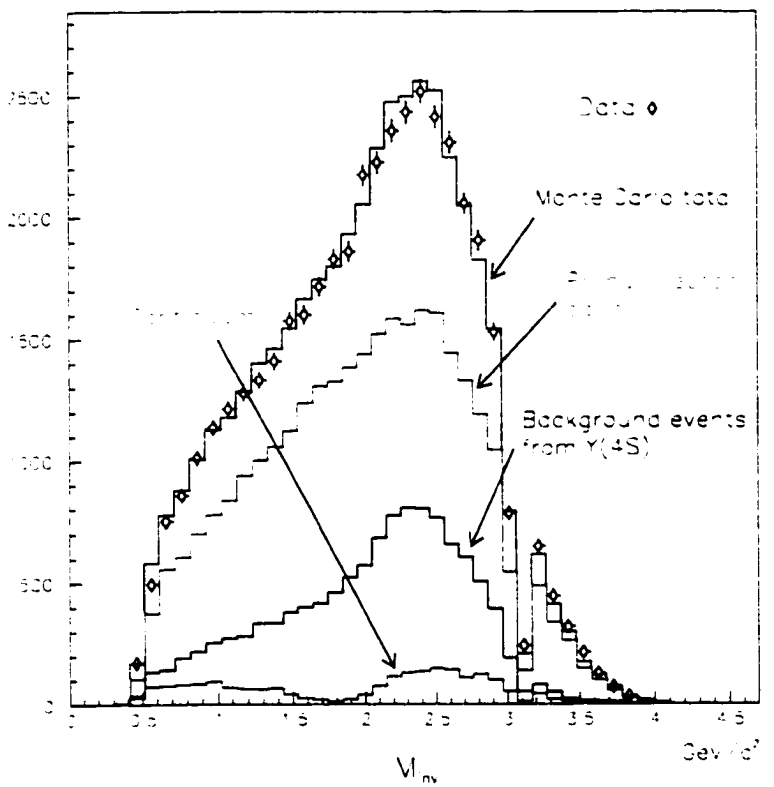

Figure 4.3: Comparison between Monte Carlo and data for the dilepton invariant mass of the selected events. The $\Upsilon(4S)$ background is defined as events with at least one secondary or fake lepton.

where $\beta\gamma = 0.425$ is the Lorentz boost factor for the Belle experiment. The effect of the B motion in the $\Upsilon(4S)$ rest frame and the systematic error associated with this approximation are discussed in Sec. 4.13.

The error in Δt_{proper}[5] depends on the accuracy of the determination of the decay vertices of the B pairs.

In the dilepton mode there is no reconstruction of the event or vertexing, with the strict definition of the term. Instead, for the B decay vertex determination, we use the intersection of the fitted lepton track with a three dimensional ellipsoid known to contain the B meson decay vertices [36]. The dimensions of this volume ($\sigma_x \sim 100-120\ \mu$m, $\sigma_y \sim 20\ \mu$m, $\sigma_z \sim 2000-3000\ \mu$m) are calculated by convolving the interaction point (IP) "profile" with the average unboosted[6] flight length of the B meson, σ_B:

$$\begin{aligned}
\sigma_x^2 &\simeq (\sigma_x^{\text{IP}})^2 + \sigma_B^2 \\
\sigma_y^2 &\simeq (\sigma_y^{\text{IP}})^2 + \sigma_B^2 \\
\sigma_z^2 &\simeq (\sigma_z^{\text{IP}})^2 + \sigma_B^2
\end{aligned}$$

The center and the dimensions of the IP profile ($\sigma_x^{\text{IP}} \sim 100 - 120\ \mu$m, $\sigma_y^{\text{IP}} \sim 5\ \mu$m, $\sigma_z^{\text{IP}} \sim 2000 - 3000\ \mu$m) are determined by the primary vertex position distribution on a run-by-run basis (a few thousand events). The value of σ_B used is 20 μm, based on a Monte Carlo estimate. The B motion has a small effect on the x width of the ellipsoid, a more substantial one for the y component[7], and practically no effect on the z direction. Given the above dimensions, it becomes clear that the "vertexing" boils down to, for all practical purposes, the intersection of the leptonic track with

[5]Since only proper times are mentioned in this paper, we drop the index "proper" from now on. The convention is that distances measured in the laboratory correspond to times in the $\Upsilon(4S)$ rest frame ($B^0\overline{B}^0$ CM).

[6]This correction is meaningful only for the x and y directions. The motion of the B mesons in the z direction is negligible compared to the $\beta\gamma = 0.425$ boost from the $\Upsilon(4S)$ rest frame to the laboratory frame.

[7]The vertex measurement resolution is not, of course, as good as σ_y^{IP}: the distribution of the y component of the vertex defines the resolution.

the z axis, as this is determined for every run. The error in the IP determination further dilutes the precision of the decay vertex measurement. The smearing in Δz reduces the analysis sensitivity.

4.6 Strategy

In general, an event can be identified as "signal" or "background". A signal event is one that contains a pair of primary leptons. Background is everything that is not signal, i.e. events with at least one non-primary lepton (secondary or fake). The shape of the signal Δz distribution contains explicit information about the parameters of the fit[8]. While signal events are always correctly tagged—i.e. it is always correctly determined whether the the B pair decayed as $B^0\overline{B}^0$ ($\ell^+\ell^-$) or $B^0B^0/\overline{B}^0\overline{B}^0$ ($\ell^+\ell^+/\ell^-\ell^-$)—background events could be either correctly or wrongly tagged. If we can determine the relative numbers of the background events with correct or wrong tagging from the (different) shapes of their distributions, we can actually use this information to help determine χ_d. On a more advanced level, we would like the fit to "recognize" a Δm dependence in the background distributions. Since this problem is too hard to solve analytically, is is done numerically. In general, all distributions are normalized, whether signal or background.

In the above, we have described the possible scenarios for $\Upsilon(4S)$ events. There is a special category of background event that does not provide information for Δm or $\cos\theta$: the continuum events.

4.6.1 General event classification

The analysis is structured in the following logic steps:

1. After all cuts have been applied, we separate the selected lepton pairs in two bins: pairs of same-sign (SS) and opposite-sign (OS). The indices "SS" and

[8]In principle, so do the background distributions, but this is a much harder problem.

"OS" refer to the *observed* tag, i.e. a quantity that is determined experimentally by the sign of charge of the (putative) primary leptons.

2. A selected event can originate from

   - B^+B^- (labeled as "chd"[9])
   - $B^0\overline{B}^0$ ("unm")
   - B^0B^0 or $\overline{B}^0\overline{B}^0$ ("mix"), or
   - continuum ("cnt")

   The indices "chd", "unm", "mix" and "cnt" refer to the "true" origin of events, in other words, information that can be retrieved only with access to the Truth Table of the Monte Carlo.

3. Events from B pairs (charged and neutral) and continuum are assigned probabilities of being selected as dilepton candidates. An event originating from

   - a B^+B^- pair can give
     - a pair of primary leptons (with opposite sign) with probability $\epsilon^{\rm chd}_{\ell^+\ell^-}$,
     - an opposite-sign lepton pair, with at least one of them being fake or secondary, and with probability $\epsilon^{\rm chd}_{OS}$ (correct tagging), or
     - a same-sign lepton pair, with at least one of them being fake or secondary, and with probability $\epsilon^{\rm chd}_{SS}$ (wrong tagging).
   - a $B^0\overline{B}^0$ pair can give
     - a pair of primary leptons (with opposite sign) with probability $\epsilon^{\rm unm}_{\ell^+\ell^-}$,
     - an opposite-sign lepton pair, with at least one of them being fake or secondary, and with probability $\epsilon^{\rm unm}_{OS}$ (correct tagging), or

[9] We cannot tell neutral from charged B mesons based only on a high momentum lepton; the two terms, however, have different lifetimes and, therefore, Δz distributions.

- a same-sign lepton pair, with at least one of them being fake or secondary, and with probability $\epsilon_{SS}^{\text{unm}}$ (wrong tagging).

- a B^0B^0 or $\overline{B}^0\overline{B}^0$ pair can give
  - a pair of primary leptons (with the same sign) with probability $\epsilon_{\ell^\pm\ell^\pm}^{\text{mix}}$,
  - a same-sign lepton pair, with at least one of them being fake or secondary, and with probability $\epsilon_{SS}^{\text{mix}}$ (correct tagging), or
  - an opposite-sign lepton pair, with at least one of them being fake or secondary, and with probability $\epsilon_{OS}^{\text{mix}}$ (wrong tagging).

- A continuum event gan give
  - a same-sign (true or fake) lepton pair with probability $\epsilon_{SS}^{\text{cnt}}$, and
  - an opposite-sign (true or fake) lepton pair with probability $\epsilon_{OS}^{\text{cnt}}$.

The above probabilities ϵ's represent selection cut efficiencies, in other words, they are absolute efficiencies with the respect to the number of generated events for the specific category.

4. By defining

$$R \equiv \frac{f_\pm}{f_0} \equiv \frac{\mathcal{B}(\Upsilon(4S) \to B^+B^-)}{\mathcal{B}(\Upsilon(4S) \to B^0\overline{B}^0)} \tag{4.3}$$

and by assuming [37]

$$f_0 + f_\pm = 1 \tag{4.4}$$

we can write the fractions of neutral and charged B pairs as

$$f_0 = \frac{1}{1+R}$$

and

$$f_\pm = \frac{R}{1+R}$$

In this analysis we use the measurement

$$\frac{f_\pm}{f_0} = 1.07 \pm 0.09 \quad [38] \tag{4.5}$$

If we have $N_{\Upsilon(4S)}$ $\Upsilon(4S)$ events, then the number of generated

- B^+B^- pairs is $f_\pm N_{\Upsilon(4S)}$.
- $B^0\overline{B}^0$ pairs is $f_0 N_{\Upsilon(4S)}(1-\chi_d)$.
- B^0B^0 or $\overline{B}^0\overline{B}^0$ pairs is $f_0 N_{\Upsilon(4S)}\chi_d$.

χ_d is a function of x_d (if CPT is not conserved, of θ as well; see Eqs. (2.22) and (2.28)).

5. To describe the SS and OS Δz distributions we divide the events into the following categories:

   - Lepton pairs from $\Upsilon(4S)$ events that consist of
     - two primary leptons, which we model *analytically* by smearing a theoretical expression with explicit lifetime and mixing (and, in general, CPT violation) dependence. The shape of the distribution contains information about the parameters to be fitted: $P^{00}_{\ell^\pm\ell^\pm}(\Delta t_{\rm smear})$, $P^{00}_{\ell^+\ell^-}(\Delta t_{\rm smear})$, $P^{+-}_{\ell^+\ell^-}(\Delta t_{\rm smear})$; the subscripts "00", "+−" stand for neutral and charged B pairs, respectively.
     - (at least) one fake or secondary lepton, which we model *numerically* by using look-up tables from Monte Carlo. At this point, the shape of the distribution contains information relevant to the mixing, even though not explicitly. The disadvantage of this method is that for this part we have to rely on the Monte Carlo: $B^{\beta}_{\alpha}(\Delta t_{\rm smear})$, with β = "mix", "unm", "chd" and α = "SS", "OS". See discussion in Sec. 4.8.2 about the modeling of the background distributions.
   - contributions from continuum events: Again, we use lookup tables from Monte Carlo: $B^{\rm cnt}_{SS}(\Delta t_{\rm smear})$, $B^{\rm cnt}_{OS}(\Delta t_{\rm smear})$.

For study of the systematic errors, we need to define one last set of parameters: the selection efficiencies for primary lepton pairs, η. The link between the efficiencies ϵ and η for primary dileptons is the semileptonic branching fractions for B^0 and $B^\pm$, b_0 and $b_\pm$, respectively.

$$\epsilon^{\text{mix}}_{\ell^\pm\ell^\pm} = \eta^{\text{mix}}_{\ell^\pm\ell^\pm}\, b_0^2 \tag{4.6}$$

$$\epsilon^{\text{unm}}_{\ell^+\ell^-} \equiv \eta^{\text{unm}}_{\ell^+\ell^-}\, b_0^2 \tag{4.7}$$

$$\epsilon^{\text{chd}}_{\ell^+\ell^-} \equiv \eta^{\text{chd}}_{\ell^+\ell^-}\, b_\pm^2 \tag{4.8}$$

We now can separate the hadronic and selection cut efficiencies from the uncertainty in the semileptonic branching fraction measurements. We will actually go one step further and relate b_0 and $b_\pm$.

We will use

$$\Gamma(B^\pm \longrightarrow \ell^\pm X\nu) \simeq \Gamma(B^0 \longrightarrow \ell^\pm X\nu) \tag{4.9}$$

The arguments for this approximation are

- the isospin symmetry between B^+ (B^-) and B^0 ($\bar{B}^0$),
- the near equality of the $B^\pm$ and B^0 masses, and
- the theoretical prejudice that the light spectator quark (u for B^+ and d for B^0) should not affect the partial decay width.

Then, the semileptonic branching ratios are just

$$b^\pm \equiv \frac{\Gamma(B^\pm \longrightarrow \ell^\pm X\nu)}{\Gamma_{B^\pm}} \quad \text{and} \quad b^0 \equiv \frac{\Gamma(B^0 \longrightarrow \ell^\pm X\nu)}{\Gamma_{B^0}} \tag{4.10}$$

or

$$\frac{b^\pm}{b^0} = \frac{\tau_{B^\pm}}{\tau_{B^0}} = 1.04 \pm 0.04 \quad [5] \tag{4.11}$$

The lifetime ratio is measured with better accuracy than the two semileptonic ratios measurements separately:

$$\frac{b^{\pm}}{b^0} = \frac{(10.3 \pm 0.9)\%}{(10.5 \pm 0.8)\%} = 0.981 \pm 0.094 \quad [5] \tag{4.12}$$

4.6.2 Breaking the distributions into pieces

By using the above classification of events we can separate the different contributions to the two histograms: SS and OS. $N_{q\bar{q}}$ is the number of generated continuum events.

- <u>SS:</u>
  - B^0B^0 or $\overline{B}^0\overline{B}^0$ pairs that gave a (detected) primary lepton pair event:

$$\epsilon^{\text{mix}}_{\ell^\pm\ell^\pm}\, P^{00}_{\ell^\pm\ell^\pm}(\Delta t_{\text{smear}})\, f_0\, \chi_d\, N_{\Upsilon(4S)}$$

  - B^0B^0 or $\overline{B}^0\overline{B}^0$ pairs that gave a background event with correct tagging:

$$\epsilon^{\text{mix}}_{SS}\, B^{\text{mix}}_{SS}(\Delta t_{\text{smear}})\, f_0\, \chi_d\, N_{\Upsilon(4S)}$$

  - $B^0\overline{B}^0$ pairs that gave a background event with wrong tagging:

$$\epsilon^{\text{unm}}_{SS}\, B^{\text{unm}}_{SS}(\Delta t_{\text{smear}})\, f_0\, (1-\chi_d)\, N_{\Upsilon(4S)}$$

  - B^+B^- pairs that gave a background event with wrong tagging:

$$\epsilon^{\text{chd}}_{SS}\, B^{\text{chd}}_{SS}(\Delta t_{\text{smear}})\, f_\pm\, N_{\Upsilon(4S)}$$

  - continuum that gave a same-sign event:

$$\epsilon^{\text{cnt}}_{SS}\, B^{\text{cnt}}_{SS}(\Delta t_{\text{smear}})\, N_{q\bar{q}}$$

- <u>OS:</u>
  - B^+B^- pairs that gave a (detected) primary lepton pair event:

$$\epsilon^{\text{chd}}_{\ell^+\ell^-}\, P^{+-}_{\ell^+\ell^-}(\Delta t_{\text{smear}})\, f_\pm\, N_{\Upsilon(4S)}$$

– $B^0\overline{B}^0$ pairs that gave a (detected) primary lepton pair event:

$$\epsilon^{\mathrm{unm}}_{\ell^+\ell^-}\, P^{00}_{\ell^+\ell^-}(\Delta t_{\mathrm{smear}})\, f_0\,(1-\chi_d)\, N_{\Upsilon(4S)}$$

– B^+B^- pairs that gave a background event with correct tagging:

$$\epsilon^{\mathrm{chd}}_{OS}\, B^{\mathrm{chd}}_{OS}(\Delta t_{\mathrm{smear}})\, f_\pm\, N_{\Upsilon(4S)}$$

– $B^0\overline{B}^0$ pairs that gave a background event with correct tagging:

$$\epsilon^{\mathrm{unm}}_{OS}\, B^{\mathrm{unm}}_{OS}(\Delta t_{\mathrm{smear}})\, f_0\,(1-\chi_d)\, N_{\Upsilon(4S)}$$

– B^0B^0 or $\overline{B}^0\overline{B}^0$ pairs that gave a background event with wrong tagging:

$$\epsilon^{\mathrm{mix}}_{OS}\, B^{\mathrm{mix}}_{OS}(\Delta t_{\mathrm{smear}})\, f_0\,\chi_d\, N_{\Upsilon(4S)}$$

$$\epsilon^{\mathrm{mix}}_{OS}\, B^{\mathrm{mix}}_{OS}(\Delta t_{\mathrm{smear}})\, f_0\,\chi_d\, N_{\Upsilon(4S)}$$

– continuum that gave an opposite-sign event:

$$\epsilon^{\mathrm{cnt}}_{OS}\, B^{\mathrm{cnt}}_{OS}(\Delta t_{\mathrm{smear}})\, N_{q\bar{q}}$$

Fig. 4.4 illustrates the separation of the proper time distributions in different contributions from mixed, unmixed, charged and continuum events.

4.7 Response function

One of the most important tasks in a time-dependent analysis is the accurate modeling or parameterization of the response function $g(\Delta t - \Delta t_{\mathrm{smear}})$. The response function is a measure of the typical deviation of the measured vertex or, equivalently, calculated proper time difference from the actual value $\Delta t \equiv \Delta t_{\text{"true"}}$; in other words, the vertex resolution of the detector. We will use a suitably normalized g to smear the theoretical expressions, trying to reproduce the distributions for the measured times.

$$P(\Delta t_{\mathrm{smear}}) \;=\; \int g(\Delta t_{\mathrm{smear}} - \Delta t)\, P_{\mathrm{theory}}(\Delta t)\, d(\Delta t) \qquad (4.13)$$

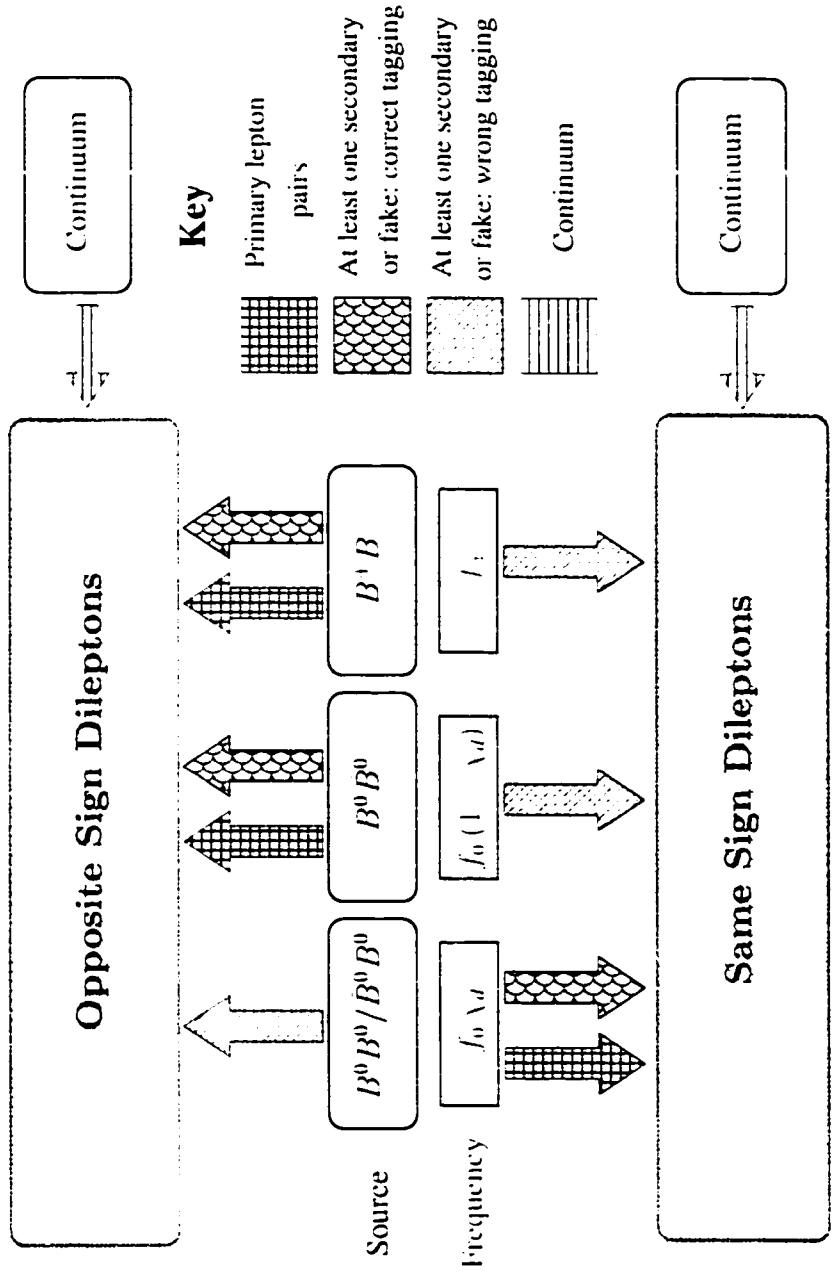

Figure 4.4: Schematic illustration of the origin of signal and background terms in the fitting and their contributions to the SS and OS histograms.

with

$$\int P(\Delta t_{\text{smear}})\, d(\Delta t_{\text{smear}}) \;=\; \int P_{\text{theory}}(\Delta t)\, d(\Delta t) \equiv 1 \qquad (4.14)$$

In principle the response function can be determined from the Monte Carlo. The problem is that for our data and the current level of analysis the vertex resolution is not as good as the Monte Carlo predicts, due to systematic errors not fully examined and understood yet. This results in larger-than-predicted RMS widths in the proper time distributions. The width of the actual response function is estimated to be larger than in the Monte Carlo by $\sim 10\%$[10]. More subtle discrepancies between Monte Carlo and data in the (less important) long tails of the response function are hard to estimate with the current statistics.

In order to solve this problem we have chosen to use the $\Delta z \equiv z_{\ell_1} - z_{\ell_2}$ distribution of $J/\psi \to \ell^+\ell^-$ decays from the same data runs. The J/ψ instantaneously decays to a lepton pair, so the Δz distribution of the J/ψ's daughters determined with the same method that we use for the dileptons (see Sec. 4.5)[11] should be a good approximation to the $g(\Delta t_{\text{smear}} - \Delta t)$ function.

In order to obtain a response function that is as realistic as possible, we select dileptons from candidate J/ψ decays from the same data runs, from events that pass the hadronic selection criteria (Sec. 4.2) and by applying the same cuts described in Sec. 4.4. The only exceptions are

- the J/ψ veto (cut #2), which was replaced by

  2′. $|M_{\ell^+\ell^-} - M_{J/\psi}| < 40$ MeV/c^2

  for obvious reasons, and

- the angle between the lepton tracks constraint (cut #4) which was omitted since it does not affect the vertex resolution and is used for suppression of

[10] The procedure for this calculation is described in Sec. 4.8.2.

[11] That is, by finding two *separate* z-vertices for the two lepton tracks.

kinematically correlated leptons. Unlike primary lepton pairs, leptons from J/ψ decays are strongly correlated.

The background events in the the mass window have, in general, different vertex resolution than the lepton pairs from J/ψ decays and, therefore, complicate the determination of the response function. The reason is that random leptons do not share a common "geometrical point" as their origin. They will therefore tend to degrade the vertex resolution. To correct for this, we follow the procedure outlined below.

The level of background contamination is estimated by fitting the dilepton invariant mass spectrum (2.75 GeV/c^2 - 3.50 GeV/c^2) to a Gaussian plus a linear function (Fig. 4.5). This method underestimates the number of J/ψ's by $\sim$ 2.7% due to the radiative tail events, which do not match the Gaussian model. After the correction, the estimated purity of the J/ψ sample is 87.4% and 82.4% for Monte Carlo and data, respectively.

For the Δz distribution of the background contained in the event sample, the $M_{\ell^+\ell^-} > 3.15$ GeV/c^2 region of the mass spectrum is used, since it does not contain radiative photon events. The distribution thus obtained is scaled to match the expected level of background contamination in the mass window, and is then subtracted from the Δz distribution of the $|M_{\ell^+\ell^-} - M_{J/\psi}| < 40$ MeV/c^2 region. The response function is the resulting Δz distribution.

To be precise, the J/ψ Δz distribution gives the $\Delta z_{\text{smear}} - \Delta z_{\text{"true"}} = \Delta z_{\text{smear}}$ deviation. This is equivalent to the $\Delta t_{\text{smear}} - \Delta t_{\text{"true"}}$ difference only if the B mesons are at rest at the $\Upsilon(4S)$ rest frame. This subtle point is discussed again in Sec. 4.13.

The hypothesis that we can use the J/ψ Δz distribution for the response function was tested using Monte Carlo: good agreement was achieved between the dilepton response function and the Δz distribution from J/ψ decays, and no significant difference in the fit results with the two methods was observed. Fig. 4.6 shows the very good matching between the two response functions.

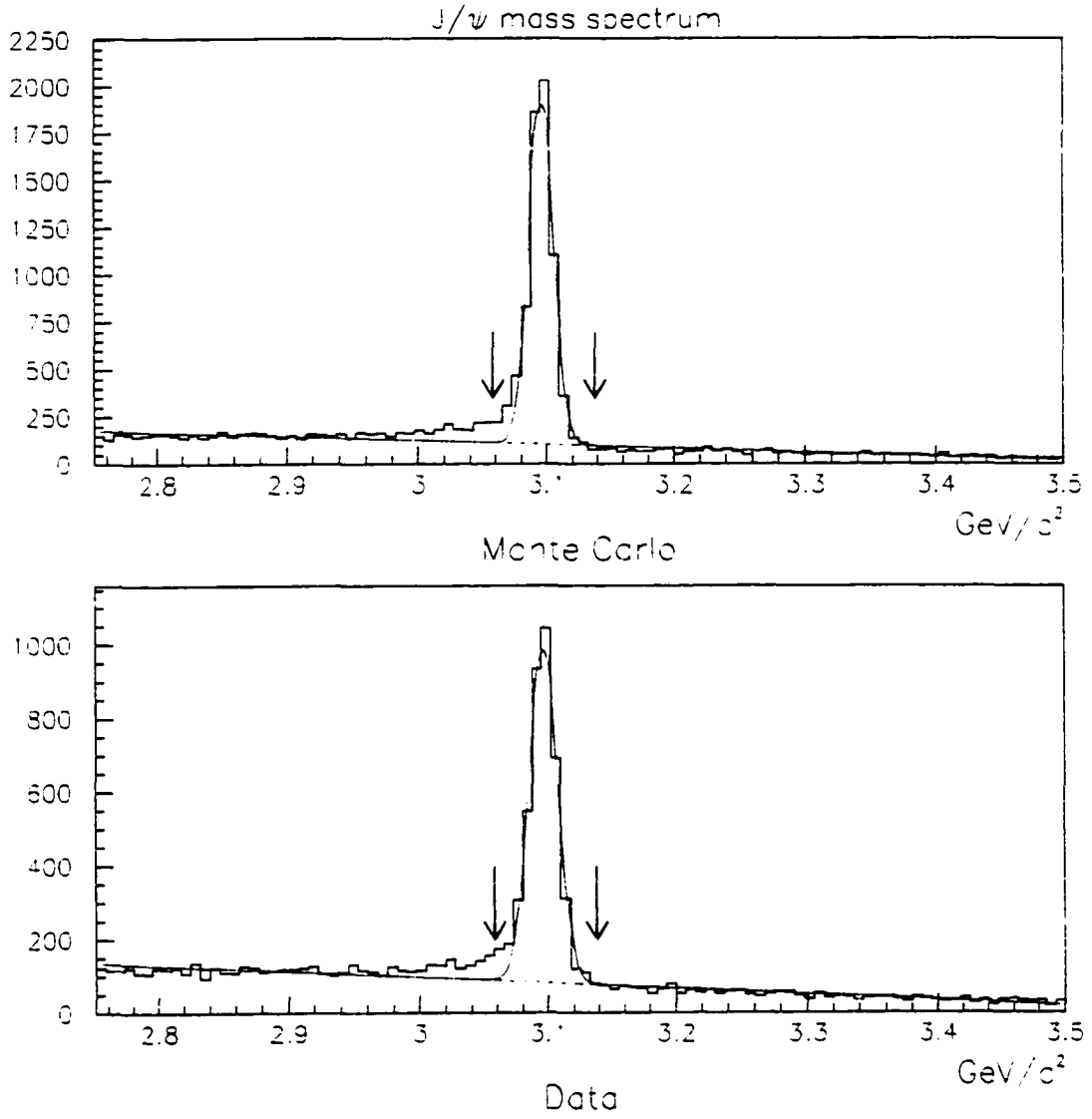

Figure 4.5: Invariant mass spectrum of dileptons for Monte Carlo (top) and data (bottom). The J/ψ peak is fitted by a Gaussian and the background by a linear form function. Candidate J/ψ's from the marked region are used for the response function.

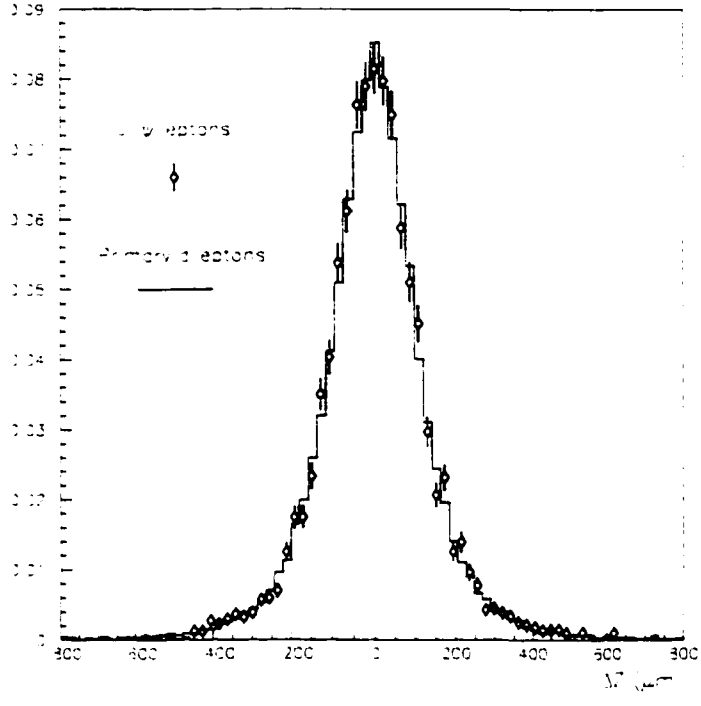

Figure 4.6: Response functions for primary dileptons (histogram) and J/ψ leptons (points) for Monte Carlo.

Fitting the J/ψ Δz distributions obtained with this method to a single Gaussian yields widths of $\sigma = 100\ \mu$m and $\sigma = 112\ \mu$m, for Monte Carlo and data, respectively. Since this is the Δz resolution, it follows that the vertex resolution for a single track is ~80 μm. The tracking performance discrepancy between Monte Carlo and data can be parameterized with an amount of extra "smearing" — specifically 50 μm — added quadratically to the expected vertex resolution: A convolution of the J/ψ Δz distribution for Monte Carlo with a single Gaussian of $\sigma = 50\ \mu$m gives a good fit to the corresponding distribution from data. Fig. 4.7 compares the distributions for Monte Carlo and data, with and without the extra smearing.

4.8 Modeling of background

For most of the pieces that go into the modeling of the background in our analysis we rely on the Monte Carlo. We categorize the different aspects of the Monte Carlo dependence in order to disentangle the different parts of the problem.

4.8.1 Sources of background - Time integrated quantities

- Secondary leptons: The major component of the background is the combination of a primary and a secondary lepton from a cascade decay ($c \to s$ or τ decay). Primary - secondary pairs from $B^0\overline{B}^0$ and B^+B^- usually end up in the SS dilepton sample. This is actually the majority of the SS events. When the secondary lepton is coming from the same B that generated the (detected) primary lepton, the lepton pair will typically be a (background) OS event (Fig. 2.2.(a) demonstrates this).

  Most of the time B mesons decay to D mesons, which, in turn, can give off leptons. The decay vertex of these (secondary by definition) leptons will be very displaced from the B decay vertex due to the non-negligible D lifetime. For this reason, secondary leptons produce a systematic shift when used for

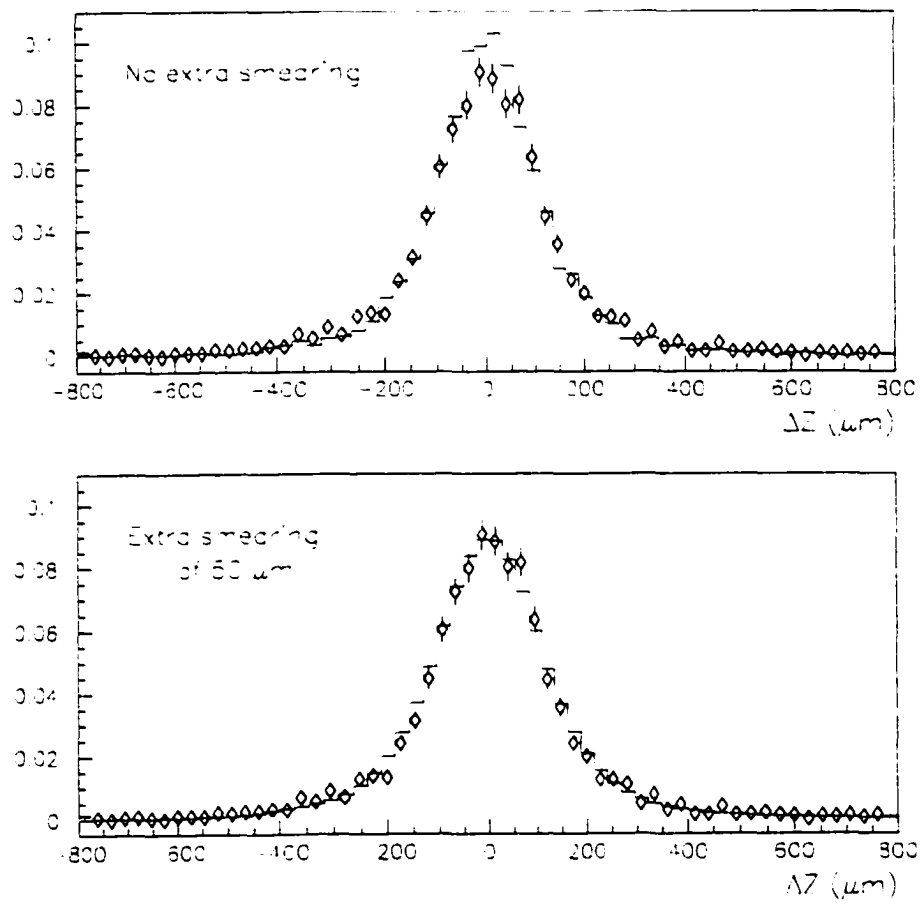

Figure 4.7: Δz distributions for dileptons from J/ψ decays for data (points) and Monte Carlo (histogram), without (top) and with extra smearing (bottom) for the Monte Carlo. The distributions are after the sideband background subtraction, as described in the text.

vertexing. As charged and neutral D mesons have significantly different lifetimes (~1.1 ps and 0.4 ps, respectively), leptons from $D^{\pm}$ and D^0 decays have unequal contributions to the mean shift. In general, they have different impacts on the Δz distributions. The relative fractions of $D^{\pm}$ and D^0 were scaled to match the branching ratios measured by CLEO [39]:

$$\mathcal{B}(B \longrightarrow D^{\pm} X) = (23.5 \pm 2.7)\%$$

$$\mathcal{B}(B \longrightarrow D^{0} X) = (63.6 \pm 2.3)\%$$

A similar effect is caused by electrons and muons from tau decays. Taus have a lifetime of ~0.3 ps and decay leptonically (to electrons and muons) 35% of the time. However, B mesons decay to τ leptons 4-5 times less often than to muons and electrons due to phase space differences [40]. So, in fact, taus have a small, albeit non negligible, contribution to the mean shift of the secondary lepton vertexing. We use again the Monte Carlo for the modeling of this component of the background.

- Fake leptons: Fake leptons are usually combined with a primary lepton and end up in either the SS or the OS sample.

  The fake rates of pions in electron and muon identification were measured using $K_S \rightarrow \pi^+\pi^-$ decays. We found the fake rates to be[12] (0.182±0.025)% for the electrons and (1.893±0.092)% for the muons. These numbers are smaller than in the Monte Carlo by (10.4±6.1)% and (2±22)%, respectively. The difference is comparable to the statistical errors of the fake rate measurements for both Monte Carlo and data. We correct for this small discrepancy.

- Continuum: The continuum contamination is estimated to be kept at relatively low levels. This is what the Monte Carlo shows and the off-resonance data

[12]Defined as: (# of fakes leptons)/(# of non-lepton tracks).

verifies.

We use 0.60 fb^{-1} of off-resonance data, taken at energies 50-60 MeV below the $\Upsilon(4S)$ peak, to determine the level of continuum in the selected hadronic sample (see Sec. 4.2). Based on a Monte Carlo simulation for the efficiencies of the hadronic cuts for $\Upsilon(4S)$ and continuum, and by using all on- and off-resonance data we determine [10]

$$\frac{\sigma(e^+e^- \to q\bar{q})}{\sigma(e^+e^- \to \Upsilon(4S))} = 3.66 \pm 0.37$$

This ratio is used for the admixture of $\Upsilon(4S)$ and continuum events in the Monte Carlo. The number of the continuum events from the off-resonance data surviving the cuts is consistent with the Monte Carlo predictions.

4.8.2 Δz distributions

In Sec. 4.6.1 (pg. 57) we described the difference in the modeling between signal and background in the fit. Whereas the signal terms are analytical functions, for the background distributions we use lookup tables, derived from the Monte Carlo and modified accordingly to the assumptions of the fit. This section describes how we partially overcome the problem of the non-flexibility and the static representation of the background distributions in a dynamic calculation, i.e. the fitting.

- Tracking performance: As we mentioned in Sec. 4.7, the vertex resolution is not as good as the Monte Carlo predicts. We therefore need to introduce some extra "smearing" in the background distributions to account for the larger RMS's. To this end, a convolution of the Δz distributions with a single Gaussian was tested. Specifically, the J/ψ Δz distribution for Monte Carlo was smeared with a single Gaussian with varying σ, and then compared with the J/ψ Δz distribution for data. For every value of σ, the confidence level of

the matching between the two distributions was obtained (see Fig. 4.8). This method determines the parameter σ of the convolving Gaussian to be

$$\sigma = 50^{+12}_{-18}\ \mu\text{m}$$

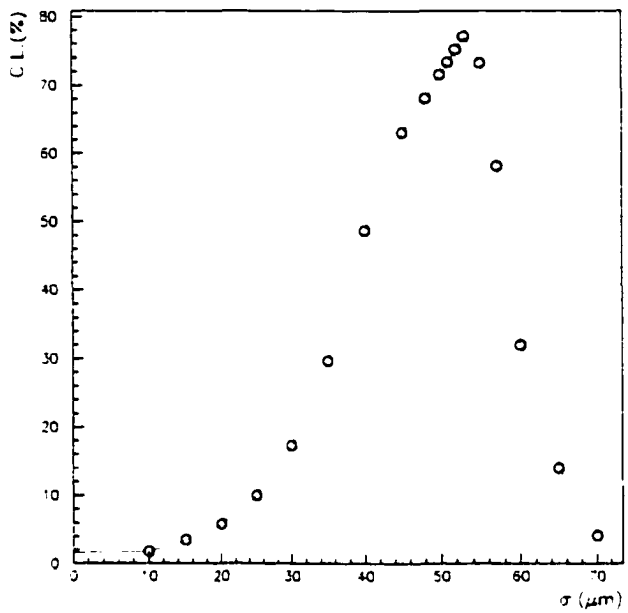

Figure 4.8: Confidence level of comparison between J/ψ Δz distribution from data and the one from Monte Carlo convolved with a single Gaussian of width σ, *vs.* σ.

Fig. 4.7 shows the comparison between Monte Carlo and data without (top) and with the extra smearing (bottom).

This procedure was verified with a different set of events. Instead of the J/ψ sample, the Δz distributions of lepton-kaon and lepton-pion pairs for data and the fake lepton Monte Carlo distributions (with the extra smearing) were compared. Good agreement between the two sets of distributions was achieved. This verifies that a convolution with a single Gaussian is a satisfactory method for getting the background distributions from the Monte Carlo with a vertex resolution similar to that of the data.

We do not use the lepton-kaon and lepton-pion pair Δz distributions from the data for the fake lepton background distributions in the fit since the method that we use in this analysis requires different contributions from mixed (B^0B^0 and $\bar{B}^0\bar{B}^0$), unmixed ($B^0\bar{B}^0$) or charged (B^+B^-) B mesons. Only Monte Carlo can provide this information.

- Δm dependence: The background distributions have an implicit Δm dependence that becomes apparent when we compare distributions generated with different values of Δm (see Fig. 4.9). Since the background is a sum of contributions from different sources (secondary or fake leptons, correct or wrong tagging, particles from the same or different B mesons), which have very different vertex resolution, it is very difficult to encode the mixing dependence and derive an analytical expression. Despite the small differences between the Monte Carlo background distributions generated with $\Delta m = 0.423$ ps^{-1} or $\Delta m = 0.464$ ps^{-1} (shown in Fig. 4.9), the choice of the Δm value for the background distributions used in the analysis has an important impact on the fit. In other words, if the "wrong" background distribution is given to the fit, we end up with a non-negligible systematic shift in the measurement.

  One way of solving this problem is to give the fitting algorithm a choice of background shape. We use a linear interpolation of background distributions generated with $\Delta m = 0.423$ ps^{-1} and $\Delta m = 0.464$ ps^{-1}:

$$B(\Delta m; \Delta z) = (1-\gamma)\, B(\Delta m = 0.423 \text{ ps}^{-1}; \Delta z) \;+\; \gamma\, B(\Delta m = 0.464 \text{ ps}^{-1}; \Delta z)$$

  where

$$\gamma \equiv \frac{\Delta m - 0.423}{0.464 - 0.423} \qquad (\Delta m \text{ in ps}^{-1})$$

  Although the relationship between γ and Δm is not necessarily strictly linear, for Δm's over a small enough interval, the linearity will be sufficiently good. Monte Carlo studies show that this is the case for the [0.423 ps^{-1}, 0.464 ps^{-1}] interval employed in this analysis.

With this method we allow the fitting to take into account small variations in Δm and adjust accordingly the shape of the background distributions in "real-time", i.e. in a dynamic way. Furthermore, with this manipulation we explicitly involve Δm in the background distributions (through $\tau = \tau(\Delta m)$), which actually serves to increase the fit's sensitivity to Δm, and indirectly the CPT violating parameters.

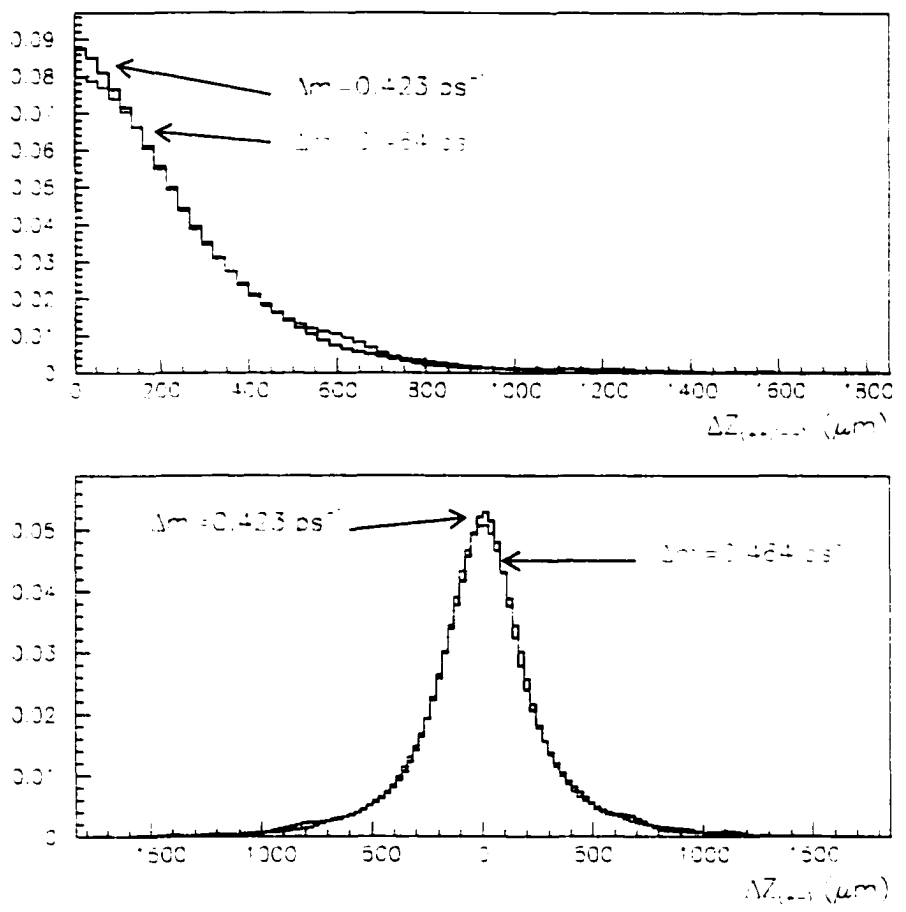

Figure 4.9: Monte Carlo background distributions from neutral B pairs, generated with $\Delta m = 0.423$ ps^{-1} and $\Delta m = 0.464$ ps^{-1} for SS (top) and OS (bottom) dileptons.

- B lifetime dependence: In this dilepton analysis we do not separately measure

the charged and neutral B lifetimes since without partially reconstructing the B mesons we cannot tell charged B pairs from neutral B pairs. Moreover, the dilepton sample is contaminated with backgrounds arising from secondary and fake leptons. However, using a value for the ratio $\tau_{B^\pm}/\tau_{B^0}$ from other measurements we can fit for the B^0 lifetime. This procedure was carried out mainly to crosscheck the modeling of the vertexing.

The B lifetime defines the steepness of the exponential form of the proper time distributions. Therefore, the B lifetime dependence of the background distributions is probably more important than that for Δm. We initially assume that the B lifetimes are measured with adequate precision. The study of the systematic errors, however, shows that the uncertainty in the B lifetime measurements is the main limitation to accurate Δm and $\mathrm{Re}(\cos\theta)$ measurements. Therefore, when we vary the lifetimes in the analytical expressions of the signal terms for the determination of the relevant systematic shift, we also have to do so for the background distributions. We do this "off-line", i.e. the B lifetime dependence of the background is static and does not change during the fit.

- Continuum: Because of the poor statistics of the off-resonance data, we rely on the Monte Carlo for the continuum Δz distributions in the dilepton sample. The problem of the vertex resolution difference between Monte Carlo and data is dealt with in the same way that it is for the $\Upsilon(4S)$ background: through a convolution with a single Gaussian of $\sigma = 50\ \mu$m.

4.9 Signal and background populations

We calculate the expected numbers for signal and background terms in the dilepton sample for the cuts discussed in Secs. 4.2 and 4.4. For this Monte Carlo study we use

3M charged and 3M neutral B events, generated with $x_d = 0.723$ ($\Delta m = 0.464$ ps^{-1}), and 24M continuum events.

For this calculation we use the following measurements

- $f_\pm/f_0 = 1.07 \pm 0.09$ [38]
- $b_\pm/b_0 = \tau_{B^\pm}/\tau_{B^0} = 1.04 \pm 0.04$ (Sec. 4.6)
- $\mathcal{B}(B \longrightarrow D^\pm X) = (23.5 \pm 2.7)\%$
  $\mathcal{B}(B \longrightarrow D^0 X) = (63.6 \pm 2.3)\%$ [39]
- Muon fake rate : $(1.893 \pm 0.092)\%$
  Electron fake rate : $(0.182 \pm 0.025)\%$ (Sec. 4.8.1)
- $\dfrac{\sigma(e^+e^- \to q\bar{q})}{\sigma(e^+e^- \to \Upsilon(4S))} = 3.66 \pm 0.37$ (Sec. 4.8.1)

and scale accordingly the different terms when there is a discrepancy with the Monte Carlo.

Table 4.2 lists the event categories in the dilepton sample as combinations of primary, secondary and fake leptons (for $\Upsilon(4S)$ events), and continuum, and their relative fractions.

The expected numbers of signal and background terms in the SS and OS dilepton samples for the same Monte Carlo sample and under the same conditions can be found in Table 4.3.

Fig. 4.10 and Fig. 4.11 display the content of Tables 4.2 and 4.3 in a pie-chart format.

The Monte Carlo shows that the ratio of SS to OS events in the dilepton sample is 0.201, assuming $x_d = 0.723$. The ratio for data is 0.208±0.003.

Table 4.2: Relative fractions of categories of events in $\Upsilon(4S)$ and total dilepton sample based on a Monte Carlo simulation. The different combinations include primary (p), secondary (s) and fake (f) leptons. True and fake leptons from continuum are collectively labeled as "continuum". The Monte Carlo sample consists of 3M charged and 3M neutral B events, generated with $x_d = 0.723$ ($\Delta m = 0.464$ ps^{-1}), assuming $f_\pm/f_0$=1.07 and $b_\pm/b_0 = \tau_{B^\pm}/\tau_{B^0}$=1.04, and 24M continuum events.

Event Categories		Fractions (%)	
		In $\Upsilon(4S)$ sample	In total sample
Signal	p-p	75.2	71.3
Background	p-s	18.2	17.2
	p-f	4.9	4.7
	s-s	1.2	1.2
	s-f	0.4	0.4
	f-f	0.1	0.1
	continuum	–	5.2

Table 4.3: Classification of different contributions to SS and OS dilepton samples and relative weights. The numbers are calculated from a Monte Carlo sample of 3M charged and 3M neutral B events, generated with $x_d = 0.723$ ($\Delta m = 0.464$ ps^{-1}), assuming $f_\pm/f_0$=1.07 and $b_\pm/b_0 = \tau_{B^\pm}/\tau_{B^0}$=1.04, and 24M continuum events.

Source	Category	Same Sign (%)	Opposite Sign (%)
B^0B^0 and $\overline{B}^0\overline{B}^0$	Primary dileptons	34.4	—
	Background, correct tag	2.8	—
	Background, wrong tag	—	1.7
$B^0\overline{B}^0$	Primary dileptons	—	32.4
	Background, correct tag	—	6.1
	Background, wrong tag	26.6	—
B^+B^-	Primary dileptons	—	46.4
	Background, correct tag	—	8.4
	Background, wrong tag	30.3	—
Continuum	Background	5.9	5.0
Total		100.0	100.0

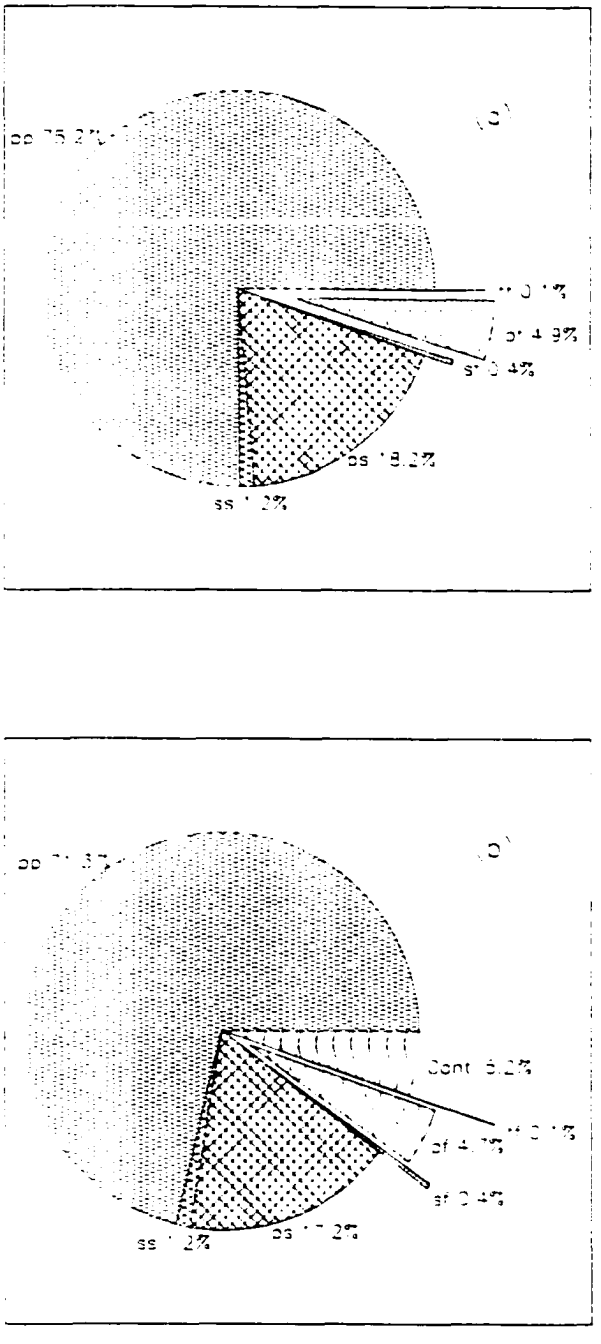

Figure 4.10: Relative fractions of combinations of leptons: primary (p), secondary (s) and fake (f), in (a) the $\Upsilon(4S)$ dilepton sample (b) the total dilepton sample (including continuum).

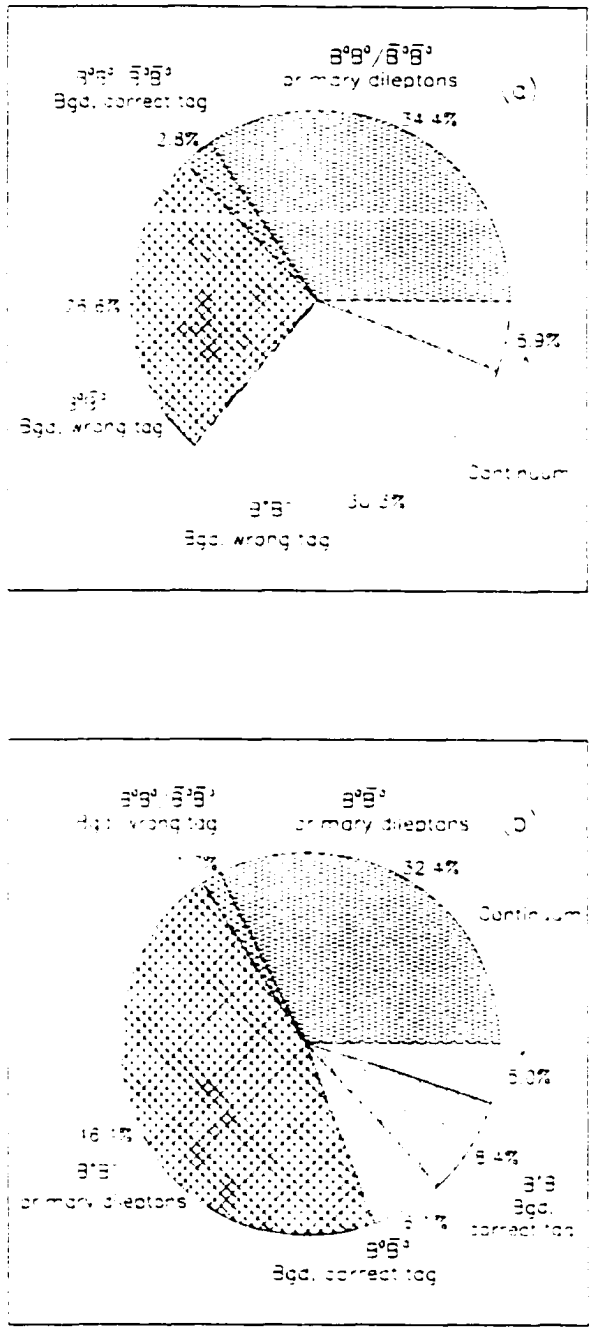

Figure 4.11: Expected representation of $B^0B^0/\overline{B}^0\overline{B}^0$, $B^0\overline{B}^0$, B^-B^- and continuum events in (a) Same-Sign and (b) Opposite-Sign Dilepton samples, separated in signal and background terms.

4.10 Time dependent analysis: signal terms

We summarize here the form in which the signal terms (primary dileptons) appear in the fitting.

We start with the theoretical expressions presented in Sec. 2.4.

The expression for the same-sign dileptons is coming from B^0B^0 and $\overline{B}^0\overline{B}^0$ pairs

$$P^{\text{mix}}_{\text{theory}}(\Delta t) = \frac{1}{c_{\ell^\pm\ell^\pm}} e^{-|\Delta t/\tau_{B^0}|}[1 - \cos(\Delta m\,\Delta t)] \tag{4.15}$$

with normalization constant

$$c_{\ell^\pm\ell^\pm} \equiv \tau_{B^0}\frac{x_d^2}{(1+x_d^2)} \tag{4.16}$$

and for the opposite-sign dileptons, the neutral B term ($B^0\overline{B}^0$)

$$P^{\text{unm}}_{\text{theory}}(\Delta t) = \frac{1}{c_{\ell^+\ell^-}} e^{-|\Delta t/\tau_{B^0}|}\Big[(1+|\cos\theta|^2) + (1-|\cos\theta|^2)\cos(\Delta m\,\Delta t) - 2\text{Im}(\cos\theta)\sin(\Delta m\,\Delta t)\Big] \tag{4.17}$$

with normalization constant

$$c_{\ell^+\ell^-} \equiv 2\tau_{B^0}\frac{(2+x_d^2+x_d^2|\cos\theta|^2)}{(1+x_d^2)} \tag{4.18}$$

and the charged B term

$$P^{\text{chd}}_{\text{theory}}(\Delta t) = \frac{1}{2\tau_{B^\pm}} e^{-|\Delta t/\tau_{B^\pm}|} \tag{4.19}$$

where $\tau_{B^\pm}$ is the $B^\pm$ lifetime, and $\tau_{B^0} = \Gamma^{-1}$, Γ being the average decay rate of the *mass* eigenstates of the B system.

The last term — Eq. (4.19) — is the contribution from charged B pairs (no mixing). The advantage of treating the primary lepton pairs from charged B mesons as signal, even though the relevant expressions contain no mixing or CPT violation information, becomes obvious here: their time evolution is known and simple ($e^{-|\Delta t/\tau_{B^\pm}|}$). Furthermore, the existence of both $\tau_{B^\pm}$ and τ_{B^0} in the theoretical expressions serves the very important goal of testing the accuracy of the response

function modeling. This point was explained earlier while discussing about the B lifetime dependence of the background distributions, in Sec. 4.8.2.

The next step is to use the response function (Sec. 4.7) to smear the theoretical expressions. By doing this we reproduce the observed distributions for the primary dilepton (measured) proper times, or equivalently, distances measured in the laboratory frame.

$$P(\Delta t_{\text{smear}}) = \int g(\Delta t_{\text{smear}} - \Delta t) P_{\text{theory}}(\Delta t)\, d(\Delta t) \tag{4.20}$$

In Eq. (4.20) we plug in

$$P_{\text{theory}}(\Delta t) = P^{\text{mix}}_{\text{theory}}(\Delta t),\ P^{\text{unm}}_{\text{theory}}(\Delta t) \text{ and } P^{\text{chd}}_{\text{theory}}(\Delta t),$$

to get

$$P(\Delta t_{\text{smear}}) = P^{00}_{\ell^\pm\ell^\pm}(\Delta t_{\text{smear}}),\ P^{00}_{\ell^+\ell^-}(\Delta t_{\text{smear}}) \text{ and } P^{+-}_{\ell^+\ell^-}(\Delta t_{\text{smear}}),$$

respectively.

4.11 Fitting

The equations for the fitting algorithm are

$$(N)_{\text{expected}} = D \times (N)_{B-\text{mesons}} + (N)_{\text{continuum}} \tag{4.21}$$

with

$$(N)_{\text{expected}} = \begin{pmatrix} N_{SS}(i) \\ \\ N_{OS}(i) \end{pmatrix}_{\text{expected}} \tag{4.22}$$

The (i)-th bin runs along the Δt_{smear} spectrum.

The D matrix is given by

$$D = \begin{array}{c} \\ SS \\ \\ OS \end{array} \begin{array}{c} \begin{array}{ccc} B^0B^0/\overline{B}^0\overline{B}^0 & B^0\overline{B}^0 & B^+B^- \end{array} \\ \begin{pmatrix} D_{SS}^{\rm mix}(i) & D_{SS}^{\rm unm}(i) & D_{SS}^{\rm chd}(i) \\ \\ D_{OS}^{\rm mix}(i) & D_{OS}^{\rm unm}(i) & D_{OS}^{\rm chd}(i) \end{pmatrix} \end{array} \tag{4.23}$$

where

$$D_{SS}^{\rm mix}(i) \equiv \epsilon_{\ell^\pm\ell^\pm}^{\rm mix}\, P_{\ell^\pm\ell^\pm}^{00}(i) + \epsilon_{SS}^{\rm mix}\, B_{SS}^{\rm mix}(i) \tag{4.24}$$

$$D_{SS}^{\rm unm}(i) \equiv \epsilon_{SS}^{\rm unm}\, B_{SS}^{\rm unm}(i) \tag{4.25}$$

$$D_{SS}^{\rm chd}(i) \equiv \epsilon_{SS}^{\rm chd}\, B_{SS}^{\rm chd}(i) \tag{4.26}$$

$$D_{OS}^{\rm mix}(i) \equiv \epsilon_{OS}^{\rm mix}\, B_{OS}^{\rm mix}(i) \tag{4.27}$$

$$D_{OS}^{\rm unm}(i) \equiv \epsilon_{\ell^+\ell^-}^{\rm unm}\, P_{\ell^+\ell^-}^{00}(i) + \epsilon_{OS}^{\rm unm}\, B_{OS}^{\rm unm}(i) \tag{4.28}$$

$$D_{OS}^{\rm chd}(i) \equiv \epsilon_{\ell^+\ell^-}^{\rm chd}\, P_{\ell^+\ell^-}^{+-}(i) + \epsilon_{OS}^{\rm chd}\, B_{OS}^{\rm unm}(i) \tag{4.29}$$

The rows of D correspond to events stored in the SS and OS histograms, whereas the columns correspond to "true" neutral mixed (B^0B^0 or $\overline{B}^0\overline{B}^0$) and unmixed ($B^0\overline{B}^0$), and charged ($B^+B^-$) events.

The number of generated B pairs is:

$$(N)_{B-\rm mesons} = N_{\Upsilon(4S)} \begin{pmatrix} f_0\,\chi_d \\ \\ f_0\,(1-\chi_d) \\ \\ f_\pm \end{pmatrix} \tag{4.30}$$

The last contribution to the dilepton sample is from continuum:

$$(N)_{\text{continuum}} = N_{q\bar{q}} \begin{pmatrix} \epsilon_{SS}^{\text{cnt}} B_{SS}^{\text{cnt}}(i) \\ \\ \epsilon_{OS}^{\text{cnt}} B_{OS}^{\text{cnt}}(i) \end{pmatrix} \tag{4.31}$$

$N_{\Upsilon(4S)}$ and $N_{q\bar{q}}$ are not independent, but related through their experimentally measured ratio (as discussed in Sec. 4.8.1). This ratio is the only relevant external parameter in the fitting besides $N_{\text{tot}} \equiv N_{SS} + N_{OS}$, the total number of selected events (for both SS and OS dilepton spectra).

In the above expressions, signal and background distributions are normalized: $P^{00}_{\ell^\pm\ell^\pm}(\Delta t_{\text{smear}}; \Delta m)$, $P^{00}_{\ell^+\ell^-}(\Delta t_{\text{smear}}; \Delta m, \theta)$, $P^{+-}_{\ell^+\ell^-}(\Delta t_{\text{smear}})$ and $B(\Delta t_{\text{smear}})$. The background distributions $B(\Delta t_{\text{smear}})$ are not modeled with analytic shapes, but rather employ the interpolation scheme described in Sec. 4.8.2. Moreover, the fraction of mixed events, $\chi_d = \chi_d(x_d, \theta)$, controls *both* the populations of signal and background events, as Eqs. (4.21) and (4.30) indicate.

The number of free parameters can be separated into those that are of interest in terms of physics (Δm, $\cos\theta$), and those that describe the fractions of signal and background in the event samples ($\eta^{\text{mix}}_{\ell^\pm\ell^\pm}$, $\eta^{\text{unm}}_{\ell^+\ell^-}$, $\eta^{\text{chd}}_{\ell^+\ell^-}$[13] and ϵ^i_j, with i = "mix", "unm", "chd", "cnt" and j = "SS", "OS": $3 + 4 \times 2 = 11$ parameters). For this secondary set of parameters we do the following:

- We fix the ratios of the primary dilepton selection efficiencies to values determined by the Monte Carlo:

$$\frac{\eta^{\text{chd}}_{\ell^+\ell^-}}{\eta^{\text{unm}}_{\ell^+\ell^-}} = \left.\frac{\eta^{\text{chd}}_{\ell^+\ell^-}}{\eta^{\text{unm}}_{\ell^+\ell^-}}\right|_{\text{MC}} \quad \text{and} \quad \frac{\eta^{\text{chd}}_{\ell^+\ell^-}}{\eta^{\text{mix}}_{\ell^\pm\ell^\pm}} = \left.\frac{\eta^{\text{chd}}_{\ell^+\ell^-}}{\eta^{\text{mix}}_{\ell^\pm\ell^\pm}}\right|_{\text{MC}}$$

- We fix the ratios of background with correct tagging efficiencies to values determined by the Monte Carlo:

[13]The primary lepton pair selection efficiencies η are defined in Sec. 4.6.1.

$$\frac{\epsilon_{OS}^{\text{chd}}}{\epsilon_{OS}^{\text{unm}}} = \left.\frac{\epsilon_{OS}^{\text{chd}}}{\epsilon_{OS}^{\text{unm}}}\right|_{\text{MC}} \quad \text{and} \quad \frac{\epsilon_{OS}^{\text{chd}}}{\epsilon_{SS}^{\text{mix}}} = \left.\frac{\epsilon_{OS}^{\text{chd}}}{\epsilon_{SS}^{\text{mix}}}\right|_{\text{MC}}$$

- We fix the ratios of background with wrong tagging efficiencies to values determined by the Monte Carlo:

$$\frac{\epsilon_{SS}^{\text{chd}}}{\epsilon_{SS}^{\text{unm}}} = \left.\frac{\epsilon_{SS}^{\text{chd}}}{\epsilon_{SS}^{\text{unm}}}\right|_{\text{MC}} \quad \text{and} \quad \frac{\epsilon_{SS}^{\text{chd}}}{\epsilon_{OS}^{\text{mix}}} = \left.\frac{\epsilon_{SS}^{\text{chd}}}{\epsilon_{OS}^{\text{mix}}}\right|_{\text{MC}}$$

- The continuum efficiencies $\epsilon_{SS}^{\text{cnt}}$ and $\epsilon_{OS}^{\text{cnt}}$ are not independent from the $\Upsilon(4S)$ efficiencies η's and ϵ's, but scaled in a way that keeps the fraction of continuum events in the dilepton sample equal to a value determined by the Monte Carlo (as shown in Table 4.3 and Fig. 4.11).

These conditions reduce the number of secondary parameters to $11 - 4 \times 2 = 3$. The number of degrees of freedom N_{dof} is reduced by one more unit by the implied condition that N_{tot} is fixed (see Eqs. (4.21)-(4.31)). What is left completely unconstrained in the fit is the ratio of the η's and the ϵ's; that is, the ratio of the number of primary dilepton to the amount of background with correct and wrong tagging efficiencies. So, effectively the number of free-to-vary secondary parameters in the fit is two: $\eta/\epsilon_{\text{"correct tag"}}$ and $\eta/\epsilon_{\text{"wrong tag"}}$, for all categories of B pairs. Everything else is determined through the corresponding Monte Carlo ratios.

We do the fit by minimizing an "Almeida-Barbi-do Vale" χ^2 [41]:

$$\chi^2_{ABV} = \sum_i \left[2(N^{(i)}_{\text{expected}} - N^{(i)}_{\text{observed}}) + (2N^{(i)}_{\text{observed}} + 1) \ln \left(\frac{2N^{(i)}_{\text{observed}} + 1}{2N^{(i)}_{\text{expected}} + 1} \right) \right] \quad (4.32)$$

where $N^{(i)}_{\text{expected}}$ and $N^{(i)}_{\text{observed}}$ are the numbers of expected and observed events in the i-th bin of the Δt_{smear} spectrum. The fit is performed simultaneously on the SS and OS Δt_{smear} histograms.

There are three ways in which the fitting procedure is sensitive to Δm and $\cos\theta$:

- The time-dependent level: this is achieved through the signal terms for the neutral B pairs, $P^{00}_{\ell^\pm\ell^\pm}(\Delta t_{\rm smear}; \Delta m)$ and $P^{00}_{\ell^+\ell^-}(\Delta t_{\rm smear}; \Delta m, \theta)$. The ability to measure the time evolution of the B pair is the main advantage of an asymmetric e^+e^- collider.

- The time-integrated level: through the time integrated mixing probability $\chi_d = \chi_d(x_d, \theta)$. Since in this analysis the B lifetime is assumed known and fixed, a constraint on x_d is effectively a constraint on Δm.

$$\chi_d \equiv \chi_d(\Delta m, \theta) = \frac{|\sin\theta|^2 \tau^2_{B^0} \Delta m^2}{|\sin\theta|^2 \tau^2_{B^0} \Delta m^2 + (2 + \tau^2_{B^0} \Delta m^2 + \tau^2_{B^0} \Delta m^2 |\cos\theta|^2)}$$

- The interpolation of the background: By using a background that explicitly depends on Δm (through $\gamma = \gamma(\Delta m)$) in the fit, we use a large fraction of the dilepton sample ($\sim$ 24% - Fig. 4.10.b or Table 4.2) for the determination of Δm that would otherwise remain "unused". In other words, the background provides some additional information on Δm.

4.12 Results

With the assumptions listed in Sec. 4.9 and by holding the B^0 lifetime fixed to its world average value[14] ($\tau_{B^0} = 1.56 \pm 0.04$ ps [5]), we obtain the following results:

- **Fitting for Δm:**

  Assuming that CPT is a good symmetry, and therefore by keeping the CPT violation parameters fixed ($\cos\theta = 0$) we find

$$\Delta m = 0.456 \pm 0.008 \text{ ps}^{-1}$$

[14] When we test the CPT symmetry by looking for a B^0-$\bar{B}^0$ lifetime difference, τ_{B^0} stands for the average lifetime of B^0 and $\bar{B}^0$.

$$\chi^2/N_{\rm dof} = 341.6/381$$

- **Fitting for Δm and $\cos\theta \equiv \mathrm{Re}(\cos\theta) + i\,\mathrm{Im}(\cos\theta)$:**

$$\Delta m = 0.456^{+0.009}_{-0.008}\ \mathrm{ps}^{-1}$$

$$\mathrm{Im}(\cos\theta) = 0.019^{+0.031}_{-0.030}$$

$$\mathrm{Re}(\cos\theta) = 0.00 \pm 0.21$$

$$\chi^2/N_{\rm dof} = 341.2/379$$

We use the larger value of the positive and negative error bars as the statistical error.

Fig. 4.12 shows the Δz distributions for the data and the fitted Monte Carlo distributions, for the SS (top) and OS (bottom) dileptons. The Monte Carlo distributions in Fig. 4.12 are not normalized separately to the numbers of the events of the two histograms, but only to the total number of data events $N_{\rm tot} = N_{SS} + N_{OS}$. In other words, the ratio of the populations of the two distributions (SS, OS) is taken care of by the fit result for the mixing, and this is a very strong constraint on the $\chi_d(\Delta m, \theta)$ parameter.

The mixing term appears in the SS Δz distribution and is marked as "primary dileptons" (signal). That this term is zero for $\Delta z = 0$ is the Einstein-Rosen-Podolsky (EPR) paradox in the B world: B mesons that decay at the same time (equivalently, at the same z in the laboratory) cannot have the same flavor. The signal term in the SS histogram is buried under the background from $B^0\overline{B}^0$ and B^+B^-. A much

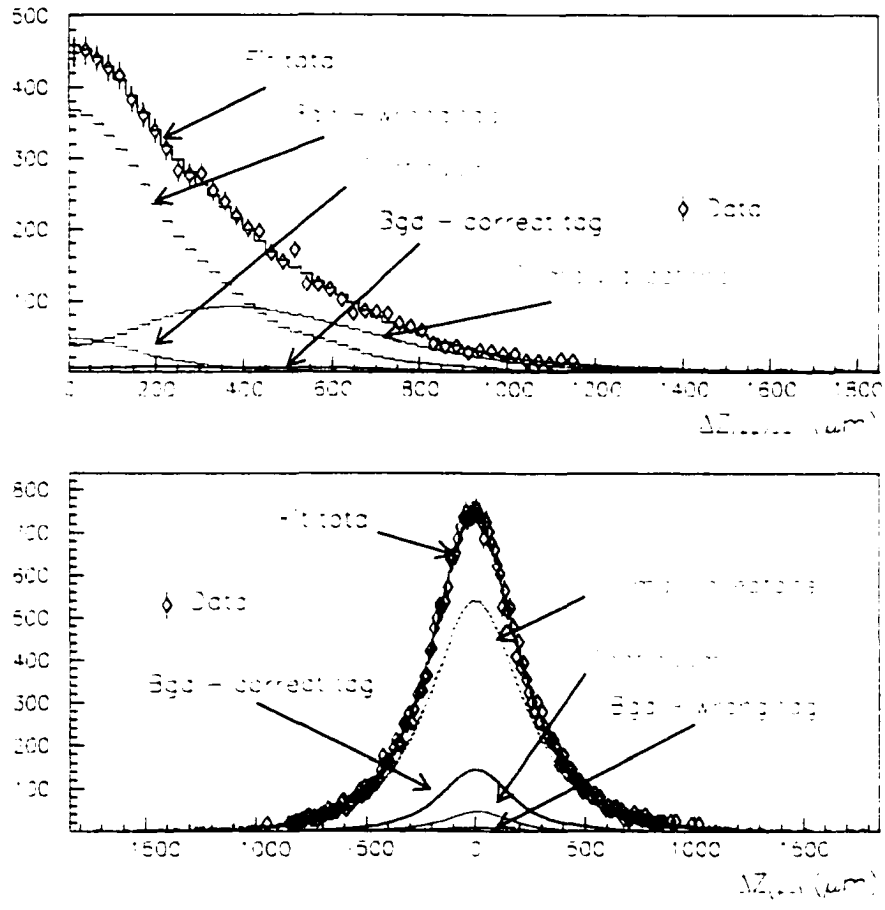

Figure 4.12: Δz distributions for same-sign (top) and opposite-sign (bottom) dileptons for data and the fitted Monte Carlo distributions. For the Monte Carlo distributions, signal and background from different categories are plotted separately.

more evocative way of displaying the $B^0 - \overline{B}^0$ oscillation is by plotting the time dependent asymmetry between SS and OS dileptons:

$$A(\Delta t) = \frac{N^{+-}(\Delta t) - N^{\pm\pm}(\Delta t)}{N^{+-}(\Delta t) + N^{\pm\pm}(\Delta t)}$$

By assuming that only primary dileptons are in the SS and OS histograms, then $A(\Delta t)$ as given by Eqs. (2.23) and (2.26) is

$$A(\Delta t) = \cos(\Delta m \Delta t)$$

The background from neutral and charged B pairs complicates this expression. To a good approximation, however, most of these extra terms cancel out, and the time dependent asymmetry reveals a clean oscillation signal, as illustrated in Fig. 4.13. In the same plot, a convolved cosine term is superimposed on the data points, with the period determined from the fit:

$$T_{\text{mixing}} = \frac{2\pi c \beta\gamma}{(\Delta m)_{\text{fit}}}$$

In Fig. 4.13, the negative Δz region for the OS histogram has been folded into the positive region for display purposes.

4.13 Systematic Errors

The systematic errors for Δm and the CPT violating parameter $\cos\theta$ arise from numerous sources. In order to estimate the effect of each of them, the associated parameter is varied, typically by $\pm\sigma$, and then the fit is repeated. The shift in the central value is taken to be the systematic error. The same data set is used for this procedure, except for some systematic errors related to the response function, for which studies are performed on the Monte Carlo. We examined the following sources of systematic errors:

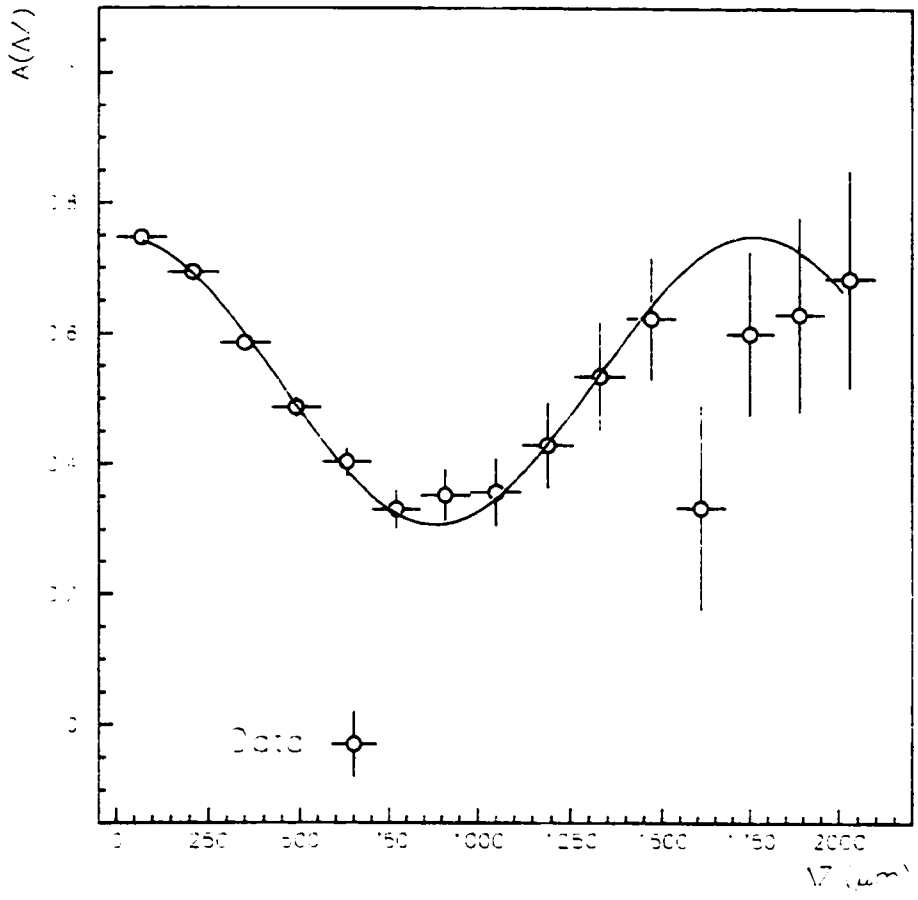

Figure 4.13: Time dependent asymmetry between SS and OS dileptons. The curve is a convoluted cosine with period determined by the fit.

- **B lifetimes and semileptonic branching ratios.** We use the current uncertainty in the B^0 lifetime, $\tau_{B^0} = 1.56 \pm 0.04$ ps, and the charged-to-neutral lifetime ratio, $\tau_{B^\pm}/\tau_{B^0} = 1.04 \pm 0.04$ [5]. We also use the relation $b^\pm/b^0 = \tau_{B^\pm}/\tau_{B^0}$ (see Sec. 4.6.1, pg. 58) to vary accordingly the semileptonic branching ratios. Varying the B^0, $B^\pm$ lifetimes affects both signal and background distributions, by changing the slope of their generic exponential form. Varying the semileptonic branching ratios affects all contributions with one or two primary leptons. Since all relevant terms are scaled according to the lifetime ratio, signal terms (two primary leptons) will be more affected than background events with one primary lepton. Finally, we check the effect of a non-zero $\Delta\Gamma$ in the proper time distributions[15]. Since no interesting constraint has been derived experimentally, we use the limit given by most theorists [4]: $\Delta\Gamma/\Gamma < 1\%$.

- **$f_\pm/f_0$ ratio.** We vary the fractions of charged and neutral B mesons at the generator level according to the measurement $f_\pm/f_0 = 1.07 \pm 0.09$ [38]. Since B^0 and $B^\pm$ have different lifetimes, changing the relative fractions affects the RMS of the total Δz distribution. This parameter change affects mostly $\chi = \chi(x_d, \theta)$, and therefore Δm and $\mathrm{Re}(\cos\theta)$ indirectly through Eq. (2.22). $\mathrm{Im}(\cos\theta)$ is not strongly affected by a χ variation, because it is the only antisymmetric term that appears in the proper time distributions, Eq. (2.16), and is therefore independent from changes in time integrated quantities.

- **Response function.**

  - **B motion in $\Upsilon(4S)$ rest frame and vertex resolution.** There are two issues here: the first one is the effect that the B motion in the $\Upsilon(4S)$ rest frame has on their time evolution, i.e. the $\Delta t_{\Upsilon(4S)}$ *vs.* ($t_{B_1}^{\mathrm{CM}}$ –

[15]The proper time distributions for $\Delta\Gamma \neq 0$ can be found in Sec. A.1.5

$t_{B_2}^{CM}$) systematic error[16], and the second one is whether primary dileptons and leptons from J/ψ decays have the same vertex resolution, i.e. the $\Delta z(\text{"}b \to c\ell\nu\text{"})/c\beta\gamma$ *vs.* $\Delta z(\text{"}J/\psi \to \ell\ell\text{"})/c\beta\gamma$ systematic error. Since we are interested in the difference of the measured vertices from the actual times that the B mesons decay at, we combine both effects by performing the fit for two cases: one by using the $g(t_{B_1}^{CM} - t_{B_2}^{CM} - \Delta z_{\text{"}b\to c\ell\nu\text{"}}/c\beta\gamma)$ and one by using the $g(\Delta z_{\text{"}J/\psi\to\ell\ell\text{"}}/c\beta\gamma)$ as the response function. Comparing the results we find the shift in Δm, Re($\cos\theta$) and Im($\sin\theta$) measurements to be negligible: $\sim$0.001 ps^{-1}, <0.01 and <0.001, respectively.

- **Statistical uncertainty in J/ψ Δz distribution.** The statistical uncertainty in the J/ψ Δz distribution (obtained from data as described in Sec. 4.7) is another source of systematic error, mainly due to the small number of J/ψ decay events surviving the cuts ($\sim$ 3700) used in the analysis. We repeat the fit by varying the numbers of events on a bin-by-bin basis for the Δz distributions from the J/ψ mass window and the sideband background, assuming Poisson distributions with mean equal to the actual number of entries for every bin. The resulting response function is used for the fit, and the central value of the fit is stored. We repeat this procedure until we obtain distributions of "mean values" for Δm, Re($\cos\theta$) and Im($\cos\theta$). The RMS of the distributions is used as the systematic error. The main effect is on Im($\cos\theta$) since this is the only asymmetric term in the distributions (see Eq. (2.16)), and a small asymmetry in the J/ψ Δz distribution is perceived by the fit as a CPT-violating Im($\cos\theta$)$\neq$ 0. This is the main contribution to the Im($\cos\theta$) systematic error.

  This parameter change affects only the signal terms.

[16]See Sec. A.1.3 for more details.

- **Other background related parameters.** The fractions of different background categories are listed in Table 4.2.

  - **Secondary lepton rates.** We examine the possibility of an inaccurate Monte Carlo in the description of B meson decays to charm mesons, especially for final states involving higher resonance states and virtual W decays. To this end, we use the current measurements of the D^0 and $D^{\pm}$ branching ratios, $\mathcal{B}(B \longrightarrow D^{\pm} X) = (23.5 \pm 2.7)\%$ and $\mathcal{B}(B \longrightarrow D^0 X) = (63.6 \pm 2.3)\%$ [39]. This change affects the populations of secondary leptons originating from D^0 and $D^{\pm}$. Since the lifetimes of charged (~1.1 ps) and neutral (~0.4 ps) D mesons are very different, the Δz distributions of the secondary lepton background are also affected.

  - **Fake lepton rates.** We conservatively take the fake rate uncertainty to be ±35%. This number takes into account the ~10% discrepancy in the fake rates between Monte Carlo and data (comparable to the statistical error of the measurements), and includes fake rates from both pions and kaons. About 5% of the selected event sample contains at least one fake lepton (~90% of the time combined with a primary lepton, ~8% of the time with a secondary lepton, and only ~2% combined with another fake lepton). So, this parameter change has only a small impact on the central values of the fit.

  - **Continuum contribution.** The uncertainty in the level of continuum contamination arises from the uncertainty in the hadronic cut efficiency (Sec. 4.2) for continuum events, but it also includes the statistical error from the Monte Carlo study. We conservatively estimate this uncertainty to be ±10%. This change affects the continuum population (~ 5% of the sample).

- **Detector resolution.** The Δz resolution of the signal terms in the SS and OS histograms is defined by the response function. The uncertainty in the modeling of the detector resolution for the signal is expressed by the variations in the response function.

  For the background from $\Upsilon(4S)$ and continuum we use a convolution of the Monte Carlo Δz distributions with a single Gaussian, in order to compensate for the vertex resolution discrepancy between Monte Carlo and data ($\sim$ 10%). The error in the determination of the width of the Gaussian ($\sigma = 50^{+12}_{-18}$ μm) is used as the uncertainty in the detector resolution for the background terms ($\sim$ 30% of the sample).

- **Binning:** Finally, we estimate the possible bias from using a binned minimization method. We repeat the fit after changing the number of bins in the SS and OS histograms and report possible shifts.

Table 4.4 summarizes the systematic errors for Δm, Re($\cos\theta$) and Im($\cos\theta$). Since the asymmetry between positive and negative errors is small ($\sim$ 15 – 20%), we conservatively take the larger of the values to be the systematic error.

4.14 Other consistency checks

As an additional check, we examine the effect of changing the value of the kinematic cuts on p^*_ℓ and $\cos\theta^*(\ell_1, \ell_2)$ (which were optimized based on Monte Carlo studies) on the central values of the fit. The small shifts are consistent with being statistical fluctuations.

A similar study is performed on sub-sets of the dilepton sample, separated according to the lepton id: ee, $\mu\mu$ and $e\mu$.

- <u>ee sample (10142 events):</u>

Table 4.4: Summary of systematic errors (a) with and (b) without invoking the CPT symmetry.

Parameter and uncertainty	Fitting for			
	(a) Δm	(b) Δm and $\cos\theta$		
	Δm (ps^{-1})	Δm (ps^{-1})	Re($\cos\theta$)	Im($\cos\theta$)
B lifetimes and semileptonic branching ratios				
$\tau_{B^0} = 1.56 \pm 0.04$ ps	$^{-0.009}_{+0.008}$	$^{-0.009}_{+0.008}$	< 0.01	± 0.003
$\tau_{B^\pm}/\tau_{B^0} = 1.04 \pm 0.04$	$^{+0.009}_{-0.012}$	$^{+0.009}_{-0.011}$	± 0.15	± 0.004
$\Delta\Gamma/\Gamma < 1\%$	< 0.001	< 0.001	± 0.02	< 0.001
Fractions of charged and neutral B mesons				
$f_\pm/f_0 = 1.07 \pm 0.09$	± 0.008	$^{+0.008}_{-0.011}$	± 0.26	$^{+0.001}_{-0.002}$
Response function				
B motion in $\Upsilon(4S)$ frame	± 0.001	± 0.001	< 0.01	< 0.001
statistics of J/ψ sample	± 0.004	± 0.004	< 0.01	± 0.061
Other background related parameters				
$\mathcal{B}(B \to D^0 X)$ ($\pm 4.6\%$)	< 0.001	< 0.001	< 0.01	± 0.001
$\mathcal{B}(B \to D^\pm X)$ ($\pm 14.3\%$)	$^{-0.000}_{+0.001}$	$^{-0.000}_{+0.001}$	< 0.01	$^{-0.002}_{+0.001}$
Fake rates ($\pm 35\%$)	± 0.003	± 0.003	< 0.01	$^{+0.002}_{-0.004}$
Continuum ($\pm 10\%$)	± 0.001	± 0.001	< 0.01	$^{-0.001}_{+0.000}$
Detector resolution ($^{+18}_{-12}$ μm)	∓ 0.001	∓ 0.001	< 0.01	$^{+0.001}_{+0.002}$
Other systematic errors				
Binning	± 0.001	± 0.001	< 0.01	± 0.001
Total	$^{\mathbf{+0.015}}_{\mathbf{-0.018}}$	$^{\mathbf{+0.015}}_{\mathbf{-0.019}}$	$\mathbf{\pm 0.30}$	$\mathbf{\pm 0.061}$

- Fitting for Δm ($\cos\theta \equiv 0$)

$$\Delta m = 0.456 \pm 0.011 \text{ ps}^{-1}$$

- Fitting for Δm and $\cos\theta$

$$\Delta m = 0.454 \pm 0.011 \text{ ps}^{-1}$$

$$\text{Im}(\cos\theta) = 0.096 \pm 0.060$$

$$\text{Re}(\cos\theta) = 0.00 \pm 0.36$$

- $\mu\mu$ sample (11624 events):

  - Fitting for Δm ($\cos\theta \equiv 0$)

$$\Delta m = 0.447 \pm 0.011 \text{ ps}^{-1}$$

  - Fitting for Δm and $\cos\theta$

$$\Delta m = 0.448 \pm 0.011 \text{ ps}^{-1}$$

$$\text{Im}(\cos\theta) = 0.032 \pm 0.070$$

$$\text{Re}(\cos\theta) = 0.00 \pm 0.22$$

- $e\mu$ sample (21309 events):

  - Fitting for Δm ($\cos\theta \equiv 0$)

$$\Delta m = 0.460 \pm 0.011 \text{ ps}^{-1}$$

  - Fitting for Δm and $\cos\theta$

$$\Delta m = 0.465 \pm 0.016 \text{ ps}^{-1}$$

$$\text{Im}(\cos\theta) = -0.023 \pm 0.042$$

$$\text{Re}(\cos\theta) = 0.16^{+0.13}_{-0.44}$$

All the results for the sub-sets are consistent with each other, within the statistical errors.

Chapter 5

Summary and discussion

5.1 Δm and $\cos\theta$ measurements

We have measured the mixing parameter (mass difference of B_H and B_L eigenstates) Δm and searched for CPT violation in the B^0 system from the time evolution of dileptons in $\Upsilon(4S)$ decays with an integrated luminosity of 5.1 fb^{-1} of data collected with the Belle detector. The run period was from January to July of 2000.

The asymmetric e^+e^- KEKB machine allows for the extraction of the B evolution from decay vertex measurements with the lepton tracks. Distributions for same-sign (SS) and opposite-sign (OS) dileptons were made. Δm and $\cos\theta$ were obtained by simultaneously fitting the SS and OS dilepton time distributions to complex functions. The modeling of the functions was done by using theoretical expressions for primary dilepton pairs (signal) and lookup tables obtained from Monte Carlo simulations for events with at least one secondary or fake lepton (background). The decay vertex resolution for the signal was estimated by the Δz distribution of $J/\psi \to \ell\ell$ decays from the same runs. A discrepancy of $\sim$10% for the resolution between Monte Carlo and data was taken into account in the modeling of the background. For $\Upsilon(4S)$ background events, recent CLEO measurements were used and a sideband study of kaons and pions was performed for the determination of the number of secondary and fake leptons, respectively. For the level of continuum contamination,

on- and off-resonance data and a Monte Carlo study for the determination of the hadronic cut efficiency were used.

Δm is determined by the signal (explicitly through the mixing oscillation period) and the background (through a linear interpolation) distributions, and the time-integrated analysis (indirectly through the fraction of mixed B pairs, or, mixing integrated probability, χ_d). Re($\cos\theta$) is mainly determined from the χ_d constraint and Im($\cos\theta$) is obtained from the analytical expressions (the only antisymmetric term).

If we invoke CPT symmetry for Δm we obtain:

$$\Delta m = 0.456 \pm 0.008 \text{ (stat)} \ \pm 0.018 \text{ (sys) ps}^{-1}$$

If we fit simultaneously for Δm and the CPT-violating parameter $\cos\theta$ we obtain:

$$\Delta m = 0.456 \pm 0.009 \text{ (stat)} \ \pm 0.019 \text{ (sys) ps}^{-1}$$

$$\text{Re}(\cos\theta) = 0.00 \pm 0.21 \text{ (stat)} \ \pm 0.30 \text{ (sys)}$$

$$\text{Im}(\cos\theta) = 0.019 \pm 0.031 \text{ (stat)} \ \pm 0.061 \text{ (sys)}$$

By using the relations (2.11) we get for the mass and the lifetime difference for B^0 and $\bar{B}^0$:

$$\frac{m_{B^0} - m_{\bar{B}^0}}{m_{B^0}} = [0.0 \pm 1.2 \text{ (stat)} \ \pm 1.7 \text{ (sys)}] \times 10^{-14}$$

or

$$\frac{|m_{B^0} - m_{\bar{B}^0}|}{m_{B^0}} < 4.1 \times 10^{-14}, \text{ at } 95\% \text{ C.L.}$$

and

$$\frac{\Gamma_{B^0} - \Gamma_{\bar{B}^0}}{\Gamma_{B^0}} = [-2.9 \pm 4.7 \text{ (stat)} \ \pm 9.2 \text{ (sys)}]\%$$

The results are consistent with CPT conservation and with the world average value $\Delta m = 0.464 \pm 0.018$ ps^{-1}.

This is the first direct measurement of Δm from a time-dependent analysis on the $\Upsilon(4S)$ resonance.

There are two previous measurements in the bibliography on $\mathrm{Im}(\cos\theta)$ from OPAL and DELPHI:

$$\text{OPAL}: \quad \mathrm{Im}(\cos\theta) = -0.040 \pm 0.032 \pm 0.012 \quad [42]$$

$$\text{DELPHI}: \quad \mathrm{Im}(\cos\theta) = -0.022 \pm 0.034 \pm 0.010 \quad [43]$$

Their statistical errors are comparable with the number from this analysis. However, our systematic error on $\mathrm{Im}(\cos\theta)$ is still large. As discussed in Sec. 4.13, this is due to the limited number of reconstructed $J/\psi \to \ell\ell$ decays that are used for the response function. With more data, we expect that the systematic error will be significantly decreased, as studies on the Monte Carlo show.

This is the first measurement on $\mathrm{Re}(\cos\theta)$, and consequently, on the mass difference between B^0 and $\overline{B}^0$.

5.2 Future prospects of mixing studies

With the current integrated luminosity, the statistical errors of our measurements are smaller than the systematic ones. The main contributions for Δm and $\mathrm{Re}(\cos\theta)$ come from the uncertainty in the measurements of the B^0 and $B^\pm$ lifetimes and the fractions of neutral and charged B mesons in the $\Upsilon(4S)$ decays. Neither of these quantities is determined by the analysis in the dilepton mode. This is because we do not fully or partially reconstruct the B mesons, and therefore, we cannot distinguish between neutral and charged ones. The necessary improvement in τ_{B^0}, $\tau_{B^\pm}/\tau_{B^0}$ and $f_\pm/f_0$ measurements for reducing the systematic errors of this analysis has to be external, i.e. it will be provided by another study. For $\mathrm{Im}(\cos\theta)$, where

the systematic error depends on the number of reconstructed $J/\psi \to \ell\ell$ (at least for this dilepton analysis), the above thoughts do not apply. On the contrary, since the $\mathrm{Im}(\cos\theta)$ appears in the only antisymmetric term in the proper time distributions, the lepton vertexing is the ideal method of proper time determination since it does not give a systematic shift[1]. A mean shift in the Δz distributions would mimic a CPT-violating $\mathrm{Im}(\cos\theta) \neq 0$ signal.

The study of dileptons is not the only method for mixing studies. The flavor tag of B mesons can be determined from full ($B \to D^{(*)}\pi$, $D^{(*)}\rho$) or partial ($B \to D^{(*)}\ell\nu$) reconstruction. Typically, one of the B mesons is reconstructed and the flavor of the other is determined by a standard tagging method (kaon(s) or high momentum lepton). This mode has very small systematic error because the reconstruction practically eliminates the background, but, for the same reason, only a small number of events is obtained and therefore the dominant contribution comes from the statistical error [44].

Dikaon events provide another mode suitable for mixing studies. If there are two kaons in the event then we can use their charge product in the same manner we do with the dileptons. Events with more than two kaons are in general more complicated cases. Kaons offer much better statistics than leptons. The disadvantage of studying the mixing with kaons is that we do not understand them as well as leptons in the B decays. This is due to the large uncertainty in the branching ratio measurements —$\mathcal{B}(B^{\pm}/B^0 \to K^+X^-) = 66 \pm 5\%$ and $\mathcal{B}(B^{\pm}/B^0 \to K^-X^+) = 13 \pm 4\%$—, but especially because of theoretical QCD uncertainties in the numerous cascade decays. Furthermore, kaons are generated from secondary, or "displaced" vertices. So, one has to take into account the dilution of the vertex measurement by the average shift, or more precisely, the uncertainty in the shift due to our limited understanding of the kaon dynamics. It is not clear how soon all these effects will be

[1]Secondary leptons do give a systematic shift when they are used for vertexing, but by taking the difference of the two decay vertices, the shift cancels out.

studied and resolved. The large number of kaons in the B decays makes this mode, however, very promising.

At this early stage of the experiment, dileptons seem to be the best choice for precise Δm and $\mathrm{Re}(\cos\theta)$ measurements. The B reconstruction mode has large statistical errors and dikaons are currently hampered by systematics. It is not obvious what the optimum mode in the future will be. This will depend on the accuracy of the lifetime and $f_{\pm}/f_0$ measurements for dileptons and the theoretical and experimental understanding of the kaons in B decays for the dikaon mode. The partial and full B reconstruction is at this point the safest mode for the future, when many tens or even hundreds of millions B pairs are available. In any case, the study of $B^0 - \overline{B}^0$ mixing at a high luminosity asymmetric e^+e^- collider offers a very sensitive interferometry and might open a window to new physics. The comparison of results obtained with different methods will be interesting.

Appendix A

A.1 Formalism

A.1.1 Mass matrix

The effective Hamiltonian of the $B^0 - \overline{B}^0$ mixing can be described with 4 complex parameters. In the approach generally followed in the bibliography [1] one solves the time-independent Schrödinger equation using the most general Hamiltonian for a B at rest. In the $B^0 - \overline{B}^0$ base

$$H = \begin{pmatrix} E\cos\theta - iD & E\sin\theta e^{-i\phi} \\ E\sin\theta e^{i\phi} & -E\cos\theta - iD \end{pmatrix} \tag{A.1}$$

Clearly, one can see that CP invariance requires $|\langle\overline{B}^0|H|B^0\rangle| = |\langle B^0|H|\overline{B}^0\rangle|$, i.e. $\mathrm{Im}\phi = 0$, and CPT invariance — which is the focus of this analysis — implies $\langle B^0|H|B^0\rangle = \langle\overline{B}^0|H|\overline{B}^0\rangle$, or $\cos\theta = 0$[2].

[1] For a more detailed description of the mixing phenomenology in B physics, see Ref. [45]-[46]. For the formalism of CPT violation, Ref. [7] is the original paper that one should read. We use the notation found in [8]-[9] that has also been used in a Belle paper before [47].

[2] This is because if we define $CPT \equiv \zeta$, then $(\zeta|\alpha\rangle, \zeta|\beta\rangle) = \langle\beta|\alpha\rangle$. By setting $\alpha \equiv HB^0$ and $\beta \equiv B^0$, we have $\langle B^0|H|B^0\rangle = (\zeta H|B^0\rangle, \zeta|B^0\rangle)$. If CPT is a good symmetry then $\zeta H\zeta^{-1} = H^\dagger$, or $\langle B^0|H|B^0\rangle = (H^\dagger\zeta|B^0\rangle, \zeta|B^0\rangle) = \langle\overline{B}^0|H|\overline{B}^0\rangle$.

The mass eigenstates of the Hamiltonian are:

$$|B_H\rangle \equiv |B_1\rangle = \frac{1}{\sqrt{|p_1|^2 + |q_1|^2}} (p_1 |B^0\rangle + q_1 |\bar{B}^0\rangle)$$
$$|B_L\rangle \equiv |B_2\rangle = \frac{1}{\sqrt{|p_2|^2 + |q_2|^2}} (p_2 |B^0\rangle - q_2 |\bar{B}^0\rangle) \tag{A.2}$$

with eigenvalues

$$\lambda_{H,L} \equiv \lambda_{1,2} = \pm E - iD \equiv m_{1,2} - \frac{i}{2}\Gamma_{1,2} \tag{A.3}$$

where E and D are complex in general and $m_{1,2}$ and $\Gamma_{1,2}$ are defined to be real.

The p_1, q_1, p_2 and q_2 are simply related to the parameters ϕ and θ via the following equations:

$$\frac{q_1}{p_1} = e^{i\phi} \tan\frac{\theta}{2}, \quad \frac{q_2}{p_2} = e^{i\phi} \cot\frac{\theta}{2} \tag{A.4}$$

Since CPT invariance demands

$$\frac{q_1}{p_1} = \frac{q_2}{p_2} = e^{i\phi} \tag{A.5}$$

one naturally uses a quantity such as $\cos\theta$ or

$$\cot\theta = \frac{1}{2}\left(\frac{q_2}{p_2} - \frac{q_1}{p_1}\right) e^{-i\phi} \tag{A.6}$$

to probe for CPT violation. A CPT-violating $\cos\theta \neq 0$ corresponds to differences in the mass and the lifetime of a particle and those of its antiparticle. From Eqs. (A.1), (A.3) it follows that

$$\cos\theta = \frac{H_{11} - H_{22}}{\lambda_1 - \lambda_2} = \frac{\Delta m_0 - (i/2)\,\Delta\Gamma_0}{\Delta m - (i/2)\,\Delta\Gamma} \tag{A.7}$$

where $\Delta m_0 \equiv m_{B^0} - m_{\bar{B}^0}$, $\Delta\Gamma_0 \equiv \Gamma_{B^0} - \Gamma_{\bar{B}^0}$, $\Delta m \equiv m_1 - m_2 \equiv m_H - m_L$ and $\Delta\Gamma \equiv \Gamma_1 - \Gamma_2 \equiv \Gamma_H - \Gamma_L$. So

$$\mathrm{Re}(\cos\theta) = \frac{4\Delta m_0\,\Delta m + \Delta\Gamma_0\,\Delta\Gamma}{4(\Delta m)^2 + (\Delta\Gamma)^2}$$
$$\mathrm{Im}(\cos\theta) = \frac{2(\Delta m_0\,\Delta\Gamma - \Delta\Gamma_0\,\Delta m)}{4(\Delta m)^2 + (\Delta\Gamma)^2} \tag{A.8}$$

By using the approximation[3] [4]

$$\frac{\Delta\Gamma}{\Gamma} \simeq 0 \tag{A.9}$$

we obtain

$$\mathrm{Re}(\cos\theta) \simeq \frac{\Delta m_0}{\Delta m}, \qquad \mathrm{Im}(\cos\theta) \simeq \frac{-\Delta\Gamma_0}{2\Delta m} \tag{A.10}$$

or

$$\frac{m_{B^0} - m_{\bar{B}^0}}{m_{B^0}} \simeq \mathrm{Re}(\cos\theta) \times \frac{\Delta m}{m_{B^0}} \quad \text{and} \quad \frac{\Gamma_{B^0} - \Gamma_{\bar{B}^0}}{\Gamma_{B^0}} \simeq -\frac{2\,\Delta m}{\Gamma_{B^0}} \times \mathrm{Im}(\cos\theta) \tag{A.11}$$

Using the bibliography values for $\Delta m = 3.1 \times 10^{-13}$ GeV$/c^2$, $m_{B^0} = 5.28$ GeV$/c^2$ and $\Gamma_{B^0} = 4.1 \times 10^{-13}$ GeV$/\hbar$ [5] we obtain

$$\frac{m_{B^0} - m_{\bar{B}^0}}{m_{B^0}} \sim 5.8 \times 10^{-14} \times \mathrm{Re}(\cos\theta) \quad \text{and} \quad \frac{\Gamma_{B^0} - \Gamma_{\bar{B}^0}}{\Gamma_{B^0}} \sim -1.5 \times \mathrm{Im}(\cos\theta) \tag{A.12}$$

So, a non-zero $\mathrm{Re}(\cos\theta)$ corresponds to a mass difference between B^0 and $\bar{B}^0$, whereas a non-zero $\mathrm{Im}(\cos\theta)$ corresponds to a lifetime difference between the two flavor states.

A.1.2 Decay rate of $B^0\bar{B}^0$ pair

An initially ($t = 0$) pure flavor eigenstate ($|B^0\rangle$ or $|\bar{B}^0\rangle$) will evolve in time: this time-dependence will not be the same for the two states because of different contributions from the two mass eigenstates [reverse of Eq. (A.2)]. The proper time evolution at the rest frame of the B mesons is given by:

$$\begin{aligned} |B^0(t)\rangle &= (c^2 e^{-i\lambda_1 t} + s^2 e^{-i\lambda_2 t})\,|B^0\rangle + sce^{i\phi}\,(e^{-i\lambda_1 t} - e^{-i\lambda_2 t})\,|\bar{B}^0\rangle \\ |\bar{B}^0(t)\rangle &= sce^{-i\phi}\,(e^{-i\lambda_1 t} - e^{-i\lambda_2 t})\,|B^0\rangle + (s^2 e^{-i\lambda_1 t} + c^2 e^{-i\lambda_2 t})\,|\bar{B}^0\rangle \end{aligned} \tag{A.13}$$

where $c \equiv \cos\frac{\theta}{2}$ and $s \equiv \sin\frac{\theta}{2}$.

Using the mean mass $m \equiv (m_1 + m_2)/2$, mean width $\Gamma \equiv (\Gamma_1 + \Gamma_2)/2$, mass difference $\Delta m \equiv m_1 - m_2 \equiv m_H - m_L$ and width difference $\Delta\Gamma \equiv \Gamma_1 - \Gamma_2 \equiv \Gamma_H - \Gamma_L$.

[3] Γ is the mean width of the mass eigenstates: $\Gamma \equiv (\Gamma_1 + \Gamma_2)/2$.

the equations take the more familiar form:

$$\begin{aligned} |B^0(t)\rangle &= e^{-(im+\frac{\Gamma}{2})t}\left[g_+(t)\,|B^0\rangle \;+\; \tilde{g}_+(t)\,|\bar{B}^0\rangle\right] \\ |\bar{B}^0(t)\rangle &= e^{-(im+\frac{\Gamma}{2})t}\left[\tilde{g}_-(t)\,|B^0\rangle \;+\; g_-(t)\,|\bar{B}^0\rangle\right] \end{aligned} \tag{A.14}$$

where

$$\begin{aligned} g_\pm(t) &= c^2 e^{\mp(i\Delta m+\frac{\Delta\Gamma}{2})\frac{t}{2}} + s^2 e^{\pm(i\Delta m+\frac{\Delta\Gamma}{2})\frac{t}{2}} \\ \tilde{g}_\pm(t) &= sce^{\pm i\phi}\left[e^{-(i\Delta m+\frac{\Delta\Gamma}{2})\frac{t}{2}} - e^{(i\Delta m+\frac{\Delta\Gamma}{2})\frac{t}{2}}\right] \end{aligned} \tag{A.15}$$

Notice that the condition $\mathrm{Im}\phi \neq 0$ (CP violation) is equivalent to $|\tilde{g}_+(t)| \neq |\tilde{g}_-(t)|$ and that CPT violation originates from the inequality $g_+(t) \neq g_-(t)$.

We consider the wavefunction describing the $B^0\bar{B}^0$ pair:

$$|\Psi(t)\rangle \;=\; \frac{1}{\sqrt{2}}\left[|B^0(\vec{k},t)\rangle\,|\bar{B}^0(-\vec{k},t)\rangle \;+\; C\,|B^0(-\vec{k},t)\rangle\,|\bar{B}^0(\vec{k},t)\rangle\right] \tag{A.16}$$

The *strong* decay $\Upsilon(4S) \longrightarrow B^0\bar{B}^0$ conserves the quantum numbers $J^{PC} = 1^{--}$, so the charge conjugation of the $B^0\bar{B}^0$ pair is $C = -1$. In Eq. (A.16), $\pm\vec{k}$ are the wavelength/momentum vectors of the neutral B pair at the $B^0\bar{B}^0$ CM frame. The meson pair undergoes a *coherent* mixing until one of them decays, say as B^0. At that point the wavefunction $|\Psi(t)\rangle$ collapses and the other meson is tagged as a $\bar{B}^0(t)$ evolving according to the second of Eqs. (A.14). This quantum-mechanical approach eventually leads to a new version of the EPR paradox: how does the second meson know *what* the first meson decayed as and *when*? To avoid this nuisance we can take an alternative path and calculate directly the amplitude for the *joint* decay. We call t_a and t_b the times that the B mesons (whether B^0 or $\bar{B}^0$ is irrelevant) with momenta $\vec{k}$ and $-\vec{k}$ decay to the final states f_a and f_b, respectively.

$$\begin{aligned} \Upsilon(4S) &\longrightarrow B\bar{B}, \quad \text{at } t=0 \\ B(\vec{k}) &\longrightarrow f_a, \quad \text{at } t=t_a \\ B(-\vec{k}) &\longrightarrow f_b, \quad \text{at } t=t_b \end{aligned}$$

The most general probability density for the joint decay is:

$$\begin{aligned} &P(\Psi \to f_a,t_a;f_b,t_b) \;\sim\; |\langle f_a,t_a;f_b,t_b|T|\Psi\rangle|^2 = \\ &\quad \frac{1}{2}\Big|\langle f_a|T|B^0(\vec{k},t_a)\rangle\,\langle f_b|T|\bar{B}^0(-\vec{k},t_b)\rangle - \langle f_b|T|B^0(-\vec{k},t_b)\rangle\,\langle f_a|T|\bar{B}^0(\vec{k},t_a)\rangle\Big|^2 \end{aligned} \tag{A.17}$$

Defining

$$A_i \equiv \langle f_i|T|B^0\rangle, \quad \bar{A}_i \equiv \langle f_i|T|\bar{B}^0\rangle, \qquad i = a, b \tag{A.18}$$

and by using the approximation (A.9) we obtain after a few steps:

$$P(f_a, t_a; f_b, t_b) \sim \frac{1}{4} e^{-\Gamma(t_a+t_b)}[A(\theta)\cos[\Delta m(t_a - t_b)] + B(\theta)\sin[\Delta m(t_a - t_b)] + C(\theta)] \tag{A.19}$$

where

$$\begin{aligned} A(\theta) &\equiv |A_a\bar{A}_b - \bar{A}_a A_b|^2 - |\sin\theta|^2|e^{-i\phi}A_a A_b - e^{i\phi}\bar{A}_a\bar{A}_b|^2 - |\cos\theta|^2|A_a\bar{A}_b + \bar{A}_a A_b|^2 \\ &\quad +2\,\mathrm{Re}[\cos\theta^* \sin\theta(e^{-i\phi}A_a A_b - e^{i\phi}\bar{A}_a\bar{A}_b)(A_a\bar{A}_b + \bar{A}_a A_b)^*] \end{aligned} \tag{A.20}$$

$$\begin{aligned} B(\theta) &\equiv 2\,\mathrm{Im}[\sin\theta(e^{-i\phi}A_a A_b - e^{i\phi}\bar{A}_a\bar{A}_b)(A_a\bar{A}_b - \bar{A}_a A_b)^* \\ &\quad - \cos\theta(A_a\bar{A}_b + \bar{A}_a A_b)(A_a\bar{A}_b - \bar{A}_a A_b)^*] \end{aligned} \tag{A.21}$$

$$\begin{aligned} C(\theta) &\equiv |A_a\bar{A}_b - \bar{A}_a A_b|^2 + |\sin\theta|^2|e^{-i\phi}A_a A_b - e^{i\phi}\bar{A}_a\bar{A}_b|^2 + |\cos\theta|^2|A_a\bar{A}_b + \bar{A}_a A_b|^2 \\ &\quad - 2\,\mathrm{Re}[\cos\theta^* \sin\theta(e^{-i\phi}A_a A_b - e^{i\phi}\bar{A}_a\bar{A}_b)(A_a\bar{A}_b + \bar{A}_a A_b)^*] \end{aligned} \tag{A.22}$$

The last equations are significantly simplified for the dilepton mode.

A.1.3 Motion of the B mesons in the $\Upsilon(4S)$ rest frame

In section A.1.2 we assumed that the two B mesons are still in the $\Upsilon(4S)$ rest frame ($B^0\bar{B}^0$ CM) and that the time appearing in the equations represents time measured in that frame. However, the decay of $\Upsilon(4S) = \Upsilon(10580)$ offers energy to each of the $B(5279)$ mesons that is slightly more than their mass. So, technically speaking, identifying the $B^0\bar{B}^0$ CM with the B^0 and the $\bar{B}^0$ rest frames is not accurate. In practice, the B mesons are born quite nonrelativistic ($|\vec{\beta}| \simeq 0.064$) so that the only result of this discrepancy is a minor Doppler shift in their energy and time distributions. We will examine what changes in the formalism when we consider a B in motion.

The time dependence of the B mesons — e.g. in Eq. (A.13) — includes the invariant phase and the decay probability

$$e^{-iMt}e^{-\frac{\Gamma}{2}t}$$

In the above expression the time t should not be taken as the time elapsed at the $\Upsilon(4S)$ rest frame, but as the time at a frame where the B is still. Therefore $t = t_{\text{proper}}$ is the time elapsed in the B rest frame and the above expression is accurate for both B mesons, but calculated at their proper times and at their rest frames [48].

So, we need to redefine the frame of the times t_1, t_2 and $\Delta t \equiv t_1 - t_2$ in the equations. The general rule is that all times are times elapsed at the respective rest frames for the two B mesons. Therefore, the products $\Delta m\,\Delta t$ and $\Gamma\Delta t$ in the proper time distributions are defined with $\Delta t \equiv t^*_{B_1} - t^*_{B_2}$. In order to estimate the error that we introduce by assuming a common rest frame we need the (weak) boost from the B rest frame to the $B^0\bar{B}^0$ CM

$$\gamma = \frac{M_{\Upsilon(4S)}}{2M_{B^0}} \simeq 1.002, \quad \text{or} \quad |\vec{\beta}| \simeq 0.064$$

Since

$$t_{\text{proper}} \simeq \frac{t_{\Upsilon(4S)\ \text{rest frame}}}{\gamma}$$

it follows that we overestimate the time difference by 0.2%.

A.1.4 Proper time distribution of Dileptons

The Standard Model gives in the leading order

$$\begin{aligned} A_{\ell^+} &\equiv \langle \ell^+ X^- | T | B^0 \rangle \neq 0, \quad \bar{A}_{\ell^-} \equiv \langle \ell^- X^+ | T | \bar{B}^0 \rangle \neq 0 \\ A_{\ell^-} &\equiv \langle \ell^- X^+ | T | B^0 \rangle = 0, \quad \bar{A}_{\ell^+} \equiv \langle \ell^+ X^- | T | \bar{B}^0 \rangle = 0 \end{aligned} \tag{A.23}$$

We will elaborate on the same-sign and opposite-sign cases separately.

Same-sign dileptons

Setting $|f_a\rangle = |f_b\rangle = |\ell^\pm X^\mp\rangle$ in Eqs. (A.20)-(A.22), we have:

- for $|\ell^+ X^-\rangle$: $(A_a)_{\ell^+\ell^+} = (A_b)_{\ell^+\ell^+} = A_{\ell^+}$ and $(\bar{A}_a)_{\ell^+\ell^+} = (\bar{A}_b)_{\ell^+\ell^+} = 0$, and
- for $|\ell^- X^+\rangle$: $(A_a)_{\ell^-\ell^-} = (A_b)_{\ell^-\ell^-} = 0$ and $(\bar{A}_a)_{\ell^-\ell^-} = (\bar{A}_b)_{\ell^-\ell^-} = \bar{A}_{\ell^-}$

Then

$$A(\theta)_{\ell^\pm\ell^\pm} = -C(\theta)_{\ell^\pm\ell^\pm} = -|\sin\theta e^{\mp i\phi}|^2 \times \begin{cases} |A_{\ell^+}|^4, & \text{if } \ell^+\ell^+ \\ |\bar{A}_{\ell^-}|^4, & \text{if } \ell^-\ell^- \end{cases} \tag{A.24}$$

and

$$B(\theta)_{\ell^+\ell^+} = 0 \tag{A.25}$$

So

$$\begin{aligned} P(\ell^+, t_a; \ell^+, t_b) &\sim \frac{|A_{\ell^+}|^4}{4}|\sin\theta e^{-i\phi}|^2 e^{-\Gamma(t_a+t_b)}\left[1 - \cos[\Delta m\,(t_a - t_b)]\right] \\ P(\ell^-, t_a; \ell^-, t_b) &\sim \frac{|\bar{A}_{\ell^-}|^4}{4}|\sin\theta e^{+i\phi}|^2 e^{-\Gamma(t_a+t_b)}\left[1 - \cos[\Delta m\,(t_a - t_b)]\right] \end{aligned} \tag{A.26}$$

Since in the Belle experiment we don't have the luxury of *independently* knowing t_a and t_b but only their relative difference $\Delta t = t_a - t_b$, we need to modify Eqs. (A.26). We integrate over $t' = t_a + t_b$, for Δt held fixed[4], yielding in a distribution that is a function solely of Δt.

$$\begin{aligned} P(\ell^+\ell^+, \Delta t) &\sim \frac{|A_{\ell^+}|^4}{4\Gamma}|\sin\theta e^{-i\phi}|^2 e^{-\Gamma|\Delta t|}\left[1 - \cos(\Delta m\,\Delta t)\right] \\ P(\ell^-\ell^-, \Delta t) &\sim \frac{|\bar{A}_{\ell^-}|^4}{4\Gamma}|\sin\theta e^{+i\phi}|^2 e^{-\Gamma|\Delta t|}\left[1 - \cos(\Delta m\,\Delta t)\right] \end{aligned} \tag{A.27}$$

One notices that there is no way to tell apart the experimentally indistinguishable same sign leptons by putting labels "a" or "b" since $|f_a\rangle = |f_b\rangle$. Therefore, it only makes sense to talk about $|\Delta t|$ and treat the (symmetric) distribution of Eqs. (A.27) as a function of Δt expanding from 0 to ∞.

$P(\ell^+\ell^+, \Delta t)$ and $P(\ell^-\ell^-, \Delta t)$ have the same dependence on Δt, regardless of the value of θ.

To get the total number of same-sign dileptons we integrate Eqs. (A.27):

$$N^{++} \equiv N(\ell^+\ell^+) = 2\int_0^\infty P(\ell^+\ell^+, \Delta t)\, d(\Delta t) \sim \frac{|A_{\ell^+}|^4}{2\Gamma^2}|\sin\theta e^{-i\phi}|^2 \frac{x_d^2}{1+x_d^2} \tag{A.28}$$

[4]The limits of integration are $t'_{\min} = |\Delta t|$ and $t'_{\max} = \infty$.

$$N^{--} \equiv N(\ell^-\ell^-) = 2\int_0^\infty P(\ell^-\ell^-, \Delta t)\, d(\Delta t) \sim \frac{|\bar{A}_{\ell^-}|^4}{2\Gamma^2} |\sin\theta e^{i\phi}|^2 \frac{x_d^2}{1+x_d^2} \quad (A.29)$$

where

$$x_d \equiv \frac{\Delta m}{\Gamma} = 0.723 \pm 0.035 \quad [5] \quad (A.30)$$

is the mixing parameter for the $B^0\overline{B}^0$ pair.

Opposite-sign dileptons

Taking $|f_a\rangle = |\ell^\pm X^\mp\rangle$, $|f_b\rangle = |\ell^\mp X^\pm\rangle$ we similarly find

- for $|f_a; f_b\rangle = |\ell^+X^-; \ell^-X^+\rangle$:

  $(A_a)_{\ell^+\ell^-} = A_{\ell^+}$, $(\bar{A}_b)_{\ell^+\ell^-} = \bar{A}_{\ell^-}$ and $(A_b)_{\ell^+\ell^-} = (\bar{A}_a)_{\ell^+\ell^-} = 0$, and

- for $|f_a; f_b\rangle = |\ell^-X^+; \ell^+X^-\rangle$:

  $(A_b)_{\ell^-\ell^+} = A_{\ell^+}$, $(\bar{A}_a)_{\ell^-\ell^+} = \bar{A}_{\ell^-}$ and $(A_a)_{\ell^-\ell^+} = (\bar{A}_b)_{\ell^-\ell^+} = 0$

Then

$$A(\theta)_{\ell^\pm\ell^\mp} = |A_{\ell^+}\bar{A}_{\ell^-}|^2(1 - |\cos\theta|^2), \quad C(\theta)_{\ell^\pm\ell^\mp} = |A_{\ell^+}\bar{A}_{\ell^-}|^2(1 + |\cos\theta|^2) \quad (A.31)$$

and

$$B(\theta)_{\ell^\pm\ell^\mp} = \mp|A_{\ell^+}\bar{A}_{\ell^-}|^2| \times 2\,\mathrm{Im}(\cos\theta) \quad (A.32)$$

The coefficient of the only antisymmetric — with respect to $\Delta t = t_a - t_b$ — contribution, $B(\theta)$, has unlike sign for $\ell^+\ell^-$ and $\ell^-\ell^+$, resulting in an identical distribution for the two cases. So, we define $|f_a\rangle = |\ell^+X^-\rangle$, $|f_b\rangle = |\ell^-X^+\rangle$ and we obtain

$$P(\ell^+, t_a; \ell^-, t_b) \sim \frac{|A_{\ell^+}\bar{A}_{\ell^-}|^2}{2\Gamma} e^{-\Gamma(t_a+t_b)}\Big[(1 - |\cos\theta|^2)\cos(\Delta m\,\Delta t) - 2\,\mathrm{Im}(\cos\theta)\sin(\Delta m\,\Delta t) + (1 + |\cos\theta|^2)\Big] \quad (A.33)$$

Here Δt is always defined as $t_{\ell^+} - t_{\ell^-}$.

Integrating as before over $t' = t_a + t_b$

$$P(\ell^+\ell^-, \Delta t) \sim \frac{|A_{\ell^+}\bar{A}_{\ell^-}|^2}{2\Gamma} e^{-\Gamma|\Delta t|}\Big[(1 - |\cos\theta|^2)\cos(\Delta m\,\Delta t) \\ -2\,\mathrm{Im}(\cos\theta)\sin(\Delta m\,\Delta t) + (1 + |\cos\theta|^2)\Big] \quad (A.34)$$

The total number of opposite-sign dilepton events from neutral B pairs is

$$N^{+-} \equiv N(\ell^+\ell^-) = \int_{-\infty}^{\infty} P(\ell^+\ell^-, \Delta t)\, d(\Delta t) \\ \sim \frac{|A_{\ell^+}\bar{A}_{\ell^-}|^2}{\Gamma^2} \frac{(2 + x_d^2 + x_d^2|\cos\theta|^2)}{1 + x_d^2} \quad (A.35)$$

In Eq. (A.35) N^{+-} refers to final products independent of order, corresponding to both $\ell^+\ell^-$ and $\ell^-\ell^+$ cases.

Combining Eqs. (A.28), (A.29) and (A.35) we obtain the fraction of the same-sign dileptons in the total sample of neutral B pairs

$$\chi_d \equiv \frac{N^{++} + N^{--}}{N^{++} + N^{--} + N^{+-}} \quad (A.36)$$

For

$$|A_{\ell^+}|^2 = |\bar{A}_{\ell^-}|^2 \quad (A.37)$$

we obtain

$$\chi_d = \frac{|\sin\theta|^2(|e^{-i\phi}|^2 + |e^{i\phi}|^2)\, x_d^2}{|\sin\theta|^2(|e^{-i\phi}|^2 + |e^{i\phi}|^2)\, x_d^2 + 2(2 + x_d^2 + x_d^2|\cos\theta|^2)} \quad (A.38)$$

At this point we will use the theoretical expectation that CP violation in the B mixing should not be large[5] and use a small $\mathrm{Im}(\phi)$ to approximate

$$|e^{\pm i\phi}|^2 \simeq 1 \mp 2\mathrm{Im}(\phi) \quad (A.39)$$

So

$$\chi_d \equiv \chi_d(x_d, \theta) = \frac{|\sin\theta|^2\, x_d^2}{|\sin\theta|^2\, x_d^2 + (2 + x_d^2 + x_d^2|\cos\theta|^2)} \quad (A.40)$$

If CPT is a good symmetry

$$\chi_d \overset{CPT}{=} \frac{x_d^2}{2(1 + x_d^2)} \quad (A.41)$$

[5] A quantitative argument for this can be found in Ref. [3].

A.1.5 Proper time distribution of Dileptons if $\Delta\Gamma \neq 0$

For completeness, we give the proper time distributions if the lifetime difference of the physical states $\Delta\Gamma \equiv \Gamma_1 - \Gamma_2 \equiv \Gamma_H - \Gamma_L$ is non zero.

Same-sign dileptons

$$P(\ell^+\ell^+, \Delta t) \sim \frac{|A_{\ell^+}|^4}{4\Gamma} |\sin\theta e^{-i\phi}|^2 e^{-\Gamma|\Delta t|} [\cosh(\Delta\Gamma\,\Delta t/2) - \cos(\Delta m\,\Delta t)]$$
$$P(\ell^-\ell^-, \Delta t) \sim \frac{|\bar{A}_{\ell^-}|^4}{4\Gamma} |\sin\theta e^{+i\phi}|^2 e^{-\Gamma|\Delta t|} [\cosh(\Delta\Gamma\,\Delta t/2) - \cos(\Delta m\,\Delta t)] \tag{A.42}$$

$$N^{++} \equiv N(\ell^+\ell^+) \sim \frac{|A_{\ell^+}|^4}{2\Gamma^2} |\sin\theta e^{-i\phi}|^2 \frac{(x_d^2 + y_d^2)}{(1 + x_d^2)(1 - y_d^2)} \tag{A.43}$$

$$N^{--} \equiv N(\ell^-\ell^-) \sim \frac{|\bar{A}_{\ell^-}|^4}{2\Gamma^2} |\sin\theta e^{i\phi}|^2 \frac{(x_d^2 + y_d^2)}{(1 + x_d^2)(1 - y_d^2)} \tag{A.44}$$

where

$$y_d \equiv \frac{\Delta\Gamma}{2\Gamma} \tag{A.45}$$

Opposite-sign dileptons

$$P(\ell^+\ell^-, \Delta t) \sim \frac{|A_{\ell^+}\bar{A}_{\ell^-}|^2}{2\Gamma} e^{-\Gamma|\Delta t|} \Big[(1 - |\cos\theta|^2)\cos(\Delta m\,\Delta t) - 2\,\mathrm{Im}(\cos\theta)\sin(\Delta m\,\Delta t)$$
$$+2\,\mathrm{Re}(\cos\theta)\sinh(\Delta\Gamma\,\Delta t/2) + \cosh(\Delta\Gamma\,\Delta t/2)(1 + |\cos\theta|^2)\Big] \tag{A.46}$$

The total number of opposite-sign dilepton events from neutral B pairs is

$$N^{+-} \equiv N(\ell^+\ell^-) \sim \frac{|A_{\ell^+}\bar{A}_{\ell^-}|^2}{\Gamma^2} \frac{[2 + x_d^2 - y_d^2 + (x_d^2 + y_d^2)\,|\cos\theta|^2]}{(1 + x_d^2)(1 - y_d^2)} \tag{A.47}$$

The fraction of the same-sign dileptons in the total sample is

$$\chi_d = \frac{|\sin\theta|^2(|e^{-i\phi}|^2 + |e^{i\phi}|^2)\,(x_d^2 + y_d^2)}{|\sin\theta|^2(|e^{-i\phi}|^2 + |e^{i\phi}|^2)\,(x_d^2 + y_d^2) + 2(2 + x_d^2 - y_d^2) + 2(x_d^2 + y_d^2)\,|\cos\theta|^2} \tag{A.48}$$

For a small $\mathrm{Im}(\phi)$

$$\chi_d \equiv \chi_d(x_d, y_d, \theta) = \frac{|\sin\theta|^2\,(x_d^2 + y_d^2)}{(2 + x_d^2 - y_d^2) + (x_d^2 + y_d^2)\,(|\cos\theta|^2 + |\sin\theta|^2)} \tag{A.49}$$

A.2 The CPT violating event generator

The QQ memxcp.F subroutine does the mixing and the CP violation through the mixing mechanism. The modified memxcp.F can handle CPT violation for coherent mixing with charge conjugation $C = -1$ ($B^0\bar{B}^0$) for CP – e.g. $J/\Psi K_S$ – or non-CP – e.g. dileptons – decay modes[6]. Even though the memxcp.F can be used with $\cos\theta = 0$, in case of CPT conservation a flag switches to the original version of memxcp.F to eliminate the effect of a hypothetical bug in the code.

The parameter of the program that controls the CPT violation is the complex angle $\theta \equiv \theta_1 + i\,\theta_2$, given in degrees and converted to radians. Then, $\cos\theta$ and $\sin\theta$ are calculated using the relations

$$\mathrm{Re}(\cos\theta) = \cos\theta_1 \frac{e^{-\theta_2} + e^{\theta_2}}{2}, \qquad \mathrm{Im}(\cos\theta) = \sin\theta_1 \frac{e^{-\theta_2} - e^{\theta_2}}{2} \tag{B.1}$$

$$\mathrm{Re}(\sin\theta) = \sin\theta_1 \frac{e^{-\theta_2} + e^{\theta_2}}{2}, \qquad \mathrm{Im}(\sin\theta) = \cos\theta_1 \frac{e^{\theta_2} - e^{-\theta_2}}{2} \tag{B.2}$$

A.2.1 Dilepton mode

First the generator decides for mixing or non-mixing by comparing a random number RAN(0) to the ratio

$$R = \frac{N^{++} + N^{--}}{N^{+-}} = \frac{|\sin\theta|^2 x_d^2}{(2 + x_d^2 + x_d^2 |\cos\theta|^2)}$$

- **<u>If RAN(0)< R:</u>** then mixing occurs. 50% of the time the generator will produce a B^0B^0 pair and 50% of the time a $\bar{B}^0\bar{B}^0$ pair. The B mesons will decay with a time difference $\Delta t \equiv t^*_{B_1} - t^*_{B_2}$ from each other[7], with Δt following the distribution

$$P(B^0B^0/\bar{B}^0\bar{B}^0, \Delta t) \sim e^{-|\Delta t/\tau_{B^0}|}[1 - \cos(\Delta m\,\Delta t)]$$

[6]The subroutine can be found on the bwg03 machine, under ~leo/work/safe/
[7]See Sec. A.1.3.

- **If RAN(0)> R:** then there is no mixing. A $B^0\overline{B}^0$ pair is generated and the B mesons decay with a time difference $\Delta t \equiv t^*_{B_1} - t^*_{B_2}$ from each other, with $\Delta t \equiv t_{B^0} - t_{\overline{B}^0}$ following the distribution

$$P(B^0\overline{B}^0, \Delta t) \sim e^{-|\Delta t/\tau_{B^0}|}\Big[(1 - |\cos\theta|^2)\cos(\Delta m\,\Delta t) - 2\,\mathrm{Im}(\cos\theta)\sin(\Delta m\,\Delta t) + (1 + |\cos\theta|^2)\Big]$$

A.2.2 CP Eigenstate mode

For completeness, we give the proper time distributions and the relative fractions of B^0 and $\overline{B}^0$ flavor tags for CP modes[8]. Here we use the "gold-plated" $J/\psi K_S$ mode as an example. $\sin 2\beta$ is the CP violation parameter.

The generator decides if the B^0 or the $\overline{B}^0$ will decay to the CP eigenstate by comparing a random number RAN(0) to the ratio

$$R_{CP} \equiv \frac{N^+}{N^-} = \frac{2 + x_d^2\left[1 + |\sin\theta|^2 + |\cos\theta|^2 - 2[\mathrm{Im}(\cos\theta^*\sin\theta)\sin 2\beta + \mathrm{Re}(\cos\theta^*\sin\theta)\cos 2\beta]\right]}{2 + x_d^2\left[1 + |\sin\theta|^2 + |\cos\theta|^2 - 2[\mathrm{Im}(\cos\theta^*\sin\theta)\sin 2\beta - \mathrm{Re}(\cos\theta^*\sin\theta)\cos 2\beta]\right]}$$

- **If RAN(0)< R_{CP}:** then one B decays as B^0 (f_+ generic decay). The other B decays to the CP eigenstate (traditionally, it decays as $\overline{B}^0$, even though this is only a technicality that does not affect the generation process).

- **If RAN(0)> R_{CP}:** then one B decays as $\overline{B}^0$ (f_- generic decay) and the other B decays to the CP eigenstate (as B^0).

The B mesons will decay again with a time difference $\Delta t \equiv t^*_{B_1} - t^*_{B_2}$ from each other, with $\Delta t \equiv t_{J/\Psi K_S} - t_{f_\pm} \equiv t_{CP} - t_{f_\pm}$ following the distribution

$$P(J/\Psi K_S, f_\pm; \Delta t) \sim e^{-|\Delta t/\tau_B|}\Big[\Big\{1 - |\sin\theta|^2 - |\cos\theta|^2 + 2[\mathrm{Im}(\cos\theta^*\sin\theta)\sin 2\beta \mp \mathrm{Re}(\cos\theta^*\sin\theta)\cos 2\beta]\Big\}\cos(\Delta m\,\Delta t)$$

[8]The formalism for the CP modes including CPT violation can be found in Ref. [49].

$$+ 2\Big\{\mathrm{Im}(\sin\theta)\cos 2\beta \pm \mathrm{Re}(\sin\theta)\sin 2\beta \pm \mathrm{Im}(\cos\theta)\Big\}\sin(\Delta m\,\Delta t)$$

$$+ 1 + |\sin\theta|^2 + |\cos\theta|^2 - 2[\mathrm{Im}(\cos\theta^*\sin\theta)\sin 2\beta \mp \mathrm{Re}(\cos\theta^*\sin\theta)\cos 2\beta]\Big]$$

Bibliography

[1] J.W. Herb *et al.*, Phys. Rev. Lett. **39** 252 (1977).

[2] M. Kobayashi and T. Maskawa, Prog. of Theo. Phys. **49** 652 (1973).

[3] I.I. Bigi and A.I. Sanda, "CP Violation", Cambridge University Press, 2000.

[4] The conservative estimation used by most phenomenologists is $\frac{\Delta\Gamma}{\Gamma} \leq 10^{-2}$. There is no experimental evidence for $\Delta\Gamma \neq 0$. See also the discussion in [5], pg. 555-557.

[5] Review of Particle Physics, C. Caso *et al.*, Eur. Phys. J. **C3** (1998) 1.

[6] See for example:

D. Colladay and A. Kostelecky, Phys. Rev. **D58** 116002 (1998).

D. Colladay and A. Kostelecky, Phys. Lett. **B344** 259-265 (1995).

[7] M. Kobayashi and A.I. Sanda, Phys. Rev. Lett. **69**, 3139 (1992).

[8] P. Colangelo and G. Corcella, Eur. Phys. J. C **1**, 515-522 (1998).

[9] Z.Z. Xing, Phys. Rev. D **50**, 2957-2961 (1994).

[10] B. Casey, "Measurement of the number of $B^0\overline{B}^0$ events in Experiment 5 Data and the $B^0\overline{B}^0$ Cross Section at KEKB", Belle Note # 296, March 1st, 2000, and

B. Casey, private communication (updates on continuum contamination).

[11] `http://www.lns.cornell.edu/public/lab-info/upsilon.html`

[12] "KEKB - *B* Factory Design Report", KEK-Report 95-7, 1995.

[13] Technical Design Report, KEK-Report 95-1, April 1995.

The Belle Collaboration, "Belle Progress Report, April 1996 - March 1997", March 1997.

[14] G. Alimonti *et al.*, KEK preprint 2000-34.

[15] R. Fruhwirth, Nucl. Instrum. Methods, **A 262**, 444 (1987).

R. Harr, IEEE Transactions on Nuclear Science, Vol. 42, No. 23, 134 (1995).

[16] H. Hirano *et al.*, KEK preprint 2000-2, submitted to Nucl. Instrum. Methods.

M. Akatsu *et al.*, DPNU-00-06, submitted to Nucl. Instrum. Methods.

[17] T. Iijima *et al.*, Proceedings of the 7th International Conference on Instrumentation for Colliding Beam Physics, Hamamatsu, Japan, Nov. 15-19, 1999.

[18] H. Kichimi *et al.*, submitted to Nucl. Instrum. Methods.

[19] V.M. Aulchenko *et al.*, Nucl. Instrum. Meth. **A379** 491-494 (1996).

[20] H. Ikeda *et al.*, Nucl. Instrum. Methods, **411**, 401 (2000).

[21] Credit for Figs. 3.8 and 3.9 goes to Sven Vahsen, graduate student at Princeton University.

[22] A. Abashian *et al.*, Nucl. Instrum. Methods, **A449** 112 (2000).

[23] R. Itoh, "BASF User's Manual", Belle Note # 161, December 23rd, 1996.

[24] R. Itoh and S. Ichizawa, "Prototype of Belle data processing framework", Proceedings of International Conference on Computing in High Energy Physics, CHEP95, Rio de Janeiro (1995).

R. Itoh and S. Ichizawa, "Status of Belle event processing framework based on a SMP-server cluster system", Belle Note # 97, November 6th, 1995.

[25] S. Nagayama, "PANTHER User's manual", Belle Note # 130, November 6th, 1996.

S. Nagayama, "PANTHER reference manual", Belle Note # 131, November 6th, 1996, and

S. Nagayama, "PANTHER Primer", Belle Note # 132, November 6th, 1996.

[26] H. Ozaki, "Mini-DST Tables V0.0", Belle Note # 146, October 1st, 1996, and

H. Ozaki, "Mini-DST Tables V0.60", November 9th, 1998.

[27] R. Itoh, "QQ Quick Reference for Belle", version 1.04 (internal Belle documentation).

[28] `http://www.lns.cornell.edu/public/CLEO/soft/QQ/qq_expert.html`

[29] T. Sjostrand, Computer Physics Commun. **82**, 74 (1994).

[30] B. Anderson *et al.*, Phys. Rep. **97**, 31 (1983).

[31] `http://consult.cern.ch/writeup/geant/`

[32] H. Ozaki, "FSIM V5.0", Belle Note # 99, February 23rd, 1996.

[33] K. Hanagaki *et al.*, "Status of Electron Identification", Belle Note # 312, May 1st, 2000.

[34] L. Piilonen *et al.*, "Belle muon identification", Belle Note # 338, June 28th, 2000.

[35] G. Fox and S. Wolfram, Phys. Rev. Lett. **41** 1581 (1978).

[36] For a description of the vertex fitting subroutine, see

J.I. Tanaka, "KFitter Usage and Effect", version 3.1, Belle Note # 193, April 20th, 2000.

[37] The current experimental limit for non-B meson decays of the $\Upsilon(4S)$ is less than 4% at the 95% confidence level. See

CLEO Collaboration, B. Barish *et al.*, Phys. Rev. Lett. **76**, 1570 (1996).

[38] CLEO Collaboration, J.P. Alexander *et al.*, CLNS 00-1670, CLEO 00-7, hep-ex/00006002, submitted to Phys. Rev. Lett.

We use the CLEO measurement and the B lifetime ratio from [5].

[39] CLEO Collaboration, L. Gibbons *et al.*, Phys. Rev. **D56** 3783 (1997).

[40] DELPHI Collaboration, P. Abreu *et al.*, CERN-EP/99-162, submitted to Phys. Lett. B.

[41] F.M.L. Almeida Jr., M. Barbi and M.A.B. do Vale, Nucl. Instrum. Meth. **A449** (2000), 383-395.

[42] The OPAL Collaboration, K. Ackerstaff *et al.*, Z. Phys. **C76**, 401-415 (1997).

[43] DELPHI Collaboration, M. Feindt, C. Kreuter and O. Podobrin, DEPHI 97-98, CONF 80, July 20th, 1997.

[44] J. Suzuki, "Measurement of $B^0\bar{B}^0$ oscillation frequency Δm and wrong flavor tag fraction using $B^0 \to D^{(*)}\ell\nu$", Belle Note # 313, April 27th, 2000.

[45] Y. Grossman, B. Kayser and Y. Nir, Phys. Lett. B **415** (1997) 90-96.

[46] B. Kayser, FERMILAB-CONF-96/429-T, Proceedings of the Summer School in High Energy Physics and Cosmology, International Center for Theoretical Physics, Trieste, Italy, June-July 1995, hep-ph/9702264.

[47] A. Mohapatra *et al.*, Phys. Rev. D **58** (1998).

[48] B. Kayser, FERMILAB-CONF-95-287-T, Proceedings of the Moriond Workshop on Electroweak Interactions and Unified Theories, Les Arcs, France, March 1995, hep-ph/9509386.

B. Kayser, NSF-PT-96-3, Proceedings of the 28th International Conference on High Energy Physics, Warsaw, July 1996, hep-ph/9702327.

[49] C. Leonidopoulos, "CPT violation in the $B^0 - \overline{B}^0$ mixing: A Simulation Study for the Belle experiment", Belle Note # 279, January 4th, 2000.