Measuring $\sin(2\beta)$ in the Golden Channel at BaBar.

Master thesis submitted

8th of February 2001

for the degree of

Cand. Scient.

by

Troels C. Petersen

Department for Experimental High Energy Physics

Niels Bohr Institute

University of Copenhagen

This master thesis is submitted for the title of cand. scient. thereby concluding the physics studies of Troels C. Petersen at University of Copenhagen. It was commenced in February 2000 in Paris and finalized in February 2000 in Copenhagen.

The work was mainly supported by the Danish Ministry of Education and the Danish Ministry of Research. Additional funds were granted through the Socrates-Erasmus program, by Direktør Einar Hansen og hustru Vera Hansens fond c/o, Observatør mag. scient. Julie Marie Vinter Hansens Rejselegat and most generously by Lørup Scholarstipendium.

Funds for the collaboration meetings in Annecy, France (March 2000) and San Francisco, USA (April 2000) were provided by the BaBar-group at Laboratoire de Physique Nucléaire et des Hautes Energies (LPNHE), and funds for the CP-conference in Ferrara, Italy (September 2000), the CP-school in Prerow, Germany (October 2000) and the collaboration meeting in San Francisco, USA (December 2000) were provided by the HEP-group at the Niels Bohr Institute (NBI).

Supervisors:

Lektor John Renner Hansen
Niels Bohr Institute, HEP
Blegdamsvej 17, Copenhagen Ø
Denmark
renner@nbi.dk

Prof. François Le Diberder
LPNHE, IN2P3 - CNRS, Paris VI
Tour 33, 4 Place Jussieu, Paris 75005
France
diberder@in2p3.fr

Submitted 8$^{\text{th}}$ of February 2001

Troels C. Petersen

This document is set in LaTeX, figures made using `xfig` and graphs using `PAW`.

Abstract

The CKM matrix and CP violation in the B-system are some of the tasks that remains in constraining and thereby checking the Standard Model (SM). Asymmetric B-factories play an essential role in this task and have despite their infancy already contributed significantly.

In the following the theory of B-meson phenomenology and CP violation is reviewed and the PEP-II collider and the BaBar detector are described. Using $9.0\mathrm{fb}^{-1}$ of data collected at the $\Upsilon(4S)$-resonance by the BaBar detector at SLAC, $\sin(2\beta)$ is measured using the golden channels $B^0 \to J/\psi K_S^0$ and $B^0 \to \psi(2S)K_S^0$.

This measurement requires that the b-quark flavor can be determined (tagged), which is done by observing the flavor of the other coherently produced B-meson. The ability to tag has to be measured, as the uncertainty in the tagging efficiency propagates directly to the $\sin(2\beta)$ result. A measurement of the tagging efficiencies using fully reconstructed hadronic and semileptonic events is presented and the sources of backgrounds are estimated in particular the non-mixing background from charged B-mesons.

Both in the extraction of the tagging efficiency and in the measurement of $\sin(2\beta)$ an unbinned likelihood fit is normally used. However, as the analysis is new, checks with simpler methods are needed. Such methods, nearly optimal and less vulnerable to systematic errors, have been developed and applied for computing the results.

The value of the CP asymmetry obtained was: $\sin(2\beta) = 0.12 \pm 0.37$ (stat.) $\pm\, 0.09$ (syst.)

Il faut de toute necessité que des actions dissymetriques president pendant la vie à l'élaboration des vrais principes immédiats naturels dissymetriques. Quelle peut être la nature de ces actions dissymetriques? Je pense, quant à moi, qu'elles sont d'ordre cosmique. L'univers est un énsemble dissymetrique et je suis persuadé que la vie, telle qu'elle se manifeste à nous, est fonction de la dissymetrie de l'univers ou des consequences qu'elle entraîne. L'univers est dissymetrique.

It is inescapable that asymmetric forces must be operative during the synthesis of the first asymmetric natural products. What might these forces be? I, for my part, think that they are cosmological. The universe is asymmetric and I am persuaded that life, as it is known to us, is a direct result of the asymmetry of the universe or of its indirect consequences. The universe is asymmetric.

[Louis Pasteur 1822-1895]

Contents

List of Figures

List of Tables

1 Introduction

1.1 What is B-physics?

B-physics is the study of the bottom quark, b, which is done through its bound hadronic states in B-mesons. The main interest is measuring the couplings between the different quarks, thereby testing the Standard Model (SM). Furthermore, B-mesons provides basis for tests and development of Quantum Chromo Dynamics (QCD), since the heaviness of the b quark allows for perturbative calculations.

Since the discovery and the following experimental accessibility of the b-quark, B-physics has become a major field in particle physics. Dedicated B-experiments along with theoretical developments have laid the ground, and the newly build B-factories will probe the field much further in the coming years.

1.1.1 History of CP-violation, the CKM-matrix and B-physics

Since the birth of physics, Nature was thought to be symmetric in time and space, and after the discovery of antiparticles, these too were considered as exact opposites (Charge conjugates, C) of each other.

In 1956, after reviewing the experimental data then available, Lee and Yang concluded that spatial inversion (Parity, P) was not conserved in weak interactions [1]. The year after (1957) Wu *et al.* discovered parity violation in the β-decay of Co^{60} to Ni^{60*} [2], and this was confirmed by others [3]. However CP was still believed to be a good symmetry of Nature.
In 1964, CP-violation was discovered by Christenson *et al.* in the decay of the longlived neutral kaon [4], which occasionally (2×10^{-3}) decayed into a state with opposite CP, thereby breaking CP-invariance.

Already in 1963, Cabibbo proposed mixing between the d and the s quark [5], and in 1970 Glashow, Iliopoulos and Maiani suggested a fourth quark (the charm quark, c) in this scheme to cancel an unwanted (i.e. unobserved) $\Delta S = 1$ neutral current [6]. An extension of this quark scheme to six quarks was proposed in 1972 by Kobayashi and Maskawa [7], since a 3×3 complex mixing matrix, now called the Cabibbo, Kobayashi and Maskawa (CKM) matrix, had the possibility of CP-violation built into it.

With the discovery of the b-quark in the Υ resonance in 1977 [8], the extension to three generations was verified, and a new field of physics – B-physics – was born. The goal is among other things to determine the CKM-matrix in various ways to check the consistency of the Standard Model, thereby testing it.

1.1.2 CP-violation in cosmological context

In 1966, two years after the discovery of CP-violation and just one year after the discovery of the microwave background radiation predicted by Big Bang theory, Sakharov wrote a paper on the possibility of explaining the apparent charge asymmetry of the Universe in terms of particle theory [9]. He argued that three conditions (now known as the Sakharov conditions) had to be satisfied in order to obtain a baryon asymmetry in the Universe:

- *Baryon non-conservation.* If the total baryon number of the Universe were initially zero and baryon number was conserved, then the Universe would remain symmetric.
- *C and CP violation.* If C or CP were conserved, the rate of reactions would be the same for particles and antiparticles, and an initially symmetric Universe would remain symmetric.
- *Deviations from thermal equilibrium.* In thermal equilibrium there is no time dependence, and consequently an initial baryon number of zero would remain that way.

There are of course ways around Sakharov's conditions. It could be that the Universe is only locally asymmetric, but this has now been ruled out locally (20 Mpc) by direct search [10] and globally (visible Universe) by indirect methods [11]. More exotic models also exist, but Sakharov's conditions are generally believed to be true, which makes CP violation extremely interesting also from a existential point of view. In addition, detailed models of baryogenesis implies that there are sources of CP violation beyond the Standard Model [12].

1.2 Preface

Before reading this thesis, it is recommendable to read below, as parts of the thesis is influenced by external factors, which are described there. Writing customs are pointed out, the experimental situation is summarized, and the variety of subjects treated in this thesis is explained.

1.2.1 About this thesis

In order to benefit from this thesis, the reader should have a basic knowledge of particle physics, and also some knowledge of B-physics is advisable. An introduction to particle physics can be found at [13], general references in [14],[15], and a reference on B-physics at [16].

Three parts of the thesis were originally written for other purposes, and have been included in rewritten form. The first part is on B-mixing, for the class 'Weak Interactions' at Orsay. The second part is the section on OneBin, which was BaBar Analysis Document (BAD) #51 [17], and last part is on the D^{**} background, which was written for BAD #22 [18] and BAD #151 [71].

In the following, abbreviations will be used extensively as is custom in particle physics. At first appearance they will be typed out and the abbreviation written afterward in parenthesis. However, the names of accelerators and experiments will not be written out, and the unfamiliar reader will have to look them up in the list of abbreviations in the appendix.

The convention $c = \hbar = 1$ will be used as is customary, which means that all masses and momenta will be measured in terms of GeV and not GeV/c^2 and GeV/c respectively. Times and distances will however be kept in normal units, most often ps and μm since they are the relevant scales for B-physics. The symbol B^0 will refer to the B_d-meson, and charge conjugate modes are implied throughout the thesis unless otherwise stated[1]. Four vectors are written with normal font, while three vectors are written in bold, as are matrices. Most but far from all quotations are taken from [19].

1.2.2 Experimental situation

The experimental situation up to the commencement of BaBar was that ALEPH and OPAL at LEP and CDF at the Tevatron each had a $\sin(2\beta)$ measurement, while CLEO-II and to a certain extend CDF had the best limits on B decays.

The next generation of experimental B-physics besides BaBar includes, Belle, CDF/D0 Run-II, CLEO-III and HERA-B. However, HERA-B have many problems, CLEO-III still have no boost and Run-II at the Tevatron will not start until 2001, and so Belle has been the largest competitor. Belle started data taking before BaBar, and even though BaBar has surpassed Belle with respect to amount of data, Belle have had more time to understand their detector and data.

The goal of BaBar is to produce at least 10^8 $B^0\overline{B}^0$ pairs in order to constrain the third generation entries in the CKM-matrix, thereby studying CP-violation and testing the SM. A $\sin(2\beta)$ measurement is one of the primary goals, but also measurements of $\sin(2\alpha)$ and $\sin(2\beta+\gamma)$ are planned. Additionally, rare decays, semileptonic and radiative penguins, hadronic B- and C-physics and Quantum Electro Dynamics (QED) with τ-physics will be studied.

[1] e.g. $B^0 \to D^-\pi^+$ implicitly includes the conjugate decay $\overline{B}^0 \to D^+\pi^-$

1.2.3 Subjects treated

The work on BaBar was started in Paris in February 2000 without any prior work on the subject. At the time, $\sin(2\beta)$ was the benchmark measurement BaBar had to present at the ICHEP2000 conference. Therefore it was natural to work on this subject (with François Le Diberder & Lydia Roos mainly, and Bruno Durin) along with the measurement of the tagging efficiencies (with Stephane Plaszczynski, Marie-Helene Schune and François Le Diberder). This subject takes up most of this thesis. The idea of using just one bin in the TagMix analysis was improved to include smearing and biases, and the idea was extended to also include the CP-fitting. For the evaluation of the charged background in the semileptonic mixing sample, an extensive toy Monte Carlo (MC) was built, and the contamination in the TagMix data sample was estimated (with Stephane Plaszczynski).

Jumping in subject, the decay $B \to D^{(*)} K_S^0 \pi$ was proposed by Roy Aleksan to measure $\sin(2\beta + \gamma)$, and though the work alone would suffice for a thesis, some efforts were made on this channel. A Monte Carlo with a threebody generator was build and the signal efficiency and background rejection potential was estimated using the BaBar Monte Carlo (with Jan Stark). However, since there was no lack of material for this thesis, this subject was omitted.

Parts of the analysis were carried out for the XXXth International Conference on High Energy Physics (ICHEP2000), held from the 27$^{\text{th}}$ of July to the 2$^{\text{nd}}$ of August 2000 in Osaka, Japan. The next analysis is aiming at the Winter Conferences in late February 2001, and since no results will be published before then, the analysis presented in this thesis is in accordance with the one for Osaka, from where the *results are all preliminary*. However, both the amount of data and the analysis has improved much since then, and consequently new features in the analysis will be mentioned occasionally, and at the end a few comments will be devoted to the progress. Last but not least the new result for $\sin(2\beta)$, to be presented at a SLAC seminar the 12$^{\text{th}}$ of February, has been included.

1.3 Outline

This thesis consists of three parts. The first is a theoretical introduction to B-physics, which goes through mixing and CP violation both in general and in the cases specific for this thesis. In addition, the current knowledge of the CKM-matrix and CP-violation is summarized.

The second part describes the PEP-II accelerator and the BaBar detector along with the acceptance and triggers. Given the time period of the thesis, direct work with the detector has not been possible, so this part will mostly be a description based on references.

The third part contains the analysis. After an outline of the analysis the selection of the data samples is presented, and then the tagging algorithms and the resolution function are described. After that follows a thorough discussion of the analysis and the fitting methods behind the dilution measurement (the TagMix analysis), and both the mixing and CP results are then presented and discussed.

Part I
Physics

Man masters Nature not by force but by understanding. That is why science has succeeded where magic failed: Because it has looked for no spell to cast on Nature.

[Jacob Bronowski 1908-1974]

2 The CKM matrix and *CP* violation

With the discovery of the b quark, a third generation of quarks was established. Extending quark mixing to three generations, parametrized by the Cabbibo-Kobayashi-Maskawa (CKM) mixing matrix, V_{CKM}, gives rise to a CP-violating phase, which provides an elegant explanation of the well-established CP violation observed in K_L^0 decay. However, studies of CP violation in the neutral kaon system does not allow for the crucial test whether CP violation is described by the CKM mixing matrix alone. This gives B-physics a central role, to be investigated in the following.

First, the CKM matrix will be accounted for on the basis of the Standard Model (SM) and the symmetries of the theory investigated. Secondly, the various parametrizations of the CKM matrix and CP violation in general will be discussed, and finally a short overview of experimental determination of the CKM matrix will be given. This section is based on [14], [16] and [20] unless explicitly stated.

2.1 Why the CKM matrix?

The Standard Model, so far accurately describing all experimental data in particle physics, consists of the strong force (an $SU(3)_C$ gauge theory) and the ElectroWeak (EW) force (a broken $SU(2)_L \times U(1)_Y$ gauge theory). The Lagrangian of this theory is highly restricted by the requirement of renormalizability and gauge invariance. In order to write mass terms, i.e. assign masses to the particles, without breaking gauge invariance, one has to introduce Higgs couplings[2]. These couplings to quarks do *not* follow from a gauge principle and will in general mix the various flavors.

Denote by u_L^i and d_L^i the left handed u-type and d-type quarks in their flavor basis and let $u_L'^i$ and $d_L'^i$ be the quarks in their physical basis (that is the basis that diagonalize the Higgs couplings). The two bases are then related by a unitary transformation:

$$u_L^i = (u_L, c_L, t_L), \quad u_L'^i = (u_L', c_L', t_L'), \quad \longrightarrow \quad u_L^i = \mathcal{U}_u^{ij} u_L'^j \tag{1}$$

$$d_L^i = (d_L, s_L, b_L), \quad d_L'^i = (d_L', s_L', b_L'), \quad \longrightarrow \quad d_L^i = \mathcal{U}_d^{ij} d_L'^j \tag{2}$$

When rewriting the SM Lagrangian from flavor to physical states, one introduces these matrices into the theory. In Quantum Chromo Dynamics (QCD) the quarks have identical couplings, and so the $\mathcal{U}$-matrices cancel out, since they commute with this coupling. Commutation with the covariant derivative makes the $\mathcal{U}$-matrices drop out of the kinetic terms, and then only the $SU(2)_L \times U(1)_Y$ couplings remains. In the ElectroMagnetic (EM) couplings they also vanish, as can be seen from a typical term $eA_\mu \bar{u}_L^i \gamma^\mu u_L^i = eA_\mu \bar{u}_L'^i \gamma^\mu (\mathcal{U}_u^\dagger \mathcal{U}_u) u_L'^i = eA_\mu \bar{u}_L'^i \gamma^\mu u_L'^i$, and equivalently for the Z^0-terms. This is experimentally well founded since flavor changing neutral currents are not observed at tree level.

The only place were the $\mathcal{U}$-matrices enters is in the flavor changing charged current interactions mediated by the W-bosons, which looks as follows:

$$\mathcal{L}_W = \frac{g}{\sqrt{2}} \bar{u}_L^i \gamma^\mu d_L^i W_\mu^+ = \frac{g}{\sqrt{2}} \bar{u}_L'^i \gamma^\mu \mathcal{U}_u^\dagger \mathcal{U}_d d_L'^i W_\mu^+ \equiv \frac{g}{\sqrt{2}} \bar{u}_L'^i \gamma^\mu V_{\text{CKM}} d_L'^i W_\mu^+, \tag{3}$$

where, in the Lagrangian, Hermitian conjugates are implicit. The matrix, $V_{\text{CKM}} \equiv \mathcal{U}_u^\dagger \mathcal{U}_d$, enters the Charged Currents (CC) as follows:

$$J_{\text{CC}}^\mu = \sum_{\text{color}} \begin{pmatrix} \bar{u} \\ \bar{c} \\ \bar{t} \end{pmatrix}_L \gamma^\mu \tfrac{1}{2}(1-\gamma_5)\, V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L . \tag{4}$$

[2]Other mechanisms for generating mass exists, but they seem in conflict with observation.

Thus, as a result of the fact that the physical (mass) eigenstates are not the same as the weak (flavor) eigenstates, one has to introduce a rotation matrix between the two. This rotation matrix, named the CKM matrix after its inventors Cabbibo-Kobayashi-Maskawa, relates the flavor and mass eigenstates of the three quark doublets, as can be parametrized by rotation of the d-type quarks as follows:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{5}$$

Local gauge invariance and baryon number conservation requires V_{CKM} to be unitary:

$$V_{\mathrm{CKM}} V_{\mathrm{CKM}}^{\dagger} = V_{\mathrm{CKM}}^{\dagger} V_{\mathrm{CKM}} = 1. \tag{6}$$

Given N_g generations of particles, the CKM matrix has a priori $2N_g^2$ degrees of freedom (dof.). Out of these N_g^2 dof. are removed by the unitarity condition and $2N_g - 1$ dof. (number of *relative* phases between quarks) can be absorbed by redefining the quark fields. Thus, one is left with $N_{\mathrm{dof}} \equiv (N_g - 1)^2$ degrees of freedom. In a world with only two generations, $N_{\mathrm{dof}} = 1$, which is just a rotation angle in a real unitary 2×2 matrix. The SM has three generations, making $N_{\mathrm{dof}} = 4$, which gives the three angles of a 3×3 rotation matrix (the Euler angles) and one complex phase[3]. *It is this complex phase that gives rise to CP violation in the SM.*

2.1.1 Mixing among leptons

Since the CKM matrix came from the breaking of $SU(2)_L \times U(1)_Y$, which resulted in a difference between the flavor eigenstates and the mass eigenstates and thus a mixing matrix between quarks, an equivalent matrix should exist for the leptons, as they also couple to the weak force.

Originally, it was thought that mixing and CP violation was absent in the lepton sector, since neither lepton number violation nor neutrino masses had been observed. If neutrinos did not couple to the Higgs particle, the mass matrix would commute with the weak eigenstates and there would be no mixing matrix in the lepton sector.

But given the non-zero results on neutrino mixing from (Super)-Kamiokande [21], a mixing matrix for leptons is required and a new source of CP violation introduced[4]. The entries of the Maki-Nakagana-Sakata (MNS) lepton mixing matrix [22] are very poorly known, as only one parameter can be constrained by solar neutrino data into either a small or a large mixing angle, but once ν-factories start running, the other MNS entries will surely be pursued.

2.2 *C*, *P* and *T* transformations

> *Exact observation of reality shows, that spatial relations — the phenomena of symmetry — lie at the basis of all the phenomena we have studied.*
>
> [Vladimir Ivanovich Vernadskii, 1863-1945]

Three discrete operators are potential symmetries of field theories. Two are space-time operators (T, P), which are part of the Poincaré group, while the third is an intrinsic transformation (C):

- C – Charge conjugation. Transforms a particle into its antiparticle, $P(\vec{p}, \vec{s}) \to \overline{P}(\vec{p}, \vec{s})$.
- P – Parity change. Flips the parity (handedness) of space, $P(\vec{p}, \vec{s}) \to P(-\vec{p}, \vec{s})$.
- T – Time reversal. Interchanges the forward and backward lightcone, $P(\vec{p}, \vec{s}) \to P(-\vec{p}, -\vec{s})$.

[3] In general, N_g generations will give $\frac{1}{2}N_g(N_g - 1)$ rotation angles and $\frac{1}{2}(N_g - 1)(N_g - 2)$ phases.
[4] Given the current data, CP violation will be insignificant and thus not to worry about in the following.

In an unbroken gauge theory, both C, P and T are conserved separately, and therefore the strong and the EM force respect these symmetries. The weak force, however, being of chiral (i.e. handed) form, maximally violates C and P. This is because each of the operators interchanges particles which couple to the weak force (left-handed particles and right-handed antiparticles) with particles that don't (left-handed antiparticles and right-handed particles). The combined symmetry CP would be conserved, though, was it not for the breaking of the $SU(2)_L \otimes U(1)_Y$ symmetry. This introduces Higgs couplings, mixing between generations and thus CP violation.

From Nöther's Theorem, which states that *for each symmetry there exist exactly one corresponding conservation law*, CP symmetry induces an abstract quantity, CP. Invariance of CP would imply that this quantum number was conserved in all reactions.

CP violation is not implied by mixing between generations in itself, but by the occurrence of complex couplings in the Lagrangian, which are not possible to remove by any choice of phase redefinitions, as is the case for the CKM matrix — $V_{\text{CKM}} \neq V^*_{\text{CKM}}$. This is because CP transformation of the Lagrangian corresponds to Hermitian conjugation, which changes the sign of any imaginary enteties. Thus, for any complex Lagrangian, CP will generally not be a good symmetry (see section 3.4).

2.2.1 The strong *CP* problem

Even with the powerful restrictions of renormalizability and gauge invariance, one can in principle add terms to the Lagrangian, which violate P and T. For the $SU(2)_L$ and $U(1)_Y$ fields this has no observable effect, but for $SU(3)_C$ it implies an electric dipole moment for the neutron, which is experimentally excluded to an impressive precision.

To solve this problem, known as the strong CP problem, one has either to constrain the Higgs couplings or add additional structure to the Higgs sector [23] introducing a goldstone boson (the axion), which by now is almost excluded by experiment.

2.3 Parametrizations of the CKM matrix

The four parameters of the CKM matrix can be parametrized in many ways of which two will be presented here. The first one is exact and resembles ordinary rotation matrices much. The second is only approximate but emphasizes the hierarchy of the mixing matrix and is thus very intuitive. It should be kept in mind that the parametrization has no influence on the physics.

2.3.1 Standard parametrization

The most obvious way of expressing the CKM matrix is in terms of three rotation angles, $\theta_{12}, \theta_{13}, \theta_{23}$ and one complex phase, δ_{13} as follows [24][5]:

$$V_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \quad (7)$$

where $c_{ij} \equiv \cos(\theta_{ij})$ and $s_{ij} \equiv \sin(\theta_{ij})$, $i, j = 1, 2, 3$ labeling the three generations. An advantage of this representation is that if an angle, θ_{ij}, vanishes, then so does the mixing between the generations i and j. The form can be much simplified by using that $|V_{ub}| \sim 0.003$ is small, so that to an excellent approximation ($\mathcal{O}(10^{-5})$) $c_{13} \simeq 1$ and s_{13} can be neglected compared to terms of order unity.

[5]Kobayashi and Maskawa [7] originally proposed a parametrization closer to that of Cabibbo [5].

2.3.2 Wolfenstein parametrization

Noting that the diagonal terms in the CKM matrix are close to unity, while the couplings between different generations diminishes as one moves away from the diagonal, one can choose to emphasize this hierarchy by writing the entries in powers of $\lambda \equiv s_{12} = 0.2205 \pm 0.0018$, the sine of the Cabibbo angle [5], as follows:

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{8}$$

where A, ρ and η are real numbers intended to be of order unity. This approximation is called the Wolfenstein parametrization [25], and it is valid up till fourth order of λ, which is accurate enough for most purposes ($\mathcal{O}(10^{-3})$).

2.4 The unitary triangle(s)

Using Eq. (6) the nine constrains on the CKM matrix mentioned in section 2.1 can be written explicitly as:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} V_{ud}^* & V_{cd}^* & V_{td}^* \\ V_{us}^* & V_{cs}^* & V_{ts}^* \\ V_{ub}^* & V_{cb}^* & V_{tb}^* \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{9}$$

Three of these nine linearly independent equations are normalizations and six are complex equations written below, divided according to their origin with respect to quark generations:

$$\begin{array}{ccc} \underline{1 \otimes 2 \text{ generation}} & \underline{1 \otimes 3 \text{ generation}} & \underline{2 \otimes 3 \text{ generation}} \\ (\mathcal{O}(\lambda)+\mathcal{O}(\lambda)+\mathcal{O}(\lambda^5) = 0) & (\mathcal{O}(\lambda^3)+\mathcal{O}(\lambda^3)+\mathcal{O}(\lambda^3) = 0) & (\mathcal{O}(\lambda^4)+\mathcal{O}(\lambda^2)+\mathcal{O}(\lambda^2) = 0) \\ V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0 & \boxed{V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0} & V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \\ V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* = 0 & V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* = 0 & V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* = 0 \end{array} \tag{10}$$

These six equations can each be represented by triangles in the complex plane, since each of the three complex VV^* terms will correspond to a side in the triangle. From their Wolfenstein-structure one can see that the triangles are pairwise very alike, which is due to the high degree of symmetry of the CKM matrix. One can also see that the ones combining first and second generation ($1 \otimes 2$) and second and third ($2 \otimes 3$) are very "flat", as one side is much smaller (λ^4 and λ^2 respectively) than the two others, which can be seen in Fig. 1.

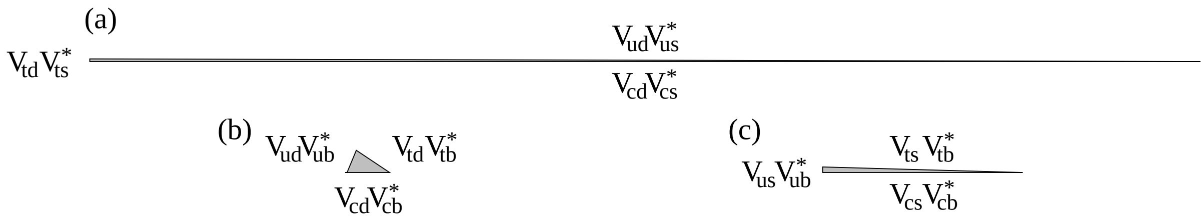


Figure 1: Unitary triangles. The three triangles are geometric interpretations of the three equations $\sum_{q=u,c,t} V_{qa}V_{qb}^* = 0$ with (a) $a = d, b = s$ ($1 \otimes 2$), (b) $a = d, b = b$ ($1 \otimes 3$), (c) $a = s, b = b$ ($2 \otimes 3$). The triangles have equal areas and are here approximately drawn to scale in order to show their shape.

Equation (10) and its triangle representation have the advantage that they are invariant under redefinition of quark fields, i.e. the lengths and angles of the triangles do not change with phase transformation[6] $V_{ij} \to e^{i\phi_i} V_{ij} e^{-i\phi_j}$, and are thus physical observables. Moreover, it is these angles that are directly measured in branching ratios, and the geometrical interpretation is thus natural.

Figure 1 reveals that only the unitary condition from combining first and third generation couplings of the CKM matrix results in a triangle with comparable sides and angles. The others are very hard to measure, since they contain small angles and sidelengths. For historical reasons the triangle corresponding to the equation $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$ (framed in Eq. (10)), is known as *the* unitary triangle; a terminology that will be used here and onwards. However, the second $1 \otimes 3$ triangle is very alike, as none of the side lengths differs by more than 5 %, and it will therefore be a while before the two can be experimentally distinguished. In Fig. 2 below the unitary triangle is shown:

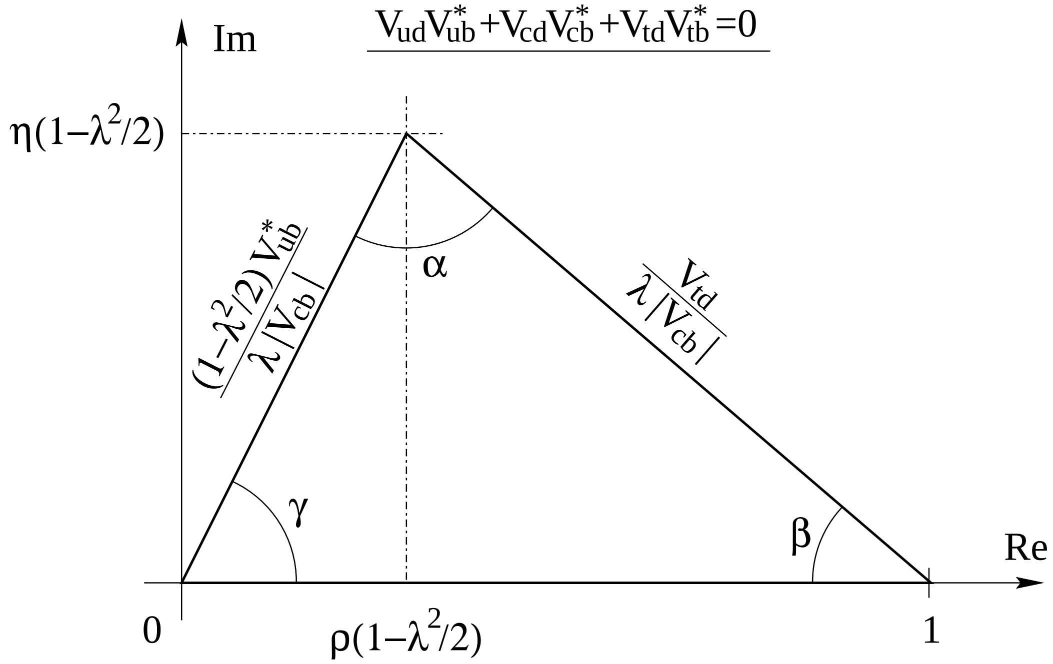


Figure 2: The unitary triangle. The triangle corresponding to the equation $V_{ud}V_{ub}^*+V_{cd}V_{cb}^*+V_{td}V_{tb}^* = 0$, is for historic reasons known as *the* unitary triangle. Note that the baseline has been normalized by $V_{cd}V_{cb}^*$ to equal one.

The three angles of the unitary triangle expressed in terms of CKM matrix entries are:

$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \qquad \beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \qquad \gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{11}$$

These physical observables can be measured by CP asymmetries in various B-decays, and consistency of the measurements provide in part tests of the SM. The real test, though, is to fit all experimental and theoretical CKM related quantities with *only four parameters*, be it in the standard or the Wolfenstein representation. It is this test, which is the real goal of CKM-fitting, where B-physics play a central role.

[6]It is such phase redefinitions that took away $2N_g - 1$ degrees of freedom when counting these in section 2.1.

2.4.1 Requirements for *CP* violation and the Jarlskog parameter

CP is not necessarily violated given three generations. In addition one requires that none of the masses of same-type quarks are equal, that none of the three mixing angles are 0 or $\pi/2$, and that the phase is not 0 or π. The last eight conditions are elegantly summarized in a parameter-independent manner as follows [26]:

$$\mathcal{I}m(V_{ij}V_{kl}V_{il}^{*}V_{kj}^{*}) = J\sum_{m,n=1}^{3}\epsilon_{ikm}\epsilon_{jln}, \quad J \neq 0. \tag{12}$$

The areas of the unitary triangles all equals $|J|/2$, and are thus a measure of the amount of CP violation in the SM. The maximal value J can assume is $1/(6\sqrt{3}) \simeq 0.1$, but the value suggested by data (see section 4.2) is many orders of magnitude smaller, which gives meaning to the statement that CP violation in the SM is small.

2.4.2 Effects of physics beyond Standard Model

Mixing and CP violation are sensitive probes of physics beyond the SM, because SM predictions may be altered by new physics. There are several possibilities, the most interesting being:

- *A fourth generation of quarks*, implying that the unitary triangle should be replaced by a unitary square, such that $\alpha + \beta + \gamma \neq \pi$, and also the amount of mixing and CP violation would be modified.
- Z^0*-mediated flavor changing neutral currents*, which would drastically change the unitary triangle. Moreover, one would in addition need to introduce new couplings between same-type quarks.
- *Super Symmetry (SUSY)*, which would, in its minimal form, only change $B^0\overline{B}^0$ mixing slightly ($\mathcal{O}(20\,\%)$), but less restricted versions could strongly affect both mixing and CP asymmetries.

However, in the following only the SM will be considered, and the above cases will not be treated or discussed.

3 Mixing and CP violation in the B-system

The B-meson is the hydrogen atom of QCD.

[Common saying, though first used by Kurt Gottfried (about charmonium!)]

Like the neutral K-mesons (see section 3.6.1), the neutral B-mesons are complicated by the fact that the flavor eigenstates, with quark composition $\bar{b}d$ (B^0-meson) and $b\bar{d}$ ($\overline{B}^0$-meson)[7], are not the same as the physical states, which have definite masses and lifetimes. Therefore the flavor eigenstates B^0 and $\overline{B}^0$ mix as they propagate through space, and this is what is referred to as $B^0\overline{B}^0$-mixing, which will be investigated next. First the diagrams will be considered followed by a fairly detailed discussion of the Hamiltonian. The eigenstates are found and time developed, and finally various types of CP violation are considered. The discussion is model independent unless otherwise stated.

3.1 The mixing diagrams

The Feynman diagrams responsible for the leading contribution to mixing are so-called *box diagrams*, which are shown in Fig. 3 below:

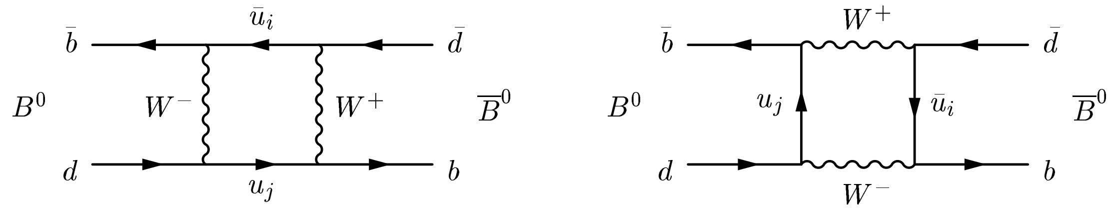


Figure 3: Diagrams for B-mixing. The two dominant diagrams for B-mixing shown above, are suppressed by the GIM mechanism (see appendix B), and get their dominant contribution from top quarks in the loop.

The amplitude for mixing can be directly written down from the two diagrams in Fig. 3, and it looks as follows:

$$\begin{aligned}\mathcal{M}_{\text{Mixing}} = \frac{g^2}{8} \sum_{i,j=u,c,t} V_{u_i d} V^*_{u_i b} V_{u_j d} V^*_{u_j b} [&(\bar{v}_{\bar{b}} \gamma_\mu W_-^{\mu\nu} v_{\bar{u}_i})(\bar{v}_{\bar{u}_i} \gamma_\rho W_+^{\rho\sigma} v_{\bar{d}})(\bar{u}_b \gamma_\sigma u_{u_j})(\bar{u}_{u_j} \gamma_\nu u_d) + \\ &(\bar{u}_b \gamma_\mu W_-^{\mu\nu} v_{\bar{u}_i})(\bar{v}_{\bar{u}_i} \gamma_\rho W_+^{\rho\sigma} v_{\bar{d}})(\bar{v}_{\bar{b}} \gamma_\sigma u_{u_j})(\bar{u}_{u_j} \gamma_\nu u_d)], \quad (13)\end{aligned}$$

where g is the weak coupling constant, V_{ij} are CKM matrix entries, u and v are *lefthanded* spinors (subscripts have been omitted in order not to complicate the formula), γ_μ the Dirac matrices and $W^{\alpha\beta} = -ig^{\alpha\beta}/(p_W^2 - m_W^2)$ are the W-propagators. When integrating over the loop momentum, k, each boson propagator has dimensions k^{-2} and each fermion propagator k^{-1}, and so the integral over d^4k is finite.

If the u-type quark masses were equal, the sum over the diagrams would cancel due to unitarity of the CKM matrix. However, the masses are far from equal, and so the amplitude for mixing is merely suppressed, which is called the GIM mechanism (see appendix B). The SM prediction for the B_d^0 mixing rate, which is exactly the mass difference, Δm_{B_d}, between the two physical B-states (see section 3.2), has been calculated to be [20]:

$$\mathcal{M}_{\text{Mixing}} \simeq \frac{1}{2}\Delta m_{B_d} = \frac{G_F^2}{3\pi^2} \eta_{B_d} m_{B_d} m_W^2 (B_{B_d} F_{B_d}^2) S_0(x_t) |V_{td} V_{tb}^*|^2, \quad (14)$$

[7] The notation convention is that the B^0 is the isospin partner of the B^+, which contains a $\bar{b}$ quark.

where G_F is the Fermi constant, η_{B_d} is a QCD correction factor, $x_q \equiv m_q^2/m_W^2$ and $S_0(x)$ is a kinematic function including additional EW loop contributions, but without QCD corrections (the Inami-Lim function [27]). Due to its large mass, the top quark, t, completely dominates the two mixing diagrams (see appendix B for discussion of contributions), which explains the CKM factors. Since mixing is dominated by short distance contributions, S_0 can be calculated fairly accurately.

The main source of theoretical uncertainty comes from the B^0 and $\overline{B}^0$ states, parametrized by the constants B_B and f_B, which can be estimated by lattice calculations[8] to be $B_{B_d} = 1.30 \pm 0.17$ and $f_{B_d} = 200 \pm 30$. Using these numbers, the SM prediction of the mixing rate is $\Delta m_{B_d} = 0.50 \pm 0.08$, which is in agreement with the measured value (see section 12.1.2), and this in turn gives a bound on $|V_{tb}^* V_{td}|$ (see section 4.2).

The above discussion can be generalized to include the B_s-system, simply by changing d to s. The change in mixing frequency from Δm_{B_d} to Δm_{B_s} is quite large, since the CKM entries are changed by a factor of 30, making Δm_{B_s} very hard to measure, since the typical mixing distance is then smaller than most detector resolutions. A measurement of Δm_{B_s} is desirable, as much of the theoretical uncertainty cancels out in the ratio $\Delta m_{B_d}/\Delta m_{B_s}$, which constrains the CKM matrix.

3.2 The Hamiltonian

The Hamiltonian for a system of mixing and decaying particles, is *not* Hermitian, but can be written as a sum of two Hermitian matrices $\mathbf{M}$ (the dispersive part) and $\mathbf{\Gamma}$ (the absorptive part):

$$\mathbf{H} \equiv \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} = \begin{pmatrix} \langle B^0|H|B^0\rangle & \langle B^0|H|\overline{B}^0\rangle \\ \langle \overline{B}^0|H|B^0\rangle & \langle \overline{B}^0|H|\overline{B}^0\rangle \end{pmatrix}, \tag{15}$$

where the indices $1,2$ refer to whether the initial and final states were B^0 or $\overline{B}^0$. The two parts of the Hamiltonian can be written as:

$$M_{ij} = m_B\delta_{ij} + \langle i|T^{\Delta B=2}|j\rangle + \sum_v \frac{\langle i|T^{\Delta B=1}|v\rangle\langle v|T^{\Delta B=1}|j\rangle}{m_B - E_v} \quad \sim \quad \begin{pmatrix} B^0 \to B^0 & B^0 \to \overline{B}^0 \\ \overline{B}^0 \to B^0 & \overline{B}^0 \to \overline{B}^0 \end{pmatrix}, \tag{16}$$

$$\Gamma_{ij} = 2\pi \sum_f \langle i|T^{\Delta B=1}|f\rangle\langle f|T^{\Delta B=1}|j\rangle \; \delta(E_f - m_B) \quad \sim \quad \begin{pmatrix} B^0 \to f & B^0 \to f \to \overline{B}^0 \\ \overline{B}^0 \to f \to B^0 & \overline{B}^0 \to f \end{pmatrix}, \tag{17}$$

where v and f are virtual and real states respectively and T denotes the transition operator. The last two parentheses illustrate what processes the matrix entry describes; the dispersive part, $\mathbf{M}$, governs mixing through virtual states while the absorptive part, $\mathbf{\Gamma}$, governs decay to real states. It is thus H_{12} and H_{21} that describes mixing between the two B-states.

The diagonal terms of the Hamiltonian matrix $\mathbf{H}$ are mostly due to strong interaction, whereas the non-diagonal terms are generated by the transitions between B^0 and $\overline{B}^0$ states, thus governed by the weak interaction, and for this reason the diagonal terms H_{11} and H_{22} dominate. In the B-system the dispersive part of the off-diagonal elements of the Hamiltonian, M_{12}, is only expected to get short distance contributions, i.e. from high loop momentum, $k \gg \Lambda_{\mathrm{QCD}}$, making it a good probe for heavy intermediate states from new physics. The opposite is the case for Γ_{12}, where long distance effects are expected to be dominant, and it is thus poorly predicted and not a good probe for new physics.

3.2.1 Hamiltonian properties under *CPT* and *CP* transformation

Conservation of CP and CPT constrains the Hamiltonian, $\mathbf{H}$. CPT invariance ensures that particles and antiparticles have equal masses and decay rates, and it is a general requirement of

[8] B_B can be measured in $B^\pm \to \ell^\pm \nu_\ell$ decays, but this is a very tough measurement, currently out of reach.

relativistic field theory (see Appendix A). CP invariance requires in addition that *partial* decay rates of CP-conjugate processes are equal.

Working in the flavor basis, that is with B^0 and $\overline{B}^0$, one has that $CP|B^{\rm in}\rangle = |\overline{B}^{\rm in}\rangle$ and $T|B^{\rm in}\rangle = |B^{\rm out}\rangle^\dagger \equiv \langle B^{\rm out}|$, where $\dagger$ stands for Hermitian conjugation (the charge index has been omitted not to clutter the equations). The incoming and outgoing states, $|B^{\rm in}\rangle$ and $|B^{\rm out}\rangle$ have to be separated, as they belong to two different spaces [28]. For the off-diagonal terms, the Hermiticity of $\mathbf{M}$ and $\mathbf{\Gamma}$, i.e. that $\mathbf{M} = \mathbf{M}^\dagger$ and $\mathbf{\Gamma} = \mathbf{\Gamma}^\dagger$, implies that $M_{12} = M_{21}^*$ and $\Gamma_{12} = \Gamma_{21}^*$.

The constrains from CPT invariance can be obtained from transforming the terms in the last parenthesis of Eq. (15), and one obtains:

$$\langle B^{\rm out}|M|B^{\rm in}\rangle \overset{CPT}{=} \langle B^{\rm out}|(CPT)^\dagger M (CPT)|B^{\rm in}\rangle \overset{\rm def.}{=} \langle \overline{B}^{\rm out}|M|\overline{B}^{\rm in}\rangle \quad\Rightarrow\quad M_{11} = M_{22} \equiv m_B, \tag{18}$$

$$\langle B^{\rm out}|\,\Gamma\,|B^{\rm in}\rangle \overset{CPT}{=} \langle B^{\rm out}|(CPT)^\dagger\,\Gamma\,(CPT)|B^{\rm in}\rangle \overset{\rm def.}{=} \langle \overline{B}^{\rm out}|\,\Gamma\,|\overline{B}^{\rm in}\rangle \quad\Rightarrow\quad \Gamma_{11} = \Gamma_{22} \equiv \Gamma_B, \tag{19}$$

and from the Hermeticity of $\mathbf{M}$ and $\mathbf{\Gamma}$ follows that M_{11}, M_{22}, Γ_{11} and Γ_{22} are real ($\in \mathbb{R}$). CPT invariance is believed to be an exact symmetry of Nature, since it is a general requirement of relativistic field theory, and we will therefore in the following assume CPT invariance.

When considering CP conservation, one has in addition to the CPT requirements that:

$$\langle B^{\rm out}|M|\overline{B}^{\rm in}\rangle \overset{CP}{=} \langle B^{\rm out}|(CP)^\dagger M (CP)|\overline{B}^{\rm in}\rangle \overset{\rm def.}{=} \langle \overline{B}^{\rm out}|M|B^{\rm in}\rangle \quad\Rightarrow\quad M_{12} = M_{21}, \tag{20}$$

$$\langle B^{\rm out}|\,\Gamma\,|\overline{B}^{\rm in}\rangle \overset{CP}{=} \langle B^{\rm out}|(CP)^\dagger\,\Gamma\,(CP)|\overline{B}^{\rm in}\rangle \overset{\rm def.}{=} \langle \overline{B}^{\rm out}|\,\Gamma\,|B^{\rm in}\rangle \quad\Rightarrow\quad \Gamma_{12} = \Gamma_{21}\,. \tag{21}$$

Along the Hermiticity requirements this implies that the off-diagonal elements of the Hamiltonian, H_{12} and H_{21}, are equal and real. However, CP is not an exact symmetry of Nature, and violation of this is exactly the general requirement for CP violation.

3.2.2 Hamiltonian eigenvalues and eigenstates

The eigenvalues λ_L and λ_H are obtained solving the equation:

$$\det(\mathbf{H} - \lambda\mathbf{I}) = \begin{vmatrix} H_{11} - \lambda & H_{12} \\ H_{21} & H_{22} - \lambda \end{vmatrix} = 0, \tag{22}$$

and it is straightforward to obtain the two eigenvalues:

$$\lambda_H \;\equiv\; H_{11} + \sqrt{H_{12}H_{21}} \;=\; (M_{11} - \tfrac{i}{2}\Gamma_{11}) + \sqrt{(M_{12} - \tfrac{i}{2}\Gamma_{12})(M_{12}^* - \tfrac{i}{2}\Gamma_{12}^*)} \;\equiv\; m_H - \tfrac{i}{2}\Gamma_H, \tag{23}$$

$$\lambda_L \;\equiv\; H_{11} - \sqrt{H_{12}H_{21}} \;=\; (M_{11} - \tfrac{i}{2}\Gamma_{11}) - \sqrt{(M_{12} - \tfrac{i}{2}\Gamma_{12})(M_{12}^* - \tfrac{i}{2}\Gamma_{12}^*)} \;\equiv\; m_L - \tfrac{i}{2}\Gamma_L, \tag{24}$$

with m and Γ real. As is customary in the literature, the following intuitive definitions are introduced:

$$m_B \equiv \frac{m_H + m_L}{2}, \quad \Delta m_B \equiv m_H - m_L \quad\Leftrightarrow\quad m_{H,L} \equiv m_B \pm \frac{\Delta m_B}{2}, \tag{25}$$

$$\Gamma_B \equiv \frac{\Gamma_H + \Gamma_L}{2}, \quad \Delta\Gamma_B \equiv \Gamma_H - \Gamma_L \quad\Leftrightarrow\quad \Gamma_{H,L} \equiv \Gamma_B \pm \frac{\Delta\,\Gamma_B}{2}. \tag{26}$$

As the subscripts of the physical B-states in Eqs. (25)+(26) suggest the mass difference, Δm_B, is dominating. Although the width difference, $\Delta\Gamma_B$, has not been measured yet, it is expected to be very small[9], since it is produced by decay channels common to B^0 and $\overline{B}^0$, which have branching ratios of $\mathcal{O}(10^{-3})$ that contribute with alternating sign. It is thus a model independent

[9]The SM predicts $\Delta\Gamma_B < 0$, i.e. the heavy state has the shortest lifetime $\Gamma_{B_L} < \Gamma_{B_H}$, but this is usually ignored.

and safe approximation, that the sum does not exceed the individual level much, and one has that $\Delta\Gamma_B/\Gamma \lesssim \mathcal{O}(10^{-2})$.

The mass difference Δm_B has been measured and is of the same order as the width, from which one deduces that $\Delta\Gamma_B/\Delta m_B = \mathcal{O}(10^{-2})$. It should be noted that this is not necessarily the case for $B_s\overline{B}_s$-system, where the relative width difference is estimated to be of order 10 % due to the larger size of V_{ts} compared to V_{td}.

In order to obtain the physical states, the matrix $\mathbf{H}$ must be diagonalized, which is also done by using Eq. (22) and inserting the eigenvalues. Expressing the physical eigenstates B_L^0 (L for light) and B_H^0 (H for heavy)[10] in terms of the flavor eigenstates B^0 and $\overline{B}^0$ one gets:

$$|B_L^0\rangle = p|B^0\rangle + q|\overline{B}^0\rangle = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{1+|\epsilon|^2}}((1+\epsilon)|B^0\rangle + (1-\epsilon)|\overline{B}^0\rangle), \tag{27}$$

$$|B_H^0\rangle = p|B^0\rangle - q|\overline{B}^0\rangle = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{1+|\epsilon|^2}}((1+\epsilon)|B^0\rangle - (1-\epsilon)|\overline{B}^0\rangle), \tag{28}$$

where p and q are complex coefficients obeying the normalization condition $|p|^2+|q|^2=1$. The notation with p and q is standard in B-physics, while the ϵ notation is mostly used in K-physics to express the CP violation explicitly. For this reason, ϵ will be used in the part on CP violation (section 3.2.4), and the notation with p and q used elsewhere. The relation between the two notations and the Hamiltonian is:

$$\frac{p}{q} = \frac{1+\epsilon}{1-\epsilon} = \sqrt{\frac{H_{12}}{H_{21}}} = \sqrt{\frac{M_{12}-\frac{i}{2}\Gamma_{12}}{M_{12}^*-\frac{i}{2}\Gamma_{12}^*}}. \tag{29}$$

The phase of p/q is arbitrary, which will be commented in section 3.2.3. Along with the eigenstates, one also obtains relations between the mass- and lifetime differences and the off diagonal elements of the Hamiltonian:

$$(\Delta m_B)^2 - (\tfrac{1}{2}\Delta\Gamma_B)^2 = 4|M_{12}|^2 - |\Gamma_{12}|^2, \quad \Delta m_B\Delta\Gamma_B = 4\mathcal{R}e(M_{12}\Gamma_{12}^*). \tag{30}$$

Using this and the fact that $\Delta\Gamma_B$ is small compared to Δm_B, it can be deduced that a difference in orders of magnitude is also the case for the two parts of the off-diagonal elements, namely that $|M_{12}| \gg |\Gamma_{12}|$, and the above expressions then simplify to:

$$\Delta m_B \simeq 2|M_{12}|, \quad \Delta\Gamma_B \simeq 2\mathcal{R}e(M_{12}\Gamma_{12}^*)/|M_{12}|. \tag{31}$$

3.2.3 Phase conventions

He who can properly define and divide is to be considered a god.

[Plato ca. 429-347 bc.]

The two B-states, B^0 and $\overline{B}^0$, are related through CP transformation, as follows:

$$CP|B^0\rangle = e^{2i\xi}|\overline{B}^0\rangle, \quad CP|\overline{B}^0\rangle = e^{-2i\xi}|B^0\rangle, \tag{32}$$

where the phase ξ is *arbitrary*, i.e. without physical meaning, so a phase transformation:

$$CP|B_\zeta^0\rangle = e^{-i\zeta}|\overline{B}^0\rangle, \quad CP|\overline{B}_\zeta^0\rangle = e^{+i\zeta}|B^0\rangle, \tag{33}$$

has no physical effects. In the new basis, indicated with the phase as super- or subscript, CP transformation will look as follows:

$$(CP)_\zeta|B_\zeta^0\rangle = e^{2i(\xi-\zeta)}|\overline{B}_\zeta^0\rangle, \quad (CP)_\zeta|\overline{B}_\zeta^0\rangle = e^{-2i(\xi-\zeta)}|B_\zeta^0\rangle. \tag{34}$$

[10]In the K-system (see section 3.6.1) $\Delta\Gamma_B \ggg \Delta m_B$, hence the terminology K_{short} and K_{long}.

Various quantities, the definition of which includes $|B^0\rangle$ and/or $|\overline{B}^0\rangle$ will change accordingly, for example the decay amplitude $\mathcal{A}_f \equiv \langle f|H|B^0\rangle$ and the matrix element $H_{12} \equiv \langle B^0|H|\overline{B}^0\rangle$ will change like $(\mathcal{A}_f)_\zeta = e^{-i\zeta}\mathcal{A}_f$ and $(H_{12})_\zeta = e^{-2i\zeta}H_{12}$ respectively. For this reason one has to be careful when considering phases and construct quantities with convention-independent phases, as only these will have physical meaning. One such phase invariant quantity is $\lambda \equiv \frac{q}{p}\frac{\overline{\mathcal{A}}}{\mathcal{A}}$, which plays a central role when discussing CP violation (see section 3.3.3).

3.2.4 *CP* eigenstates

The CP eigenstates can be determined by imposing the conditions of CP invariance (see Eqs. (20-21)), which translates into $p = q = \frac{1}{\sqrt{2}}$, and they are:

$$|B^0_+\rangle = \frac{|B^0\rangle + |\overline{B}^0\rangle}{\sqrt{2}}, \qquad |B^0_-\rangle = \frac{|B^0\rangle - |\overline{B}^0\rangle}{\sqrt{2}}. \tag{35}$$

From this, one can express the physical states in terms of the CP eigenstates:

$$|B^0_H\rangle = \frac{1}{\sqrt{1+|\epsilon|^2}}\left(|B^0_+\rangle + \epsilon|B^0_-\rangle\right), \qquad |B^0_L\rangle = \frac{1}{\sqrt{1+|\epsilon|^2}}\left(\epsilon|B^0_+\rangle + |B^0_-\rangle\right). \tag{36}$$

Here one sees that ϵ is a measure for CP violation[11], since any value different from zero implies that the physical states are not CP eigenstates. In the K-system this value is small ($\mathcal{O}(10^{-3})$) (see section 3.6.1), but it doesn't necessarily have to be equally small in the B–system, in fact it is expected to be much larger.

3.3 Time development of *B*-states

All composite things decay. Strive diligently. [Buddha ca. 563 - 483 bc., Last words]

Quantum mechanical time evolution is described by the Schrödinger equation, $i\frac{d|\ \rangle}{dt} = H|\ \rangle$, which applies for physical states. In order to obtain the time development of the flavor states, one has to write these in terms of physical states, time develop them and write the result back into flavor states:

$$\left.\begin{array}{l} |B^0_L\rangle = p|B^0\rangle + q|\overline{B}^0\rangle \\ |B^0_H\rangle = p|B^0\rangle - q|\overline{B}^0\rangle \end{array}\right\} \iff \left\{\begin{array}{l} |B^0\rangle = \frac{1}{2p}(|B^0_L\rangle + |B^0_H\rangle) \\ |\overline{B}^0\rangle = \frac{1}{2q}(|B^0_L\rangle - |B^0_H\rangle) \end{array}\right. . \tag{37}$$

Time development of the physical states gives:

$$|B^0_L(t)\rangle = e^{-iHt}|B^0_L\rangle = e^{-i\lambda_L t}|B^0_L\rangle, \quad |B^0_H(t)\rangle = e^{-iHt}|B^0_H\rangle = e^{-i\lambda_H t}|B^0_H\rangle, \tag{38}$$

which expressed in terms of the flavor states corresponds to:

$$\begin{aligned} |B^0_{\text{phys}}(t)\rangle &= \frac{1}{2p}(e^{-i\lambda_L t}|B^0_L\rangle + e^{-i\lambda_H t}|B^0_H\rangle) = \frac{1}{2}\left[e^{-i\lambda_L t}(|B^0\rangle - \frac{q}{p}|\overline{B}^0\rangle) + e^{-i\lambda_H t}(|B^0\rangle + \frac{q}{p}|\overline{B}^0\rangle)\right] \\ &= e^{-i\lambda_B t}\left[\cos(\tfrac{1}{2}\Delta\lambda t)|B^0\rangle + \frac{q}{p}i\sin(\tfrac{1}{2}\Delta\lambda t)|\overline{B}^0\rangle\right], \end{aligned} \tag{39}$$

$$|\overline{B}^0_{\text{phys}}(t)\rangle = e^{-i\lambda_B t}\left[\frac{p}{q}i\sin(\tfrac{1}{2}\Delta\lambda t)|B^0\rangle + \cos(\tfrac{1}{2}\Delta\lambda t)|\overline{B}^0\rangle\right], \tag{40}$$

with $\lambda_B \equiv \frac{1}{2}(\lambda_H + \lambda_L)$ and $\Delta\lambda \equiv \lambda_H - \lambda_L \equiv \Delta m - \frac{i}{2}\Delta\Gamma$. By using the short hand notation

$$g_+(t) \equiv e^{-i\lambda_B t}\cos(\tfrac{1}{2}\Delta\lambda t), \quad g_-(t) \equiv e^{-i\lambda_B t}i\sin(\tfrac{1}{2}\Delta\lambda t), \tag{41}$$

[11] In this case *indirect* (in mixing) CP violation, as opposed to *direct* (in decay) CP violation (see section 3.4).

one can rewrite the time dependent physical states into the following form:

$$|B^0_{\rm phys}(t)\rangle = g_+(t)|B^0\rangle + \frac{q}{p}\, g_-(t)|\overline{B}^0\rangle, \tag{42}$$
$$|\overline{B}^0_{\rm phys}(t)\rangle = \frac{p}{q}\, g_-(t)|B^0\rangle + g_+(t)|\overline{B}^0\rangle. \tag{43}$$

Mixing arises from the difference between the (time dependent) phases of the $|B^0_L\rangle$ and $|B^0_H\rangle$, while the common phase e^{-im_Bt} has no physical significance[12]. From this one can obtain expressions for the probability that a B^0-meson at time $t = 0$ will be observed at time, t, as a B^0-meson or a $\overline{B}^0$-meson respectively:

$$\begin{aligned} P(B^0(0) \to B^0(t)) &= |\langle B^0|B^0_{\rm phys}(t)\rangle|^2 = \tfrac{1}{4}\Gamma_B\, e^{-\Gamma_B|t|}\,[\cosh(\tfrac{1}{2}\Delta\Gamma t) + \cos(\Delta m t)] \\ &\simeq \tfrac{1}{4}\Gamma_B\, e^{-\Gamma_B|t|}\,[1 + \cos(\Delta m t)], \end{aligned} \tag{44}$$
$$P(B^0(0) \to \overline{B}^0(t)) \simeq \tfrac{1}{4}\Gamma_B\, e^{-\Gamma_B|t|}\,[1 - \cos(\Delta m t)], \tag{45}$$

where the approximation, which will be used in the following, is valid to $\mathcal{O}(10^{-3})$, since $\Delta\Gamma$ is expected to be small. These time distributions are used when measuring the tagging efficiency (see section 8). The opposite case, namely that an initially tagged B^0-meson is observed at time, t, as a $\overline{B}^0$-meson has an opposite sign in front of the cosine term, and the time-integrated probability of a mixed event is then:

$$\chi_d \equiv \int_{-\infty}^{\infty} P(B^0_d(0) \to \overline{B}^0_d(t))\, dt \simeq \int_{-\infty}^{\infty} \tfrac{1}{4}\Gamma_B\, e^{-\Gamma_B|t|}\,[1 - \cos(\Delta m t)]\, dt = \frac{1}{2}\frac{x_d^2}{1+x_d^2}. \tag{46}$$

Since $x_d \equiv \Delta M_{B_d}/\Gamma_{B_d}$ is small compared to one, this time-independant way of measuring the mass difference in the B_d-system is excellent, assuming knowledge of the lifetime. In the B_s-system, however, x_s is expected to be much higher than one, and so the idea of the method vanishes, since it will be very hard to distinguish χ_s from $\frac{1}{2}$.

3.3.1 Classical analogue

Every system is its own best analogue. [Andrew Donald Booth, 1918-]

The Hamiltonian Eq. (15) has a classical analogue in two equal mass coupled pendula. Without coupling they each have oscillation frequency $m \equiv \frac{1}{2\pi}\sqrt{g/l}$ and damping constant Γ and correspond to meson X^0 and antimeson $\overline{X}^0$. Coupled with a spring with spring constant m_{12} and damping constant Γ_{12} (both real and positive) the two pendula have two eigenstates: One low damping (i.e. long lifetime) low frequency (i.e. low mass) state $X_{\rm long,light}$ where the pendula oscillate in phase (Fig. 4a) and one high damping (i.e. short lifetime) high frequency (i.e. high mass) state $X_{\rm short,heavy}$ where the pendula oscillate in opposite phase (Fig. 4b). The differences in frequency and damping between these to eigenstates are $\Delta m = 2m_{12}$ and $\Delta\Gamma = 2\Gamma_{12}$, respectively. These two states much resemble the two physical states in the neutral meson system.

When one pendulum is excited, it will transfer energy to the other pendulum and vice versa. This beating between two states of only one pendulum oscillating correspond to the oscillation between a meson X^0 and antimeson $\overline{X}^0$. The beat frequency is Δm corresponding to the mixing frequency derived in Eq. (45).

However, the analogue has its limits. In a classical system m_{12} and Γ_{12} both have to be positive, thus $m_{12}\Gamma_{12} > 0$, opposite of the situation for mixing mesons. Furthermore, the lack of non-trivial phases prohibits CP violation in the classical system.

[12]Well, almost! It can be used to test whether gravity repels particles and antiparticles, which is not the case.

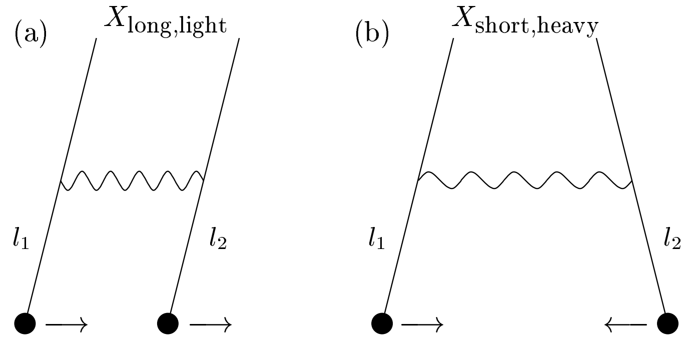


Figure 4: Classical analogue to mixing. A difference in the frequency between two coupled systems results in beating between two eigenstates. In this respect two coupled pendula resemble the mixing between neutral mesons.

3.3.2 Construction of coherent B$\overline{\mathbf{B}}$-state

At e^+e^- colliders, B_d-mesons are produced from the decay of the $\Upsilon(4S)$ resonance, which is the lightest $b\bar{b}$ resonance above the threshold for decaying into two B_d^0-mesons. They have zero spin, and are therefore produced in a coherent $L = 1$ state[13], since $J_{\Upsilon(4S)} = 1$. In order to construct such a state, one considers the parity operator $P|B^0\overline{B}^0\rangle = |B^0\overline{B}^0\rangle$. From the spin-statistics theorem for bosons, $P = (-1)^L \overset{L=1}{=} -1$, one finds that the state must be antisymmetric, which can be obtained with a normalized coherent state as follows:

$$
\begin{aligned}
|B_1(t_1)B_2(t_2)\rangle \overset{L=1}{=} & \frac{1}{\sqrt{2}}\left[|B^0_{\text{phys}}(t_1)\overline{B}^0_{\text{phys}}(t_2)\rangle - |\overline{B}^0_{\text{phys}}(t_1)B^0_{\text{phys}}(t_2)\rangle\right]\sin(\theta) \quad \text{(antisymmetric state)} \\
= & \frac{1}{\sqrt{2}}e^{-\Gamma_B(t_1+t_2)}\Big[\cos(\tfrac{1}{2}\Delta m(t_1-t_2))[B_1^0\overline{B}_2^0 - \overline{B}_1^0 B_2^0] \\
& -i\sin(\tfrac{1}{2}\Delta m(t_1-t_2))[\tfrac{p}{q}B_1^0B_2^0 - \tfrac{q}{p}\overline{B}_1^0\overline{B}_2^0]\Big]\sin(\theta),
\end{aligned}
\tag{47}
$$

where θ is the angle between the decays and the beam axis. As can be seen from Eq. (47), the result of the coherency is that the two B-mesons each evolve in time as a single B-meson (described above) *but in phase*, so that at all times, until one B-meson decays, there is exactly *one* B^0-meson and *one* $\overline{B}^0$-meson[14]. However, after the decay of one B-meson, the other continues to oscillate. Therefore a $B^0\overline{B}^0$-event can be either *unmixed*, i.e. the flavor of the two decayed B-mesons were opposite, or *mixed*, i.e. the flavor of the two decayed B-mesons were equal, as described by Eq. (45).

3.3.3 Coherent decay and tagging

To determine the flavor of a decayed B-meson (to *tag* it, in common terminology), one needs a decay, which is only allowed by one of the flavors. In principle no such flavor specific decays exist by the 'λ^4-theorem' [29], which states that for an unsuppressed decay of one flavor, the decay to the same final state for the other flavor is always possible but suppressed by $\lambda^4 \simeq 0.0016$ and possibly also EW suppressed[15]. So to a good approximation one can consider a flavor specific decay, that is:

$$
\mathcal{A}_{\overline{B}^0\to f} = 0 \;\Rightarrow\; f \text{ is a } B^0 \text{ tag}, \qquad \mathcal{A}_{B^0\to f} = 0 \;\Rightarrow\; f \text{ is a } \overline{B}^0 \text{ tag}, \tag{48}
$$

[13]The coherent production is characteristic of the B-factories, whereas $b\bar{b}$ jet fragmentation is not coherent.
[14]This is just another case of the classic Einstein-Podolsky-Rosen (EPR) situation.
[15]Given current statistics this is an insignificant effect, but in a larger data sample, corrections may be needed.

where $\mathcal{A}_{X\to Y} \equiv \langle Y|H|X\rangle$ is the amplitude. The tagging of one B-meson identifies the flavor of the other *at that time*, and the flavor can be infered at all times, if the time difference between the two decays is known.

Consider a coherent $B^0\overline{B}^0$ pair, where one B-meson decays into a CP eigenstate f_{CP} at time t_{CP}, and the other into a tagging mode $f_{\rm tag}$ at time $t_{\rm tag}$. Let for example $f_{\rm tag}$ be a flavor specific decay of the $\overline{B}^0$, thus $\mathcal{A}_{\rm tag} \equiv \langle f_{\rm tag}|H|B^0\rangle = 0$, and write the matrix element corresponding to observing this final state:

$$\begin{aligned}
\mathcal{M} &\equiv \langle f_{CP} f_{\rm tag}|H|B_1(t_{CP})B_2(t_{\rm tag})\rangle \\
&= \frac{1}{\sqrt{2}}\left[\langle f_{CP}|H|B^0_{\rm phys}(t_{CP})\rangle\, \langle f_{\rm tag}|H|\overline{B}^0_{\rm phys}(t_{\rm tag})\rangle - \langle f_{CP}|H|\overline{B}^0_{\rm phys}(t_{CP})\rangle\, \langle f_{\rm tag}|H|B^0_{\rm phys}(t_{\rm tag})\rangle\right] \\
&= \frac{\overline{\mathcal{A}}_{\rm tag}}{\sqrt{2}}\left[\left(g_+(t_{CP})\mathcal{A}_{f_{CP}} + \frac{q}{p} g_-(t_{CP})\overline{\mathcal{A}}_{f_{CP}}\right)\left(\frac{p}{q} g_-(t_{\rm tag})\mathcal{A}_{\rm tag} + g_+(t_{\rm tag})\overline{\mathcal{A}}_{\rm tag}\right) - \right. \\
&\qquad \left.\left(\frac{p}{q} g_-(t_{CP})\mathcal{A}_{f_{CP}} + g_+(t_{CP})\overline{\mathcal{A}}_{f_{CP}}\right)\left(g_+(t_{\rm tag})\mathcal{A}_{\rm tag} + \frac{q}{p} g_-(t_{\rm tag})\overline{\mathcal{A}}_{\rm tag}\right)\right], \\
&= \frac{\overline{\mathcal{A}}_{\rm tag}}{\sqrt{2}}\, e^{-i\lambda_B(t_1+t_2)}\left[\cos(\tfrac{1}{2}\Delta m t_{CP})\cos(\tfrac{1}{2}\Delta m t_{\rm tag})\mathcal{A}_{f_{CP}} + \frac{q}{p} i \sin(\tfrac{1}{2}\Delta m t_{CP})\cos(\tfrac{1}{2}\Delta m t_{\rm tag})\overline{\mathcal{A}}_{f_{CP}} - \right. \\
&\qquad \left. i^2 \sin(\tfrac{1}{2}\Delta m t_{CP})\sin(\tfrac{1}{2}\Delta m t_{\rm tag})\mathcal{A}_{f_{CP}} - \frac{q}{p} i \cos(\tfrac{1}{2}\Delta m t_{CP})\sin(\tfrac{1}{2}\Delta m t_{\rm tag})\overline{\mathcal{A}}_{f_{CP}}\right], \\
&= \frac{\mathcal{A}_{f_{CP}}\overline{\mathcal{A}}_{\rm tag}}{\sqrt{2}}\, e^{-i\lambda_B(t_1+t_2)}\left[\cos(\tfrac{1}{2}\Delta m(t_{f_{CP}} - t_{\rm tag})) + \lambda_{f_{CP}} i \sin(\tfrac{1}{2}\Delta m(t_{CP} - t_{\rm tag}))\right].
\end{aligned} \tag{49}$$

where in the last line the following definition is introduced[16]:

$$\lambda_{f_{CP}} = \frac{q}{p}\frac{\overline{\mathcal{A}}_{f_{CP}}}{\mathcal{A}_{f_{CP}}}. \tag{50}$$

Contrary to most other quantities, *the phase of $\lambda_{f_{CP}}$ is independent of convention and thus physical significant* (see section 3.2.3). In section 3.4 it will be shown, that any deviation of $\lambda_{f_{CP}}$ from one is equivalent to CP violation.
By squaring the amplitude calculated above, one obtains the matrix element for the decay:

$$\begin{aligned}
|\mathcal{M}|^2 &= \tfrac{1}{2}|\mathcal{A}_{f_{CP}}|^2|\overline{\mathcal{A}}_{\rm tag}|^2 e^{-\Gamma(t_{\rm tag}+t_{CP})}\left[\cos^2(\tfrac{1}{2}\Delta m t) + |\lambda_{f_{CP}}|^2\sin^2(\tfrac{1}{2}\Delta m t) + \right. \\
&\qquad \left. i\sin(\tfrac{1}{2}\Delta m t)\cos(\tfrac{1}{2}\Delta m t)(\lambda_{f_{CP}} - \lambda^*_{f_{CP}})\right] \\
&\propto \tfrac{1}{2}\, e^{-\Gamma|t|}\left[1 + |\lambda_{f_{CP}}|^2 + (1 - |\lambda_{f_{CP}}|^2)\cos(\Delta m t) + 2\,{\rm Im}\,(\lambda_{f_{CP}})\sin(\Delta m t)\right], \\
&\stackrel{|\lambda|=1}{=} e^{-\Gamma|t|}\left[1 + {\rm Im}\,(\lambda_{f_{CP}})\sin(\Delta m t)\right],
\end{aligned} \tag{51}$$

where $t \equiv t_{CP} - t_{\rm tag}$ in the following[17]. Had the tagging mode, $f_{\rm tag}$, been a flavor specific decay of the B^0, thus $\overline{\mathcal{A}}_{\rm tag} \equiv \langle f_{\rm tag}|H|\overline{B}^0\rangle = 0$, the sign in front of the trigonometric functions would have been opposite.

The type of $B^0\overline{B}^0$-decay described above measures the imaginary part of $\lambda_{f_{CP}}$, which expresses a CP-asymmetry. This is the most accessible form of CP violation in the B-system (see section 3.4), and therefore the primary measurement to carry out.

3.4 Types of *CP* violation

CP violation can manifest itself in three different ways:

[16] The notation is standard but a bit unfortunate, as λ is also used for eigenvalues and as Wolfenstein parameter.
[17] The CP decay time is used as reference ($t_{CP} \equiv 0$), which makes the natural notation t instead of Δt.

- *CP violation in decay.* The amplitude for a decay (charged or neutral) and its CP conjugate have different magnitudes, that is $|\overline{\mathcal{A}}_{\overline{f}}/\mathcal{A}_f| \neq 1$.
- *CP violation in mixing.* Two neutral mass eigenstates, that are CP conjugates of each other, cannot be chosen to be CP eigenstates, that is $|q/p| \neq 1$.
- *CP violation in interference between decays with and without mixing.* Decays of neutral CP-conjugates into final CP eigenstate exhibits differences, that is $\mathrm{Im}\lambda_{f_{CP}} \neq 0$.

Each of the three types of CP violation can occur by itself, but they are by no means mutually exclusive. Possible phase redefinitions of quark fields complicate the discussion, and have to a large extent been omitted in the following. The reader is referred to [16] and [30].

3.4.1 *CP* violation in decay — direct *CP* violation

Direct CP violation occurs when CP conjugate decay rates differ:

$$\boxed{|\overline{\mathcal{A}}_{\overline{f}}/\mathcal{A}_f| \neq 1 \quad \Longrightarrow \quad \textbf{Direct } CP \text{ violation}} \tag{52}$$

and it is a result of interference between phases. There are two types of phases:

- *strong phases*, δ, from intermediate on-shell states in the decay, which appear with the same sign in $\mathcal{A}_f$ and the CP-conjugated amplitude $\overline{\mathcal{A}}_{\overline{f}}$.
- *weak phases*, ϕ, from the complex parameters (the CKM phases in the SM) in the Lagrangian, which will appear complex conjugated (opposite sign) in $\overline{\mathcal{A}}_{\overline{f}}$.

The amplitude of a process can be written as a sum of amplitudes, one for each diagram contributing to the process:

$$\mathcal{A}_f = \sum_i |\mathcal{A}_i| e^{i(\delta_i + \phi_i)}, \quad \overline{\mathcal{A}}_{\overline{f}} = \sum_j |\overline{\mathcal{A}}_j| e^{i(\delta_j - \phi_j)}. \tag{53}$$

If only one diagram contributes to an amplitude, the change of sign of the weak phase will not alter the absolute value of this amplitude. Direct CP violation requires that at least two amplitudes with different weak phases acquire different strong phases, which can be seen from considering the difference between the square of the amplitudes:

$$|\mathcal{A}|^2 - |\overline{\mathcal{A}}|^2 = -2 \sum_{i,j} \mathcal{A}_i \mathcal{A}_j \sin(\delta_i - \delta_j) \sin(\phi_i - \phi_j). \tag{54}$$

For neutral B-mesons, direct CP violation results in $\lambda_{f_{CP}} \neq 1$, which can be measured by fitting to Eq. (51). However, the neutral case requires tagging, lowering the statistics, and the charged case is thus more straightforward, since one at the same time expects higher asymmetries:

$$a_f = \frac{\Gamma(B^+ \to f) - \Gamma(B^- \to \overline{f})}{\Gamma(B^+ \to f) + \Gamma(B^- \to \overline{f})} = \frac{1 - |\overline{\mathcal{A}}_{\overline{f}}/\mathcal{A}_f|^2}{1 + |\overline{\mathcal{A}}_{\overline{f}}/\mathcal{A}_f|^2}. \tag{55}$$

The decay channels of interest are $B \to K\pi$, $B \to KK$, $B \to \rho K$, $B \to DD^*$ and lately $B^\pm \to J/\psi\pi^\pm$, but also observing a difference in the number of $B^0/\overline{B}^0$ tags would constitute evidence for direct CP violation, though one has to take great care that it is not due to differences between the efficiency of determining the flavor of B^0 and $\overline{B}^0$-mesons.

The possible asymmetries are, contrary to the K-system (see section 3.6.1), expected to be large (10 %). However, these channels all have low branching ratios ($\mathcal{O}(10^{-4})$ to $\mathcal{O}(10^{-5})$)[18]. Furthermore, calculation of the strong phases includes long distance QCD, leading to model dependence and significant theoretical uncertainties.

[18] The "uncertainty principle" of direct CP violation: Branching ratio $\times$ asymmetry $\lesssim$ $\mathcal{O}(10^{-6})$ [Jim Smith].

3.4.2 *CP* violation in mixing — indirect *CP* violation

CP conservation requires that the physical eigenstates must be CP eigenstates for a certain value of an arbitrary phase, that is q/p is a pure phase. Indirect CP violation occurs if this is not the case:

$$\boxed{|q/p| \neq 1 \quad \Longrightarrow \quad \textbf{Indirect } CP \text{ violation}} \tag{56}$$

that is if the mixing rates between the two eigenstates are not equal. Indirect CP violation has been observed unambiguously in the neutral kaon system (see section 3.6.1), and (eventually) measurable effects are expected in the neutral B-system, for example in semileptonic decays:

$$a_{sl} = \frac{\Gamma(\overline{B}^0_{\text{phys}}(t) \to \ell^+\nu X) - \Gamma(B^0_{\text{phys}}(t) \to \ell^-\bar{\nu} X)}{\Gamma(\overline{B}^0_{\text{phys}}(t) \to \ell^+\nu X) + \Gamma(B^0_{\text{phys}}(t) \to \ell^-\bar{\nu} X)} = \frac{1-|q/p|^4}{1+|q/p|^4} \tag{57}$$

which follows from the opposite ratio of q/p in the amplitudes (assume no direct CP violation):

$$\langle \ell^-\bar{\nu}X|H|B^0_{\text{phys}}(t)\rangle = (q/p)g_-(t)A^*, \quad \langle \ell^+\nu X|H|\overline{B}^0_{\text{phys}}(t)\rangle = (p/q)g_-(t)A. \tag{58}$$

However, since $|q/p|$ depends on Γ_{12}, theoretical uncertainties from long distance QCD effects are once again expected to complicate the extraction of CKM parameters from this asymmetry. To first order only the top quark loop contributes, and q/p is just a pure phase. Contributions from charm quark loops are expected to be $\mathcal{O}(\frac{m_c}{m_t})$, and so any CP violation in mixing will be of this order.

3.4.3 *CP* violation in the interference between decays with and without mixing

Conservation of CP implies in addition to $|\overline{\mathcal{A}}_{\overline{f}_{CP}}/\mathcal{A}_{f_{CP}}| = 1$ (no direct CP-violation), and $|q/p| = 1$ (no indirect CP violation), that the relative phase between $\overline{\mathcal{A}}_{\overline{f}_{CP}}/\mathcal{A}_{f_{CP}}$, and q/p vanishes. CP violation in the interference between decays with and without mixing, sometimes abbreviated CP violation in the interference between mixing and decay, terms this type of CP violation, and the condition is:

$$\boxed{\operatorname{Im}\lambda_{CP} \neq 0 \quad \Longrightarrow \quad CP \text{ violation } \textbf{in interference}} \tag{59}$$

In the neutral B-system this type of CP violation is observed in the time dependent difference between B^0 and $\overline{B}^0$ decays into final CP eigenstates, f_{CP}, which is expressed in the following asymmetry:

$$a_{f_{CP}} = \frac{\Gamma(B^0_{\text{phys}}(t) \to f_{CP}) - \Gamma(\overline{B}^0_{\text{phys}}(t) \to f_{CP})}{\Gamma(B^0_{\text{phys}}(t) \to f_{CP}) + \Gamma(\overline{B}^0_{\text{phys}}(t) \to f_{CP})}. \tag{60}$$

Inserting Eq. (51), where the conjugate expression has opposite sign in the trigonometric terms, one obtains:

$$a_{f_{CP}} = \frac{(1-|\lambda_{f_{CP}}|^2)\cos(\Delta mt) + 2\operatorname{Im}\lambda_{f_{CP}}\sin(\Delta mt)}{1+|\lambda_{CP}|^2} \overset{|\lambda|=1}{=} \operatorname{Im}\lambda_{f_{CP}}\sin(\Delta mt). \tag{61}$$

However, since f_{CP} is equally accessible for the two neutral states, flavor tagging of the other B-meson is needed, which dilutes and complicates the extraction (see section 8). Yet, the expected asymmetries are large, and in the case of $b \to c\bar{c}s$ decays directly related to CKM parameters. Therefore, this type of CP violation is the most accessible in the B-system and, thus, the first one to test experimentally. As the next section will show,

this is exactly what the sin(2β) measurement will do.

3.5 *CP* asymmetries and the angles of the unitary triangle

3.5.1 Determination of β

The angle β in the unitary triangle is most easily determined from decays of the type $b \to c\bar{c}s$, which at hadronic level becomes e.g. $B^0 \to J/\psi K_S$ (see Fig. 5). This decay channel is called "the golden channel", since the dominating penguin diagrams (loop diagrams contributing to the same amplitude, see Fig. 6), has the same CKM phase as the tree diagram, thus not polluting the extraction of β. Diagrams with different CKM phases contributing to this channel are either Cabbibo or mass (x_q) suppressed. Consequently, this channel is to a very good approximation theoretically clean.

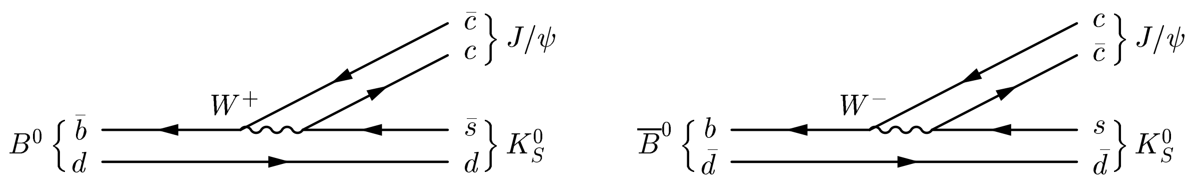


Figure 5: Tree diagram for $B^0/\overline{B}^0 \to J/\psi K_S$ decay. This "golden" decay channel provides a theoretically clean extraction of the angle $\sin(2\beta)$.

From comparing the two diagrams in Fig. 5, which are complex conjugates of each other, the expression for λ_{CP} in the $B^0 \to J/\psi K_S$ case in terms of ratios of CKM parameters is obtained:

$$\lambda(B_d \to J/\psi K_S) \quad = \quad -\left(\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}\right)\left(\frac{V_{cs}^* V_{cb}}{V_{cs} V_{cb}^*}\right)\left(\frac{V_{cd}^* V_{cs}}{V_{cd} V_{cs}^*}\right). \tag{62}$$

The three ratios originate from processes that involve complex (CKM) parameters, where the two states will have oppositely complex conjugated entries. The first ratio is from $B_d^0\overline{B}_d^0$ mixing (see Fig. 3), the second from the ratio of decay amplitudes $\overline{\mathcal{A}}_{f_{CP}}/\mathcal{A}_{f_{CP}}$ and the third from $K\overline{K}$ mixing. The minus sign originates from $CP(J/\psi K_S^0) = -1$. This leads to[19]:

$$\boxed{\mathrm{Im}\,\lambda(B^0 \to J/\psi K_S^0) = \sin(2\beta)} \tag{63}$$

The higher $c\bar{c}$-resonance $\psi(2S)$ is equally good, and the resonances can in addition to the K_S^0 be combined with K_L^0 and π^0, if taking into account that these states have CP=+1, i.e. any possible CP asymmetry will have the opposite sign. Furthermore, K^{0*} can be used, though this requires angular analysis, as both the J/ψ and K^{0*} are vector particles.

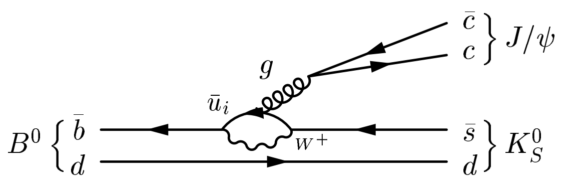


Figure 6: Penguin diagram for $B^0 \to J/\psi K_S^0$ decay. With an up quark in the loop, this diagram has different CKM factors than the tree diagram but it is Cabibbo suppressed, and so to $\mathcal{O}(10^{-2})$ the $J/\psi K_S^0$ channel should provide a theoretically clean extraction of the angle $\sin(2\beta)$.

[19]Essentially all entries cancel their conjugate part as they are real, except V_{td}, which is propotional to $e^{-\beta}$.

3.5.2 Determination of α

In principle, $\sin(2\alpha)$ can be determined in almost the same way as $\sin(2\beta)$ just considering decays of the type $b \to u\bar{u}d$, hadronically being neutral B-decays to $\pi\pi$, $\rho\pi$ and $a_0\pi$. From analysing the CKM parameters involved in the tree level diagram, one obtains:

$$\mathrm{Im}\,\lambda(B^0 \to \pi^+\pi^-) \quad = \quad \sin 2\alpha. \tag{64}$$

However, contrary to the golden $J/\psi K_S^0$ channel, penguin diagrams with different CKM parameters are not Cabibbo suppressed but of order 30 % of the total branching ratio [31]. Due to non-pertubative long distance effects the relative branching fractions can not be calculated, as neither heavy quark effective theory nor lattice QCD can be applied. For this reason, measurements of CP asymmetries in these channels can not be directly related to $\sin(2\alpha)$ and thus CKM parameters.

The problem can be solved by isospin analysis [32]. The key observation is that the dominant penguin diagrams are purely $\Delta I = \frac{1}{2}$, since they involve a gluon ($\Delta I = 0$), while the tree level diagram have $\Delta I = \frac{3}{2}$ and $\Delta I = \frac{1}{2}$ components. If the $\Delta I = \frac{3}{2}$ part can be isolated, it can be used to determine $\sin 2\alpha$ without theoretical dilution. This procedure requires different final states, as one effectively subtracts the penguins. For this reason the $B \to \rho\pi$ channel is more suitable, since it has three final states. Recent studies [33],[34] indicate that this type of analysis, with a few years of running at BaBar, will yield reasonable precision, and the subject is therefore being intensely pursued [35].

3.5.3 Determination of γ

The angle γ in the unitary triangle is generally considered to be the most difficult to measure directly, and so far γ has only been estimated by indirect methods, which constrain the unitary triangle. In principle $b \to u\bar{c}s$ and $b \to c\bar{u}s$ decays, hadronically $B \to DK$, have no penguins and extracts γ in a theoretical clean way. However, the experimental obstacles are many, and since the triangle relations for extracting γ are rather "flat" (like four of the unitary triangles), the required statistics make this an almost unfeasible method. Other methods involving $B_{u,d} \to K\pi$ and $B_d \to D^{(*)}\pi$ decays and $B_s\overline{B}_s$ mixing have been considered.

Recently the decay $B^0/\overline{B}^0 \to D^{\pm(*)}K_S\pi^\mp$ was proposed [36]. The branching ratio for this decay can be estimated by considering similar decay channels with measured branching ratios. One such is the $B \to D^{(*)\pm}\pi^0\pi^\mp$ decay channel, which has a measured branching ratio of $\mathcal{B}r(B \to D^{*\pm}\pi^0\pi^\mp) = (1.5 \pm 0.5)\,\%$ [37]. The only difference between this and the $B \to D^{*\pm}K_S^0\pi^\mp$ decay is their V_{CKM}-factors, and the fact that only half of the neutral kaons are K_S^0. The estimate of the branching ratio for the $B \to D^{*\pm}K_S^0\pi^\mp$ channel is then:

$$\mathcal{B}r(B \to D^{*\pm}K_S^0\pi^\mp) \quad \sim \quad \frac{1}{2}\frac{V_{us}^*}{V_{ud}^*}\mathcal{B}r(B \to D^{*\pm}\pi^0\pi^\mp) \quad = \quad (0.3 \pm 0.1) \times 10^{-4}. \tag{65}$$

Due to the larger number of states (three spin states), the production of D^* is in general three times larger than that of D, disregarding phase space and various other minor effects. Thus, by including the $B \to D^\pm K_S^0\pi^\mp$ channel, the total branching ratio reaches 4×10^{-4}, which is clearly within reach of the BaBar experiment. However, this number is a very rough estimate, and the error could be a factor of two.

The extraction of $\sin(2\beta+\gamma)$ would be based on Dalitz plots[20], as shown in Fig. 8. Essentially, one compares the number of events from B^0 and $\overline{B}^0$ in the same (small) region of the Dalitz plot, since the strong phases, δ, are then equal, and thereby extract $\sin(2\beta+\gamma+\delta)$. The same is done for a small neighboring region, where $\sin(2\beta+\gamma+\delta')$ then is measured, and by assuming a

[20]Scatter plots of the invariant mass between pairs of particles. Used for characterizing threebody decays.

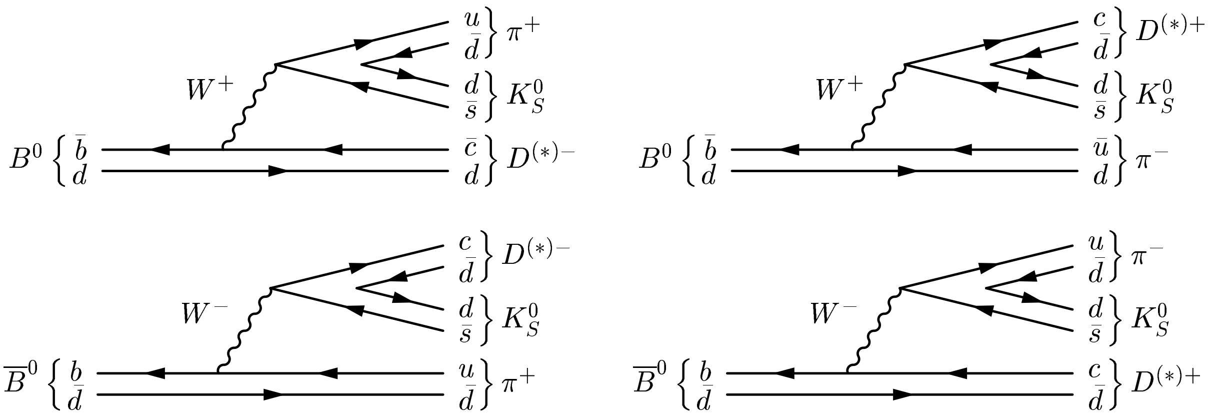


Figure 7: Diagrams for $B \to D^{(*)\pm} K_S^0 \pi^\mp$ decay. *The four diagrams shown above measures the angle* $\sin(2\beta+\gamma)$ *and do not have penguin contributions. Their branching ratio is estimated to be* $\mathcal{O}(4\times10^{-4})$.

continuous variation of the strong phases, $\sin(2\beta+\gamma)$ can be extracted. This would of course be done with a global fit, which includes the assumed slow and continuous variation of the strong phases. The assumption is reasonable as long as one stays clear of the resonances from $D^*K_S^0$, $D_s^*\pi$ and DK^*, which can be seen as bands in Fig. 8.

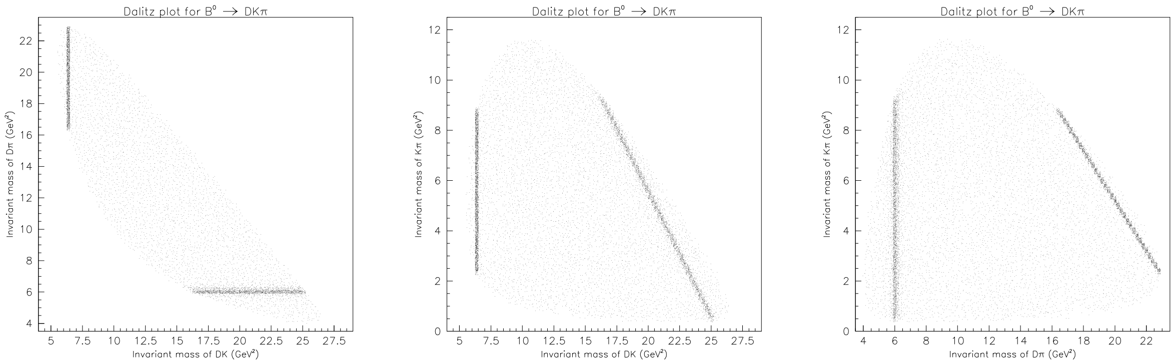


Figure 8: Dalitz plot of decay $B^0 \to D^\pm K_S^0\, \pi^\mp$. Each pair of invariant masses between particles gives rise to a plot. The events were generated by a toy Monte Carlos with a flat q^2 distributions and the resonant bands added.

3.6 *CP* violation in other systems

The short lifetime of the top quark does not allow for it to form bound states, leaving the B-system as the only possibility for directly studying the third generation of the CKM matrix. However, other quark systems can provide information about the CKM matrix, namely the $K^0\overline{K}^0$ system based on the s-quark and the $D^0\overline{D}^0$ based on the c-quark. Since the neutral pion is an equal mixture of $u\bar{u}$ and $d\bar{d}$, it is its own antiparticle, and mixing is therefore meaningless.

3.6.1 The *K*-system

The K-system is so far the only place where CP violation has been observed directly. The physical states with CP and CPT violating parameters, ϵ and Δ respectively written explicitly, look as follows:

$$|K_S^0\rangle = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{1+|\epsilon+\Delta|^2}}((1+(\epsilon+\Delta))|K^0\rangle + (1-(\epsilon+\Delta))|\overline{K}^0\rangle), \tag{66}$$

$$|K_L^0\rangle = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{1+|\epsilon+\Delta|^2}}((1+(\epsilon-\Delta))|K^0\rangle - (1-(\epsilon-\Delta))|\overline{K}^0\rangle), \tag{67}$$

where the notation S (Short) and L (Long) steams from the dominating lifetime difference ($\tau_{K_S^0} = (0.8935 \pm 0.0008) \times 10^{-10}s$, $\tau_{K_L^0} = (5.17 \pm 0.04) \times 10^{-8}s$ [37]). This feature makes it experimentally easy to separate the two K-states, while in the B-system one considers B^0_{phys} and $\overline{B}^0_{\text{phys}}$ and tag them according to flavor specific decays.

The K-mixing is very similar to B-mixing in having u-type quarks in the loops of the mixing diagram (see Fig. 3). However, in K-mixing top-charm and charm-charm loop contributions cannot be neglected, and this makes it difficult to extract information from observables in the K-system, even though they are well measured ($\Delta m_K = (3.489 \pm 0.008) \times 10^{-15}$ GeV). The measured CP-violating quantities in the K-system are the following:

$$\eta_{00} \equiv \mathcal{A}(K_L^0 \to \pi^0\pi^0)/\mathcal{A}(K_S^0 \to \pi^0\pi^0) = (2.262 \pm 0.017) \times 10^{-3}, \tag{68}$$

$$\eta_{+-} \equiv \mathcal{A}(K_L^0 \to \pi^+\pi^-)/\mathcal{A}(K_L^0 \to \pi^+\pi^-) = (2.276 \pm 0.017) \times 10^{-3}, \tag{69}$$

$$\epsilon_K = (p-q)/(p+q) = (2.271 \pm 0.017) \times 10^{-3}, \tag{70}$$

$$\epsilon'/\epsilon_K \stackrel{CPT}{=} \mathcal{R}e(\epsilon'/\epsilon_K) = \tfrac{1}{3}(1-|\eta_{00}/\eta_{+-}|) = (19.2 \pm 2.4) \times 10^{-4}, \tag{71}$$

where ϵ_K measures CP violation in mixing, while ϵ' (historically written without subscript) measures direct CP violation, thus superweak theory[21] (which requires $\epsilon' = 0$) is excluded.
As can be seen in Fig. 9, CPT invariance gives a correlation between η_{00}, η_{+-} and ϵ_K, and therefore CPT invariance can be tested by considering:

$$a_{\text{CPT}} = \frac{P(\overline{K}^0 \to \overline{K}^0(t)) - P(K^0 \to K^0(t))}{P(\overline{K}^0 \to \overline{K}^0(t)) + P(K^0 \to K^0(t))} = \frac{H_{11}-H_{22}}{H_{11}+H_{22}} = 4\mathcal{R}e(\Delta), \tag{72}$$

where P is the (time dependent) probability. Experimental data indicates no CPT violation, since the current measurements of Δ gives:

$$\mathcal{R}e(\Delta) = (2.9 \pm 2.7) \times 10^{-4}, \quad \mathcal{I}m(\Delta) = (-0.8 \pm 3.1) \times 10^{-3}. \tag{73}$$

The rare decays $K^\pm \to \pi^\pm\nu\bar{\nu}$ and $K_L^0 \to \pi^0\nu\bar{\nu}$ are of special interest, since large CP violation is expected with clean interpretation, which expressed in terms of ρ and η looks as follows [39]:

$$\Gamma(K^\pm \to \pi^\pm\nu\bar{\nu}) = 8.33 \times 10^{-6}|V_{cb}|^4(X(x_t))^2\left[\eta^2 + (\rho_0^2 - \rho^2)\right] = (3.1 \pm 1.3) \times 10^{-11}, \tag{74}$$

$$\Gamma(K_{L,S}^0 \to \pi^0\nu\bar{\nu}) = 3.29 \times 10^{-5}|V_{cb}|^4(X(x_t))^2\eta^2 = (8.2 \pm 3.2) \times 10^{-11}, \tag{75}$$

[21] A theory explaining CP violation as a consequence of an additional superweak force [38].

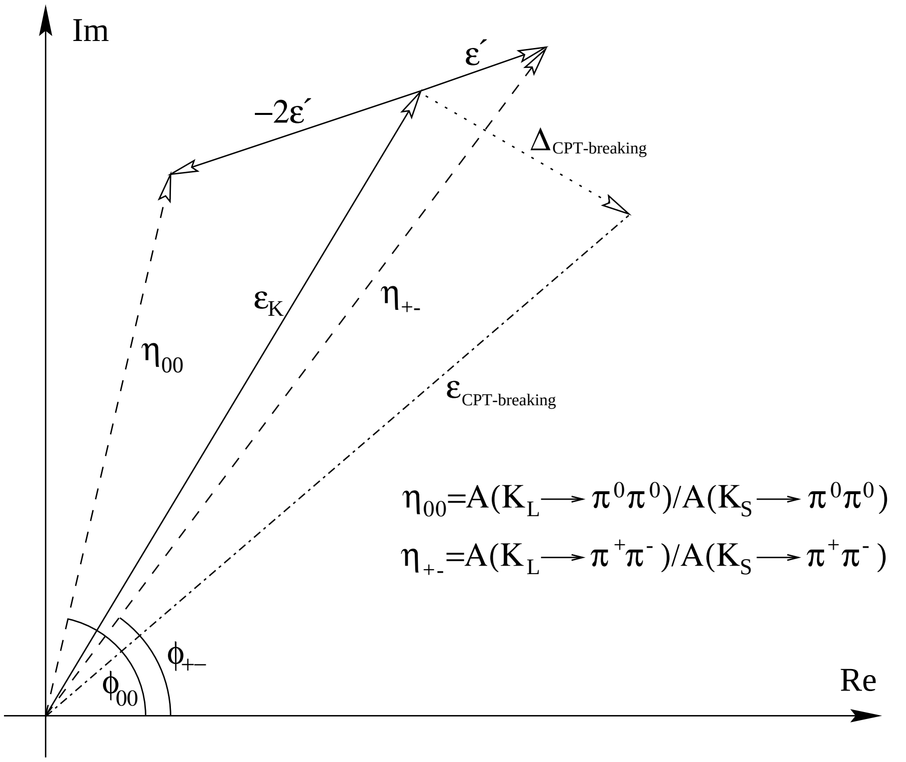


Figure 9: CP violating parameters ϵ_K and ϵ'. Geometrical representation of the CP violating parameters ϵ_K and ϵ' in the complex plane. Also shown is the CPT violating parameter Δ.

with $\rho_0 = 1 + \frac{P_0(X)}{X(x_t)} \frac{\lambda^4}{|V_{cb}|^2} \simeq 1.4$, where $X(x_t)$ and $P_0(X)$ are EW loop contributions in next to leading order for the top and the charm quark, respectively. Evidently, ρ and η can be cleanly extracted, and so the experimental challenge of the low branching ratios is therefore met with proposals of dedicated experiments[22]. E787 at BNL have seen one charged event [40].

3.6.2 The D-system

The D-system is fundamentally different from the K and the B-systems in that the mixing loop (see Fig. 3) contains d-type quarks instead of u-type quarks, so the heaviest quark in the loop is b instead of t. Since $m_b \ll m_t$ and since the contribution from the heaviest quark, b, is suppressed by λ^2 ($|V_{ub}V_{cb}|$ compared to $|V_{td}V_{tb}|$ for the B-system), mixing in the D-system is expected to be very small. In the SM one expects $x_D \equiv \Delta m_D/\Gamma_D \lesssim 0.002$, i.e. the mixing time is 500 times larger than the lifetime, which makes the probability of observing mixing equal to $\chi_D \equiv x_D^2/2(1+x_D^2) \simeq 4 \times 10^{-6}$.

Most decays of the D-meson have flavor specific final states, resulting in a very small expected lifetime difference. Furthermore, both direct and indirect CP violation in the D-system is expected to be insignificant, leaving only CP violation in the interference between mixing and decay as a possible measurement, for example in the $D \to K\pi$ channel.

The current experimental limit on the probability of mixing in the D-system is $\chi_D < 5.0 \times 10^{-3}$ at 90 % confidence level (CL) by the E791 collaboration, which is still three orders of magnitude above the SM prediction. Thus, if D-mixing is observed around the current experimental level, it is a sign of new physics.

[22]Kopio (BNL) and Kami (FNAL) for the neutral mode and E949 (BNL) and CKM (FNAL) for the charged.

4 Current knowledge of the CKM matrix

First, the current knowledge of the CKM matrix elements is reviewed [37]. Assuming unitarity of the CKM matrix, this matrix only contains four independent parameters, and the overconstrain is used to further improve the knowledge of the parameters and, more importantly, to test the consistency of the various measurements.

4.1 Experimental constrains of $V_{\mathbf{CKM}}$

The following CKM matrix elements have been directly measured:

- $|\mathbf{V_{ud}}| = 0.9735 \pm 0.0008$ from nuclear beta decays and highly polarized cold neutrons.
- $|\mathbf{V_{us}}| = 0.2196 \pm 0.0023$ from $K \to \pi e \nu$ decays, and to a lesser extent hyperon decays.
- $|\mathbf{V_{cd}}| = 0.224 \pm 0.016$ from charm production in neutrino scattering off valence d-quarks.
- $|\mathbf{V_{cs}}| = 0.99 \pm 0.09$ from $D \to K e \nu$ decays and charm tagged W-decays.
- $|\mathbf{V_{cb}}| = 0.0402 \pm 0.0019$ from exclusive $B \to \overline{D}^* \ell \nu$ and inclusive $B \to X_c \ell \nu$ decays.
- $|\mathbf{V_{ub}}| = 0.090 \pm 0.025$ from $b \to u \ell \nu$ decays with high p_ℓ and exclusive $B \to (\pi/\rho)\ell\nu$ decays.
- $|\mathbf{V_{tb}}| = 0.99 \pm 0.29$ from $t \to b \ell \nu$ decays compared to other top quark decays.

Additional constrains of both single parameters and combinations exist. From unitarity of the first two columns, hadronic W-decays constrain the least well determined CKM entry of the six: $|V_{cs}| = 0.989 \pm 0.016$. Comparing the measured value of Δm_{B_d} with the predicted dictates that $|V_{td} V_{tb}^*| = 0.0083 \pm 0.0016$. The upper limit on $\Delta m_{B_s} > 14.3 ps^{-1}$ in turn constrains $|V_{td}| < 0.010$, and the measurement of $\mathrm{Br}(b \to s\gamma)$ results in limits on $|V_{ts}^* V_{tb}/V_{cb}| = 0.976 \pm 0.010$.

4.2 Combining measurements

As previously stated, the goal of combining the various CKM related quantities is both to obtain the most accurate values of the four parameters, whichever parametrization one chooses, but also, and more importantly, to test whether CP violation can be described within the SM. Many groups have pursued this task ([41], [42], [43]), but the following short description will be based on [44] and [45].

The idea is to make a global fit to all measurements constraining the four independent CKM parameters, and then set limits on the parameters of interest (call them a), which in the case of the Wolfenstein parametrization are ρ and η, leaving the remaining parameters (call them μ) fixed. If the errors involved were Gaussian, the global χ^2-fit could be done with $\chi^2(a,\mu) = \sum (X_{\text{meas.}} - X_{\text{meas.}}(a,\mu))/\sigma^2_{\text{meas.}}$. However, most of the errors are theoretical and therefore not Gaussian. In order to manage these complications, the χ^2 is expressed using probability density functions (pdf); $\chi^2(a,\mu) = \sum -2\ln(\text{pdf})$. In this way theoretical errors, confidence limits, etc. can be treated properly, and at the same time possible incompatibilities in data can be incorporated.

Once the minimum, χ^2_{min}, has been found, and consequentlyt all parameters simultaneously fitted, the μ-parameters are fixed and the confidence level of the a-parameters determined $CL(a) = Prob(\Delta\chi^2, N_a)$ with $\Delta\chi^2 = \chi^2(a) - \chi^2_{min}$ varying a only. The result is shown in Fig. 10 below. Since $\Delta\chi^2$ will always equal zero at the minimum, the CL will always reach one. The blue color (the outermost colored region in b/w verions) approximately equals to the 95 % CL. The current situation is, that *all measured quantities sensible to the CKM matrix can be fitted with four parameters and are thus compatible.* The fit suggests $J = (2.70 \pm 0.72) \times 10^{-5}$, (95 %

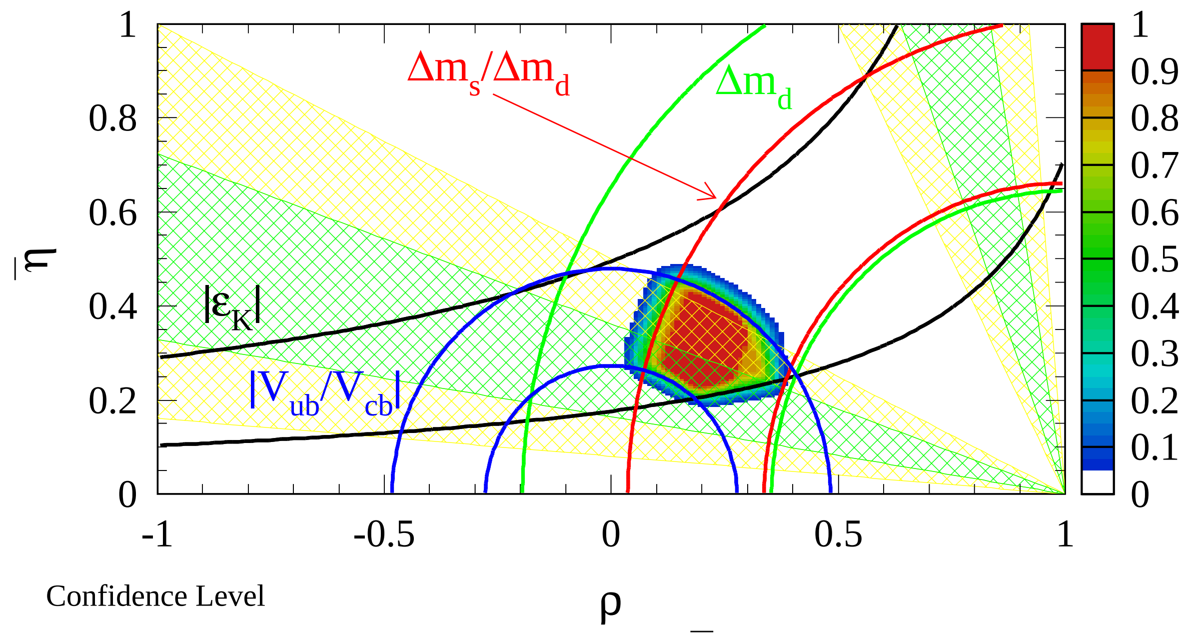


Figure 10: Fit to CKM parameters in the $\rho - \eta$ plane. The figure shows the result of fitting all parameters that contains information about the CKM matrix elements. The resulting allowed region depends on the assumed probability density functions, since most uncertainties originate from theoretical calculations and are therefore not Gaussian. The superimposed green and yellow areas correspond to the world average $\sin(2\beta)$ value including one and two standard deviations respectively. See text and reference for further explanation [44].

CL.), in accordance with the now well established CP-violation in the K-system, and so CP violation is naturally small.

The largest inconsistency in the CKM matrix is the unitarity of the first column, since $\sum_i |V_{ui}|^2 = 0.9965 \pm 0.0014$. This is a 2.2σ effect, but obtained by three doubling the theoretical errors in radiative corrections in superallowed β-decays of N^{14}, since they were not compatible. A new experiment with ultracold neutrons at LANL also gives “low” values, and it has lower theoretical errors due to the lack of radiative corrections [46], so perhaps there is a problem with unitarity of the CKM matrix, though nothing can be concluded at the present stage.

Part II
Accelerator and Detector

You know the definition of the perfectly designed machine... The perfectly designed machine is one in which all its working parts wear out simultaneously.

[Frederick Alexander Lindemann 1886-1957]

There are three roads to ruin; women, gambling and technicians. The most pleasant is with women, the quickest is with gambling, but the surest is with technicians.

[Georges Pompidou, 1911-1974]

5 The PEP-II accelerator

In order to pursue the program of studying CP violation in the B-system, asymmetric e^+e^- colliders operating at the $\Upsilon(4S)$ resonance, so-called B-factories, were proposed. The $\Upsilon(4S)$ resonance is the lightest $b\bar{b}$ state above the $B_d\overline{B}_d$ threshold, and the asymmetry of the collider allows for time dependent measurements, as the long lifetime of the B-meson (order 1.5 ps) results in resolvable decay lengths (order 250 μm).

As the branching ratio for interesting B-decays is of order $\mathcal{O}(10^{-4})$ or less, at least 10^7 $B\overline{B}$ pairs are needed, requiring a very high luminosity and a large acceptance in the Center-of-Mass (CM) frame. The wishes for fully reconstructed events with precise time information demands good Particle Identification (PID) and good vertex resolution. The PEP-II collider and the BaBar detector were designed for exactly those purposes.

In the following, first the PEP-II accelerator and second, and more thoroughly, the BaBar detector will be described. The accelerator description will include some of the relevant physics and comparison with other experiments. In the detector description, each subdetector will be described separately, and then, with extensive use of abbreviations for the subdetectors, the performance of the entire detector will be described.

Since the understanding of the BaBar detector is still developing, some of the information given in the following may be somewhat outdated, though large discrepancies should not occur. The section is based mainly on the Technical Design Report (TDR) [47], the BaBar Physics Book [16] and a preliminary draft of a paper to the journal for Nuclear Instrumentation Methods [48].

5.1 General accelerator design

The PEP-II accelerator is situated at SLAC in California, USA. It consists of an about three kilometer long linear accelerator, which injects 9.0 GeV electrons into a High Energy Ring (HER) and 3.1 GeV positrons into a Low Energy Ring (LER) circulating in the opposite direction. The two rings are housed in the former PEP tunnel of 800 meter diameter. The SLAC accelerator site is sketched in Fig. 11 below.
PEP-II was designed to have 1658 particle bunches in each ring, consisting of 2.1×10^{10} electrons for the HER and 5.9×10^{10} positrons for the LER, resulting in a bunch crossing every 2.1 ns at the interaction point, where BaBar is located. The rings are recharged by the linear accelerator when the luminosity drops below 90 % of its peak value, approximately every 9/3 hours for the HER/LER respectively.

Parameter	Unit	Design	Best achieved	Typical
Number of Bunches (HER+LER)		1658	1658	553-829
Total Beam Current (HER)	A	0.75 (1.0)	0.92	0.70
(LER)	A	2.14	2.14	1.10
Beam Lifetime (HER)	hours	4 (@ 1.0A)	11 (@ 0.90A)	9 (@ 0.70A)
(LER)	hours	4 (@ 2.0A)	3.3 (@ 1.40A)	3 (@ 1.10A)
Luminosity	$\mathrm{cm}^{-2}\mathrm{s}^{-1}$	3×10^{33}	3.1×10^{33}	2.0×10^{33}
Hor. Spot Size Σ_X	μm	220		147
Ver. Spot Size Σ_Y	μm	6.7		5.0
Fill Time	minutes	6	8	10
Integrated Luminosity	ps^{-1}/day	135	174	140
Total Int. Lum. (Run-I)	fb^{-1}		25.3	

Table 1: Parameters and performance of PEP-II. The numbers are from Run-I.

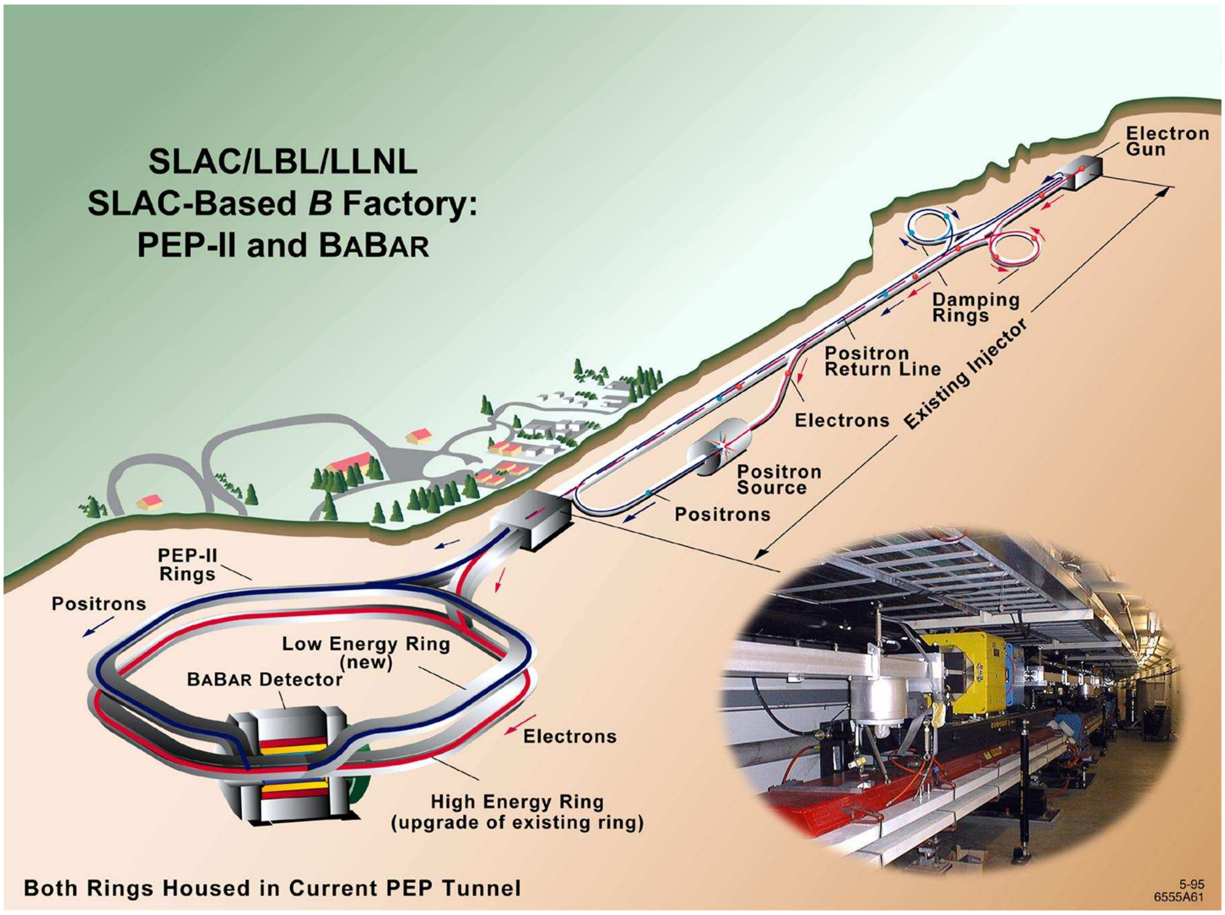


Figure 11: Layout of the SLAC accelerator site. The linear accelerator fills the PEP-II rings with positrons respectively, which are brought to collide at the interaction point of the BaBar detector.

PEP-II has delivered beam from 26$^{\text{th}}$ of May 1999 until 31$^{\text{st}}$ of October 2000 (Run-I), starting again 1$^{\text{st}}$ of February 2001. The luminosity in Run-I was averagely 72 % of its design value, which is $3 \times 10^{33}\ cm^{-2}s^{-1}$, resulting in a total integrated luminosity of 25.3fb^{-1} of which the BaBar detector logged 21.22fb^{-1} on the $\Upsilon(4S)$ resonance and 2.66fb^{-1} off resonance. In Table 5.1 are listed various beam parameters (anno 2000) and luminosity performances [47].

5.2 *B*-factory vs. other experiments

The advantages of an e^+e^- collider at the $\Upsilon(4S)$ resonance are many:

- High ratio of signal to background ($\sigma_{b\bar{b}}/\sigma_{\text{total}} = 0.28$).
- Clean events (average charged multiplicity ~ 11, $B\overline{B}$-decays are exclusive.).
- Kinematic constrains ($p_{\Upsilon(4S)}$ and $|\mathbf{p}_B|^{[CM]}$ are known).
- Acceptable $b\bar{b}$ cross-section ($\sigma(e^+e^- \to \Upsilon(4S)) = 1.05nb$).
- π^0 and radiative decay capability through γ detection.

The clean events, especially the fact that no other particles are produced along with a $B\overline{B}$-pair, makes reconstruction and in particular tagging highly efficient compared to other types of experiments. The limiting factor for the B-factories is the luminosity. In order to make a 1 % measurement of a channel with a branching ratio of $\sim 10^{-4}$, which is typical for the $B \to \pi^+\pi^-$

and the $B \to J/\psi K_S$ channels (measuring $\sin(2\alpha)$ and $\sin(2\beta)$ respectively), one needs $\sim 10^9$ $B\overline{B}$-decays. The current luminosity of PEP-II, 3×10^{33}, allows for reasonable statistics ($\sim 10^8$) given a few years of running. So in order to reach the 1 % level, further luminosity upgrades are needed, which is also planned for (see section 7.3.3). If one wishes to go beyond 1 % precision, even higher luminosities are needed. Such "super" B-factories with luminosities beyond 10^{34} are currently under investigation [49].

Other types of experiments with B-physics potential can be divided into three categories: *LEP experiments*, *HERA-B*, and *Tevatron/LHC experiments*. The four *LEP experiments* had the advantage of a somewhat high $b\bar{b}$ cross-section (6 nb) with a large boost ($\beta\gamma \simeq 9$) in a fairly clean environment, but the luminosity of LEP did not suffice for adequate statistics. *HERA-B* have about the same parameters (cross-section of 10 nb and boost of $\beta\gamma \simeq 20$), but a much higher luminosity. However, the very low signal to background ratio requires an extremely good background rejection and complicates the experiment. The *Tevatron/LHC experiments* have even higher $b\bar{b}$ cross-section and boost, and the ratio of signal to background is fair. For this reason they will make the ultimate precision measurement in B-physics.

All these other types of experiments have the advantage of B_s capabilities, which can only be done at B-factories by going to the much broader and therefore less advantageous $\Upsilon(5S)$, which lies above the $B_s\overline{B}_s$ threshold. This possibility has been investigated, and depending on results from other experiments, it is feasible, but only after several years of running at the $\Upsilon(4S)$ resonance.

5.3 Cross-sections at the $\Upsilon(4S)$ resonance

In order to know the physics backgrounds[23] of the experiment, it is necessary to know what is produced at the $\Upsilon(4S)$ resonance. The cross-section for production of fermion pairs from e^+e^- collisions with CM energy at the $\Upsilon(4S)$ resonance is given in Table 2 below[24] [16]:

$f\bar{f}$	Cross-section (nb)
$b\bar{b}$	1.05
$c\bar{c}$	1.30
$s\bar{s}$	0.35
$d\bar{d}$	0.35
$u\bar{u}$	1.39
$\tau^+\tau^-$	0.94
$\mu^+\mu^-$	1.16
e^+e^-	~ 40

Table 2: Cross-sections for e^+e^- at $\sqrt{s} = m_{\Upsilon(4S)}$.

Apart from the cross-section to e^+e^- (mainly Bhabha scattering), having been corrected for detector acceptance as it is highly angular dependent, these numbers reflect to zeroth order the classic cross section for e^+e^- collisions at high energy:

$$\sigma(f\bar{f}) = N_c Q^2 \frac{4\pi\alpha^2}{3E_{\rm CM}^2}\sqrt{1-\frac{m_f^2}{E^2}}\left(1+\frac{1}{2}\frac{m_f^2}{E^2}\right) \overset{E_{CM}=m_{\Upsilon(4S)}}{\simeq} N_c Q^2\ 0.78\text{nb}. \tag{76}$$

Taking initial state radiation into account, the *actual* CM energy has a long tail into the lower energy region. The result is that fermion pairs are also created at lower CM energies, where

[23]Backgrounds that originate from physics processes — thus not related to running the accelerator.
[24]These cross-sections were calculated with the Jetset 7.4 event generator.

the cross-section is higher due to the inverse scaling with E^2_{CM}, thus making the effective cross-section higher. Radiative corrections are expected to be largest for the light quarks, which is the reason for the larger $u\bar{u}$ than $c\bar{c}$ cross-section.

5.4 Resonance vs. Continuum running

In order to make precision measurements, one needs to determine the backgrounds from the continuum. This is done by running "off peak", i.e. 40 MeV below the $\Upsilon(4S)$ resonance, where the $B\overline{B}$ signal has dropped to $s\Gamma_{e^+e^-}(s)\Gamma(s)/((s-m_{\Upsilon(4S)})^2)^2+s\Gamma^2(s)) \simeq 3\%$, but the continuum is essentially unchanged. For decay channels with little background (e.g. $J/\psi K_S$) the optimal fraction of continuum running will be low. But for channels with more background (e.g. $\pi^+\pi^-$), the uncertainty in the background becomes significant, and continuum running is needed.
In order to determine the optimal fraction of continuum running, c, for a specific channel with background of signal, $b \equiv B/S$, one considers the number of events on resonance and in continuum for a fixed amount of integrated luminosity, $\mathcal{L}$:

$$N_{\Upsilon(4S)} = (1-c)\mathcal{L}(1+b)S, \qquad N_{\mathrm{continuum}} = c\mathcal{L}bS. \tag{77}$$

Isolating S and determining its error gives the following expression:

$$\sigma(S) \quad = \quad \sqrt{\frac{S}{\mathcal{L}}}\sqrt{\frac{c+b}{c(1-c)}}. \tag{78}$$

Not surprisingly, the error is proportional to the square root of the signal over the integrated luminosity, which just serves as normalization constants. The minimization of the second square root yields the optimal fraction of continuum running. For no background ($b=0$) the minimum is obviously at $c=0$, but with just 5 % background, 18 % continuum running is optimal, and for $b=1$ the optimal fraction of continuum running reaches 41 %.
The first goal of BaBar is to measure $\sin(2\beta)$ in the $B \to J/\psi K^0_S$ decay channel, which is expected to have very low backgrounds (2−3%), thus initially most running will be on resonance. However, once other measurements, which include channels with higher backgrounds, are wanted such as measuring $\sin(2\alpha)$ in the $B \to \pi^+\pi^-$ channel, a larger fraction of continuum running is needed.

5.5 Backgrounds from the accelerator

The requirement of high luminosity and asymmetric beam energies necessarily leads to high levels of machine-related backgrounds in the interaction region. The largest background comes from focusing and bending of the beams in the interaction region, which causes Synchrotron Radiation (SR), limiting the lifetime of the Silicon Vertex Tracker (SVT) and the electromagnetic calorimeter (EMC). While maintaining the large acceptance, the detector must be protected from radiation without lowering the luminosity. The SR from the closest bending magnets is designed to pass through the interaction region without interacting with the beam pipe, while other sources of SR are masked away.
The bending of the beams near the interaction point also causes off-energy beam particles to enter the detector. Coulomb scattering of beam-gas molecules, bremsstrahlung, radiative Bhabha scattering, and beam-beam interactions give rise to similar background, which either enters the detector directly or causes secondary particles that does. This also causes radiation damage as well as occupancy, effectively lowering the efficiency of the detector.

6 The BaBar detector

The measurement of CP violation in the B-system essentially requires good capabilities of:

- reconstructing B-decays into exclusive final states,
- tagging the other B-mesons in each event, and
- measuring the relative time between the two B-decays.

The ability to fulfill the above with a high efficiency for decay channels with branching ratios of order $\mathcal{O}(10^{-4})$ or below, requires large acceptance, good momentum resolution, clean particle identification and precise vertex resolution. It is the optimization of these demands with respect to cost and efficiency that has dictated the design of the BaBar detector. After a general overview, the various subdetectors will be treated separately, and their function then described in the framework of the entire detector.

6.1 General detector design

The mass of the $\Upsilon(4S)$ resonance is $21.2 \pm 3.6\ MeV$ above the $B\overline{B}$ threshold [37], which gives each B-meson a boost of $\beta\gamma \simeq 0.06$ in the rest frame of the $\Upsilon(4S)$. Given the lifetime of the neutral B-meson, $\tau_{B^0} = 1.548 \pm 0.032\ ps$, this results in an average flight distance of 30 μm, which is not resolvable by todays particle detector technology. For this reason the PEP-II collider has, unlike other colliders, been designed with asymmetric beam energies, giving the produced particles a boost of $\beta\gamma \simeq 0.56$. For the $\Upsilon(4S) \to B\overline{B}$ decay, this increases the average flight distance to about 260 μm, well within reach of modern silicon vertex detector technology (see section 6.2). In Fig. 12 is shown the so-called detector protractor, which is the correlation between polar angles in the lab and in the CM reference frame for massless particles. Due to the asymmetry of the collider, the BaBar detector has also been designed asymmetrical, reaching into the forward direction in order to maximize the acceptance.

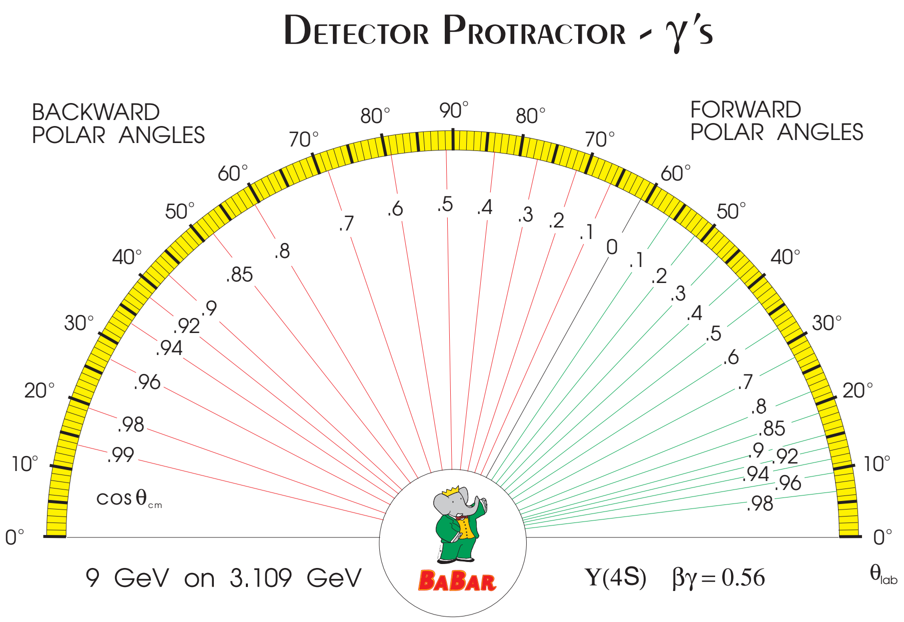


Figure 12: Detector protractor. The correlation between polar angles in the laboratory and the CM reference frame at $\beta\gamma = 0.56$ for massless particles, e.g. a photon emitted perpendicular to the z-axis in the CM frame, will in the lab be detected at about 60° in the *forward* direction.

From inside and out the detector consists of a:

- *Silicon Vertex Tracker (SVT)* for precise charged track information,
- *Drift CHamber (DCH)* for momentum measurement of charged particles and partial PID,
- *Detector of Internally Reflected Čerenkov light (DIRC)* for charged hadron identification,
- *ElectroMagnetic Calorimeter (EMC)* for photon, electron, and charged particle detection,
- *Superconducting coil*, providing a 1.5 T solenoidal magnetic field, and
- *Instrumented Flux Return (IFR)*, for muon and neutral hadron identification,

as can be seen in Fig. 13 below. The above description of functions is only approximate, as most subdetectors provide additional information. Thus, to obtain optimal measurements, information from various subdetectors is combined.

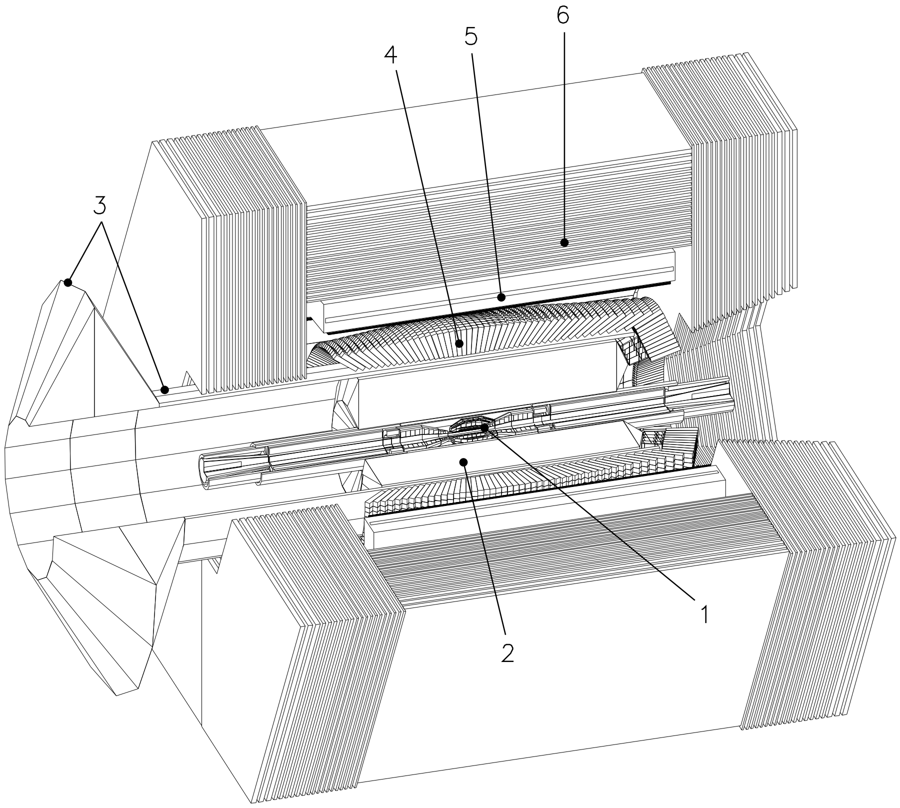


Figure 13: Layout of the BaBar detector. 1. Silicon Vertex Tracker (SVT), 2. Drift CHamber (DCH), 3. Detector of Internally Reflected Čerenkov light (DIRC), 4. ElectroMagnetic Calorimeter (EMC), 5. Superconducting coil, 6. Instrumented Flux Return (IFR).

Like most other particle detectors, the BaBar detector consists of layers of different independent subdetector systems with complimentary functions. Since the particle tracks are smeared by multiple scattering as they pass through material (e.g. the detector), the innermost detector parts should be the ones which require highest spatial precision, and they should consist of least possible material in order not to degrade the measurements of the outer subdetectors. All parts of the detector should also be radiation hard, as the high luminosity of PEP-II causes high levels of radiation.

6.2 The Silicon Vertex Tracker (SVT)

When charged particles pass through semiconducting material they create electron/hole pairs along their path, which are collected by an electric field and thus detected. The energy required to create electron-hole pairs is an order of magnitude smaller than to ionize gas, which gives good spatial resolution, and since they require little space, they are often the innermost subdetector.

6.2.1 SVT functions

The main task of the SVT is to measure the decay vertices of the two B-mesons in order to provide the crucial time information for the CP measurements. In order to efficiently fit the CP asymmetries[25], the determination of separation between the two B-decays should be better than half the average separation of 260 μm, corresponding to 90 μm resolution for each vertex. This precision can be reached, but since better resolution improves background rejection and the statistical power in the CP fit, the aim is the best resolution achievable. The point resolution is limited by multiple scattering to $10-15$ μm for the inner layers and $30-40$ μm for the outer ones, thus impact parameter and angular information from the innermost points, provided by the SVT, are the best.
Due to the strong magnetic field, charged particles with transverse momentum less than 100 MeV will not reach the DCH, and in this case the SVT provides the only tracking information. This is of particular importance for the slow pions from the decay of the D^*, as the D^* plays a central role in the measurement of the tagging efficiency (see section 9). In addition the SVT must be efficient for particles that decays within the active volume, such as K_S^0 (for the $J/\psi K_S$ channel), and it also contributes to pattern recognition and background rejection.
To perform the above functions the SVT should have small segmentation, cause little multiple scattering, and cover much solid angle. As the innermost detector, the SVT should in addition be somewhat radiation hard, in order to stay reliable in the PEP-II environment for many years.

6.2.2 SVT design

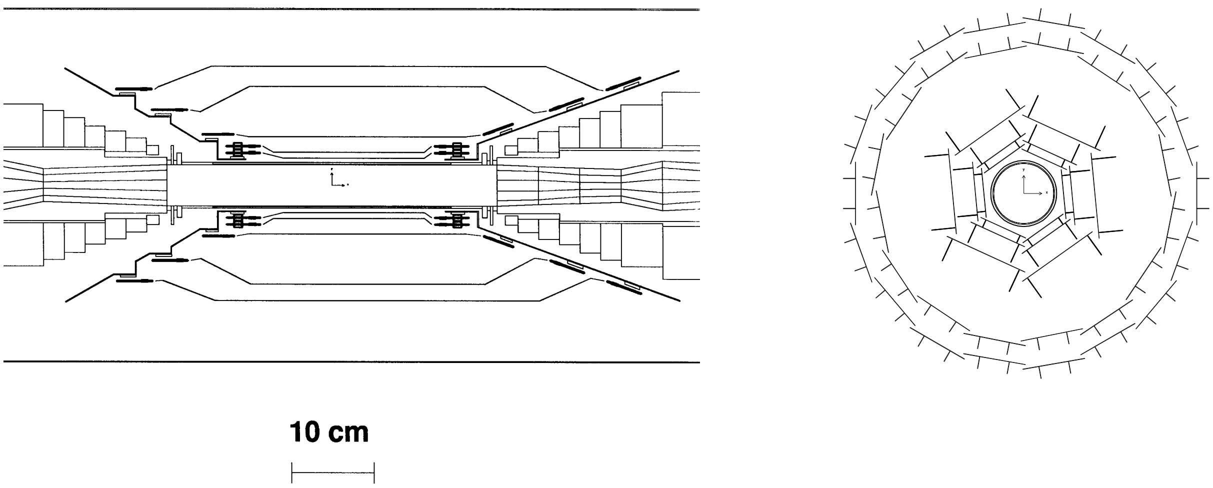


Figure 14: The Silicon Vertex Tracker (SVT). The left figure shows the SVT in the plane of the beam. The right figure shows the SVT in the plane perpendicular to the beam.

The SVT consists of five layers of double-sided silicon strip detector; the three inner layers are in standard concentric cylinder design while the two outer layers are in a novel "arch" design in

[25]What is meant by "efficiently fit" is that 90 % of the statistical power is preserved in the CP asymmetry fits.

which the barrel ends are folded inwards toward the beam pipe, as can be seen in the left figure of Fig. 14. This design maximizes the polar acceptance ($17.2° < \theta_{lab} < 150°$) while minimizing the track incidence angles and the amount of material and thereby the multiple scattering. The modules of the inner layers mutually overlap each other, while the modules on the outer layers are divided into type "a" and type "b" with slightly different radii in order to allow for overlap between the modules. On each module silicon strips are placed both parallel (z-strips, on the inner side) and perpendicular (ϕ-strips, on the outer side) to the beam line, giving both a z (polar) and a ψ (azimuthal) measurement. The readout electronics are placed outside the active region, and the modules are supported on ribs placed on the end cones. The SVT covers a total area of 0.94 m^2, and has about 150.000 readout channels. The dimensions and parameters for the SVT can be found in Table 3 below:

Parameter	Layer 1	Layer 2	Layer 3	Layer 4a/b	Layer 5a/b
Radius (mm)	32	40	54	124/127	140/144
Readout channels	18432	24567	30720	32768	36864
Modules/Layer	6	6	6	8	9
Resolution (μm) z	10	10	10	10-12	10-12
Resolution (μm) ϕ	12	12	12	25	25

Table 3: Dimensions and parameters for the SVT. The resolution quoted is the intrinsic one at 90° incident assuming signal-to-noise of 20 : 1, but multiple scattering makes the actual one higher. The outer layers have a radius at the end of 91 mm and 114 mm respectively.

6.2.3 SVT performance

During the first year of running, the performance of the SVT has reached all the major design goals. The average hit reconstruction efficiency is 98 %, the resolution of the impact parameter is less than 50 μm for particles with $p_\perp > 1$ GeV, and the hit resolution for incident angles of 90° is around 13 μm and increasing with dip angle in accordance with expectations, as can be seen in Fig. 15. The figure shows the Layer 1 hit resolution vs. incident track angle in z (upper plot) and ϕ (lower plot) direction.

Thus the inner and outer layers satisfactoryly fulfill their function of performing angle and impact parameter measurements, and pattern recognition and low $p_\perp$ tracking, respectively.

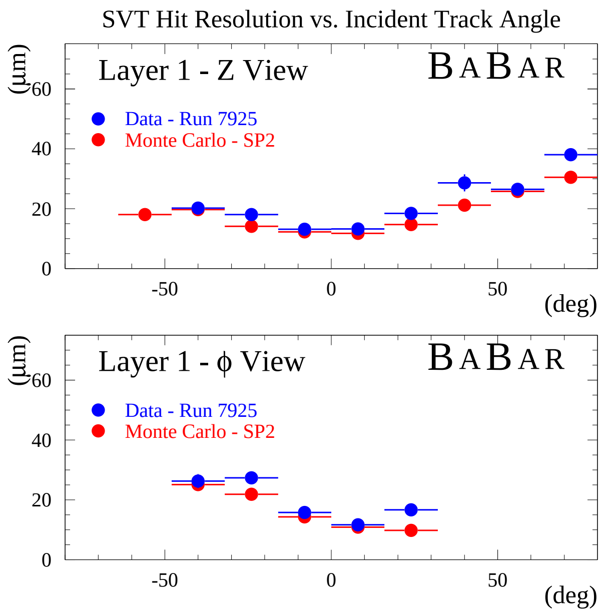


Figure 15: The SVT performance.

6.3 The Drift Chamber (DCH)

Drift chambers use the ionization of gas (or liquid) to reconstruct tracks of charged particles. Anode wires are surrounded by cathode wires in a chamber filled with a suitable gas such that ions and electrons will drift. The distance from the various wires can be deduced from the drift time, and the path of the ionizing particle established.

6.3.1 DCH functions

The DCH plays a central part in the BaBar detector, as it serves several important functions required for exclusive B-reconstruction essential for CP studies. The DCH is the main tracking device, and must therefore have very good spatial resolution and large angular coverage, especially in the forward direction (due to the boost) in order to provide good momentum resolution and high reconstruction efficiency for particles with transverse momentum down to 100 MeV. Good momentum resolution minimizes the background, and the DCH should have a resolution of $\sigma_{p_\perp}/p_\perp = 0.3\,\% \times p_\perp$ for tracks above 1 GeV and $\sigma_{p_\perp}/p_\perp = 0.3 - 0.4\,\%$ down to 100 MeV, which requires a spatial resolution of 140μm in average. Like the SVT, the DCH should remain efficient for particles decaying in the active volume (such as K^0_S).
In addition the DCH provides a measurement of dE/dx with 7 % precision[26] from ionization energy loss. This reliably discriminates between pions and kaons up to momenta around 700 MeV, after which PID is mainly provided by the DIRC (see section 6.4).
Finally, the DCH provides one of the two Level 1 (L1) triggers, and must therefore be able to trigger on charged tracks with a latency of no more than $(9.5 \pm 0.5)\mu$s, constrained by the SVT data buffering. The DCH should be able to perform the above functions in a 1.5 T magnetic field, and with the backgrounds expected to pass the SVT.

6.3.2 DCH design

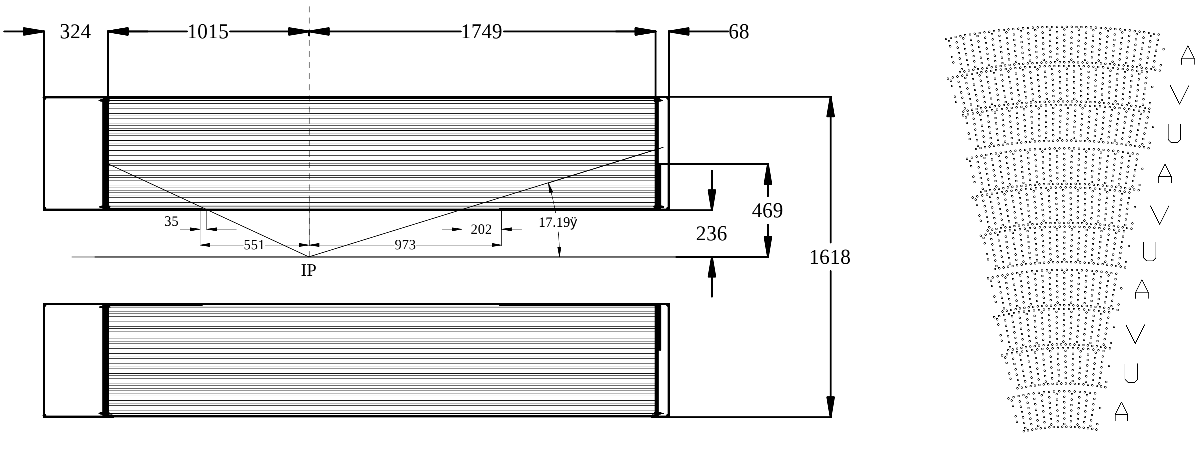


Figure 16: The Drift CHamber (DCH). (left) The DCH in the plane of the beam (the dimensions are in mm). (right) The cell layer of the DCH perpendicular to the beam.

In order to perform the above functions, a small-cell, low-mass drift chamber design was chosen. The DCH shown in Fig. 16 left is a 2.8m long cylindrical chamber with an inner radius of 23.6cm and an outer radius of 80.9cm. It consists of 40 layers of small hexagonal cells arranged in 10 axial (A) (in r-direction), and stereo (U,V) (mainly in z-direction and slightly in ϕ-direction) super-layers following the alternating pattern AUVAUVAUVA (Fig. 16 right). The stereo angles vary from 45 mrad in the innermost super-layer to 75 mrad in the outermost. The 7104 cells have a 20 μm gold-plated tungsten-rhenium sense wires through them, and are divided by 80/120 μm gold-plated aluminum field wires, typically giving each cell dimensions of $12 \times 18\text{mm}^2$. The super-layer design is due to the L1 trigger requirement of fast reduction of input via segment finding. The DCH gas is a mixture of helium (80 %) and isobutane (20 %), since this gives fairly short drift times and good spatial and dE/dx resolution, while at the same time minimizing the amount of material.

[26] Assuming passage of 40 DCH layers, which is the case when the PID by dE/dx is difficult ($p \sim 700$ MeV).

To further minimize the amount of material in the forward direction, the forward endplate (12 mm) was made thinner than the rear endplate (24 mm), and all read-out electronics were mounted on the rear endplate. The latter also allows for placing the DCH in the forward direction with respect to the interaction point as close as possible to the machine parts, maximizing the angular acceptance in the forward direction ($-0.92 < \theta_{lab} < 0.96$).

6.3.3 DCH performance

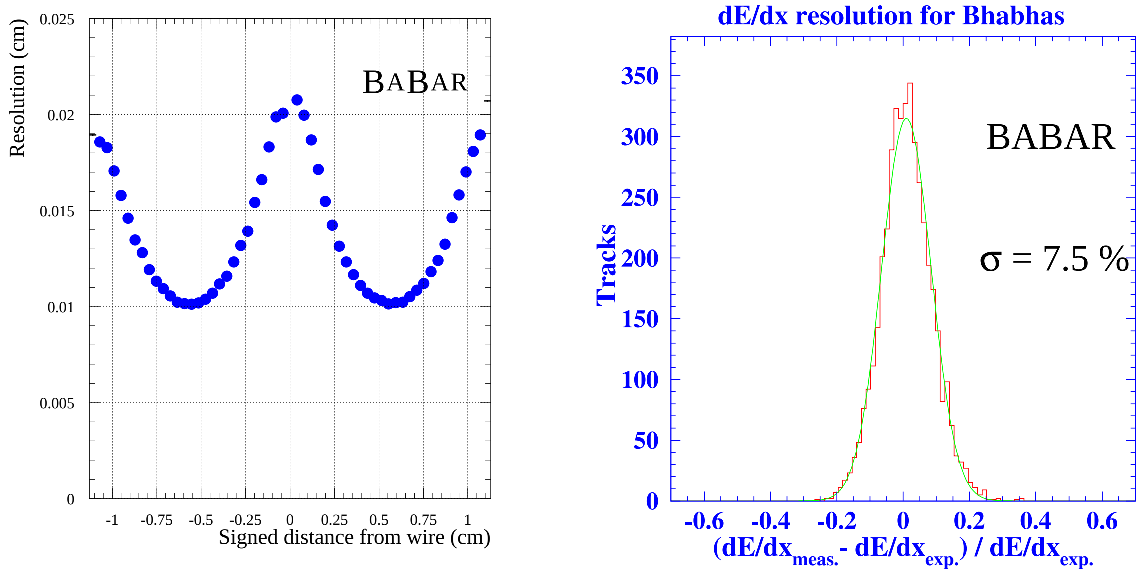


Figure 17: DCH performance. (left) The DCH spatial resolution as a function of drift time. (right) The dE/dx capabilities of the DCH for Bhabha events.

From an ensemble of charged tracks, the single cell resolution was determined as a function of distance from the sense wire shown in Fig. 17 left. The weighted average resolution is 125μm, which is below the 140μm design resolution requirement. Increasing the voltage further improves the resolution, but also induces background and shortens the lifetime of the DCH, and so the voltage setting is a trade-off between the two.
To establish a reliable measurement of dE/dx, polar angle, path length, and individual wire and layer gain has to be corrected for. The resulting measurement of truncated mean dE/dx for Bhabha scattering events is shown in Fig. 17 right. As can be seen, the observed resolution is 7.5 %, and through further correction it is expected to decrease to the design value of 7 %.

6.4 The Detector of Internally Reflected Čerenkov light (DIRC)

When particles exceed the speed of light in a medium, Čerenkov light is emitted in a cone, analogue to the Mach cone created when traveling faster than the speed of sound. As is sketched in Fig. 18 left one can from Huygens principle calculate the light intensity at time t_0 from a particle passing through the origin at time $t = 0$ with velocity $v = \beta c$ in z-direction by summing the amplitudes:

$$A = \int_{-\infty}^{t_0} \frac{e^{iw(\frac{n}{c}\sqrt{(\beta ct)^2+r^2}-t)}}{\sqrt{(\beta ct)^2 + r^2}}\, dt, \tag{79}$$

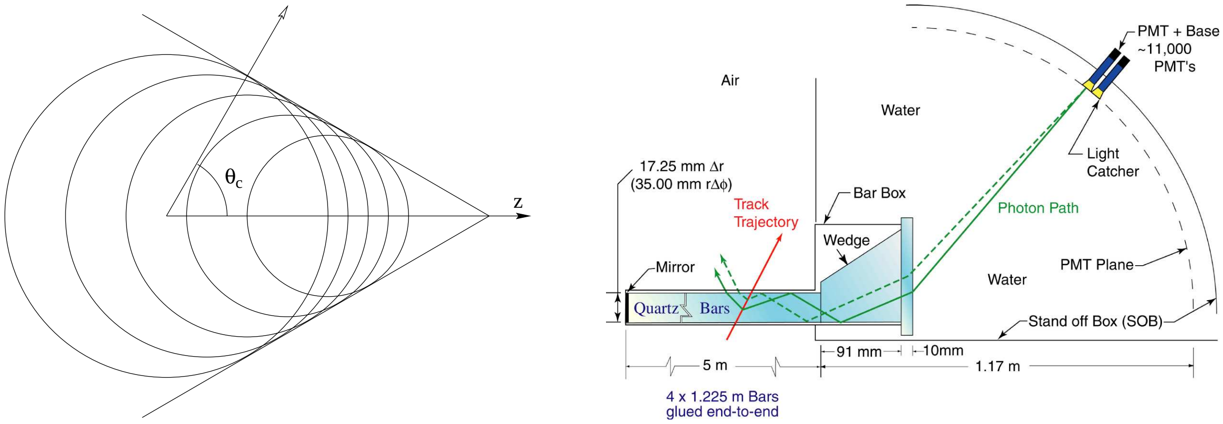


Figure 18: The Čerenkov cone and the principle of the DIRC subdetector. (left) From Huygens principle the direction of constructive interference can be deduced. (right) Sketch of the working principle behind the DIRC subdetector.

where r is the direction orthogonal to z and n is the refraction index of the medium. This integral converges to zero, unless

$$\frac{n}{c}\sqrt{(\beta ct)^2 + r^2} = t \quad \Rightarrow \quad \cos\theta_c = \frac{1}{\beta n}. \tag{80}$$

Thus, given a known material one can from measuring the angle of the emitted Čerenkov light determine the velocity β of a particle, which together with momentum determination is used for identifying particles.

6.4.1 DIRC functions

The DIRC subdetector is devoted to particle identification (PID), primarily separating kaons from pions at high momentum, but also identifying muons at lower momentum and PID in general. The kaon identification is important, both for tagging purposes (see section 10.1.2), but also for CP studies, as the $B^0 \to \pi^+\pi^-$ channel for measuring the angle α has to be separated from the $B^0 \to K^+\pi^-$ channel. Thus, the DIRC must be able to separate kaons from pions up to 4.0 GeV, since the dE/dx measurement of the DCH (see previous section) is only reliable up to 700 MeV. The DIRC should also contribute to the muon identification below 750 MeV, where the IFR (see section 6.6) is not fully efficient.

6.4.2 DIRC design

The DIRC consists of 144 quartz bars (refraction index $n = 1.474$), each 17mm thick (Δr), 35mm wide ($r\Delta\phi$), and 4.9m long (Δz), grouped in 12 barboxes arranged around the drift chamber forming a cylinder with radius 0.85m around the beam. The photons emitted by particles above the Čerenkov threshold, $n\beta > 1$, are through internal reflection guided to the water filled Stand-Off-Box (SOB) at the back end of the bars, where 10752 photomultipliers (PMT) measure the angle of the light. With this design the angular coverage is $-0.84 < \cos\theta_{lab} < 0.90$, and the principle of the DIRC is sketched in Fig. 18 right.

6.4.3 DIRC performance

From the track information the Čerenkov angle can be deduced, but due to the internal reflection of the photons within the quartz bars there are several solutions when associating a DIRC

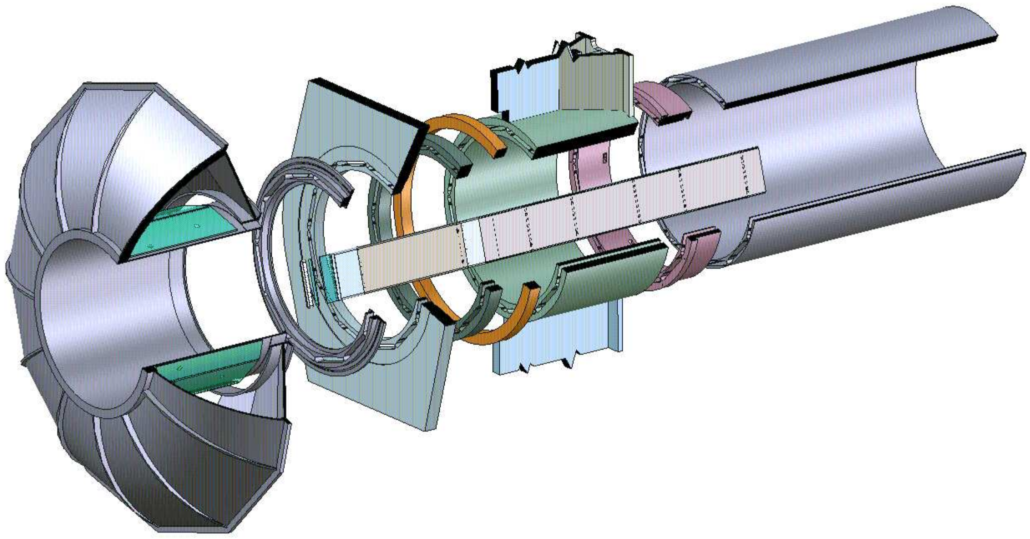


Figure 19: Layout of the DIRC. The DIRC detector with its bars, light guides and Stand-Of-Box.

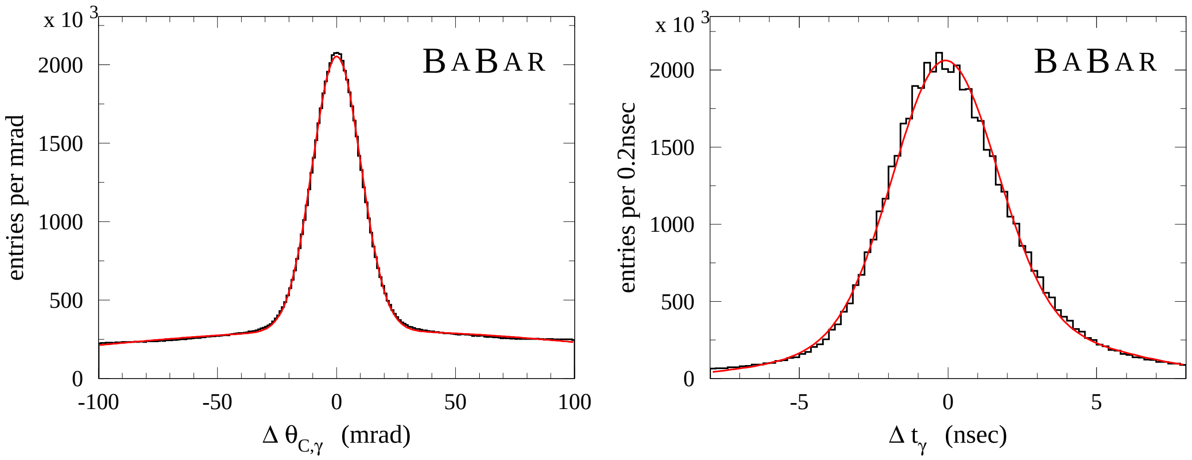


Figure 20: The DIRC performance. (left) Single photon resolution of the Čerenkov angle. The almost flat background is due to several hit-to-track solutions. (right) Time resolution $t_{\text{measured}} - t_{\text{expected}}$.

PMT-hit with a track. Each track produces on average 30 photons. Thus with a single photon resolution of 10.2 mrad (see Fig. 20 left) the track resolution becomes 2.8 mrad, which enables a 3σ separation of charged kaons and pions up to 3.0 GeV. The time resolution of the PMT is 1.7ns shown in Fig. 20 right, which effectively suppresses the incorrect photon solutions as well as background photons. Recent improvement in alignment and fitting has increased the track resolution to 2.55 mrad and removed a slight bias in the reconstructed angle.

6.5 The ElectroMagnetic Calorimeter (EMC)

The detection of photons allow for π^0-reconstruction as well as measuring radiative decays, and for this energy measurement of photons and other particles interacting electromagnetically, “doped” crystals are used. When they pass through crystal, the atoms are excited and emit photons in the ultraviolet spectrum, that through a wavelength shifter (the “doped” impurities, which absorb light at one frequency and reemit it at another one) are turned into frequencies compatible with standard PhotoMultiplier Tube (PMT) acceptance.

6.5.1 EMC functions

In order to efficiently tag and reconstruct CP eigenstates, which is the primary goal of BaBar, an electromagnetic (EM) calorimeter with excellent energy and angular resolution and high efficiency is needed. The resolution requirements are to improve the π^0 and B mass resolution, in turn improving the background rejection. The required EMC-resolutions in energy and angle respectively are at 90° polar angle:

$$\sigma_E/E(\mathrm{GeV}) = \frac{1.0\,\%}{\sqrt[4]{E(\mathrm{GeV})}} \oplus 1.2\,\%, \qquad \sigma_{\theta,\phi} = \frac{3\mathrm{mrad}}{\sqrt[2]{E(\mathrm{GeV})}} \oplus 2\mathrm{mrad}. \tag{81}$$

Good efficiency is required in order not to degrade the already low branching ratios typical for CP eigenstates ($\mathcal{O}(10^{-3})$), which often have photons in their final state. Therefore, the calorimeter should cover as much solid angle as possible, and efficiently detect photons down to $10-20$ MeV, as generic B-decays on average contain 5.5 photons, with energies peaking below 100 MeV.

The EMC should also supply lepton identification, efficiently separating electrons from pions and muons, since lepton tagging requires a rejection of pions misidentified as electrons to the level of $\sim 10^{-3}$. Contrary to the previous detectors, the EMC should not minimize material, as the particles detected by the outermost subdetector (the IFR, see next section) are only slightly affected by the EMC crystals. The total radiation length, X_0, from the interaction point to the EMC is 0.25, mostly due to the DIRC. This is an acceptable level, considering that less material would imply degrading the inner subdetectors, and it allows for the desired energy resolution.

6.5.2 EMC design

The EMC consist of 6580 cesium iodide (CsI) crystals doped with 0.1 % thallium (Tl) iodide, arranged in a barrel section and a forward conic endcap. The dimensions of the crystals are 47-60 mm $\times$ 47-60 mm $\times$ 297.6-325.5 mm decreasing in transverse dimension, and increasing in length toward the forward direction, the lengths corresponding to 16.1-17.6 radiation lengths, X_0. The barrel has inner/outer radii of 91cm/136cm respectively, and is located asymmetrically around the interaction point, while the endcap extends down to 55.3cm from the beam, see Fig. 21. This design gives a polar angle coverage of $-0.78 < \cos\theta_{lab} < 0.96$, though the effective coverage is slightly less due to shower leakage in the end zones.

Crystals with energy deposit above 1 MeV are read out at the rear end with two independent 2cm^2 photodiodes, because of the strong magnetic field surrounding the EMC. The photodiodes reside together with preamplifiers in a shielded housing also used for heat removal. The noise from readout electronics is expected to be around 150keV, which is entirely negligible compared to the background noise.

To reach the desired energy resolution for the EMC, constant calibration of the energy scale is crucial, in order to monitor both short- and long-term variations. This is done using the following techniques:

- *Charge injection into front end of amplifiers*, in order to monitor the precise response function (0.1 % level) of the read-out electronics for each channel.
- *Light pulsar system* injecting light via fiberoptics into the rear end of each crystal for tracking short term changes to better than 0.15 % (with daily light pulsar runs).
- *Radioactive source* (N^{16}) producing 6.1 MeV photons for setting the energy scale precisely (> 0.5 %), and monitoring long term variations.

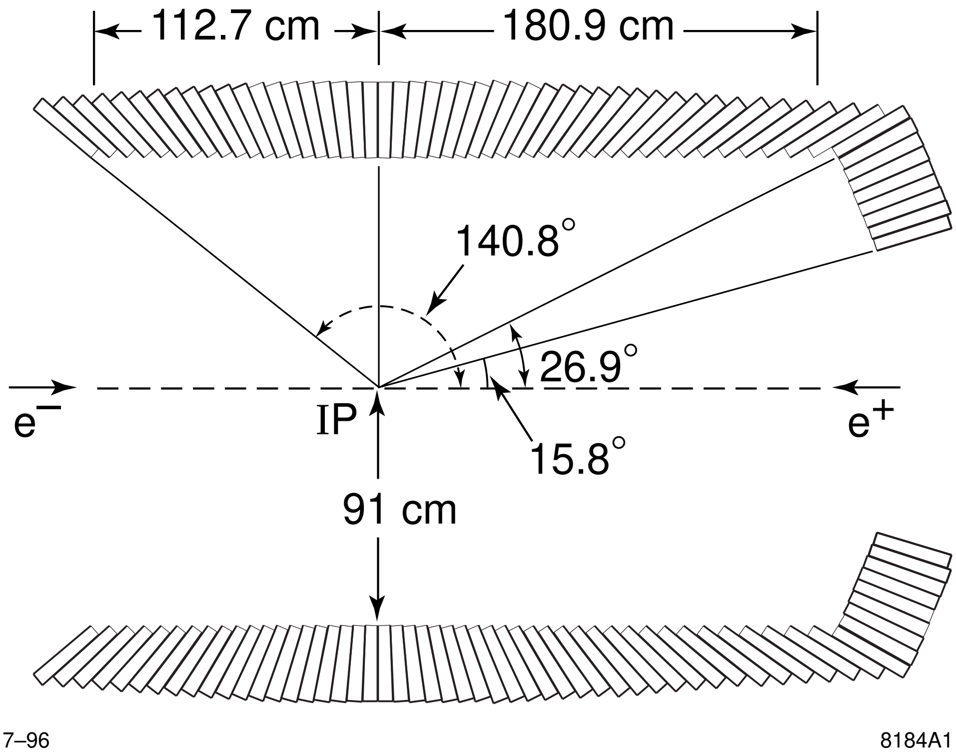


Figure 21: Layout of the EMC. Sideview of the EMC, showing the crystal arrangement of the barrel and the endcap.

- *Known events* (Bhabha, π^0, $\gamma\gamma$ and $\mu^+\mu^-$ events) adding calibration points. Together with the radioactive source this information is updated weekly to measure the small changes in light yield with integrated radiation.

Parameter	Barrel	Endcap
Number of Crystals ($17.6X_0$)	840	820
Number of Crystals ($17.1X_0$)	840	0
Number of Crystals ($16.6X_0$)	840	0
Number of Crystals ($16.1X_0$)	3240	0
Total Volume (m^3)	5.2	0.7
Number of Readout Channels	11760	1640

Table 4: Parameters for the EMC.

6.5.3 EMC performance

The energy resolution is determined by fitting data from the N^{16} source, pi^0, η and Bhabha events, while the angular resolution is measured with equal energy pi^0 and η decays to two photons. The results are:

$$\sigma_E/E(\mathrm{GeV}) = \frac{2.55 \pm 0.30\,\%}{\sqrt[4]{E(\mathrm{GeV})}} \oplus 1.51 \pm 0.22\,\%, \qquad \sigma_{\theta,\phi} = \frac{3.7 \pm 0.3\mathrm{mrad}}{\sqrt[2]{E(\mathrm{GeV})}} \oplus 0.31 \pm 0.27\mathrm{mrad}. \quad (82)$$

Property	CsI(Tl)
Radiation Length	1.85 cm
Density	4.53 g/cm^3
$dE/dx\|_{\text{Min. Ionizing part.}}$	5.6 MeV/cm
Light Yield	$\sim$ 5000 γ/MeV
Peak Wavelength	565nm
Decay Time	940ns

Table 5: Properties of Thallium-doped CsI.

The angular resolution is slightly better than the design requirements. The obtained energy resolution is in agreement with recent MC studies, but 20-60 % worse than TDR expectations. This is thought to be due to induction of signal from one channel to the next (cross talk), such that the noise from read-out electronics is 250 keV after signal processing. It degrades the ΔE resolution for B-decays involving neutrals, but the EMC is still performing well for physics, and the TDR discrepancy is thought to be due to unrealistic simulations. Digital filters are considered, as they would decrease the noise from electronics by a factor of two.
For photons with $E_\gamma > 20$ MeV the efficiency exceeds 96 %, and the separation of electrons from hadrons is also satisfactory (see section 7.2). The loss of light from radiation damage follows most optimistic expectations, and if this favorable behavior continues the crystals could last the entire experiment without major replacements.

6.6 The Instrumented Flux Return (IFR)

Particles reaching the IFR, such as muons and neutral hadrons, require much material to be stopped, and thus have their energy determined. To do so, alternating layers of metal and scintillater plates are used, since the showering in the metal will create charged particles, detectable in the scintillation layers. Charged hadrons are also detected in the IFR, but since they are also visible to the much better inner detectors, the IFR information is rarely used for that.

6.6.1 IFR functions

The main tasks of the IFR are muon identification and neutral hadron detection. Both reconstruction and tagging of B-decays require good muon identification over a large range of momenta, since the muons originate from both primary decay vertices as well as secondary, lower-momentum cascade decays (18 % of B-decays contain at least one muon). The IFR should be able to identify muons with momenta ranging from 450 MeV (250 MeV at low angle) up to several GeV. For muon tagging purposes, the upper limit for a hadron to be misidentified as a muon should be below 5 %.
The identification of neutral hadrons mainly concerns the longlived neutral kaon, K_L^0, as their detection opens the possibility of studying the $J/\psi K_L$ channel for CP-violation studies. Also, many interesting rare decays involve K_L^0. In addition, the IFR is used to veto on events with missing hadronic energy, and provides a trigger on cosmic ray events used for calibration and vetoing. Finally, the iron layers of the IFR comprise the flux return for the solenoidal magnet, thus the name: Instrumented Flux Return.

6.6.2 IFR design

The IFR shown in Fig. 22 below, is the outermost subdetector of BaBar, and consists of alternating layers of iron and Resistive Plate Counters (RPC) arranged to form a barrel and two

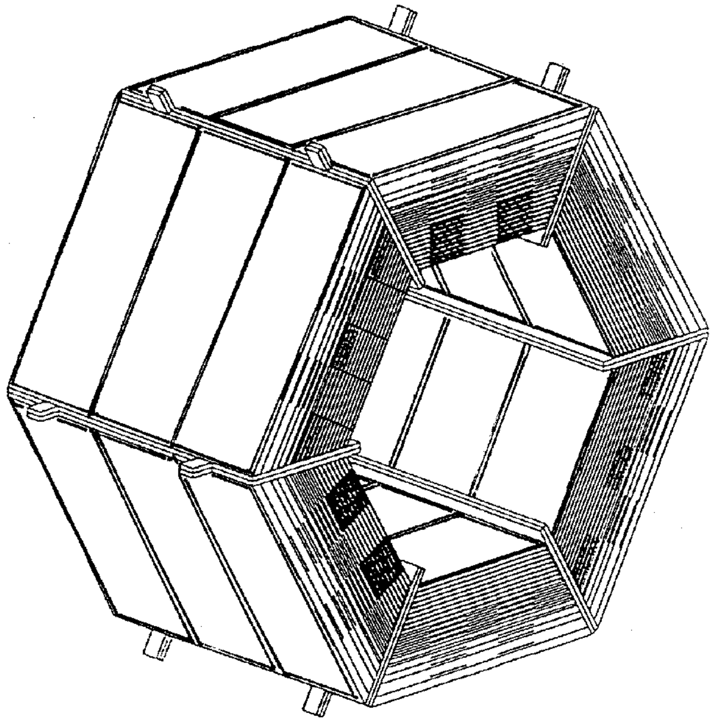


Figure 22: Layout of the IFR. The barrel and the two endcaps give almost 4π angular coverage.

endcaps. The iron plate thickness increases from 2 cm for the innermost plates to 10cm for the outermost for a total of 65cm (60cm) of iron and 19 (18) layers of RPC for the barrel (endcaps). In addition, two RPC layers are placed in the barrel region between the EMC and the cryostat for the solenoidal magnet.

The barrel extends radially from 1.78m to 3.01m and is divided into sextants, each 3.75m long and 1.88m to 3.23m wide. The endcaps are hexagonal with a central hole for the beam components, giving the IFR an angular coverage of $-0.92 < \cos\theta_{lab} < 0.96$. The iron flux return is water-cooled to reduce temperature variations larger than 1°C. The RPC layers are divided into strips in alternating direction, such that they provide ϕz-position in the barrel, and xy-position in the endcaps.

6.6.3 IFR performance

The performance of the RPCs are determined using both collision and cosmic ray data, and the average efficiency for the 2000 run was 75 % for the barrel and 81 % for the endcaps, slowly decreasing with time. This is partly due to high-voltage problems with a small fraction of the RPCs; problems which are currently being investigated.

6.7 The superconducting solenoidal magnet

The magnetic field needed for measuring the momentum of charged particles is provided by a thin 1.5 T superconducting solenoidal magnet. The inner radius and length of the magnet is determined by the four inner subdetectors, and the thickness by the required muon and K_L^0 efficiencies. The magnitude and uniformity specifications are derived from DCH momentum resolution requirements that seem to suggest, that a 1.5 T field with 2 % uniformity in the tracking region will yield a momentum resolution of 21MeV for the most energetic events, such as $B \to \pi^+\pi^-$.

6.8 The trigger

The trigger makes the decision of which events, observed by the detector, to record for later physics analysis, and which ones, to discard as uninteresting or background. The trigger consists of two levels: a hardware level (called Level 1 (L1)), and a software level (called Level 3 (L3))[27]. The L1 trigger is designed to reduce the initial PEP-II event rate from up to several MHz to 2 kHz. A further reduction to 100 Hz (required by the online computing system) is achieved by online reconstruction in the L3 trigger. This also measures the luminosity to about 2 % precision, and the data is then stored for analysis. Since the signal event rates at PEP-II are relatively small, the efficiency should be close to 100 % for all processes of interest.

6.8.1 Level 1 trigger

The L1 trigger is designed to achieve very high and well understood efficiencies. It comprises information from the DCH (charged particles), the EMC (neutral particles), and the IFR (cosmic ray event veto) to form a decision within 11-12 μs after the corresponding interaction. The requirement for a charged track trigger is at least two tracks in the DCH, one passing all 10 superlayers with $p_\perp > 150$ MeV and one passing 5 superlayers with $p_\perp > 120$ MeV. The neutral trigger requires two energy clusters, one above 100 MeV, and one above 300 MeV in the EMC. Each of these triggers are more than 99 % efficient for $B\overline{B}$ events, allowing for cross-calibration and combined efficiency above 99.9 %.

A global L1 triggering unit attempts to match the angular information from the charged and the neutral triggers, and includes a veto on cosmic ray events from the IFR for the final L1 trigger, which is both general and flexible. In a typical run the L1 trigger rate is 700 ± 70 Hz, which significantly reduces deadtime and leaves room for the expected increase in luminosity. The 700 Hz L1 trigger rate comes from lost particles interacting with the beam pipe (~ 60 %), Bhabha events (~ 20 %), and cosmic ray events (~ 20 %), while actual physics events (see Tab. 2) still comprise only a small fraction of the triggers (> 5 %).

6.8.2 Level 3 trigger

Further reduction of the rate without loss of signal requires reconstruction of the event, which is done in the Online Event Processing (OEP) consisting of 32 farm nodes running in parallel. Table 6 shows the different output rates from the L1 trigger. The L3 trigger is a logical OR of two orthogonal filters using the reconstruction information based on lookup tables. The track filter requires either one track coming from the interaction point with $p_\perp > 800$ MeV or two tracks with less strict vertex cuts and $p_\perp > 250$ MeV. The cluster filter requires either high multiplicity or large energy deposits in the EMC with high invariant mass. The $B\overline{B}$ efficiency is 99 % for the track filter, and 94 % for the cluster filter, leading to a good efficiency. While vetoing on most Bhabha events, some are logged along with random triggers for calibration, and the typical logging rate for $\mathcal{L} = 2 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ is ~ 85 Hz, performing beyond design.

6.8.3 Online Prompt Reconstruction

Online Prompt Reconstruction (OPR) of the 100 Hz event rate is performed with 150 unix machines, which makes "rolling calibration" (realtime update of calibration constants) possible. In addition, good online data quality monitoring can be made, and the data is available shortly after running. A second identical farm reprocesses older data, whenever new and improved algorithms become available.

[27] A level 2 trigger, which is a fast software trigger on partial detector information, was not needed for BaBar.

Process	L1 output rate
$\Upsilon(4S)$	4 Hz
$q\bar{q}$	8 Hz
$\mu^+\mu^-$	2 Hz
$\tau^+\tau^-$	2 Hz
$\gamma\gamma$ (signal)	3 Hz
Bhabhas (recorded)	10 Hz
Random triggers	5 Hz
Signal	$\sim$ 35 Hz
$\gamma\gamma$ (background)	45 Hz
Bhabhas (total)	120 Hz
Lost particles	380 Hz
Beam gas	15 Hz
Cosmic rays	130 Hz
Total	$\sim$ 700 Hz

Table 6: Signal and background rates from the L1 trigger [50]. The expected rates are for a luminosity of $\mathcal{L} = 3 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$.

7 Detector performance

The full detector performance is the result of an interplay between the various subdetectors, which provide both complementary and supporting measurements. In the following, the performance of the full detector in terms of track, vertex, energy resolution, and particle identification as well as general issues will be presented.

7.1 Tracking and vertexing

The tracking algorithm searches for all possible tracks in the DCH, and then do a refined Kalman fit [51], which takes the detector material and magnetic field variations into account. The tracks are then projected into the SVT, and refitted if consistent.

The tracking efficiency is measured using several methods, where one has knowledge of charged tracks without using SVT and DCH information. Low multiplicity high $p_\perp$ Bhabha scattering events based on EMC information cover the high momentum range while $D^0 \to K^-\pi^+\pi^+\pi^-$ compared to $D^0 \to K^-\pi^+$ measures the lower momenta. Also $\tau^+\tau^-$ events, multiplicity distribution, and SVT/DCH comparison are used as checks, and the result is a very conservative 2.5 % systematic error per track with $p_\perp > 200$ MeV. Further checks show that tracking and vertex errors are estimated correctly within 10-20 %.

7.1.1 Slow pion reconstruction

The decay $D^{*\pm} \to D^0\pi^\pm_{\text{slow}}$ is important for the tagging efficiency measurement (see section 9), and since slow pions with $p_\perp < 120$ MeV are not detected by the DCH, the SVT has been designed to have stand-alone slow pion reconstruction capabilities in the $50 - 120\,\text{MeV}$ $p_\perp$ range. Figure 23 left shows the transverse momentum spectrum both for data and MC, and Fig. 23 right the efficiency obtained from MC, since the data is still statistically limited. As can be seen, the SVT provides good reconstruction down to 60 MeV.

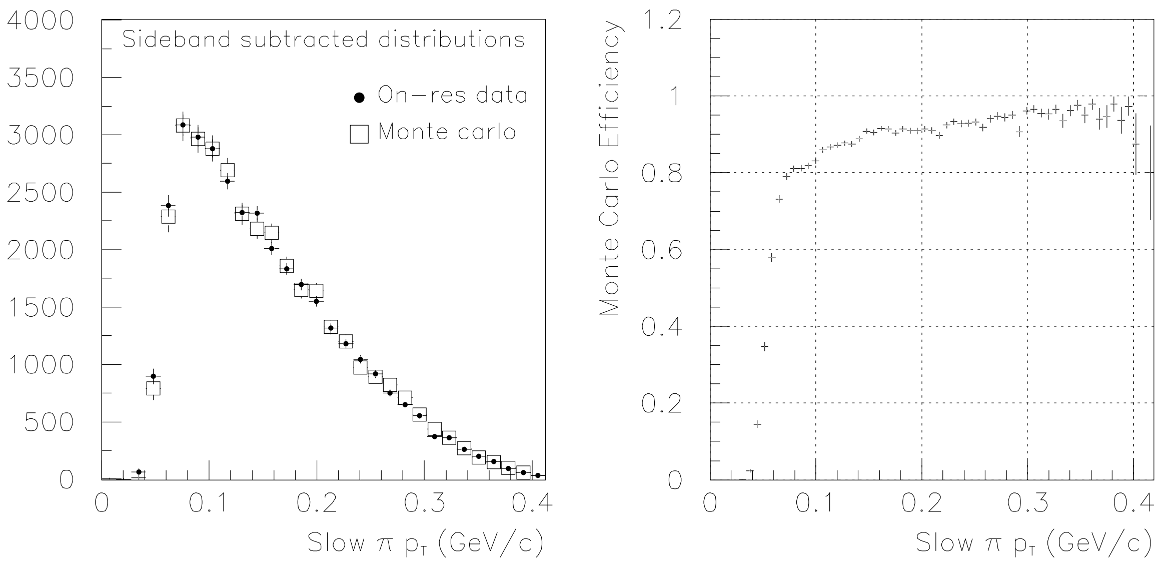


Figure 23: Slow pion reconstruction capabilities. (left) Transverse momentum spectrum of slow pions obtained from data (dots) and MC (squares). (right) Slow pion reconstruction efficiency obtained from MC, since data is still statistically limited.

7.1.2 Beam spot and vertex resolution

The collision point is determined from tracks in order to monitor movements of the beam with respect to BaBar. The size of the beam spot is $150\,\mu\text{m} \times 4\,\mu\text{m} \times 1$ cm in xyz, and the primary vertex resolution is $\sim 70\mu$m in x, y and z for hadronic events. Biases in determining vertex positions due to alignment problems are at the level of few microns.

7.1.3 Angular acceptance

The angular acceptance for charged tracks is defined by the coverage of the full SVT, and stretches over the range $-0.82 < \cos\theta_{lab} < 0.92$, corresponding to 410 mrad (602 mrad) between acceptance edge and beam in the forward (backward) direction. For neutral tracks, the acceptance in the backward direction is reduced due to the lack of an EMC endcap. The resulting angular acceptance is $-0.74 < \cos\theta_{lab} < 0.92$, i.e. 732 mrad from the beam axis.

7.2 Particle Identification

Particle Identification efficiencies can be determined by samples obtained from kinematic information only. These samples should be as pure as possible, and cover the entire phase space both in momentum and angle. The two performance measures of importance for PID are *efficiency* and *purity*. In general, the two are correlated, and the chosen level[28] thus reflects the type of analysis one wishes to make.

7.2.1 Electrons

Electrons are identified by the DCH, DIRC, and in particular the EMC, using E/p, transverse shower shape and various moments as discriminating variables. The PID efficiency is obtained

[28]The general terminology for selections is: very loose, loose, tight and very tight.

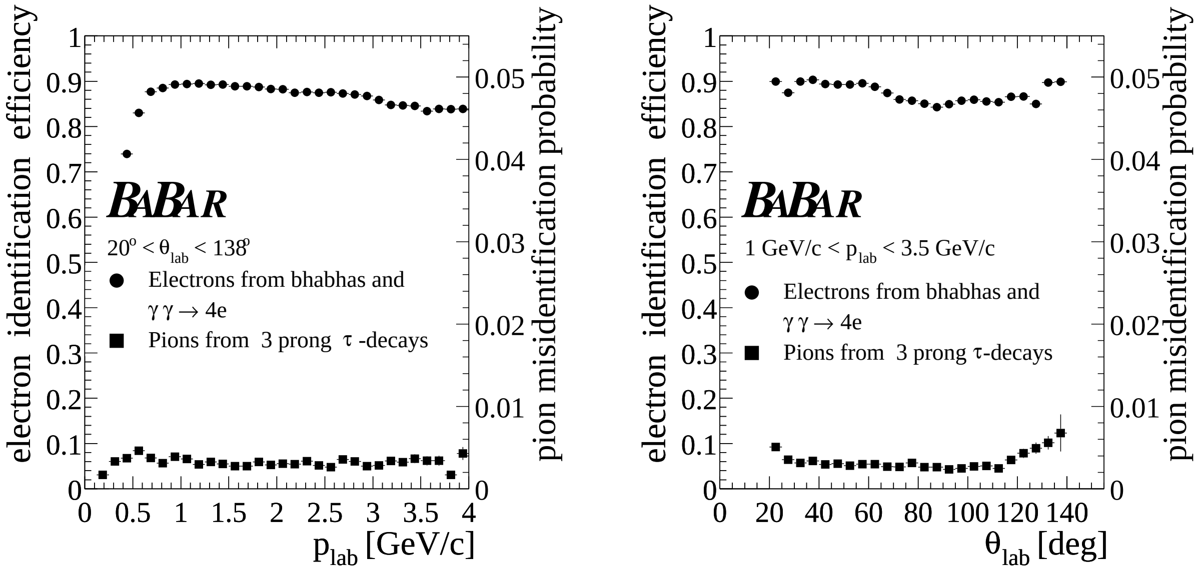


Figure 24: Electron identification capabilities. Efficiency and pion misidentificattion probability as a function of (left) momentum and (right) angle.

using electrons from Bhabha and $\gamma\gamma \to e^+e^-$ events, while the misidentification rate is measured using $\tau \to \nu_\tau 3\pi$ events. The results are 94.8 % efficiency and 1.2 % probability for misidentification as pion ("tight" selection), and 88.1 % efficiency and 0.3 % probability for misidentification ("very tight" selection), as shown as a function of momentum and angle in Fig. 24, thus TDR design expectations are reached.

7.2.2 Muons

The main subdetector for muon identification is the IFR, where hits are identified with extrapolations of charged tracks from the DCH. For the PID evaluation of the IFR hits one uses penetration depth, χ^2 of expected vs. actual hit position and number of stripes fired. The later is powerful for μ/π-discrimination, as the hadronic shower is much broader than the leptonic. The samples for measuring the muon PID efficiency are $\mu\mu ee$ and $\mu\mu\gamma$, while $\tau \to \nu_\tau 3\pi$, and $K_S^0 \to \pi^+\pi^-$ is used to determine the pion misidentification rate. The results are shown in Fig. 25 as a function of momentum, and the average muon efficiency is 75 % for momentum above 1 GeV associated with a pion misidentification rate of 2.6 % in the range $1\text{GeV} < p < 3\,\text{GeV}$.

7.2.3 Charged hadrons

Charged hadron identification is essentially a matter of kaon identification and K/π-separation, since this is the most demanding as well as important. A likelihood fit is done to the Čerenkov angle and the number of photons provided by the DIRC, and to the dE/dx information for the SVT and the DCH separately. Each detector contributes in a certain momentum range, and depending on the level of efficiency and purity required, various cuts are applied, e.g. a *tight* selection is defined as:

- SVT: $0.025 < p < 0.7$ GeV and dE/dx measurements > 2.
- DCH: $0.090 < p < 0.7$ GeV and dE/dx measurements > 10.
- DIRC: $p > 0.6$ GeV and number of expected photons > 0.

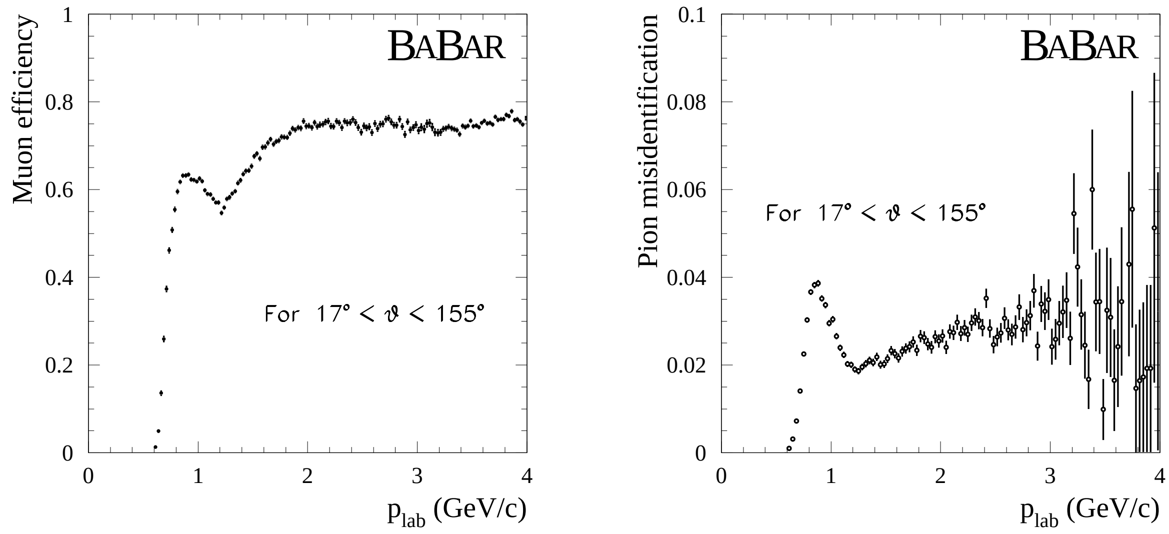


Figure 25: Muon identification capabilities. (left) Efficiency in muon identification as a function of momentum. The average is 75 % for momentum above 1 GeV. (right) Pion misidentification probability as a function of momentum. The average is 2.6 % in the range 1 GeV $< p <$ 3 GeV.

The $D^{*\pm} \to D^0\pi^\pm(D^0 \to K^\mp\pi^\pm)$ decay provides a fairly pure sample of kaons (and pions), where the purity is about 90 % due to combinatorial background (see Fig. 26 left). The sample is used to determine the average kaon efficiency, which is 90 %, and the average pion misidentification rate, which is 2.5 %, (see Fig. 26 right).

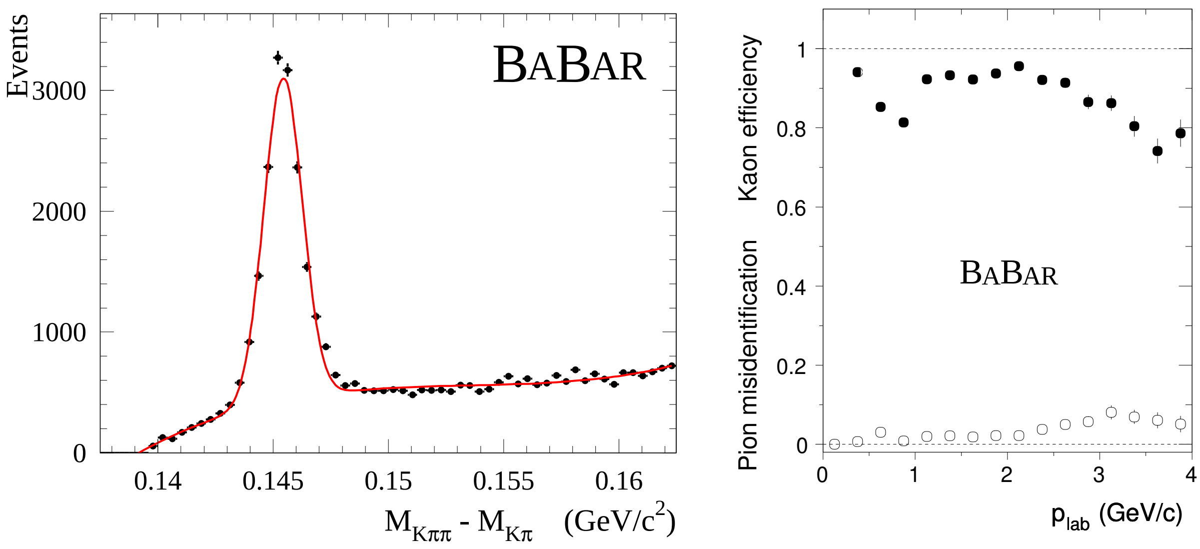


Figure 26: Kaon identification. (left) $D^{*\pm} \to D^0\pi^\pm(D^0 \to K^\mp\pi^\pm)$ signal for kaon sample of 90 % purity. (right) Kaon efficiency and pion misidentification rate as a function of momentum for a tight selection determined from the D^* event sample.

7.2.4 Photons detection and π^0 and η^0 reconstruction

Photons are detected by the EMC, where all energy deposits above 1 MeV are grouped in clusters. If a cluster has more than one local maximum, the cluster is divided into two, sharing the cluster energy deposit assuming EM shower shape. Clusters with no DCH track leading to them and with compatible shower shapes are accepted as photon candidates, if the energy is above 30 MeV.

Neutral pion and eta candidates are constructed from pairs of photon candidates. The invariant mass of the π^0 or η candidate can be calculated by assuming origin in the interaction point. The distribution of invariant masses from photon pairs can be seen in Fig. 27 for (upper figure) photons satisfying $E_\gamma > 30$ MeV, $E_{\gamma\gamma} > 300$ MeV and (lower figure) photons satisfying $E_\gamma > 100$ MeV, $E_{\gamma\gamma} >$ 1 GeV. The π^0 and η peaks are clearly seen, and the resolutions are 6.5 MeV and 16 MeV, respectively. In hadronic events the combinatorial background can be reduced by a factor of 5 at the cost of 25 % signal events. The π^0 and η results are in very good agreement with simulation.

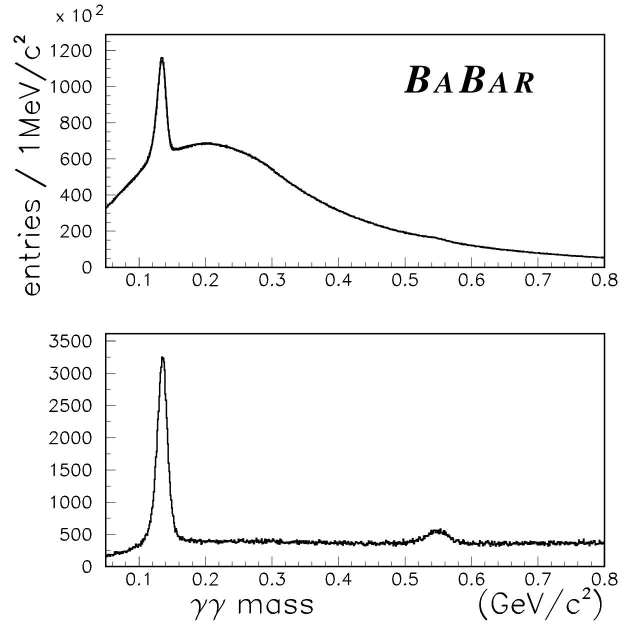


Figure 27: Invariant mass of photon pairs.

7.2.5 Neutral kaons

The two very different lifetimes of the neutral kaons make their detection equally different, as the K^0_S mostly decays before reaching the detector, while the K^0_L rarely decays before exiting (or hitting the IFR). Furthermore, the neutral and the charged decays of the K^0_S involve different detectors, and thus each require different efforts.

$\mathbf{K^0_S} \to \pi^+\pi^-$ Any pair of oppositely charged tracks from the same vertex are considered for reconstruction of the shortlived neutral kaon. In addition the K^0_S momentum is required to match the direction between the primary vertex and the K^0_S vertex within 45 mrad, and the transverse momentum with respect to this direction should be between 0.11 and 0.22 GeV to discriminate against combinatoric and $\Lambda^0 \to p\pi$ background. The $K^0_S \to \pi^+\pi^-$ spectrum can be fitted with a double Gaussian, and the central one containing 70 % has a width of 2.8MeV (see Fig. 28 left).

$\mathbf{K^0_S} \to \pi^0\pi^0$ Neutral pion candidates, which do not come from common photons, are combined into $K^0_S \to \pi^0\pi^0$ candidates, if their energy is above 800MeV and their invariant mass (assuming the interaction point as vertex) is within 200MeV of m_{K^0}. From the events primary vertex, the most probably K^0_S vertex is found and then used to recompute the invariant mass of the two neutral pions. A vertex resolution of 5 cm is considered realistic (obtained from simulation), and the mass resolution of better than 10 MeV can be seen in Fig. 28 right.

$\mathbf{K^0_L}$ The longlived neutral kaons are reconstructed from neutral clusters in the EMC and the IFR without associated tracks. In the EMC, γ and π^0 backgrounds are rejected by analysing the

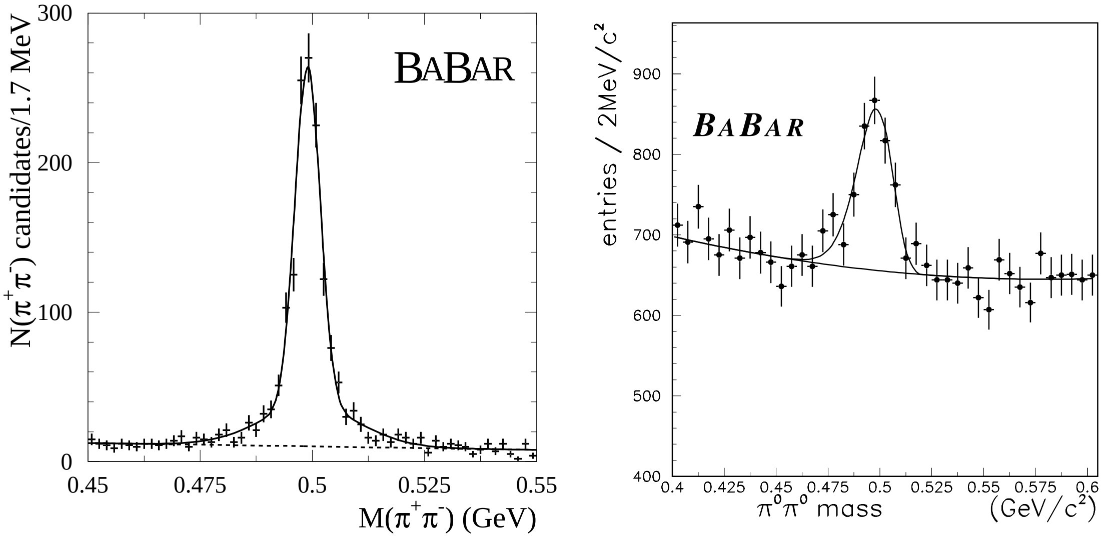


Figure 28: K_S^0 reconstruction for (left) $K_S^0 \to \pi^+\pi^-$ and (right) $K_S^0 \to \pi^0\pi^0$.

shower shape, which is efficient down to 200MeV. The IFR requires hits in two or more layers of resistive plate counters (RPC) to reduce detector noise, and rejects K_L^0 candidates with nearby tracks due to background from "splitoffs" from charged hadronic showers. The efficiency is a little better than 50 %, increasing with momentum.

7.3 General issues and prospects

Apart from the tracking and particle identification performance, the ability to monitor the luminosity and beam energy is important. In addition, the increasing luminosity requires foresight in order to upgrade succesfully.

7.3.1 Determination of luminosity

The luminosity is determined offline using Bhabha and muon pair events, and both methods are currently limited to 2 % precision by systematic errors. The results are in agreement, but the muon method should be improved to order 0.5 % with more MC statistics.

7.3.2 Definition of frames

Due to the boost of the $\Upsilon(4S)$ frame, the BaBar experiment operates with two systems of coordinates:

- **Laboratory (LAB)** frame, which defines the z-axis along the DCH axial wires, the y-axis upward, and the x-axis outward of the ring, such that it forms a righthanded coordinate system.

- **Center-of-Mass (CMS)** frame, which is defined such that the two beams have equal energy (z-axis in the electron beam direction, the y-axis upward, and the x-axis outward).

The LAB-frame z-axis is tilted slightly (19 mrad) in the horizontal plane with respect to the beam axis, thus the transformation from LAB to CMS involves a slight rotation aligning the two z-axis' and then a boost. Variations in the beam energy can be monitored to 1 % precision.

7.3.3 Planned upgrades

There is still more in this machine. [Uli Wienands, PEP-II Run Coordinator]

The successful running of PEP-II have induced belief, that the luminosity can be increased fivefold over the next four years, and by an additional factor of two by the end of 2005. Such massive upgrades have major impact on the BaBar detector, especially the trigger and the data acquisition. The L1 trigger will require hardware for using the z-coordinate from the DCH to reduce background not from interaction point, while the L3 trigger will demand more computing power.

The innermost subdetectors, especially the SVT and the DCH, will be affected by the higher radiation. A replacement of the SVT around 2002 is already planned, and though the DCH should last longer, radiation degrading from high luminosity may force a replacement. Increased luminosity will also cause inefficiency in the DIRC electronics, and the front cap of the EMC might eventually have to be replaced. Finally, the interaction region will have to be redesigned, when luminosity reaches 3×10^{34} cm^{-2}s^{-1}.

All these matters are being addressed continuously, as failure to see upcoming problems might result in long periods without collecting data. So far, the estimated prognoses of integrated luminosity have been always increasing as the PEP-II has shown its potential. The current belief is, that BaBar will have 55 fb^{-1} by the end of 2001, 100 fb^{-1} by the end of 2002 and up to 200 fb^{-1} in 2003. Most subdetectors have reached design goals, and BaBar is thus an excellent place to study B-physics.

Part III
Analysis

It is a capital mistake to theorize before one has data.

[Sir Arthur Conan Doyle, 1859-1930]

8 Outline of the sin(2β) measurement

The following short chapter gives an overview of the analysis associated with measuring $\sin(2\beta)$ in the golden $J/\psi K^0_S$ channel. The concepts involved in the analysis will be introduced along with some commonly used terminology.

8.1 The *CP* analysis

As was shown in section 3.5.1, $\sin(2\beta)$ can be measured in $b \to c\bar{c}s$ decays such as the $B^0 \to J/\psi K^0_S$ channel by observing an asymmetry in the lifetime distributions for B^0 and $\overline{B}^0$ respectively:

$$f_{\pm}(t) \quad \simeq \quad \tfrac{1}{4}\Gamma\, e^{-\Gamma|t|}[1 \pm \sin(2\beta)\sin(\Delta m t)]. \tag{83}$$

However, since $J/\psi K^0_S$ is a CP eigenstate, one can by definition not know the original B-flavor (i.e. if the $J/\psi K^0_S$ came from a B^0 or a $\overline{B}^0$), as the B^0 and the $\overline{B}^0$ decay into this final state with equal probability. To obtain this crucial information, the flavor of the other B-meson in the $\Upsilon(4S)$ decay has to be determined along with the time between the two decays, since the remaining B-meson continues to oscillate, thus making the correlation between the two flavors time dependent (see section 3.3). As only a small fraction of the B-decays are fully reconstructed, one has to determine these two quantities from incomplete information[29], which of course induces errors.

This situation is schematically illustrated in Fig. 29. From an asymmetric e^+e^- collision at the $\Upsilon(4S)$ resonance, one creates a boosted $B^0\overline{B}^0$ pair. One of these decays into $J/\psi K^0_S$, and from the remaining tracks and clusters both the flavor and the decay vertex of the other B-meson has to be determined. The drawing is simplified, as generic B-decays on average have 5.5 charged and an equivalent number of neutral particles in the final state [16].

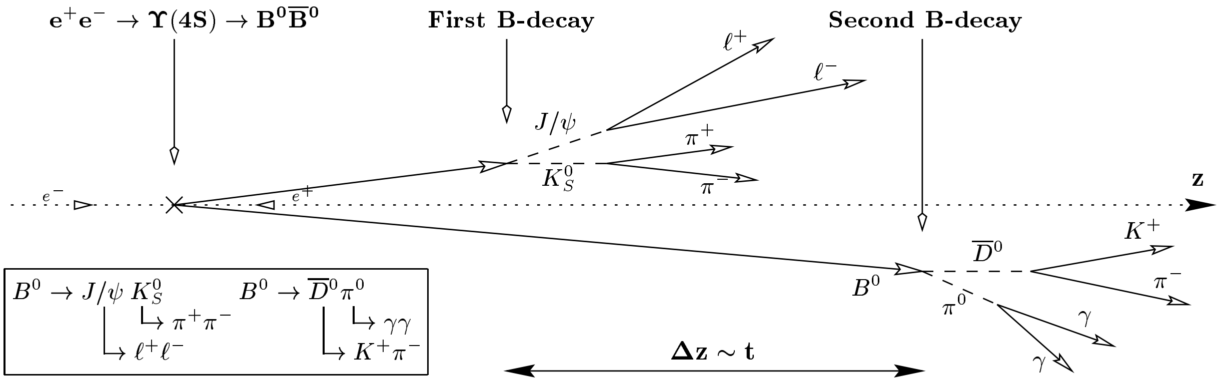


Figure 29: $B^0\overline{B}^0$ decay with $J/\psi K^0_S$ final state at asymmetric B-factory. To fit the CP asymmetry in the lifetime distributions, both the flavor and the decay vertex of the other (tagging side) B-meson has to be determined. Note that the drawing is schematic and not to scale.

8.2 Flavor tagging

The ability to *tag*, i.e. determine the flavor of the non-reconstructed B-meson (commonly referred to as the tagging side B-meson) plays a central role in the $\sin(2\beta)$ measurement. Due to incomplete informations, the tagging efficiency is far from perfect and a certain fraction of the

[29] Mostly due to particles outside the acceptance, but also particle misidentification and accelerator noise.

B-mesons, w_{tag} (called the wrong tag fraction), will wrongly be tagged as having the opposite flavor. This dilutes the CP-asymmetry Eq. (83) as follows:

$$F_{\pm}(t) \quad \equiv \quad (1-w_{\text{tag}})f_{\pm}(t) + w_{\text{tag}}f_{\mp}(t) \;\; = \;\; \tfrac{1}{4}\Gamma\, e^{-\Gamma|t|}[1 \pm D\sin(2\beta)\sin(\Delta m t)], \tag{84}$$

where $D \equiv 1 - 2w_{\text{tag}} \in [0, 1]$. Thus, in order to obtain $\sin(2\beta)$, a high but also well determined value of D is required, as its uncertainty propagates directly to the $\sin(2\beta)$ result.

The value of D is determination by considering decays where the flavor of one B-meson is known with certainty. This is the case for fully reconstructed flavor eigenstate decays, either hadronic or semileptonic, such as the $B^0 \to D^{*-}\ell\nu_\ell$ decay sketched in Fig. 30.

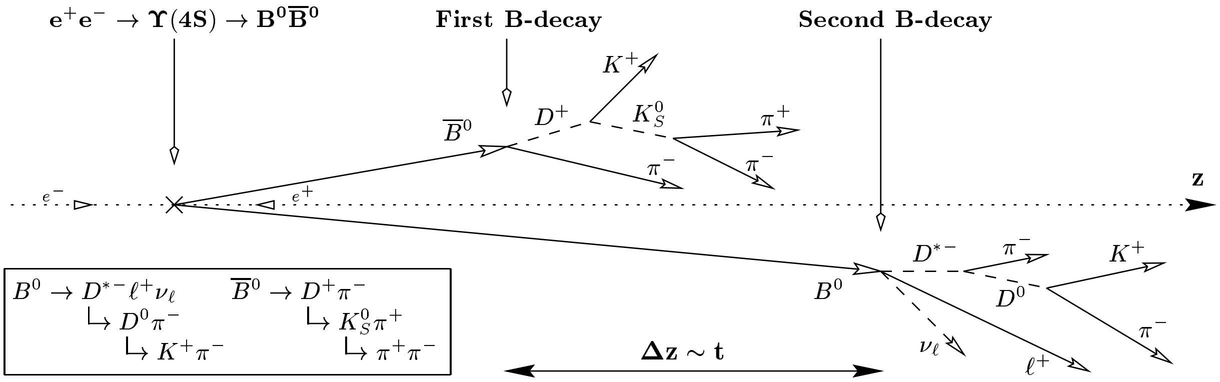


Figure 30: $B^0\overline{B}^0$ decay with flavor tagging final state $D^{*-}\ell^+\nu_\ell$. Using the knowledge of the flavor of one B-meson, and measuring the time difference between the two decays, the theoretical expectation of the other B-mesons flavor can be compared with the output of a tagging algorithm, and the efficiency of the tagging algorithm then be determined.

The mixing between two B-mesons originating from the $\Upsilon(4S)$ resonance (see section 3.3) is to a very good approximation (which will be used in the following) governed by Eq. (45). Imperfect tagging, parametrized by D, enters in the same way as for the CP-asymmetry, yielding:

$$H_{\pm}(t) \quad \equiv \quad \tfrac{1}{4}\Gamma\, e^{-\Gamma|t|}[1 \pm D\cos(\Delta m t)], \tag{85}$$

Thus, one considers events where one B-meson is reconstructed and the flavor of it known, and then observes the result of a tagging algorithm on the remaining tracks from the tagging side B-meson. The difference between the theoretical mixing rate and the one actually observed is exactly due to the wrong tag fraction, thus *from this mixing distribution one can extract D.*

The measurement of the mixing rate, Δm, and the so-called dilution, D, corresponds exactly to fitting the frequency and the amplitude respectively of the time dependent mixing rate $(N_{\text{unmixed}} - N_{\text{mixed}})/(N_{\text{unmixed}} + N_{\text{mixed}})$ (see Fig. 43). This is the reason why this analysis is commonly referred to as the TagMix analysis.

8.3 Time resolution

The time information is obtained from measuring the distance between the two decay vertices. While the vertices of the reconstructed B-mesons are fairly well determined, the vertices of the tagging side B-mesons cause difficulties. Due to particles going undetected and backgrounds, but also due to missing reconstruction of particles with a non-negligible lifetime, such as the D-meson, the tagging vertex is both smeared and biased. In order to perform a correct fit for a CP-asymmetry, this timing resolution has to be well known and correctly included (see section 10.2).

9 Mixing and *CP* data samples

The measurement of a CP asymmetry is based on a data sample of CP events (CP sample), and a data sample of tagging events (mixing sample) needed for measuring the wrong tag fractions. The selection of these data samples are described below, and the various results with respect to size, purity, etc. are discussed. Data samples for calibration of the detector are not included, since this is mostly done online, nor are control samples and samples for cross-checks discussed, as they are very similar to the ones described below.

The samples were obtained from $9.0\,\mathrm{fb}^{-1}$ on-resonance data and $0.8\,\mathrm{fb}^{-1}$ off-resonance data. Given a cross section of $\sigma_{b\bar{b}} = 1.05\,\mathrm{nb}$ for $b\bar{b}$-production at the $\Upsilon(4S)$ resonance, this corresponds to an estimated $(10.1 \pm 0.4) \times 10^6$ $B\overline{B}$ events, of which half are $B^0\overline{B}^0$ events.

9.1 General cuts for $\mathrm{B\overline{B}}$ selection

Common for all B-analyses is a cut on so-called shape variables, which are computed using charged tracks in the CM frame. The two most important variables for continuum discrimination and background fighting at B-factories are the thrust and the normalized Fox-Wolfram moment.

9.1.1 Thrust

The thrust is a measure of the degree of direction in an event. The thrust axis, $\mathbf{T}$, is defined by the axis, which maximizes the sum of the longitudinal momenta with respect to this axis. The thrust of the event is then defined by

$$T = \frac{\sum_i |\mathbf{T} \cdot \mathbf{p_i}|}{\sum_i |\mathbf{p_i}|}. \tag{86}$$

The range of T is $[0.5, 1]$, where $T \sim 1$ corresponds to a highly directional (jet-like) event, while $T \sim 0.5$ is an isotropic event. Continuum events tend to be jet-like, thus having a high thrust, while $B\overline{B}$ events are more isotropic.

9.1.2 Fox-Wolfram moments

Fox-Wolfram moments [52] are defined as momentum weighted sums of Legendre polynomials from the cosine of the angle between pairs of charged tracks, that is

$$H_l = \sum_{i,j} \frac{|\mathbf{p}_i| \cdot |\mathbf{p}_j|}{E_{\mathrm{vis}}^2} P_l(\cos\theta_{ij}), \tag{87}$$

where P_l are the Legendre polynomials, θ_{ij} the angle between particles i and j, and E_{vis} is the visible energy of the event. For 2-jet events $H_{\mathrm{even}} \sim 1$ and $H_{\mathrm{odd}} \sim 0$, and one uses the ratio $R_2 \equiv H_2/H_0$ as discriminating variable.

9.1.3 Kinematic variables

Fully reconstructed B candidates can be characterized by two almost uncorrelated kinematic variables. One is the Lorentz invariant energy difference, ΔE, between the reconstructed B-meson energy, E_B, and the beam energy, E_{beam}. The other is the beam-energy substituted mass, m_{ES}, that is the mass calculated from the beam energy, $\hat{E}_{\mathrm{beam}}$, and a kinematic fit of the B-candidate momentum, $\hat{\mathbf{p}}_B$, both evaluated in the CM frame:

$$\Delta E \equiv E_B - E_{\mathrm{beam}} \qquad m_{\mathrm{ES}} \equiv \sqrt{(E_{\mathrm{beam}}^{\mathrm{CM}})^2 - (\mathbf{p}_B^{\mathrm{CM}})^2}. \tag{88}$$

Thus, fully reconstructed candidates can be plotted in the $(\Delta E, m_{\mathrm{ES}})$-plane, and the signal should exhibit a peak, around which suitable cuts can be made.

9.1.4 General requirements

The general requirements for $B\overline{B}$ selection are designed to have a high efficiency with a low systematic uncertainty. For charged tracks the fiducial volume is defined to be 410 mrad from the beam axis in the front end, to 602 mrad from the beam axis in the back end, corresponding to $0.92 > \cos\theta_{lab} > -0.82$. Due to the missing back end cap of the EMC, the coverage for neutral clusters is limited to 732 mrad from the beam axis in the backward direction.
The selection criteria for events are based on tracks and clusters within the above defined fiducial volume, and they are:

- Minimum four tracks, at least three of which have DCH information.
- Primary vertex within 5 mm of the beam spot in x and y direction.
- The normalized second Fox-Wolfram moment, R_2, is less than 0.7.
- Total energy greater than 5 GeV.

With this selection the $B\overline{B}$ efficiency is 96 %, while backgrounds from beam gas, two photon and tau pairs are around 2 %. For the hadronic and semileptonic events one requires in addition, that R_2 is less than 0.5 and 0.45 respectively, while the thrust requirements varies with decay mode.

9.2 Mixing data samples

As discussed in the previous chapter, the wrong tag fraction is measured using $B\overline{B}$ events, where one B-meson is reconstructed in a flavor eigenstate. The determination of the resolution function (see section 10.2.2) also uses fully reconstructed events, and a large mixing sample is therefore desirable.

The typical branching fractions for B-decays are small, so in order to achieve reasonable statistics a variety of decay channels have to be considered. At the present stage, reconstructed events include[30] $D^{(*)-}\pi^+$, $D^{(*)-}\rho^+$, $D^{(*)-}a_1^+$, and the semileptonic mode $D^{*-}\ell^+\nu_\ell$ ($\ell = e, \mu$) where the neutrino is not reconstructed. For this reason, reconstruction of the B-mass and subsequent background rejection from it is not possible in the $D^{*-}\ell^+\nu_\ell$ channel, and consequently the cuts are in general tighter.

For the hadronic events, the only background is combinatorial, that is from background events, which by chance have particles with an invariant mass around that of the D^- or the D^{*-}. The size of this background can be determined rather precisely from considering the mass spectrum of D^- and D^{*-}-candidates (see Fig. 31).

As the semileptonic events can not be entirely reconstructed, and therefore not constrained as well as the hadronic ones, this channel suffers from several backgrounds. The reconstruction of D^{*-}-candidates is similar, and so the combinatorial background and that of the hadronic sample are alike. But in addition, the correctly reconstructed D^{*-}-candidates may be combined with lepton candidates, that are either misidentified or not from the same process, since there is no invariant mass requirement. Finally, semileptonic B-decays may have additional particles (mostly pions) in the final state, which go undetected, thus allowing for background from charged B-mesons. In the following the selection of hadronic and semileptonic samples and their respective backgrounds will be described.

[30] Recall the convention that charge conjugate modes are implicit, e.g. $B^0 \to D^-\pi^+$ includes $\overline{B}^0 \to D^+\pi^-$.

9.2.1 Selection of hadronic events

The D^{*-} candidates are constructed in the $\overline{D}^0\pi^-$ channel, where the $\overline{D}^0$ candidates are again constructed from decays to $K^+\pi^-$, $K^+\pi^-\pi^0$, $K^+\pi^-\pi^+\pi^-$ and $K^0_S\pi^+\pi^-$. The D^- candidates are constructed from decays to $K^+\pi^-\pi^-$ and $K^0_S\pi^-$.

All pions and kaons in the reconstruction are required to have a momentum greater than 200 MeV, with the exception of the pions from the a_1^+, where the limit is lowered to 150 MeV, and the soft pion from the D^* decay, which has a special status in the analysis. The $\overline{D}^0$ and D^- candidates are required to lie within 3σ of their nominal masses, have a momentum greater than 1.3 GeV in the $\Upsilon(4S)$ CM frame, and a χ^2 probability in the vertex fit greater than 0.1 %. In the $\overline{D}^0 \to K^+\pi^-\pi^0$ channel only the resonant $K^+\rho^-$ channel is reconstructed. Here, the $\pi^-\pi^0$ mass is required to lie within 150 MeV of the ρ mass, and the angle between the π^- and the $\overline{D}^0$ in the ρ CM, $\theta^{[\rho]}_{D\pi}$, should satisfy that $|\cos\theta^{[\rho]}_{D\pi}| > 0.4$.

The D^{*-} candidates are formed by combining $\overline{D}^0$ candidates with pions. The pion for the D^{*-} candidates have very low momentum due to the small phase space available (41 MeV), and are therefore mostly reconstructed with the SVT alone. In this analysis, the momentum of this "soft" pion is conservatively required to be above 70 MeV (see Fig. 23 right), but that requirement is under debate. The D^{*-} candidates are required to have a $\Delta m \equiv m(\overline{D}^0\pi-) - m(\overline{D}^0)$ within 2.4 MeV of its nominal value, 145.4 MeV, (within 3.3 MeV for the $\overline{D}^0 \to K^+\pi^-\pi^0$ mode). This is illustrated in Fig. 31, where a plot of the Δm distribution for the D^{*-} candidates is shown. Regions well outside the peak (the "sideband") are used as control samples for characterizing the background.

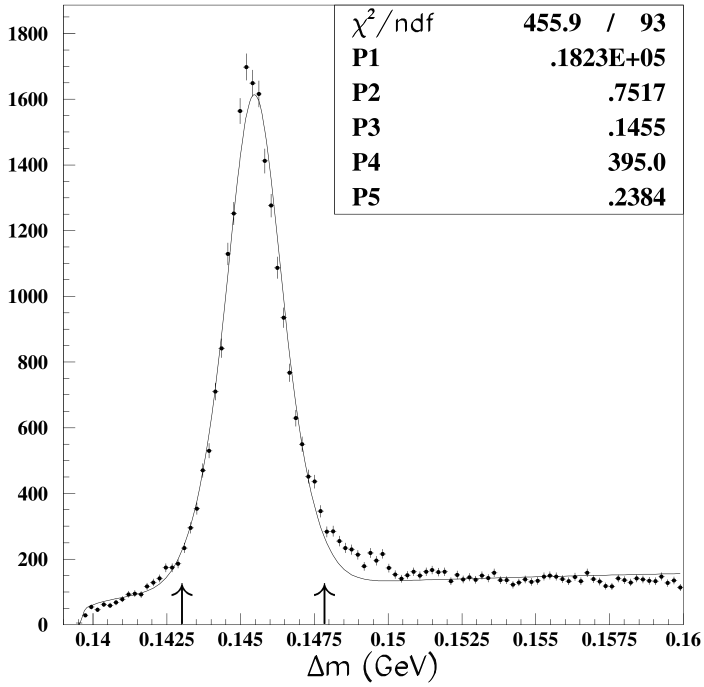


Figure 31: Mass distribution of D^{*-} candidates. The peak (within the two arrows) constitute the D^{*-} candidates, while the "sideband" regions are used for obtaining information about the background. The kink in the data to the right of the peak is mainly from the $K^+\pi^-\pi^0$ channel, and is thought to be due to imperfect photon reconstruction in the EMC.

B^0 candidates are formed by combining $D^{(*)-}$ candidates with π^+, ρ^+ and a_1^+ candidates. The π^0 from the ρ^+ decay is required to have momentum greater than 300 MeV. The a_1^+ candidates are formed by combining three charged pions with invariant mass between 1.0 and 1.6 GeV, and a χ^2 vertex fit probability greater than 0.1 %. Defining the thrust angle, θ_{th}, to be the angle between the thrust axis of the particles which form the B^0 candidate, and the thrust axis of the remaining tracks and clusters, one can further reject especially $c\bar{c}$ background, as the more jet-like continuum events tend to have a small thrust angle [53].

The signal region in the $(\Delta E, m_{\mathrm{ES}})$ plane is taken to be $5.270 < m_{\mathrm{ES}} < 5.290$ and $|\Delta E| < 3\sigma_{\Delta E}$, where the later cut varies with decay mode. If several B^0 candidates are found, the one with the smallest ΔE is chosen.
The data samples obtained are listed below. The background is entirely combinatorial, and its size can be determined from fitting the $\Delta m(D^*)$ distribution, as done in Fig. 31.

Decay mode	# of candidates	S/(S+B) [%]
$B^0 \to D^{*-}\pi^+$	622 ± 27	90
$B^0 \to D^{*-}\rho^+$	419 ± 25	84
$B^0 \to D^{*-}a_1^+$	239 ± 19	79
$B^0 \to D^-\pi^+$	630 ± 26	90
$B^0 \to D^-\rho^+$	315 ± 20	84
$B^0 \to D^-a_1^+$	225 ± 20	74
Total	2383 ± 57	85

Table 7: Hadronic B^0 candidate yields and purities for each reconstruction mode.

One may be surprised by the low number of hadronic events relative to the entire data sample. However, the BaBar detector is not hermetic, and the chance of a particle going undetected is estimated to be around 15 %. With a large number of particles in the final state, which all have to be within the acceptance, the efficiency drops. In addition none of the channels have a branching much above 1 %.

9.2.2 Selection of $\mathrm{D}^*\ell\nu_\ell$ events

The $D^{*-}\ell^+\nu_\ell$ decay channel has the largest branching ratio (4.60 ± 0.27 % [37]) and at the same time the charge of both the lepton and the D-meson contains tagging information, thus the channel is well suited for the TagMix analysis.

As in the hadronic case, D^{*-} candidates are reconstructed through their decay to $\overline{D}^0\pi^-$. The $\overline{D}^0$ is reconstructed in the three channels: $K^+\pi^-$, $K^+\pi^-\pi^0$ and $K^+\pi^-\pi^+\pi^-$ with the requirement of an invariant mass within 2.5σ of the $\overline{D}^0$ mass and a common vertex probability greater than 1 %. The additional momentum requirements are 200 MeV for the π^0 in the $K^+\pi^-\pi^0$ mode, and 200 and 150 MeV respectively for the kaons and pions in the $K^+\pi^-\pi^+\pi^-$ mode. With the exception of the $K^+\pi^-$ mode, consistant PID is required.

D^{*-} candidates are formed by combining $\overline{D}^0$ candidates with all charged tracks having a transverse momentum greater than 50 MeV and opposite charge of the kaon from the $\overline{D}^0$ reconstruction. The mass difference $m_{D^{*-}} - m_{\overline{D}^0}$ is required to lie within 2.5σ of the nominal value.

Finally, the D^{*-} candidates are combined with opposite charge lepton candidates, which are identified as such, have momentum greater than 1.2 GeV and a topological vertex fit with a χ^2 probability greater than 1 %. The lepton momentum requirement is very important, as leptons from other processes such as charm decays constitute a large background but have relatively lower momenta. In the B^0 rest frame, the D^{*-} and the lepton tend to be back to back[31], and consequently it is required that the angle between the two is greater than $\pi/2$ ($\cos\theta_{D^{*-},\ell} < 0$)

[31] The shape of the matrix element for $B^0 \to D^{*-}\ell^+\nu_\ell$ decay can be found in [54, p. 85], and is later used.

in the CM frame. In addition it is required that the event is consistent with a missing zero-mass particle:

$$0 = p_\nu^2 = (p_B - p_{(D^*\ell)})^2 \quad \Rightarrow \quad \cos\theta_{B,(D^*\ell)} = \frac{m_B^2 + m_{(D^*\ell)}^2 - 2E_B E_{(D^*\ell)}}{2|\mathbf{p}_B||\mathbf{p}_{(D^*\ell)}|}, \tag{89}$$

where the angle $\theta_{B,D^*\ell}$ is in the CM frame, and $D^*\ell$ refers to the combination of the two particles. Taking detector resolution into account, the required consistency is that $|\cos\theta_{B,D^*\ell}| < 1.1$.

Using these cuts, a sample of 7517 ± 104 $B^0 \to D^{*-}\ell^+\nu_\ell$ events is obtained. Since the neutrino is not detected, the semileptonic events cannot be plotted in the $(\Delta E, m_{\mathrm{EC}})$ plane, and so compared to the fully reconstructed hadronic modes, the semileptonic sample is less pure, due to a variety of backgrounds:

Combinatorial background

Combinatorial background is due to falsely reconstructed D^* candidates, and its fraction of the signal is estimated from the distribution of $m_{D^{*-}} - m_{\overline{D}^0}$, fitted with a Gaussian distribution and a background function, exactly as for the hadronic sample.

Wrong-lepton background

Lepton candidates may wrongly be combined with a real D^*, and there are four potential sources:

1: One background is from *misidentified leptons* (pions identified as leptons, so-called fake leptons), which is evaluated by combining D^* candidates with tracks failing the very loose lepton criteria, and weighting these with the lepton misidentification probability.

2: Real lepton from the other B-meson may wrongly be combined with D^* candidates. The background from these *uncorrelated leptons* used to be estimated by a method developed by CLEO [55], assuming that no signal failed the requirement $\cos\theta_{D^{*-},\ell} < 0$, and thus that the background, which is uncorrelated and therefore has a flat $\cos\theta_{D^{*-},\ell}$ distribution, could be estimated from the remaining region. However, this turned out not to be the case, and so the background is now (post-Osaka) estimated from Monte Carlo simulation.

3: $B^0 \to D^*DX$ decays, where the D-meson decays semileptonically, constitute *cascade lepton* background. The cut on lepton momentum rejects most of this background, and the remaining small fraction is estimated by Monte Carlo.

4: Finally, $c\bar{c}$ events can produce a D^* and lepton pair back-to-back. The size of this background is estimated from off resonance data, where the same cuts have been applied.

B$^+$ background

A non-negligible fraction of the semileptonic B-decays have additional particles in the final state, either through D^{**} resonances or non-resonant. Such $B \to D^*(n\pi)\ell\nu_\ell$ decays may originate from both neutral and charged B-mesons. The neutral decays are in this analysis considered signal, but since the charged B-mesons do not mix and have different mistag rates, they constitute a serious background to the measurement of both the dilution, D, and mixing rate, Δm.

The size of this background is difficult to estimate, as it is hard to separate from the signal. The charged background contribution is treated as a systematic error, and it is currently the largest single source of such. The size of the background can be estimated in two ways; one standard and one more involved:

1: The branching fraction, $Br(B^+ \to D^{*-}(n\pi)\ell^+\nu_\ell) = 1.25 \pm 0.16$ % has been measured by LEP [56], [57]. Using Monte Carlo simulation, the efficiency of the event selection, ϵ, for several D^{**} resonances and non-resonant states has been estimated and conservatively found to be $4.25 \pm 3.0\,\%$. This leads to an estimated charged background fraction of:

$$f(B^+) \;=\; \frac{Br(B^+) \times \epsilon(B^+)}{Br(B^0) \times \epsilon(B^0) + Br(B^+) \times \epsilon(B^+)} = 7.1 \pm 5.0\,\%. \tag{90}$$

2: The $D^{*-}(n\pi)$ part of the final state is expected to originate mainly from three different sources: non-resonant $D^*\pi$, narrow D^{**} $(D(2460)^{**}_{\Gamma=20\,MeV})$ and broad D^{**} $(D(2660)^{**}_{\Gamma=200\,MeV})$, where the mass and width of the broad D^{**} is an educated guess, which is not essential. For the determination and also the reduction of these so-called D^{**} backgrounds, one considers the variable $\cos(\theta_{B,D^*\ell})$. The distribution of this variable has been determined with a toy Monte Carlo for D^* and D^{**} events. The toy MC consists of a generator of kinematically allowed decays with a flat q^2 distribution, onto which the $B^0 \to D^*\ell\nu$ matrix element [54, page 85] (heavy quark limit assumed, which is compatible with existing data [54, pages 128-133]) is imposed, both for the D^* and D^{**} decays. The tracks are then smeared with Gaussian distributions, using $\sigma_{p_t}/p_t = 0.5\,\%$ and $\sigma_{\theta,\phi} = 2\,\mathrm{mrad}$ in accordance with detector performance. Finally the detector acceptance ($-0.88 < \cos(\theta_{lab}) < 0.94$, $|\mathbf{p}| > 75$ MeV) and the analysis cuts ($|\mathbf{p}_\ell|^{[\Upsilon(4S)]} > 1.2\,GeV$, $\cos\theta_{D^*,\ell} < 0$) are applied. The final acceptance including the $\cos\theta_{B,(D^{*(*)}\ell)}$ cut is 22% for D^*, and 6.4%, 7.8%, and 7.0% for the three types of D^{**} background. The $\cos\theta_{B,(D^{*(*)}\ell)}$ cut is almost 100% efficient for the signal, whereas it leaves only 52%, 50%, and 42% of the D^{**} backgrounds respectively. The toy MC reproduces the $\cos(\theta_{B,(D^*\ell)})$ distribution for $D^*\ell\nu$ events nicely, which can be seen in Fig. 32.

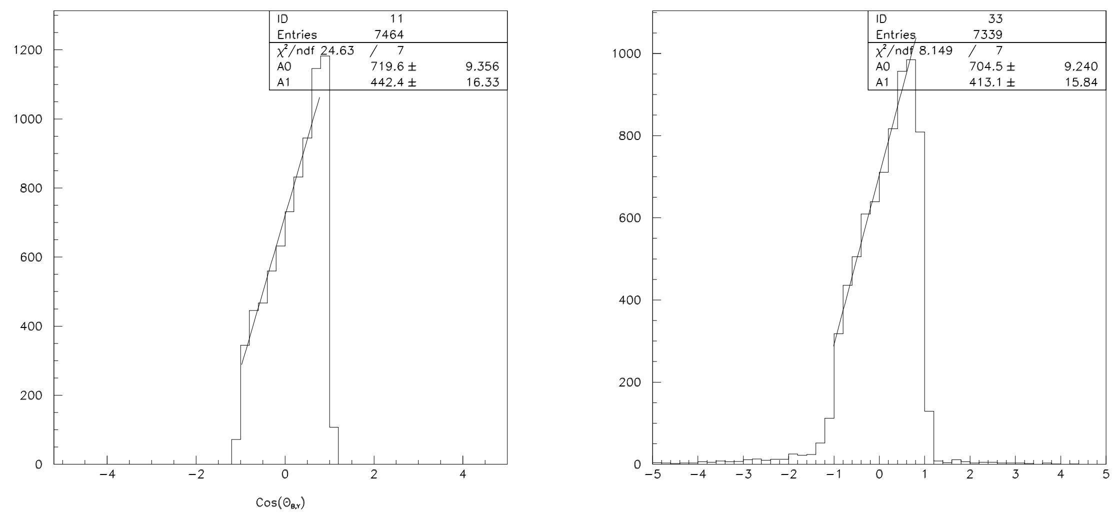


Figure 32: The distribution of $\cos\theta_{B,(D^{*(*)}\ell)}$ for toy MC (left) and BaBar MC (right). As the fit indicates, the toy MC reproduces the full BaBar MC distribution.

For the D^{**} backgrounds the distribution of $\cos\theta_{B,(D^{*(*)}\ell)}$ has a long tail in the negative direction, due to the missing pion. The shape of the three distributions are very similar (see Fig. 8), and it turns out that they can all be fitted with the following function (called the Troels-function [58]):

$$f_{fit}(x) = N\,\frac{c\,b^{\frac{a+1}{c}}}{\Gamma(\frac{a+1}{c})}\,(1-x)^a\,e^{-b(1-x)^c} \tag{91}$$

This has been done for each type of D^{**} background, and the results can be seen in Fig. 33.

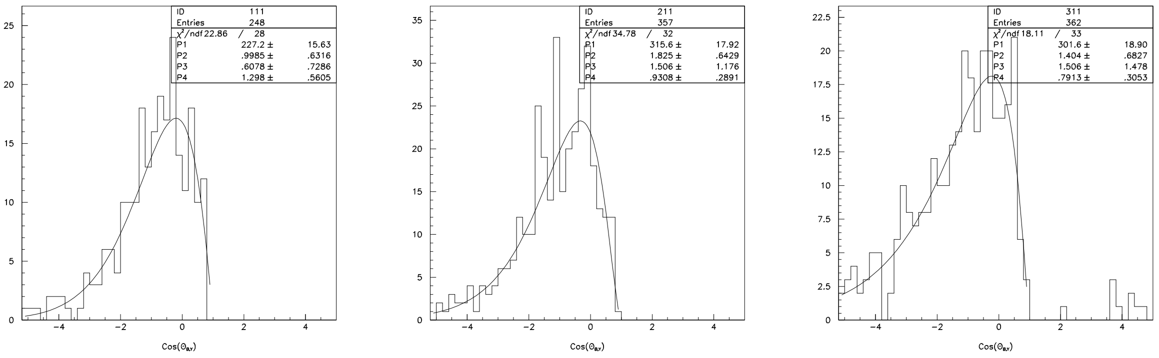


Figure 33: Distribution of $\cos(\theta_{B,(D^*\ell)})$ for the D^{**} background. The three cases of D^{**} background; (left) non-resonant $D^*\pi$, (middle) narrow D^{**} and (right) broad D^{**}. The function f_{fit} describes their essential features well.

A more realistic scenario is an equal mixture of the three [59, page 10], which is equally well fitted by the function f_{fit}, and the resulting values of parameters for the three separate D^{**} along with the equal mixture is quoted in Table 8.

Type of background	a	b	c	$\chi^2/dof.$
$D^*\pi$ non-res.	1.0 ± 0.6	0.61 ± 0.73	1.3 ± 0.6	$23/28$
D^{**} narrow	1.8 ± 0.6	1.5 ± 1.2	0.93 ± 0.29	$35/32$
D^{**} broad	1.4 ± 0.7	1.5 ± 1.5	0.79 ± 0.31	$18/33$
Mixture	1.7 ± 0.3	1.6 ± 0.6	0.88 ± 0.14	$53/37$

Table 8: The fitted parameters for the different D^{**} backgrounds.

In the data, one subtracts the contributions from the combinatorial and wrong-lepton backgrounds in order to obtain a sample containing the $D^*\ell\nu_\ell$ and $D^*(n\pi)\ell\nu_\ell$ contributions. In order to fit this data, the D^* and the D^{**} toy Monte Carlo distributions are fitted separately, with a linear fit and the function f_{fit} respectively. The shape of these fits are then fixed as well as the overall normalization of both D^* and D^{**}, and a global fit to the data is then carried out, leaving only the relative fraction, r, as a free parameter.

$$f_{global}(r) = N_{\text{events}}(rf_{fit}(x) + (1-r)(1+0.63x)). \tag{92}$$

To check that this global fit works, it has been applied to the BaBar MC distribution of $\cos\theta_{B,(D^*\ell)}$ for both D^* and D^{**}, where the relative fraction is known. The result of this can be seen along with the fit to available data (April 2000) in Fig. 34 and the corresponding numbers found in Table 9.

	N_{D^*}	$N_{D^{**}}$	$\chi^2/dof.$
BaBar Monte Carlo	1478	967	—
Fit to BaBar MC	1524 ± 38	920 ± 38	$85/48$
Fit to available data	1897 ± 62	355 ± 62	$48/28$

Table 9: The true and fitted normalizations for Monte Carlo events and the fit to available data (April 2000). The BaBar MC have more D^{**} background than data, but it suffices for testing the fit.

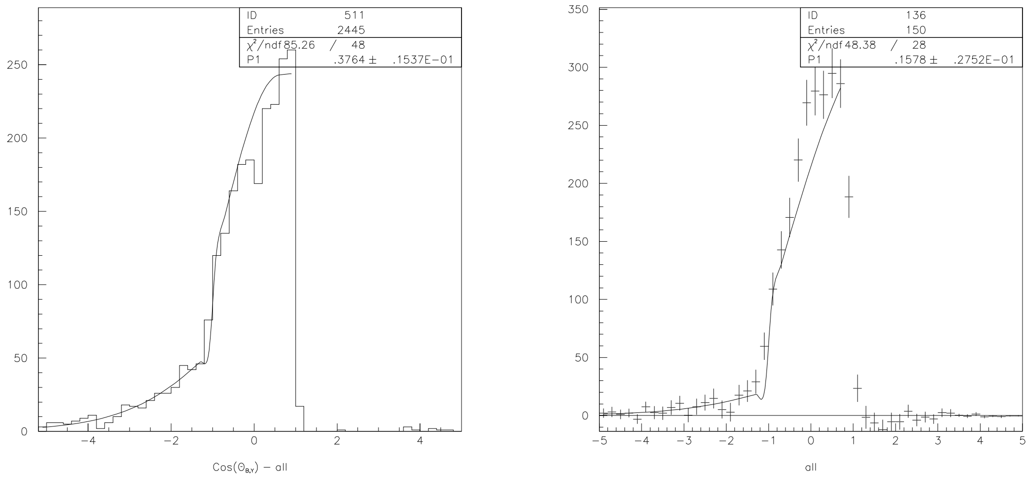


Figure 34: The fitted normalizations for Monte Carlo and data. The left figure shows the global fit to BaBar MC, where the relative fraction of D^* and D^{**} events is known, and the fit reproduces these numbers well. The global fit is then performed on the data, which is shown in the right figure.

Assuming that due to isospin symmetry $\frac{2}{3}$ of the D^{**}-mesons are from charged B-mesons, and applying the $\cos\theta_{B,(D^*\ell)}$ cut, the fit to the data implies a charged background fraction of:

$$f(B^+) \quad = \quad 5.1 \pm 0.7\,\%. \tag{93}$$

The systematic uncertainties are mainly the relative fractions of D^{**} background types and the endpoint of the fit, which approximately sums to 2.4 %.

The result is consistent with the estimate from the LEP measurement and detector simulation, and the error has been reduced. The fit is far from perfect (problems close to $\cos\theta_{B,(D^*\ell)} = 1$), and further work is required to reduce the large systematic error. However, the method is the first attempt to directly estimate the D^{**} background, and has served for the analysis of [18]. It was used as consistency check for the BaBar conference submissions Summer 2000.

Background:	Fraction of sample:
Combinatorial	0.139 ± 0.009
Fake leptons	0.033 ± 0.007
Uncorrelated leptons	0.030 ± 0.013
Continuum events ($c\bar{c}$)	0.029 ± 0.014
Charged B-mesons	0.054 ± 0.031
Signal	0.715 ± 0.038

Table 10: Signal and background fractions for the semileptonic sample.

The obtained purity of the $D^{*-}\ell^+\nu_\ell$ sample is 71.5 ± 3.8, which compared to the hadronic sample means a factor of two in background. It is, therefore, reassuring when half of the $D^{*-}\ell^+\nu_\ell$ background is found to be combinatorial, as the two channels suffer equally from this. The next largest background is the one from charged B-mesons, while the remaining ones are fairly small. The current (January 2001) background estimates can be seen in Fig. 46. However, while the combinatorial background can be precisely estimated, the charged background have large uncertainties, which is the reason for the large systematic uncertainty from the latter.

9.3 *CP* data samples

The $\sin(2\beta)$ measurement is based on the reconstruction of CP eigenstates with a charmonium resonance in, and the channels used for this measurement are currently $B^0 \to J/\psi_{(\ell^+\ell^-)} K^0_{S(\pi^+\pi^-,\pi^0\pi^0)}$ and $B^0 \to \psi(2S)_{(\ell^+\ell^-, J/\psi\ell^+\ell^-)} K^0_{S(\pi^+\pi^-)}$. The description of the control sample selection for the CP analysis has been omitted, as it is very similar.

9.3.1 Selection of $\mathbf{B^0 \to J/\psi K^0_S}$ and $\mathbf{B^0 \to \psi(2S) K^0_S}$ events

The CP selection is based on the general requirements for $B\overline{B}$ events described in the beginning of this section. The optimization of the additional selection criteria for the $J/\psi K^0_S$ and $\psi(2S)K^0_S$ events is done by maximizing $S/\sqrt{S+B}$, which is inversely proportional to the square root of the statistical uncertainty, where S is the signal determined from MC, and B is the background determined from continuum data.

For the J/ψ or $\psi(2S) \to \ell^+\ell^-$ candidates, at least one of the leptons are required to be positively identified as such. For muons the other is moreever required to be consistent with a minimum ionizing particle, if within acceptance. If both electrons are inside the EMC acceptance, an algorithm recovering bremsstrahlung from the electrons is used. The invariant mass requirements are $2.95(e)/3.06(\mu)$ GeV $< m_{\ell^+\ell^-} <$ 3.14 GeV for J/ψ candidates and $3.44(e)/3.64(\mu)$ GeV $< m_{\ell^+\ell^-} <$ 3.74 GeV for $\psi(2S)$ candidates in leptonic modes. For the $\psi(2S) \to J/\psi\pi^+\pi^-$ mode the invariant mass is required to be between 3.0 GeV and 4.1 GeV, with a mass difference between the $\psi(2S)$ and J/ψ candidates of no more than 15 MeV from the nominal value.

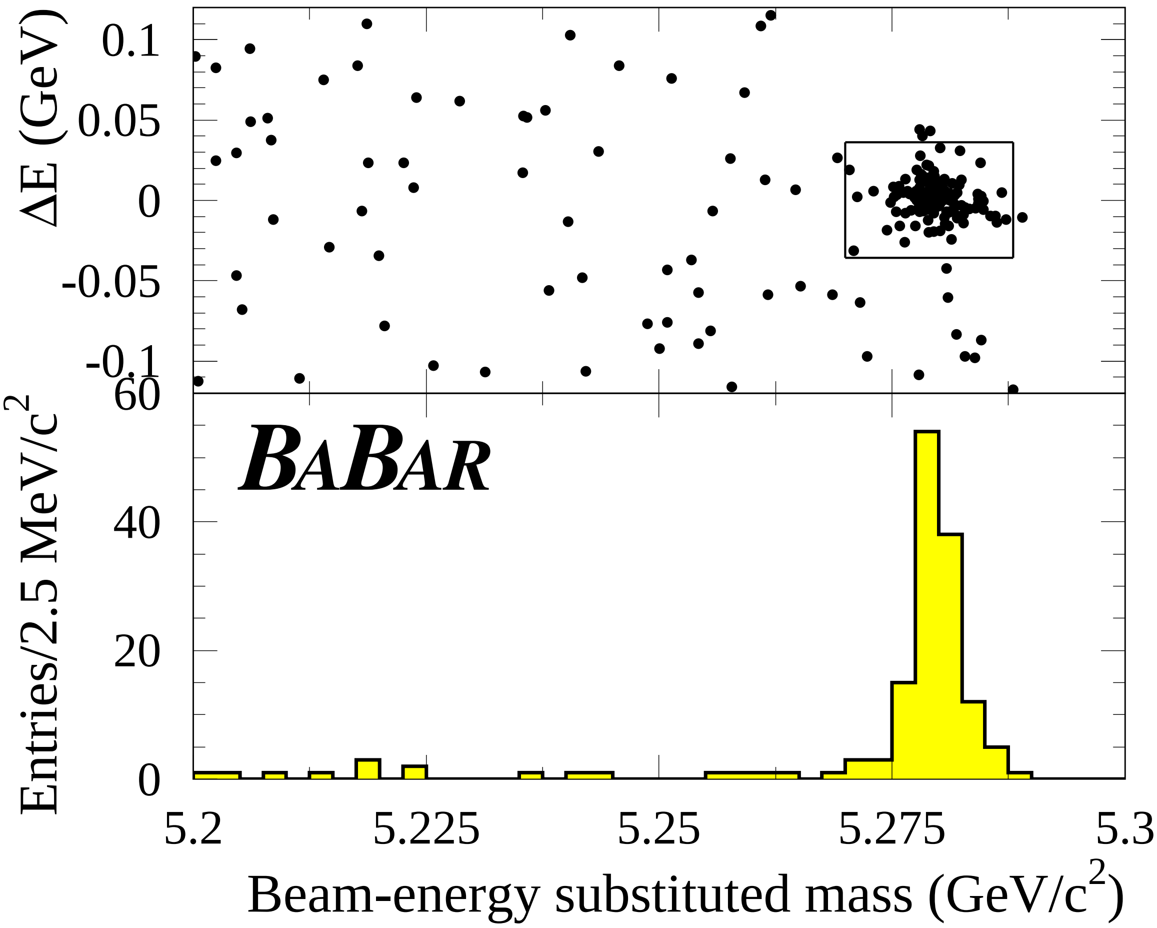
The $K^0_S \to \pi^+\pi^-$ candidates are required to have an invariant mass within 12 MeV (7 MeV) of the nominal K^0_S mass for the $J/\psi K^0_S$ ($\psi(2S)K^0_S$) selection. The $K^0_S \to \pi^0\pi^0$ channel is also considered for the $J/\psi K^0_S$ mode. The two neutral pions are required to have a total energy above 800 MeV and $|m(\pi^0\pi^0) - m(K^0_S)| <$ 200 MeV. Refitting with pion mass constrains and the most probable K^0_S decay vertex (which is required to be between -10 cm and 40 cm from the primary vertex), the new invariant K^0_S mass should be between 470 and 536 MeV.

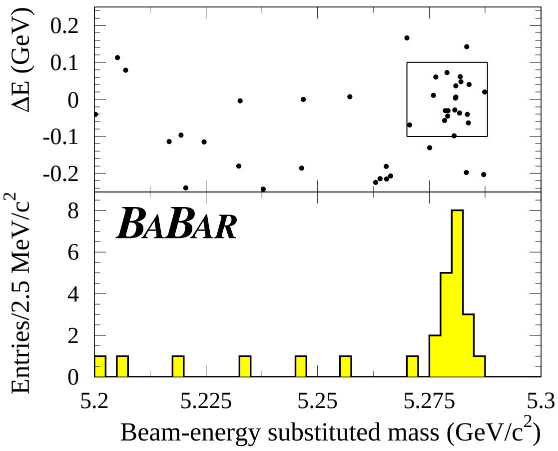
The B_{CP} candidates are formed by combining mass-constrained J/ψ and $\psi(2S)$ candidates with mass-constrained K^0_S candidates. The cosine of the helicity angle of the J/ψ or $\psi(2S)$ candidates in the B rest frame is required to be less than 0.9 (0.8 for $J/\psi \to e^+e^-$), and the K^0_S vertex from the B^0-vertex should be in the same direction as the K^0_S momentum ($\mathbf{p}_{K^0_S} \cdot \mathbf{d}(P_{K\text{-vertex}} - P_{B\text{-vertex}}) > 0$). The distance between the $\psi(2S)$ and the K^0_S vertex is required to be more than 1 mm, and in the $\psi(2S) \to J/\psi\pi^+\pi^-$ mode, the absolute value of the cosine of the angle between B_{CP} three-momentum, $\mathbf{p}_{B_{CP}}$, and the thrust vector, $\mathbf{T}$, calculated in the CM frame, must be less than 0.9.

Finally, the signal region in $(\Delta E, m_{\text{EC}})$ plane for the CP events is defined by 5.270 GeV $< m_{\text{EC}} <$ 5.290 GeV and approximately 3σ cuts on ΔE. Figure 35 shows the CP candidates plotted in the $(\Delta E, m_{\text{EC}})$ plane for each of the three modes, and Table 11 lists the yield and purity for each channel.

Channel	Yield	Purity(%)
$J/\psi K^0_S(\pi^+\pi^-)$	124 ± 12	96
$J/\psi K^0_S(\pi^0\pi^0)$	18 ± 4	91
$\psi(2S)K^0_S$	27 ± 6	93

Table 11: The CP yield and purity. The $J/\psi K^0_S$ channel dominates not surprisingly both in statistics and in purity. As the statistics grow, other channels will be added to the CP sample.





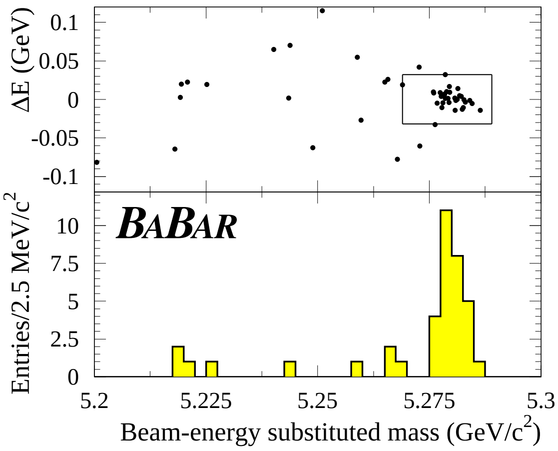


Figure 35: The CP sample in the $(\Delta E, m_{\rm ES})$-plane. The upper plot shows the $J/\psi K^0_S(\pi^+\pi^-)$ sample of 124 ± 12 events with a purity of 96 %. The lower left plot shows the $J/\psi K^0_S(\pi^0\pi^0)$ sample (18 ± 4 events with 91 % purity) and the lower right plot shows the $\psi(2S)K^0_S$ sample (27 ± 6 events with 93 % purity). From projection onto the $m_{\rm ES}$-axis one can see the low background level.

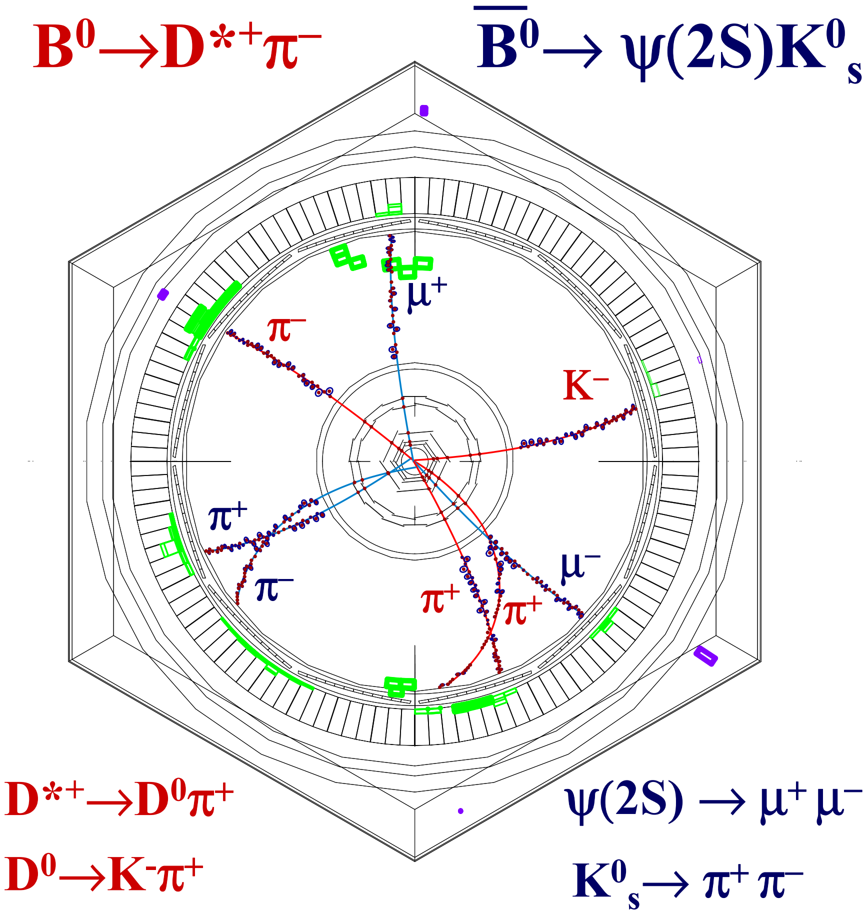


Figure 36: An extraordinary CP event, where the tagging side is fully reconstructed in a hadronic tagging decay mode. The fraction of such events is estimated to be of the order 1:5.000.000.

9.3.2 Other *CP* channels

With the increase of both data and understanding, other CP channels have become accessible. The current modes under investigation are:

- **"Golden modes":** $J/\psi(\ell^+\ell^-)K^0_S(\pi^+\pi^-, \pi^0\pi^0)$, $\psi(2S)(\ell^+\ell^-, J/\psi\pi^+\pi^-)K^0_S(\pi^+\pi^-)$.
- **"Silver modes":** $\xi_{c1}(J/\psi\,\gamma)K^0_S(\pi^+\pi^-)$. $J/\psi(\ell^+\ell^-)K^{0*}(K^0\pi^0)$(angular analysis required),
- **"Bronze mode":** $J/\psi(\ell^+\ell^-)K^0_L$.

The discussion of these channels have been omitted, as none of the analysis work is finished, and thus no results published. The silver modes suffer from requiring reconstruction with photons in the final state, while the bronze mode is oppressed by the very hard K^0_L detection. However, much effort is currently being put into using these channels, and eventually they will contribute to the $\sin 2\beta$ measurement.

10 Flavor tagging and time resolution

The theoretical lifetime distributions of mixing and CP events Eqs. (45+51) are altered by experimental conditions, such as the requirement of tagging and vertex resolution. This section describes the tagging algorithm and the resolution function.

10.1 Flavor tagging

When one B-meson is reconstructed, the determination of the b-flavor on the tagging side is done with a flavor tagging algorithm, based on four sub-algorithms, each with it's own tagging category associated. The four categories described in the following are: lepton, kaon, NT1 and NT2. The quantity that describes the effective tagging efficiency is $Q_i \equiv \epsilon_i D_i^2$, where ϵ_i is the fraction of events assigned to category i and D_i is the dilution of that category. The aim is to optimize $Q \equiv \sum Q_i$, as the statistical error on $\sin(2\beta)$ is proportional to $1/\sqrt{Q}$. Q essentially measures what fraction of the raw statistics that can actually be used as being perfectly tagged, e.g. N_{events} raw CP events will correspond to QN_{events} 100 % tagged CP events.

10.1.1 Why tagging categories?

The wrong tag fraction is expected to vary between tagging categories. This is the reason for introducing the categories in the first place, as an averaging over all tags lowers the effective tagging efficiency, Q.

Consider two types of analysis: One which uses only one tagging category, and one which uses two. Let N be the total number of CP events; N_1 in the first tagging category and N_2 in the second. The sum of the efficiencies of the two tagging categories ε_1 and ε_2 will then be the overall efficiency ε. Let w_i be the wrong tag probability of each individual event and define:

$$\text{One tagging category:} \qquad D \equiv \frac{1}{N}\sum_i^N (1-2w_i), \tag{94}$$

$$\text{Two tagging categories:} \qquad D_1 \equiv \frac{1}{N_1}\sum_{i\in N_1}(1-2w_i), \quad D_2 \equiv \frac{1}{N_2}\sum_{i\in N_2}(1-2w_i). \tag{95}$$

Obviously, $ND = N_1D_1 + N_2D_2$. The effective tagging efficiency Q for each analysis is then:

$$Q_{one} \equiv \varepsilon D^2 = \varepsilon\left(\frac{N_1D_1+N_2D_2}{N}\right)^2 = \frac{1}{\varepsilon}(\varepsilon_1^2D_1^2+\varepsilon_2^2D_2^2+2\varepsilon_1\varepsilon_2D_1D_2), \tag{96}$$

$$Q_{two} \equiv \varepsilon_1D_1^2+\varepsilon_2D_2^2. \tag{97}$$

The claim is that Q_{two} is greater than Q_{one}, which can be proven as follows:

$$\begin{aligned} Q_{two}-Q_{one} &= \varepsilon_1D_1^2+\varepsilon_2D_2^2-\frac{1}{\varepsilon}(\varepsilon_1^2D_1^2+\varepsilon_2^2D_2^2+2\varepsilon_1\varepsilon_2D_1D_2) \\ &= \frac{1}{\varepsilon}[\varepsilon_1(\varepsilon-\varepsilon_1)D_1^2+\varepsilon_2(\varepsilon-\varepsilon_2)D_2^2-2\varepsilon_1\varepsilon_2D_1D_2] \\ &= \frac{\varepsilon_1\varepsilon_2}{\varepsilon}(D_1^2+D_2^2-2D_1D_2) = \frac{\varepsilon_1\varepsilon_2}{\varepsilon}(D_1-D_2)^2 \geq 0. \quad \square \end{aligned} \tag{98}$$

By induction this argument can be extended to include an arbitrary number of tagging categories, and thus, if subdivision into classes with different dilutions is possible, it should be done. However, there are limits to this procedure. Given finite statistics, the statistical uncertainty will at a certain point make the various dilutions overlap, and one can no longer gain information by subdividing. Also, each tagging category should be well understood, which requires a certain amount of statistics. Note that the difference between Q_{one} and Q_{two} is proportional to the square of the difference between the wrong tag fraction.

10.1.2 Lepton and kaon tagging categories

The lepton and kaon tagging categories rely on the correlation between the b-flavor and the charge of a lepton (e or μ) from a primary[32] semileptonic decay or the total kaon charge. Positive charge indicates a B^0 and negative a $\overline{B}^0$. For the leptons a CM momentum greater than 1.1 GeV is required, in order to reduce the contamination of lower-momentum wrong-sign secondary leptons. The cut is chosen to maximize Q. For the kaons no such corresponding kinematic cut exists.

If the event cannot be assigned to the lepton tagging categories, the sum of the kaon charge is considered, and a kaon tag is assigned if it is different from zero. If any two tags are inconsistent, the event is rejected from the lepton and kaon tagging categories.

10.1.3 Neural network tagging categories

Events without a lepton or kaon tag can still hold information about the flavor of the original b-meson from correlation between leptons and kaons, soft pions from D^* decays, secondary leptons from charm decays, and in general the momentum spectrum of the particles. The best way to exploit this information, is with a neural network, which for each event assigns a value, x_{NT}, in the interval $[-1, 1]$ according to the degree of belief, that the flavor was a B^0 or a $\overline{B}^0$, respectively. Events with $|x_{NT}| > 0.5$ make up tagging category NT1, and events with $0.5 > |x_{NT}| > 0.2$ tagging category NT2. The remaining events are not used, as their tagging power is rather limited.

10.2 Time resolution

The reason for building asymmetric B-factories is the fact that the boost of the two B-mesons allow for measuring the distance, and thereby time difference, between their decays, and thus make time dependent measurements. However, the resolution is finite, and so the lifetime distributions will be smeared. This smearing should be included when fitting, which requires a measurement of the resolution function.

10.2.1 The relation between Δz and t

Due to the asymmetry of the beam energy at PEP-II, the $\Upsilon(4S)$ is produced with a boost, β, of:

$$\beta = \frac{\sqrt{(E_{e^+} + E_{e^-})^2 - m^2_{\Upsilon(4S)}}}{E_{e^+} + E_{e^-}} = 0.49 \quad \Rightarrow \quad \beta\gamma = 0.56. \tag{99}$$

Since the two B-mesons are produced nearly at rest in the $\Upsilon(4S)$ rest frame, they will approximately have the same boost, and from measuring the z-coordinate of the two B-decays, the difference in lifetime, t, can be inferred as follows:

$$\Delta z \quad = \quad \beta\gamma ct \tag{100}$$

This approximation was used for the Osaka conference. It ignores the small boost, that the B-mesons get from the $\Upsilon(4S)$ decay:

$$\beta_{cm} \quad = \quad \frac{\sqrt{m_\Upsilon(4S)^2 - (2m_{B^0})^2}}{m_{\Upsilon(4S)}} = 0.06, \tag{101}$$

[32]Primary meaning from the B-meson as opposed to secondary being from a decay product, such as a D-meson.

and so the correct relation between Δz and t is:

$$\Delta z = \beta\gamma ct \left(1 + \frac{\beta_{cm}}{\beta} \cos\theta_{\mathrm{CM}} \frac{t_{CP} + t_{tag}}{t_{CP} - t_{tag}}\right), \tag{102}$$

However, due to the $\sin\theta$ shape of the matrix element (see Eq. 47), the two B-mesons from the $\Upsilon(4S)$ decay are mostly emitted perpendicular to the beam direction, and consequently their difference in boost is generally small.

10.2.2 The resolution function

The inclusion of finite resolution is made by convoluting the lifetime distributions with a resolution function, $\mathcal{R}$, the parametrization of which should be flexible enough to encompass all features of the resolution.

In principle, the resolution function should be obtained from the CP sample only. However, since the time resolution is dominated by the precision of the B_{tag} vertex position, $\mathcal{R}$ can be obtained by fitting fully reconstructed B^0-decays, which have about the same vertex resolution as the CP events. From the reconstructed B-meson, the B_{tag} three-momentum can be derived, and using this information the remaining tracks are fitted to a common vertex. Neutral particles with non-negligible lifetimes (such as the D-meson and the shortlived neutral kaon) may displace the origin of tracks, and thus all reconstructed such candidates are used in the fit instead of their daughters. Any track contributing to the vertex χ^2 by more than 6 are excluded from the fit, until the fit converges, or no particles are left. Events are also rejected if $|\Delta z| > 3$mm or $\sigma_{\Delta z} > 400\mu$m. The t distribution is fitted with an exponential function, $e^{-\Gamma|t|}$ convoluted with $\mathcal{R}(t, \hat{a})$, where a suitable number of the parameters, $\hat{a}$, depending on the statistics available, are left free in the fit.

It was found, that the resolution function is well described by two Gaussian distributions. In addition, a third 1 % very wide Gaussian ($\sigma = 8$ps) has been added to account for events with extremely high time differences (so-called "outliers"), which are mostly due to vertex reconstruction problems.

The error σ_t on t, calculated from the fits to the two B vertices, is used in the likelihood fit on an event-by-event basis. However, this error is underestimated due to inclusion of particles from secondary vertices and possible underestimation of the amount of material transversed by the particles. To account for this, two scale factors, s_1 and s_2, of the Gaussian widths are introduced, such that $\sigma_1 = s_1 \times \sigma_t$ and $\sigma_2 = s_2 \times \sigma_t$. Due to low statistics, only the width (scale factor) and the bias of the core Gaussian distribution and the fraction of outliers were left free in the fit. The result of fitting the fully reconstructed hadronic sample can be seen in Fig. 37, and the parameters of $\mathcal{R}$ can be found in Table 12.

Parameter	Value	Status
$f_{core}(\%)$	75	Fixed
s_{core}(ps)	1.33 ± 0.14	From fit
δ_{core}(ps)	-0.2 ± 0.06	From fit
s_{tail}(ps)	2.1	Fixed
δ_{tail}(ps)	0.0	Fixed
$f_{outliers}(\%)$	1.6 ± 0.6	From fit

Table 12: *The parameters of the resolution function, $\mathcal{R}$.*

As a consistency check, the fitted lifetime is compared with the world average taken from the Particle Data Group (PDG). BaBar measures:

$$\boxed{\tau_{B^0} = 1.506 \pm 0.052 \text{ (stat.) } \pm 0.029 \text{ (syst.) ps}} \tag{103}$$

The PDG value is $\tau_{B^0} = 1.548 \pm 0.032$ ps [37], and the two numbers are in perfect agreement. The precision of the B^0 lifetime measurement is comparable to other experiments, but soon it will be competitive with the PDG value itself.

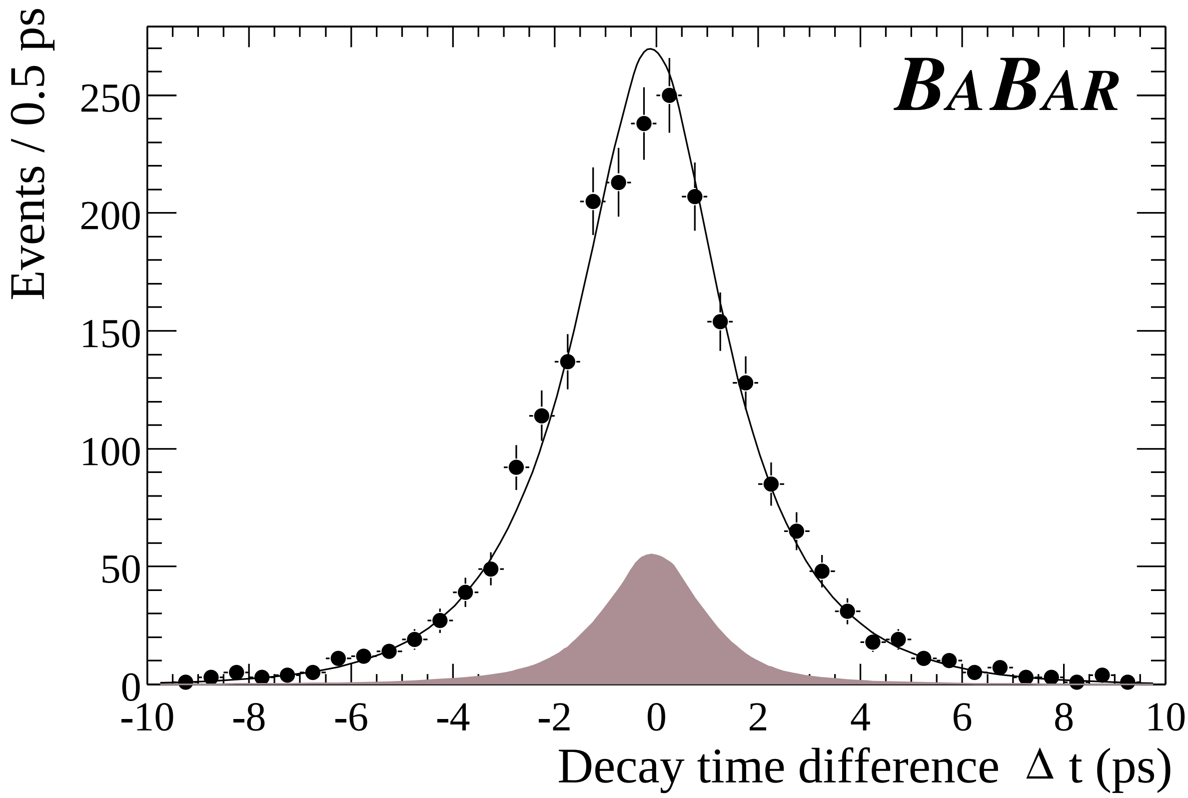


Figure 37: Fit of lifetime and resolution function to fully reconstructed events.

The correct distributions to fit the mixing and CP samples with, are then given by:

$$\mathcal{F}_\pm(t, \Gamma, \Delta m, D\sin(2\beta)), \hat{a}) \equiv F_\pm(t, \Gamma, \Delta m, D\sin(2\beta)) \otimes \mathcal{R}(t, \hat{a}), \tag{104}$$

$$\mathcal{H}_\pm(t, \Gamma, \Delta m, D, \hat{a}) \equiv H_\pm(t, \Gamma, \Delta m, D) \otimes \mathcal{R}(t, \hat{a}). \tag{105}$$

The functions $F_\pm(t)$ and $H_\pm(t)$ are given by Eqs. (84, 85) and the resolution function is:

$$\mathcal{R}(t, \hat{a}) = \sum_{i=1,3} \frac{f_i}{\sqrt{2\pi}\sigma_i} \exp((t-\delta_i)^2/2\sigma_i^2). \tag{106}$$

The distributions $\mathcal{F}_\pm(t)$ and $\mathcal{H}_\pm(t)$ are shown in Fig. 38 below, both in the unsmeared and the smeared case.

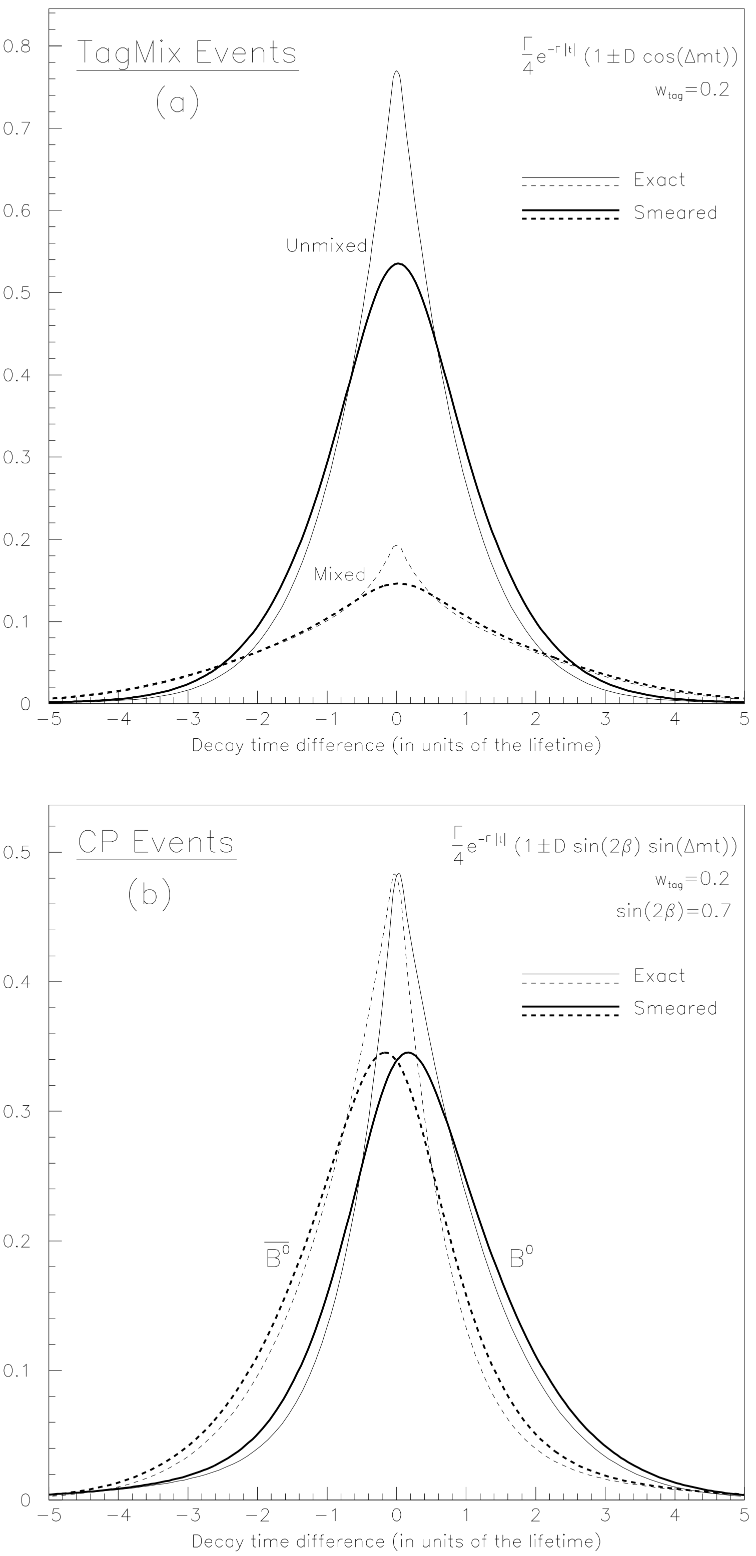


Figure 38: Distribution of (a) tagging/mixing events and (b) CP events. The thin lines show the perfect case, whereas the thick lines are when vertex resolution has been taken into account. In the TagMix distribution the events with $|t| \sim 0$ contains the most information about the wrong tag fraction. In the CP distribution the asymmetry is largest for events with $|t| \sim 2\tau$, and so in this case it is these events, which have the largest statistical power when fitting.

11 Mixing and *CP* fitting revised

There are sadistic scientists who hasten to hunt down error instead of establishing the truth.

[Marie Curie, 1867-1934]

The optimal way of extracting the dilution, $D \equiv (1-2w_{tag})$, and the CP asymmetry, $\sin(2\beta)$, along with their statistical uncertainties is provided by an unbinned likelihood fit[33]. However, fitting the full set of data with a likelihood fit is not trivial and highly non-transparent, as one cannot see how the individual events contribute to the result. Addressing this problem, this section introduces slightly less optimal but extremely simple and transparent methods of extracting the two quantities and their statistical uncertainties. The goal is not to rival the likelihood method, but to provide a clear-cut view of what is going on. The methods can serve as a check and allows for a direct computation of systematic uncertainties.

The first method applies only to the dilution/mixing measurement, and is meant as an introduction to counting techniques. The second method applies to both the TagMix and the CP fitting, which will therefore be treated simultaneously. Essentially, the method consists of using one optimally placed bin in the time distribution, hence the name OneBin. The third method weights the statistical power of each event, and provides a mean of visualizing CP violation.

The methods are described including both smearing and backgrounds, and their performance is compared to the optimal likelihood fit.

11.1 The likelihood fit

The full likelihood fit consists of maximizing the following likelihood function:

$$\ln \mathcal{L}_{CP} = \sum_i \left[\sum_{B^0\text{tag}} \ln \mathcal{F}_+(t;\Gamma,\Delta m_d,\hat{a},D_i\sin(2\beta)) + \sum_{\overline{B}^0\text{tag}} \ln \mathcal{F}_-(t;\Gamma,\Delta m_d,\hat{a},D_i\sin(2\beta)) \right], \tag{107}$$

where the outer summation is over tagging categories i. The dilutions, D_i, are determined from a similar likelihood fit to the mixing data sample:

$$\ln \mathcal{L}_{\text{mixing}} = \sum_i \left[\sum_{\text{unmixed}} \ln \mathcal{H}_+(t;\Gamma,\Delta m_d,\hat{a},D_i) + \sum_{\text{mixed}} \ln \mathcal{H}_-(t;\Gamma,\Delta m_d,\hat{a},D_i) \right]. \tag{108}$$

The functions are maximized by the program Minuit [60], providing the fitted quantities along with their statistical uncertainties.

11.2 Simple methods

The likelihood method is optimal in the sense that it minimizes the statistical uncertainty. However, if the systematic uncertainties becomes significant, then other and more simple methods may turn out to have advantages by not suffering from the same systematic uncertainties. At the same time, different methods serve as a check of the likelihood result, which is crucial in the infancy of the analysis not the least when conducting blind analysis[34]. For this reason, several other fitting methods have been considered.

[33] If an optimal fit exists (equality in the RCF inequality), this can be proven to be the likelihood fit.

[34] Blind analysis is a term for not revealing the result until the analysis is final, to avoid experimenters bias.

11.3 The CountAll method

"Can you do addition?" the White Queen asked. "What's one and one and one and one and one and one and one and one and one and one?" "I don't know", said Alice, "I lost count".

[Lewis Carroll. Through the Looking-Glass]

When no timing information is available (or used), as is the case in symmetric $e^+e^- \to \Upsilon(4S)$ experiments, only the tagging/mixing and not the CP violation in inteference measurement can be made:

$$\int_{-\infty}^{\infty} \frac{\mathcal{F}_+ - \mathcal{F}_-}{\mathcal{F}_+ + \mathcal{F}_-} \equiv 0, \qquad \int_{-\infty}^{\infty} \frac{\mathcal{H}_+ - \mathcal{H}_-}{\mathcal{H}_+ + \mathcal{H}_-} \equiv 1 - 2\chi_d. \tag{109}$$

where $\chi_d = 0.174 \pm 0.009$ is the B_d-oscillation probability defined in Eq. (46). Thus, one expects 17 % of the events to contain two equal flavored B^0-mesons (*mixed* events), and 83 % of the events to contain two opposite flavored B^0-mesons (*unmixed* events).

The idea of the CountAll method is that, knowing the value of χ_d, and comparing the time integrated rates of mixed and unmixed events (see Fig. 38a), one can extract the wrong tag fraction using the following formula:

$$\chi_{\text{obs}} = (1 - w_{\text{tag}})\chi_d + w_{\text{tag}}(1 - \chi_d) \quad \Rightarrow \quad w_{\text{tag}} = \frac{\chi_{\text{obs}} - \chi_d}{1 - 2\chi_d}, \tag{110}$$

where χ_{obs} is the measured fraction of mixed events. The inclusion of backgrounds require knowledge of the fractions f_{bkg} and effective mixing parameters χ_{bkg} of each type of background. The ways of obtaining this information were mentioned in section 9.2.1. The expression in Eq. (110) changes only slightly:

$$\chi_{\text{obs}} = f_{\text{sig}}[(1 - w_{\text{tag}})\chi_d + w_{\text{tag}}(1 - \chi_d)] + \sum_{\text{bkg}} f_{\text{bkg}}\chi_{\text{bkg}} \quad \Rightarrow \quad w_{\text{tag}} = \frac{\chi_{\text{obs}} - f_{\text{sig}}\chi_d - \sum f_{\text{bkg}}\chi_{\text{bkg}}}{f_{\text{sig}}(1 - 2\chi_d)}. \tag{111}$$

This expression does not contain the resolution function and its uncertainties, and so any systematic errors from the resolution function do not apply to results from the CountAll method.

11.3.1 OneBin method

As can be seen in Fig. 38a, the "outer" regions in the time distribution tend to dilute the ratio of mixed to unmixed events. It would, therefore, be desirable to exclude the regions which contain no information about the mixing. The idea of the OneBin method is exactly this, i.e. to find one optimized central bin, in which the wrong tag fraction can be obtained from the number of mixed and unmixed events. CP asymmetries can be fitted by choosing a bin on each side of the lifetime distribution.

In the mixing (TagMix) case, the central region of time contains most of the information about D, and so one could imagine a measurement of D by simply counting the numbers of unmixed (u) and mixed (m) events within some time interval, $I_{\text{TagMix}} \equiv [-t_0, t_0]$. This can be expressed as:

$$\boxed{D \equiv (1 - 2w_{\text{tag}}) = k_{\text{TagMix}}\, \tfrac{u-m}{u+m}} \tag{112}$$

$$u, m \equiv \int_{I_{\text{TagMix}}} \mathcal{H}_\pm \, dt, \qquad k_{\text{TagMix}} \equiv \frac{\int_{I_{\text{TagMix}}} e^{-\Gamma|t|} \otimes \mathcal{R}\, dt}{\int_{I_{\text{TagMix}}} e^{-\Gamma|t|}\cos(\Delta m t) \otimes \mathcal{R}\, dt}. \tag{113}$$

In words, one counts the number of unmixed and mixed events within some time bin, and since k_{TagMix} is just a constant, D can be directly calculated from these counts.

In the case of CP events one is considering an asymmetry, which requires a bin on each side of $t = 0$, $I^-_{CP} \equiv [-t_2, -t_1]$ and $I^+_{CP} \equiv [t_1, t_2]$. With these intervals one can define a positive (p) bin and a negative bin (n) as follows:

$$\boxed{D\sin(2\beta) = k_{CP}\ \tfrac{p-n}{p+n}} \tag{114}$$

$$p, n \equiv \int_{I^-_{CP}} \mathcal{F}_\mp\ dt + \int_{I^+_{CP}} \mathcal{F}_\pm\ dt, \qquad k_{CP} \equiv \frac{\int_{I^-_{CP}+I^+_{CP}} e^{-\Gamma|t|} \otimes \mathcal{R}\ dt}{\int_{I^-_{CP}+I^+_{CP}} e^{-\Gamma|t|}\sin(\Delta m t) \otimes \mathcal{R}\ dt}. \tag{115}$$

The expression for p and n are slightly more complicated, since the distributions are asymmetric. Thus, one has to take into account that the B^0 and the $\overline{B}^0$ have opposite asymmetries.

Essentially, the bins select the events with most statistical power, i.e. the regions in the time distributions, where the difference between the two distributions to be compared is the largest. *The place of the cuts,* $\mathbf{t_0}$ *for TagMix events and* $\mathbf{t_1}$ *and* $\mathbf{t_2}$ *for* **CP** *events, is determined by minimizing the statistical uncertainty*, and the results are shown in Fig. 39 below.

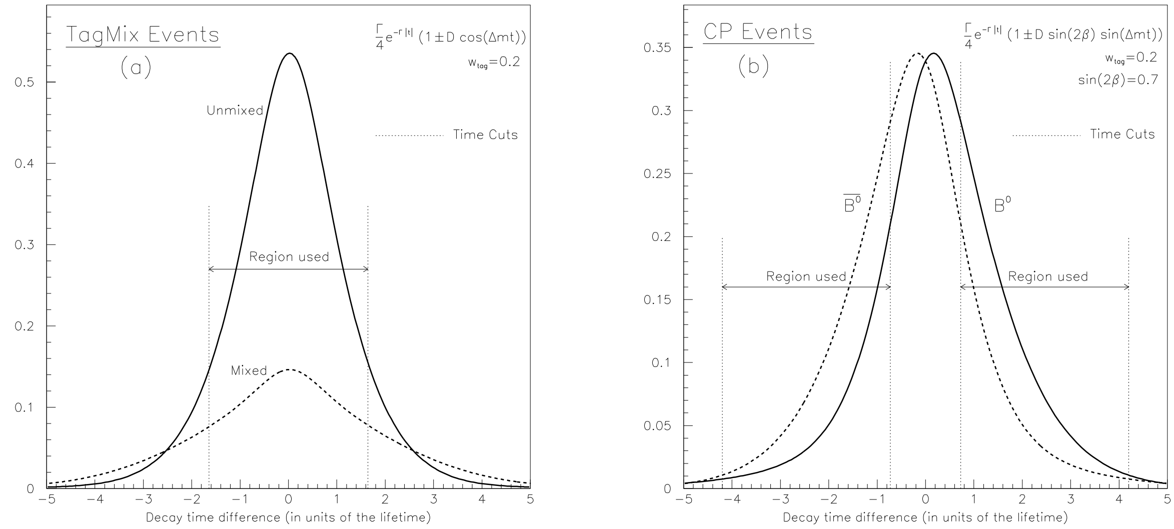


Figure 39: Optimal placing of bins for determining D and $D\sin(2\beta)$. Using only the number of events in the bins shown, D and $D\sin(2\beta)$ can be determined with almost the same statistical precision as in a likelihoodfit.

As can be seen from the figure, the bins chosen in the TagMix case are *not* at the points $|t| = 2.5\tau$, where the density of mixed events begin to exceed the density of unmixed, but somewhat before, at $|t| = 1.6\tau$. This turns out to be when the number of mixed events beyond that point matches the number of unmixed, such that this domain as a whole contains no information about mixing, i.e.

$$\int_{t_0}^{\infty} e^{-\Gamma|t|}\cos(\Delta m t) \otimes \mathcal{R}\ dt \ = \ 0 \quad \Rightarrow \quad t_0 \ \simeq \ 1.6\tau. \tag{116}$$

In the CP fitting case, it is not as easy to argue for the placing of the cuts, and one must rely on numerical optimization. However, events with $t \sim 0$ does not contain very much information about the asymmetry, and so the optimum of the inner cut lies a little away from the center, as seen on the figure. The asymmetry does not change sign until t is high ($|t| \sim 5\tau$), which makes the outer cut lie far out ($|t| \sim 4.2\tau$). Backgrounds can for the OneBin method be included in the same manner as for the CountAll method.

Other more refined methods have been considered, such as the so-called TrigBin and TwoBin methods, and at a slight cost of simplicity one can obtain statistical power even closer to that of the likelihood method. However, since the goal is to provide a simple and transparent method, the OneBin method was preferred. A short description of TrigBin and TwoBin is given in Appendix C.

11.4 Performance compared to likelihood method

As mentioned before, the likelihood method gives the optimal fit, and consequently the statistical uncertainties of other methods are necessarily larger. To see how much (in fact how little!) information is lost by using a simple method, the statistical uncertainty achieved by the various methods (including the TrigBin and TwoBin methods) relative to the statistical uncertainty of the likelihood is considered.

11.4.1 Statistical uncertainties

The statistical uncertainty of the bin methods are calculated using the following formula:

$$\sigma_\theta^{\text{Bin methods}} = \left|\frac{\partial\theta}{\partial r}\right| \sigma_r = \left|\frac{\partial\theta}{\partial r}\right| \sqrt{\frac{r(1-r)}{N}}, \tag{117}$$

where $r = u/(u+m)$ in the TagMix case and $r = p/(p+n)$ in the CP case and θ is the variable fitted for (i.e. D and $D\sin(2\beta)$). Assuming that the likelihood fit is efficient and without bias, the statistical uncertainty of it is:

$$\sigma_\theta^{\text{likelihood}} = \left(\frac{\partial^2 \log\mathcal{L}}{\partial\theta^2}\right)^{-\frac{1}{2}}, \tag{118}$$

where $\mathcal{L}$ is the likelihood function, such as defined in Eq. (107) or Eq. (108).

11.4.2 Comparison with likelihood

The ratio $\sigma^{\text{Bin method}}/\sigma^{\text{likelihood}}$ as a function of w_{tag} is shown in Fig. 40 below, for the various methods and the results summarized in Table 13.
For a mistag fraction of $w_{\text{tag}} = 0.2$, the TagMix result of OneBin yields a statistical uncertainty 5.5 % larger than the likelihood one, whereas the CP result has a 9.2 % increase. This is a small price to pay for gaining both simplicity and ability to directly compute systematic uncertainties. In Table 13 below, the results are summarized for the various methods:

In addition to the statistical uncertainty, systematic errors will eventually become important. One such is the uncertainty in the resolution function. The likelihood fit is fairly sensitive to this, while the OneBin is not affected much, and the CountAll method not at all. In addition neither of the counting methods are affected but the troublesome outliers. This means, that a change of method might allow one to "trade" systematic uncertainty for statistical, which is well understood and thus more favorable. With the inclusion of such subtleties, the likelihood method looses is status as optimal, and this should be taken into account in the analysis.

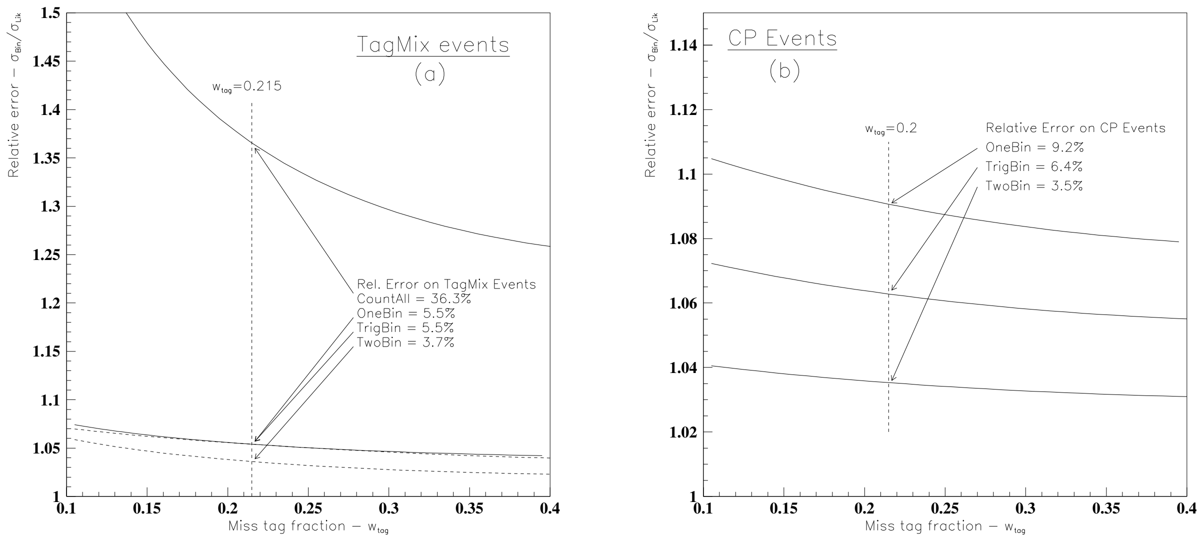


Figure 40: Statistical uncertainty of the CountAll, OneBin, TrigBin and TwoBin methods relative to the likelihood method. $\sigma^{\text{Bin methods}}/\sigma^{\text{likelihood}}$ for (a) TagMix and (b) CP fitting as a function of the wrong tag fraction w_{tag}. Note the difference in scales.

11.5 Kin method

The Kin method [61] consists of measuring the CP asymmetry by weighting each event by their statistical power, that is their place in the time distribution. For the $J/\psi K_S^0$ channel in absence of backgrounds, imperfect tagging and detector effects, the Kin variable, $\mathcal{K}$, is defined as follows:

$$\sin(2\beta)\mathcal{K} \equiv \frac{f_+ - f_-}{f_+ + f_-} \quad \Rightarrow \quad \mathcal{K} = \sin(\Delta m t). \tag{119}$$

When experimental effects are taken into account, the distributions $f_\pm$ change to $\mathcal{F}_\pm$ and the expression for $\mathcal{K}$ accordingly, but $\sin(2\beta)$ remains linear in the problem, and thus the definition remains valid. The Kin distribution:

$$\Psi(\mathcal{K}) \equiv \Psi_0(\mathcal{K})[1 + \sin(2\beta)\mathcal{K}], \tag{120}$$

has the property, that experimental effects are contained in Ψ_0, which is by construction an even function, and any asymmetry in the $\mathcal{K}$ distribution establishes CP violation. The actual

$w_{\text{tag}} = 0.215$	TagMix	CP
CountAll	36.4 %	—
OneBin	5.4 %	9.1 %
TrigBin	5.4 %	6.3 %
TwoBin	3.6 %	3.4 %

Table 13: Statistical uncertainty of the CountAll, OneBin, TrigBin, and TwoBin methods relative to the likelihood method. The quoted numbers represent the relative increase of the statistical uncertainties compared to the likelihood method from using the simpler methods described in the text, e.g. 5.5 % means that $\sigma^{\text{Bin methods}} = 1.055\ \sigma^{\text{likelihood}}$.

value and statistical error of $\sin(2\beta)$ can be determined from the ratio of the moment of the $\mathcal{K}$ distribution:

$$\sin(2\beta) = \frac{\langle\mathcal{K}\rangle}{\langle\mathcal{K}^2\rangle}, \qquad \sigma_{\sin(2\beta)} = \frac{1}{\sqrt{N\langle\mathcal{K}^2\rangle}}\sqrt{1-\sin^2(2\beta)\frac{\langle\mathcal{K}^4\rangle}{\langle\mathcal{K}^2\rangle}}. \tag{121}$$

and no likelihood maximization is needed. The Kin values fully contain the CP information, and thus it has the same error as the likelihood approach. The Kin method was used in [62] as a check of the likelihood result.
The advantage of the Kin approach is, that the assignment of a value according to statistical power allows for many different channels and experiments to be included without complicating the analysis, and at the same time the result can be presented graphically. In principle the Kin method can also be applied to the TagMix measurement, but for historical reasons it has not.

11.6 Validation of the OneBin and KIN methods

In order to study biases, systematic errors, and possible correlations of the OneBin and the KIN method, a large number of Monte Carlo experiments were generated. By studying the results compared to the likelihood ones, the validity of the two methods can be established.

11.6.1 TagMix validation

In the TagMix case, 1000 toy MC experiments were carried out for each tagging category. The statistics, the resolution, and the expected tagging fractions and efficiencies were chosen in accordance with the sample of fully reconstructed B-mesons. The actual numbers are not essential, as two methods are being compared. For each MC experiment, the result and the statistical uncertainty of the OneBin method was compared with that of the likelihood. As the results for each tagging category were very similar, only the result from the first tagging category (leptons) is shown in Fig. 41, while the rest can be found in Appendix D.
We see that the OneBin method is unbiased compared to the likelihood method, and that for the first tagging category, one expects 8 % additional error. The high number is due to the low wrong tag fraction of the lepton tagging category, and so this is in accordance with Fig. 40 as are the remaining numbers.

11.6.2 *CP* validation

For the validation of CP fitting with the OneBin and the KIN method, 2000 Monte Carlo experiments were generated. The input numbers used were[35]:

- 120 tagged CP events.
- 4 tagging categories with dilutions 0.794 (12 %), 0.602 (52 %), 0.660 (12 %) & 0.348 (24 %).
- Resolution function with two Gaussian distributions.
- World average including first BaBar result for CP-violation ($\sin(2\beta) = 0.4$).
- PDG values for B^0 lifetime and mass difference ($\tau_{B^0} = 1.548$ ps and $\Delta m_{B^0} = 0.472$ ps^{-1}).

The result of this validation can be seen in Fig. 42. Figure 42a shows that enither the KIN nor the OneBin method have biases, as the central value of the distribution of results is consistent with 0.4, which was the input value for the MC. Given the statistics, the expected

[35] The validation was carried out for the Osaka Conference, and so the numbers are from then (July 2000).

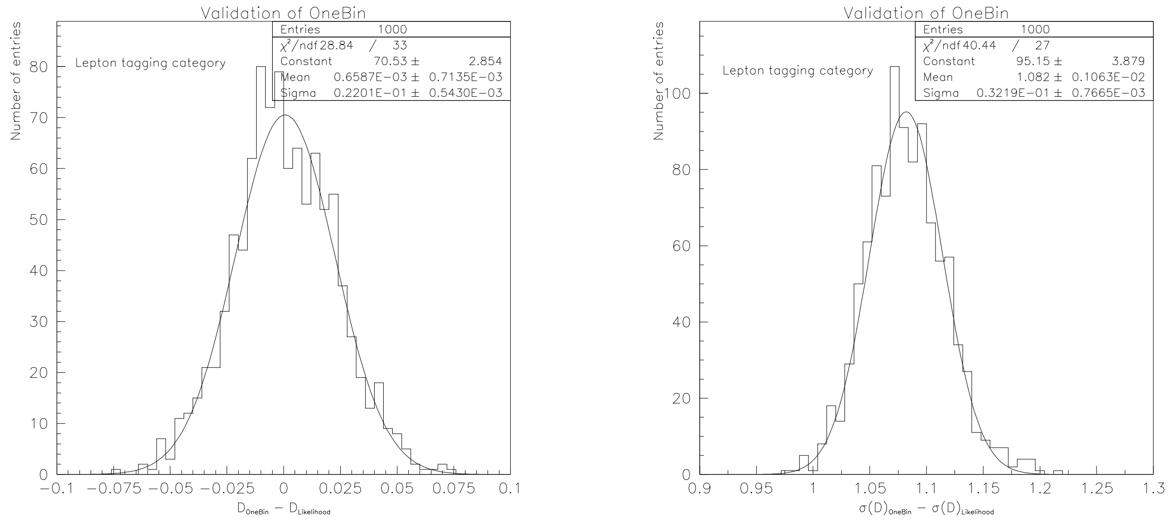


Figure 41: Validation of OneBin method in the TagMix case using 1000 MC experiments. The left figure shows the difference between the OneBin and the likelihood result, $\Delta D \equiv D_{\mathrm{OneBin}} - D_{\mathrm{likelihood}}$, and since the central value is consistent with zero, the OneBin method has no bias relative to the likelihood method. The right figure shows the ratio between the OneBin and the likelihood error, and thus one expects 8 % more error from the OneBin method in the TagMix case.

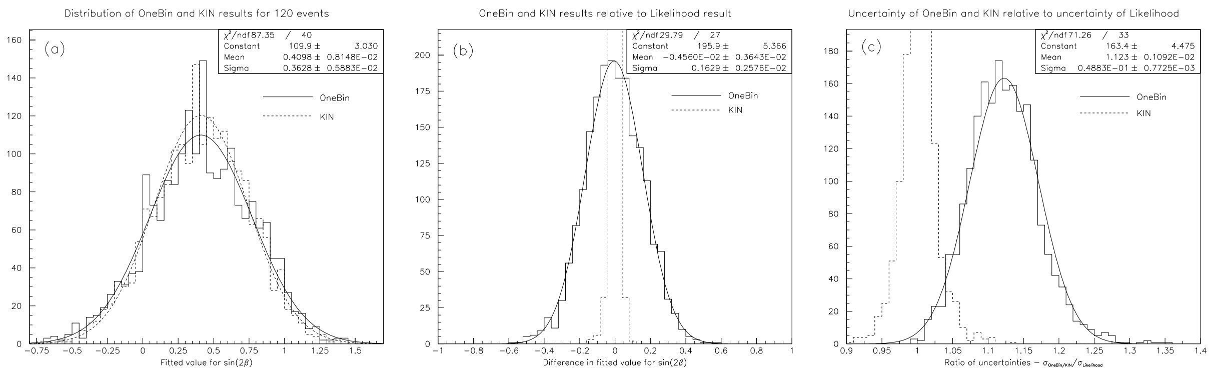


Figure 42: Validation of the OneBin and the KIN method in the CP case using 2000 MC experiments. Figure (a) shows the value obtained for $\sin(2\beta)$ for OneBin and KIN. Figure (b) shows the difference between the OneBin/KIN and the likelihood result. Figure (c) shows the statistical error of OneBin/KIN relative to the one obtained from the likelihood fit.

error on $\sin(2\beta)$ for the OneBin method is 0.36, which is to be compared with the expected likelihood error of 0.32. Figure 42b shows how close one can expect the KIN and the OneBin results to lie from the likelihood one. One sees, that the KIN result essentially reproduces the likelihood result (both in value and in uncertainty), whereas the OneBin result is allowed some range. As expected, the OneBin method is also slightly less accurate (see section 11.4.2), which can be seen from the distribution of $\sigma(\sin(2\beta))_{\mathrm{Bin\ method}}/\sigma(\sin(2\beta))_{\mathrm{likelihood}}$ in Fig. 42c. The fitted central value of 1.12 (12 % additional error) is larger than the calculation in the previous section. However, in the OneBin calculation the various tagging categories were combined, giving rise to loss of power (see section 10.1.1), and so the result seems reasonable.

A word of warning about the Gaussian fits! The distribution of OneBin results in Fig. 42 (a) is *not* expected to be Gaussian, since OneBin results only come in discrete values (the result can only be proportional to an integer divided by the total number of events, here 120). The fit in Fig. 42 (b) is trustworthy, as taking the difference smears the result and thereby allows for continuous values. The ratio of uncertainties in Fig. 42 (c) is again not expected to be Gaussian and the fit has only been made to obtain a central value.

The OneBin method averagely uses $58 \pm 5\,\%$ of the events, and further studies have shown that there is a clear and expected correlation between the fraction of events used by OneBin and the statistical error on the result.
It was further investigated if any anomalous likelihood result or error (as was actually obtained) would change the validation results, and it was found that the expected error and the distance from the likelihood result of the OneBin method simply rescale accordingly.

12 Results

Errors using inadequate data are much less than those using no data at all.

[Charles Babbage 1792-1871]

12.1 TagMix results

The results obtained for the tagging efficiency and the mixing rate are based on the mixing sample described in section 9.2. The tagging results from the OneBin method and the likelihood fit are compared and discussed. The mixing result is presented and compared to the PDG value.

12.1.1 Result for tagging

The OneBin method was applied first to the sample of fully reconstructed B mesons, which have only combinatorial background, and then to the more complicated $D^*\ell\nu$ sample. The results are compared to those of the likelihood fit and summarized in Table 14 and 15 below.

Hadronic sample	Lepton tag	Kaon tag	NT1 tag	NT2 tag
OneBin	11.0 ± 3.2	20.4 ± 2.1	15.3 ± 3.5	34.6 ± 3.7
likelihood	11.6 ± 3.2	19.6 ± 2.1	13.5 ± 3.5	31.4 ± 3.7

Table 14: Wrong tag fractions from hadronic sample using OneBin and likelihood method.

For the hadronic sample, the OneBin results are consistent with the results from the likelihood method [63], with the exception of the NT2 tagging category. From the MC validation the distribution of $\Delta D \equiv D_{\text{OneBin}} - D_{\text{likelihood}}$ has a width of 2.0 % (see the bottom left figure in Appendix D), which corresponds to a width in the Δw_{tag} distribution of 1.0 %, and so one is faced with a 3.2σ discrepancy. There is no corresponding effect in the NT2 efficiency. Other groups also using the OneBin method have obtained a similar value (34.2 % [64]), and though the results have moved closer when including the new data, the statistical uncertainty has also decreased, and a discrepancy has remained.

There can be several reasons for this, and the subject is currently under discussion. One cause could be, that events outside the time bin used by the OneBin method have well separated vertices. For this reason the tagging ability by the neural network is better here, thus the likelihood fit, which includes these events gives a lower wrong tag fraction. Since the effects have remained, a systematic effect like this seems most probable. Additional smaller discrepancies can be due to different treatments of the varies backgrounds, such as small differences in sideband definitions and fitting procedures.

The difference between the result of the two methods has been included as a systematic error, and compared to the statistical error on $\sin(2\beta)$, it does not constitute a problem. However, the source of the difference needs to be investigated in order to lower the systematic error.

As can also be seen from Table 14, the error on the OneBin result equals that of the likelihood result. This is due both to the additional systematic error from the resolution function on the likelihood result, and the fact that the likelihood fit has a floating value for Δm, while the OneBin fit uses the PDG value. The fixing of this parameter improves the statistical precision. Disregarding these two minor effects, it has been checked that the OneBin error is larger by $5 - 7$ % in perfect agreement with the calculation in section 11.4.2.

The analysis of the $D^*\ell\nu$ sample is more complicated, since many backgrounds enter. The numbers describing the backgrounds, which are obtained from data, were taken from [18]. Once again the OneBin method was applied, and the results were then compared to that of the likelihood method.

For the semileptonic sample all the OneBin results are in accordance with the likelihood results, which is encouraging, since the analysis is more complicated, but also striking as one should expect roughly the same discrepancies as in the fully reconstructed sample. The reason why this is not so, could be due to the differences in background, and since it is the likelihood fit that changes its value most, outliers could be the cause. Again the value of Δm was left floating in the likelihood fit, and the results can be seen in Table 15.

$D^*\ell\nu$	Lepton tag	Kaon tag	NT1 tag	NT2 tag
OneBin	10.2 ± 2.0	21.1 ± 1.6	21.8 ± 2.8	37.1 ± 2.6
likelihood	8.4 ± 2.0	19.9 ± 1.6	21.0 ± 2.8	36.1 ± 2.5

Table 15: Wrong tag fractions from semileptonic sample using OneBin and likelihood method.

While there is a good agreement between the hadronic and semileptonic measurements for the lepton and the kaon tagging categories, the neural network tagging categories seem to differ somewhat though not conclusively. This is also under current investigation [58], and it is interesting to notice, that the OneBin method gives more consistent results. Not surprisingly, the CountAll method supports the OneBin results.

The effective tagging efficiency (sometimes called the quality factor) of the tagging algorithm, Q, is 27.9 ± 0.5 % as can be seen from Table 16 below, and this number is in agreement with the rough TDR expectations [47].

Tagging Category	ϵ (%)	w_{tag} (%)	Q (%)
Lepton	11.5 ± 0.5	$9.6 \pm 1.7 \pm 1.3$	7.3 ± 0.3
Kaon	36.7 ± 0.9	$19.7 \pm 1.3 \pm 1.1$	13.5 ± 0.3
NT1	11.7 ± 0.5	$16.7 \pm 2.2 \pm 2.0$	5.2 ± 0.2
NT2	16.6 ± 0.6	$33.1 \pm 2.1 \pm 2.1$	1.9 ± 0.1
All	76.7 ± 0.5		27.9 ± 0.5

Table 16: Effective tagging efficiency, Q, calculated from the efficiencies and wrong tag fractions in the four tagging categories. 77 % of the CP events have some tagging information, and effectively 28 % of the CP events can be used as perfectly tagged.

12.1.2 Result for mixing

Mixing of the neutral B-meson is most apparent in the following time-dependent mixing asymmetry:

$$a(t) \equiv \frac{N_{\text{unmixed}} - N_{\text{mixed}}}{N_{\text{unmixed}} + N_{\text{mixed}}} = \frac{\mathcal{H}_+ - \mathcal{H}_-}{\mathcal{H}_+ + \mathcal{H}_-}. \tag{122}$$

However, due to the presence of backgrounds, the probability distribution functions $\mathcal{H}_\pm$ Eq. (105) have to be extended to incorporate these. Since the various backgrounds differ between tagging categories, i, and may have different lifetimes and mixing characteristics, the distribution must be split accordingly:

$$\mathcal{H}_{\pm,i} \equiv f_{\text{sig}}\mathcal{H}_\pm(t) + \sum_{\text{bkg}} f_{\text{bkg},i}\, \mathcal{B}_{\pm,\text{bkg},i}(t, b_{\pm,\text{bkg},i}), \tag{123}$$

where the background distributions, $\mathcal{B}_{\pm,\text{bkg},i}$, provide a description of the various background events in the sample. The fraction of backgrounds, $f_{\text{bkg},i}$, and the parameters describing it, $b_{\pm,\text{bkg},i}$, have been obtained from data control samples whenever possible.

For each tagging category the time distribution of the background is empirically described by three components, each with its own dilution:

- Zero lifetime component: $\mathcal{B}_{\pm,1} = (1 + D'_1) \otimes \mathcal{R}$.
- Non-zero lifetime non-mixing component: $\mathcal{B}_{\pm,2} = \frac{1}{2}\Gamma_2\, e^{-\Gamma_2|t|}(1 + D'_2) \otimes \mathcal{R}$.
- Non-zero lifetime mixing component: $\mathcal{B}_{\pm,3} = \frac{1}{2}\Gamma_3\, e^{-\Gamma_3|t|}(1 + D'_3\cos(\Delta m_3 t)) \otimes \mathcal{R}$.

A likelihood fit to the control samples determines the fraction of each time component in the background along with its characteristics. The large number of parameters is in order to describe the background shape as well as possible.

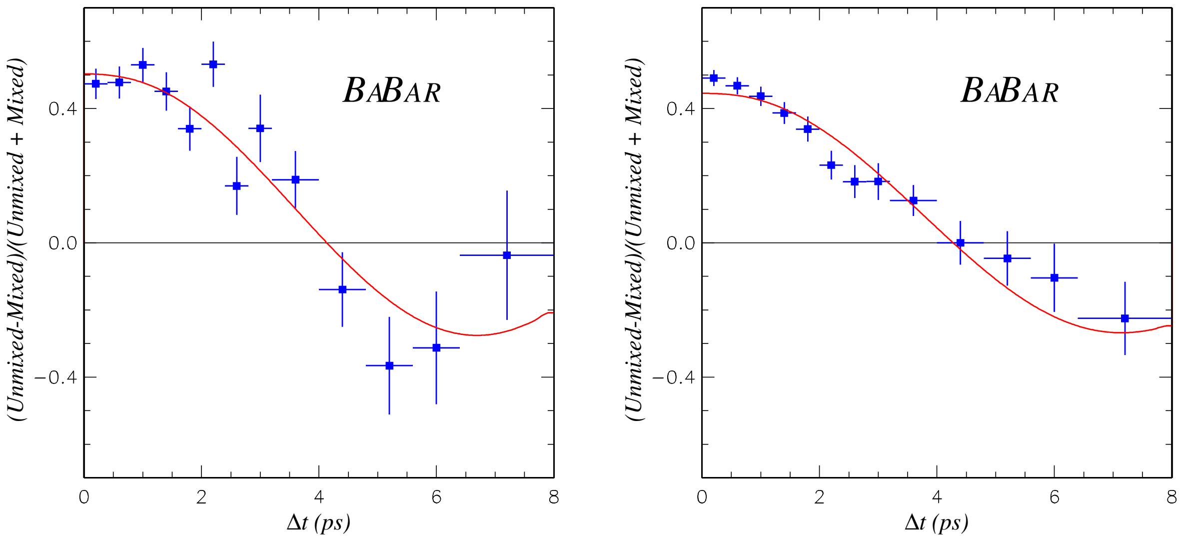


Figure 43: Time-dependent asymmetry between unmixed and mixed events. The likelihood fit to the mixing rate of the ratio $(N_{\rm unmixed} - N_{\rm mixed})/(N_{\rm unmixed} + N_{\rm mixed})$ for the hadronic sample (left) yields $\Delta m_d = 0.516 \pm 0.031$ (stat.) $\pm\, 0.018$ (syst.) ps^{-1}, while the corresponding result for the semileptonic sample (right) is $\Delta m_d = 0.508 \pm 0.020$ (stat.) $\pm\, 0.022$ (syst.) ps^{-1} [63].

The result of the likelihood fit is shown in Fig. 43, where it has been superimposed on a binned plot of the mixing asymmetry for the hadronic and semileptonic samples separately. The obtained value for Δm_d combining the two results is:

$$\boxed{\Delta m_d \;=\; 0.512 \pm 0.017\ (\text{stat.}) \pm 0.022\ (\text{syst.})\ \text{ps}^{-1}} \tag{124}$$

which is consistent with the PDG value of $\Delta m_d = 0.472 \pm 0.017\ \text{ps}^{-1}$.

12.2 *CP* results

Using the dilutions and the resolution function obtained in previous sections, the sample of tagged $B^0/\overline{B}{}^0 \to J/\psi K^0_S$ and $B^0/\overline{B}{}^0 \to \psi(2S)K^0_S$ events, described in section 9.3, has been fitted, using the maximum likelihood fit and checked with the Kin and the OneBin method. The results and their statistical uncertainty were as follows:

$$\boxed{\begin{array}{lcr} \sin(2\beta)_{\rm likelihood} & = & 0.12 \pm 0.37, \\ \sin(2\beta)_{\rm Kin} & = & 0.12 \pm 0.37, \\ \sin(2\beta)_{\rm OneBin} & = & -0.11 \pm 0.41. \end{array}} \tag{125}$$

The actual result is a little surprising in two ways. In the first place because of its low value, since the expected value, obtained from fitting the CKM matrix, is $\sin(2\beta) = 0.70 \pm 0.09$[36]. However, the obtained result, which is consistant with zero, is also fully compatible with the expected value of 0.70. The $B \to J/\Psi K_S^0$ is such a clean channel, that diluting backgrounds are hard to imagine. Both the mass spectrum of dileptons in the 3.1 GeV range and the reconstructed B-mass spectrum shows little or no background and are well understood.

Secondly, the result is surprising because of the large statistical error. Given the size of the data sample, the expected statistical uncertainty is 0.32, and the chance of getting an uncertainty of 0.37 or higher can be estimated to be 5 %, which is small but not unlikely. The chance of obtaining a lower likelihood value, given the same number of events, is 20 %.

The fact that the Kin method gives the same result is not surprising (see Fig. 42b) but reassuring. The chance of having made a mistake diminishes, when an independent calculation yields the same result. The Kin result is shown in Fig. 45, where a slight asymmetry can be seen, but the result is obviously compatible with zero.

Given the somewhat high statistical uncertainty in the likelihood fit of 0.37, the width of the distribution of difference in result between the OneBin and the likelihood fit becomes 0.19 (compare with the width 0.16 in Fig. 42b). Since the OneBin result lies within $(0.12 - (-0.11))/0.19 \simeq 1.2$ sigma of the likelihood result, they are perfectly consistant. The additional uncertainty in the OneBin method is in perfect agreement with expectations.

Thus, the OneBin method serves the purpose of checking the result of the likelihood fit. Furthermore, in the present situation of a low value of $\sin(2\beta)$, the OneBin method reassures the low result, since one by simply counting the number of events on each side of $t = 0$ obtains a negative result for $\sin(2\beta)$.

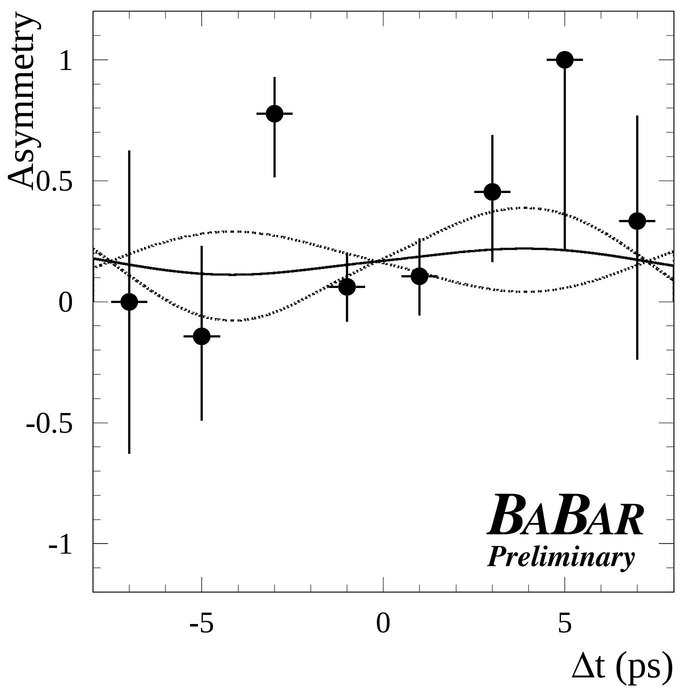


Figure 44: The asymmetry $(N_{B^0} - N_{\overline{B}^0})/(N_{B^0} + N_{\overline{B}^0})$ with binomial errors, as a function of t. The result of the likelihood fit has been superimposed along with one standard deviation curves. The offset of the curve is due to an unequal number of B^0 and $\overline{B}^0$ tagged events (70 and 50 respectively). The χ^2/dof. for the binned asymmetry is 9.2/7 [62].

[36] The error should be considered with caution, as it is mostly due to theoretical uncertainties, thus not Gaussian.

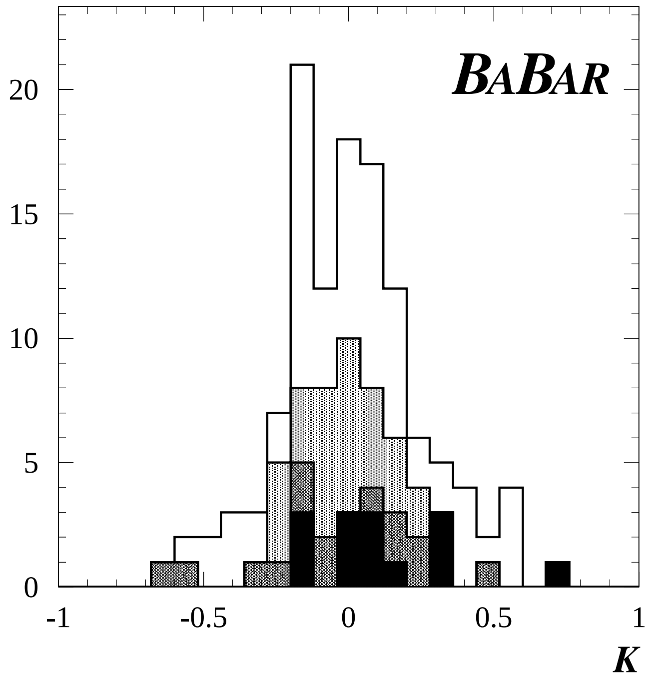


Figure 45: Distribution of the KIN variable. The slight asymmetry of the Kin distribution seen in the plot, is consistent with zero. The contributions are divided into tagging categories; black for lepton, white for kaon, dark gray for NT1, and light gray for NT2 [62].

12.2.1 Checks

Many cross checks of the fitting procedure and result have been made. Several fitting algorithms have been used, and many validation tests have been carried out. In order to check the result further, the CP sample has been subdivided into different CP channels and tagging categories, and the $\sin(2\beta)$ result for each sub-sample can be seen in Table 17. The subdivision shows no odd behavior in the result, and the χ^2/dof. values are very probable. Subdivision according to t is shown in Fig. 44, where the fitted asymmetry with errors is superimposed, and once again, no strange behavior is observed.

Sample	$\sin(2\beta)$	
Entire CP sample	$\mathbf{0.12 \pm 0.37}$	
$J/\psi K_S^0(\pi^+\pi^-)$	-0.10 ± 0.42	$\chi^2/\text{dof.} = 1.13/1.$
$J/\psi K_S^0(\pi^0\pi^0) + \psi(2S)K_S^0$	0.87 ± 0.81	
Lepton	1.6 ± 1.0	$\chi^2/\text{dof.} = 3.55/3.$
Kaon	0.14 ± 0.47	
NT1	-0.59 ± 0.87	
NT2	-0.96 ± 1.30	

Table 17: Result for $\sin(2\beta)$ when subdividing CP sample [62].

12.2.2 Control channels

As an additional check, channels which should not show any asymmetry have been fitted with exactly the same maximum-likelihood procedure as the CP sample, to see if they exhibited an apparent CP-asymmetry. The PDG values for the lifetime, τ_{B^0}, and oscillation frequency, Δm_d, were used, and the results are shown in Table 18. The control samples all have apparent CP-asymmetries consistent with zero, and the distribution of deviations is likely ($\chi^2/\text{dof.} = 4.68/4$).

Sample	Apparent CP-asymmetry
Hadronic charged B decays	0.03 ± 0.07
Hadronic neutral B decays	-0.01 ± 0.08
$J/\psi K^+$	0.13 ± 0.14
$J/\psi K^{*0}(K^+\pi^-)$	0.49 ± 0.26

Table 18: Result of fitting control channels. The control channels should not show any CP asymmetry, and it therefore serves as a check to see if their asymmetries are consistent with zero [62].

12.2.3 Systematic uncertainties

The systematic uncertainties in the $\sin(2\beta)$ measurement are naturally many, but given the modest size of the CP sample, the statistical uncertainty is by far dominating. Nevertheless, the systematic errors have to be investigated in order to assure their smallness.

Naturally, uncertainties in τ_{B^0} and Δm_d will affect the $\sin(2\beta)$ measurement, and uncertainties in the resolution function will do the same. While the Δz resolution is not crucial, control of the biases is, as these may induce apparent CP asymmetries, and for this reason the biases are carefully watched.

The largest uncertainties are from the tagging efficiencies. The measurement of dilutions is the largest single contribution to the systematic error ,but as the statistics improve, it will decrease. However, one assumes that the dilutions, D, in Eq. (105) and Eq. (104) are the same, i.e. the wrong tag fractions being the same for mixing and CP events. This turns out *not* to be the case, mostly due to the difference in the backgrounds for the two samples. Consequently, there is a systematic error from this, which is also large. To a certain extent, this can be corrected for by MC simulation. However, this is undesirable, and so other methods have been thought of, such as using the selftagging $J/\psi K^{*0}(K^+\pi^-)$ channel for measuring the dilutions, although that will arise new difficulties. Further systematic errors from differences between dilutions for B^0 and $\overline{B}^0$ and backgrounds in the CP sample have been checked, and they are small.

As the systematic uncertainties described above are believed to be independent, they have been added in quadrature. The size of each effect and the total systematic error can be seen in Table 19 below. The $\sin(2\beta)$ result is consistent with the world average [37], and currently statistically limited. The expectation is, that the error will drop significantly over the coming years, as the PEP-II accelerator is providing excellent luminosities.

Source of uncertainty	Effect on $\sin(2\beta)$
τ_{B^0}	0.002
Δm_d	0.015
Δz resolution for CP sample	0.019
Time-resolution bias for CP sample	0.047
Measurement of mistag fractions	0.053
Difference in D for CP and non-CP samples	0.050
Difference in D for B^0 and $\overline{B}^0$	0.005
Background in CP sample	0.015
Total systematic uncertainty	**0.091**

Table 19: Source of systematic uncertainty [62].

13 Progress

Since the Osaka conference in July 2000, both the understanding of the detector and the various analyses have developed significantly, aided by the increase of the data sample from 9.8 fb^{-1} to 23.9 fb^{-1}. The progress in the two analyses is shortly described below.

13.1 The TagMix analysis

The vertexing algorithm has improved the reconstruction of neutral particles with non-zero lifetime, thus lowering the biases and their uncertainties. Also the rejection of tracks contributing to the vertex χ^2 with more than 6 has been abandoned, as other more ingenious methods have been developed. This has improved the resolution function and lowered the systematic error from biases. Furthermore, a separate bias for the core Gaussian distribution is introduced for each tagging category, as the reconstruction differs somewhat.

13.1.1 The dilution measurement

The background description for the TagMix sample has improved much, and the various contributions can be seen in Fig. 46 left as a function of $\cos\theta_{B,(D^*\ell)}$. The D^{**} fit for the charged background is continuously improving and in very good agreement with Monte Carlo. The new fits can be seen in Fig. 46 right.

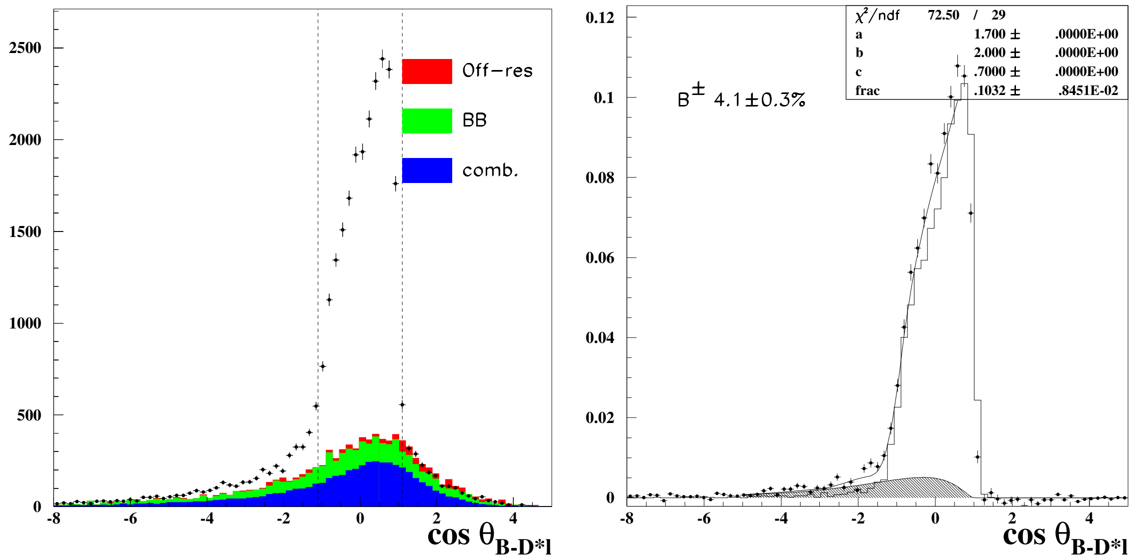


Figure 46: Semileptonic mixing sample with backgrounds as a function of $\cos\theta_{B,(D^*\ell)}$. The left plot shows the estimated amount of background from $q\bar{q}$ production, other $B\overline{B}$ channels and combinatorial background. Some background can not be accounted for by this, and is believed to be from charged B decays, which is what is fitted for in the right plot.

The current status of the dilution measurement with the OneBin method is such, that the statistical and the systematic uncertainties are about equal. One third of the systematic error is due to the resolution function, and so it has been considered going back to the CountAll method, as this would increase the statistical error by about 30 %, thus "trading" systematic error for statistical, which is better understood.

The decreasing errors have shown differences between the dilutions obtained from the hadronic and semileptonic samples, which implies that further work is needed. Also, small differences between the dilutions for B^0 and $\overline{B}^0$ are observed, and so they are fitted separately, resulting in eight instead of four dilutions.

13.1.2 The mixing measurement

The increased amount of data has reduced the statistical error, and concurrently the improved background description and bias uncertainty has reduced the systematic uncertainties. Compared to Osaka, the errors have decreased as $0.015 \rightarrow 0.010$ (stat.) and $0.022 \rightarrow 0.014$ (syst.). Both a binned and an unbinned fit is performed, providing a crosscheck of the results. There are some problems with the outliers and the fitted fraction of charged B-mesons, but almost 2 % precision seems achievable from the Run-I data, thus giving a very competitive result. An analysis weighting each event is also in progress.

13.2 The $\sin(2\beta)$ analysis

We do not care — we fit for it!

[Riccardo Faccini, coordinator of BaBar's $\sin(2\beta)$ analysis]

With the increased statistics and improved understanding of the detector and the analysis, the number of parameters to be included in a global likelihood fit[37] has increased and now encompasses:

- 8 signal dilutions (4 tagging categories $\times$ 2 flavors $(B^0/\overline{B}^0)$).
- 9 signal resolution parameters (2 widths, 2 fractions, 1 common bias and 4 different biases).
- 4 mixing, zero lifetime background fractions.
- 1 mixing, non-zero lifetime background (fraction fixed, effective lifetime fitted).
- 8 mixing background dilutions (assumed different for the two background component).
- 1 non-mixing background fraction.
- 3 background resolution parameters (only one Gaussian distribution).

Including $\sin(2\beta)$ this constitutes a 35 parameter fit, which witnesses the complexity of the analysis. The Run-I sample of golden and silver CP events is now 469 ± 23 with a purity of 95 ± 1 % and the expected statistical error on $\sin(2\beta)$ from fitting these events is 0.18.

A further improvement of the analysis can be obtained by making the purity of the sample a function of the distance between the two B-decays. A higher purity would be obtained where the asymmetry, and thus the statistical power of the events, is the largest. This especially applies to the $\psi(2S)K^0_S$ channel, where the reconstruction of the $\psi(2S)$ in the $J/\psi\pi^+\pi^-$ channel has some background (5.5 % vs. 1.5 % for the $J/\psi K^0_S(\pi^+\pi^-, \pi^0\pi^0)$ channels). Though the new $\sin(2\beta)$ fit has many free parameters, inclusion of the above feature would add information to the fit. This was discussed with the $\sin(2\beta)$ analysis coordinator during the December collaboration meeting [65], and the idea is now being pursued [66].

[37] A simultaneous fit to all the data samples extracting $\sin(2\beta)$, dilutions, resolutions, etc. at the same time.

14 Conclusion

CP conservation requires that $\lambda_{f_{CP}} = 1$. The experimentally most accessible test of this in the B-system is to search for an imaginary part of $\lambda_{f_{CP}}$ in $b \to c\bar{c}s$ decays, which at the same time can be directly related to CKM matrix elements. The requirement of a time-dependent measurement motivates an asymmetric e^+e^- collider running at the $\Upsilon(4S)$ resonance, and the subsequent analysis involves flavor tagging and reconstruction of $B \to J/\psi K^0_S$ and related decays.

The pursue of exactly this measurement has been presented. Based on the Standard Model, B-physics phenomenology and CP violation has been thoroughly described, and the relevant decay distributions calculated. Both the design and the performance of the PEP-II B-factory and the BaBar detector have then been accounted for. Finally, the complete event selection and $\sin(2\beta)$ analyses, including tagging and mixing measurements along with alternative yet validated fitting methods, have been presented.

The measured lifetime and mixing rate of the neutral B-meson:

$$\tau_{B^0} = 1.506 \pm 0.052 \text{ (stat.)} \pm 0.029 \text{ (syst.) ps,} \tag{126}$$
$$\Delta m_d = 0.512 \pm 0.017 \text{ (stat.)} \pm 0.022 \text{ (syst.) ps}^{-1} \tag{127}$$

are in agreement with the PDG values. The agreement serves as a check of consistancy for these novel measurements, but soon the analyses will have matured and the world average will be dominated by the BaBar measurements along with those from Belle. For comparison, Belle has a luminosity of $2.2 \times 10^{33}\,\text{cm}^{-2}\text{s}^{-1}$ and reached $10\,\text{fb}^{-1}$ in December 2000.

The effective tagging efficiency obtained by BaBar, $Q = 27.9 \pm 0.5\%$ (statistical error only), is a great achievement, and should be compared to the equivalent result of Belle, which is $Q \simeq 22.0\%$ [67]. The use of several methods ensures the correctness of the result, and the high efficiency preserves the statistics for the $\sin(2\beta)$ measurement. The obtained CP asymmetry (July 2000) was:

$$\boxed{\sin(2\beta) = 0.12 \pm 0.37 \text{ (stat.)} \pm 0.09 \text{ (syst.)}} \tag{128}$$

As mentioned, the statistical uncertainty is somewhat higher than expected, but this is attributed a statistical fluctuation, as all the validating tests have surported the result. The actual $\sin 2\beta$ result should be compared with those of other experiments listed in Table 20:

Collaboration:	Published:	Value for $\sin(2\beta)$:
OPAL [68]	Jan. 98	$3.2^{+2.0}_{-1.8}$(stat.) $\pm$ 0.5(syst.)
CDF [69]	Aug. 99	$0.79^{+0.41}_{-0.44}$(stat.)
ALEPH [70]	Sep. 00	$0.84^{+0.82}_{-1.04}$(stat.) $\pm$ 0.16(syst.)
Belle [67]	Dec. 00	$0.45^{+0.43}_{-0.44}$(stat.)$^{+0.07}_{-0.09}$(syst.)
World average		0.49 ± 0.23
CKM-fit [45]		0.70 ± 0.09

Table 20: Measurements of $\sin(2\beta)$.

The new $\sin(2\beta)$ result from BaBar was published at a SLAC seminar on the 12$^{\text{th}}$ of February 2001. The new value was:

$$\boxed{\sin(2\beta) = 0.34 \pm 0.20 \text{ (stat.)} \pm 0.05 \text{ (syst.)}} \tag{129}$$

Though the BaBar result is low, it is in agreement with both the world average and the value obtained from fitting the CKM matrix. However, the statistical uncertainty is soon reaching a level where possible discrepancies become visible, thus commencing the real test of CP violation in the SM framework.

With the increasing luminosity, the precision of the BaBar measurements and the subsequent constraining of the CKM matrix will improve. If the integrated luminosity predictions described in section 7.3.3 hold true, BaBar should have around 400 fb^{-1} by the end of year 2005. Simply scaling the current uncertainty on the $\sin 2\beta$ measurement, the precision will then be 0.05. Assuming that Belle will then have around 250 $>$: fb^{-1}, their error should be 0.07, again scaling from current results.

By then the Tevatron will have been running for some years, and CDF and D0 will contribute with measurements. The goal for the first two years is to reach 2 fb^{-1} (recall that the $b\bar{b}$ cross section is much higher than at e^+e^- colliders), and so 5 fb^{-1} should be reachable by 2005. Their uncertainty in measuring $\sin(2\beta)$ will then be 0.06, again obtained from scaling. Given this level of precision, the $\sin(2\beta)$ measurements will constitute a serious test of the SM.

By then, Δm_d will be measured with a precision at the per mille level. However, that in itself does not constrain the CKM matrix much, as the dominant sources of errors in relating this number to CKM parameters are theoretical. However, by then CDF and D0 will have measured the mixing rate in the B_s-system, Δm_d, and the ratio of the two gives a strong constraint. In addition, many other measurements will be made by both the B-factories and the Tevatron, further constraining the CKM matrix. Figure 47 below shows a possible 2005 scenario for the CKM fit:

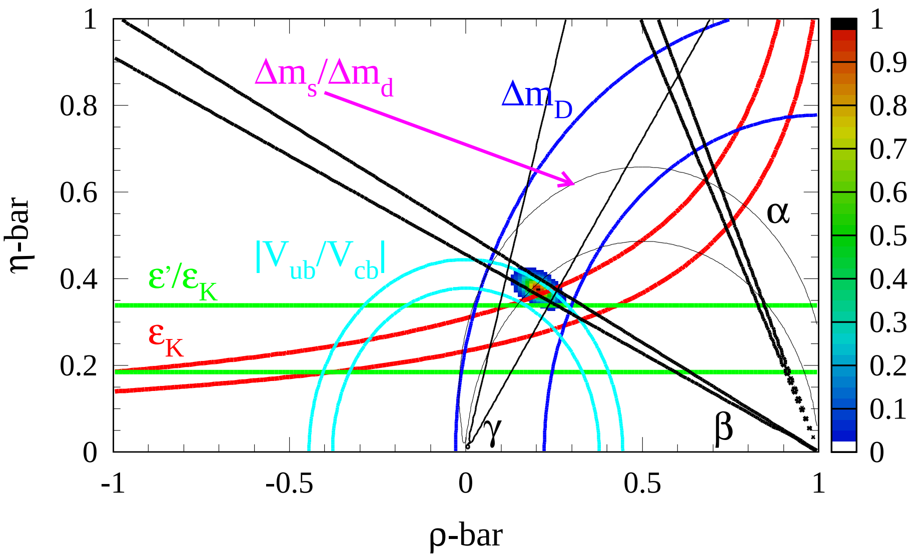


Figure 47: Possible scenario for the knowledge of the unitary triangle in 2005 [44].

Beyond 2005, the LHC experiments — LHC-b in particular — at CERN will be able to perform the ultimate measurements, as both cross section and luminosity will be unprecedented. In addition, the recently approved BTEV experiment at the Tevatron and dedicated experiments measuring the $K \to \pi\nu\bar{\nu}$ branching ratios will contribute to the accurate determination of the CKM matrix. These precision measurements will allow for the decisive test, whether CP violation can be explained within the Standard Model. Should the fit to the four free parameters of the CKM matrix fail to converge, the mixing between quarks then cannot be described by the CKM matrix alone, and new physics will have to be introduced.

15 Acknowledgement

First of all, I wish to thank my two excellent supervisors, John Renner Hansen and François Le Diberder, without whom this master thesis would never have existed. They have guided me on illuminated paths where I found only dark emptiness, and they have broadened my horizon far beyond the realms of this thesis.

Thank you, John, both for your wisdom in proposing BaBar as the subject of my master thesis and for you kind support and help in all matters. Thank you, François, not only for allowing me to work with you on BaBar, but also for your endless capacities and joyful mood, and for all the wonderful days of discussing physics at your blackboard in Paris.

It has been both a pleasure and an honor for me to be your student.

I am much indebted to the French BaBar group at LPNHE and LAL in Paris. It has been a great pleasure to work with Lydia Roos. Thank you, Lydia, for generously sharing your office with me, and for all the kindness and helpfulness you have shown me during my stay in Paris. I also thank Stephane Plaszczynski and Marie Helene Schune. It has truly been a pleasure to collaborate with you and enjoy the quality of your work.

In addition, I would like to thank Sandrine Laplace. It was great to share with you the burdens and pleasures of being a student. Thanks to Andreas Höcker and Heiko Lacker for the interesting and amusing discussions of CKM-fitting. May our collaboration not end here.

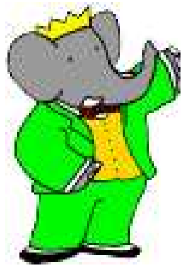
I am also indebted to Bob Cahn, who introduced me to B-physics and the BaBar experiment. Thank you, for accepting me at your office in Berkeley and for establishing the contact with Paris, which has led to all this. Thanks to the BaBar collaboration for their hospitality, and to PEP-II for providing the excellent conditions, which have made the experiment so successful.

Thanks to the HEP group at NBI for an inspiring working environment, not least to Jørn Dines Hansen, Peter Hansen, Brian Aagaard and Peter Zeiler Skands for many discussions of B-physics. I also wish to express my thanks to Christian Holm, Thomas Kittelman and Anders Wäänänen for their guidance and help in Latex and Unix matters. In addition, thanks to Morten Kjærulf, Peter Andersen and Henrik Nielsen for a most enjoyable student environment.

My thanks to friends at Regensen. I have been grateful to live among you, sharing the spirit and pleasures of our student life.

Thanks to Sebastien Fabbro, Christian Pagh and Daniel Whiteson for their great comradeship. Without you, working at LPNHE and being in Paris would never have been such a great pleasure. A special thank goes to Ane Corfitz Sørensen for the most enjoyable months in the Spring of Paris. You brought me both inspiration and happiness, for which I am endlessly indebted you.

Last but not least a grateful thanks to my wonderful family, who have not only endured my sometimes exaggerated passion for particle physics, but also accepted and supported me in pursuing the subject.



The credits are to be shared with those I've worked with, while the errors are mine alone.

References

[1] T. D. Lee and C. N. Yang. *Phys. Rev.* **104**, *254*, 1956.

[2] C. S. Wu *et al*. *Experimental Test of Parity Violation in Beta Decay*. *Phys. Rev.* **105**, *1413*, 1957.

[3] R. L. Garwin, L. M. Lederman and M. Weinrich. *Observation of the Failure of Conservation of Parity and Charge Conjugation in Meson Decays: the Magnetic Moment of the Free Muon*. *Phys. Rev.* **105**, *1415*, 1957.

[4] J. H. Christenson, J. Cronin, V. Fitch and R. Turlay. *Evidence for* 2π *decay of the* K_2^0 *meson*. *Phys. Rev. Lett.* **13**, *138*, 1964.

[5] N. Cabibbo. *Unitary symmetry and lepton decays*. *Phys. Rev. Lett.,* **10**, *531*, 1963.

[6] S. L. Glashow, J. Iliopoulos and L. Maiani. *Weak Interactions with Lepton-Hadron Symmetry*. *Phys. Rev. D2, 1285-1292*, 1970.

[7] M. Kobayashi and T. Maskawa. *CP Violation in the Renormalizable Theory of Weak Interaction*. *Prog. Theor. Phys.* **49**, *652*, 1973.

[8] S. W. Herb *et al*. *Observation of a Dimuon Resonance at* 9.5 *GeV in* 400 *GeV Proton Nucleus Collisions*. *Phys. Rev. Lett.* **39**, *252*, 1977.

[9] A. D. Sakharov. *JETP Lett.* **5**, *24*, 1967.

[10] G. Steigman. *Ann. Rev. Astr. Ap.* **14**, *339*, 1976.

[11] A. Cohen, A. DeRujula and S. Glashow. *A matter-antimatter Univers?*. *Astrophys. J.* **495**, *539*, 1998.

[12] M. Dine *et al*. *The electroweak theory and baryogenesis*. *Stanford Linear Acceleator preprint SLAC-PUB-5741*, 1992.

[13] S. L. Lloyd. `http://hepwww.ph.qmw.ac.uk/epp/lectures.html`.

[14] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Addison-Wesley Publishing Company, 1995.

[15] F. Halzen and A. D. Martin. *An Introduction Course in Modern Particle Physics*. John Wiley & Sons, Inc. 1984.

[16] BaBar Collaboration. *The BaBar Physics Book*. SLAC, 1998 . `http://www.slac.stanford.edu/pubs/slacreports/slac-r-504.html`.

[17] T. C. Petersen. *Simple methods for extracting* $(1-2w_{tag})$ *and sin(2β)*, 2000. BAD#51.

[18] S. Plaszczynski, M. H. Schune, F. L. Diberder and T. C. Petersen. *Measurement of the mistag fraction*, 2000. BAD#22.

[19] A. L. Mackay. *Dictionary of Scientific Quotations*. Institute of Physics Publishing, 1991.

[20] A. J. Buras and R. Fleischer. *Quark Mixing, CP Violation and Rare Decays After the Top Quark Discovery*. In A. J. Buras and M. Lindner, Editors, *Heavy Flavours II*. World Scientific, 1997.

[21] Y. Fukuda *et al.* *Evidence for oscillation of atmospheric neutrinos*. *Phys. Rev. Lett.* **81**, *1562-1567*, 1998.

[22] Z. Maki, M. Nakagawa and S. Sakata. *Prog. Theo. Phys.* **28**, *870*, 1962.

[23] R. D. Peccei. In C. Jarlskog, Editor, *CP violation*. World Scientific, 1989.

[24] F. J. Gilman, K. Kleinknecht and B. Renk. *The Cabibbo–Kobayashi–Maskawa mixing matrix*. *The European Physical Journal C15, 1*, 2000.

[25] L. Wolfenstein. *Parametrization of the Kobayashi-Maskawa Matrix*. *Phys. Rev. Lett.* **51**, *1945*, 1983.

[26] C. Jarlskog. *Phys. Rev. Lett.* **55**, *1039*, 1985.

[27] T. Inami and C. S. Lim. *Prog. Th. Phys.* **65**, *297*, 1981.

[28] L. Alvarez-Gaumé, C. Kounnas, S. Lola and P. Pavlopoulos. *Direct $\mathcal{T}$-violation measurements and $\mathcal{T}$-odd effects in decay experiments*. *Talk at XXXIV Rencontres de Moriond, hep-ph/9903458*, 1999.

[29] T. C. Petersen. *The λ^4-theorem*. Private note.

[30] H. Quinn and A. Sanda. *CP violation in B decay – Standard Model predictions*. *The European Physical Journal C15, 1*, 2000.

[31] M. Neubert. BaBar CKM working group meeting, Dec. 2000.

[32] A. E. Snyder and H. R. Quinn. *Measuring CP asymmetry in $B \to \rho\pi$ decays without ambiguities*. *Phys. Rev. D48*, 1993.

[33] S. Versillé. *La violation de CP dans BaBar: étiquetage des mésons B et étude de canal $B \to 3\pi$.*. Ph.D. Thesis, Université Paris XI Orsay, 1999.

[34] R. N. Cahn and T. C. Petersen. *Possible measurement of alpha in $B \to \rho\pi$ decays*, June 1999.

[35] S. Laplace. *Ph.D. thesis in progress*, Dec. 2000. Private communications.

[36] R. Aleksan. Private communications.

[37] D. E. Groom *et al.* *Review of Particle Physics*. *The European Physical Journal C15, 1*, 2000.

[38] L. Wolfenstein. *Phys. Rev. Lett.* **13**, *562*, 1964.

[39] Y. Nir. *Flavor physics and CP violation*. In *1998 European School of High-Energy Physics, CERN 99-04*.

[40] E787 Collaboration. *Evidence for the decay $K^+ \to pi^+\nu\bar{\nu}$*. 1997.

[41] A. Ali and D. London. *Eur. Phys. J, C9*, 1999.

[42] S. Plaszcynski and M. H. Schune. *hep-ph/9911280*, 1999.

[43] F. Caravaglios, F. Parodi, P. Roudeau and A. Stocchi. *hep-ph/0002171*, 2000.

[44] A. Höcker, H. Lacker, S. Laplace and F. LeDiberder. *CKM-fitter — Constraining the unitary triangle*, 2000. `http://www.slac.stanford.edu/ laplace/Workshop.html`.

[45] A. Höcker, H. Lacker, S. Laplace and F. LeDiberder. *Constraining the CKM matrix*, 2001. BAD#153.

[46] K. Kleinknecht. *Private communication.* Prerow, October 2000.

[47] BaBar Collaboration. *Technical Design Report for the BaBar Detector*. 1995.

[48] BaBar Collaboration. *The first year of the BaBar experiment at PEP-II. Preliminary Draft – submitted to the XXXth ICHEP Osaka, Japan*, July 26, 2000.

[49] *Physics at a second generation B-factory.* `http://www.physics.purdue.edu/10E34/`.

[50] E. Frank. *Architecture of the BaBar Level 3 Software Trigger*. Computing in High Energy Physics, Illinois, 1998.

[51] P. Billoir. *Nucl. Inst. Meth. 225, 352*, 1984.

[52] G. C. Fox and S. Wolfram. *Nucl. Phys. B149, 413*, 1979.

[53] D. Kirkby and C. Cheng. *Backgrounds in the $D^*\ell\nu$ Sample*, Nov. 2000 . `http://www.slac.stanford.edu/BFROOT/www/Physics/Analysis/AWG/EHBDOC/-Meetings/16Nov2000/Dstarlnu/BG.html`.

[54] J. D. Richman and P. Burchat. *Leptonic and semi-leptonic decays of charm and bottom hadrons. Rev. Mod. Phys.* **67**, *893-976*, 1995.

[55] CLEO Collaboration. *Measurement of the $B^0 \to D^*\ell\bar{\nu}$ branching ratio and $|V_{cb}|$. Phys. Rev.* **D51**, *1014-1033*, 1995.

[56] ALEPH Collaboration. *Z. Phys.* **C73** *601*, 1997.

[57] DELPHI Collaboration. *Phys. Lett.* **B475** *407*, 1999.

[58] S. Plaszczynski. *TagMix results with counting methods, Dec. 2000* . `http://www.slac.stanford.edu/BFROOT/www/Organization/CollabMtgs/2000/-detDec2000/Tues4a/plaszczynski.ps`.

[59] ALEPH, CDF, DELPHI, L3, OPAL and SLC Collaborations. *Combined results on b-hadron production rates, lifetimes, oscillations and semileptonic decays. LEPHFS note 99-02 (draft 2)*, 10th March 2000.

[60] F. James and M. Roos. *Comput. Phys. Commun. bf 10, 343*, 1975.

[61] S. Henrot-Versillé and F. L. Diberder. *Kin: the optimal variable for CP-violation. hep-ex/0007025*, 2000.

[62] BaBar Collaboration. *A study of time-dependent CP-violating asymmetries in $B^0 \to J/\psi\, K^0_S$ and $B^0 \to \psi(2S)\, K^0_S$ decays.. ICHEP 2000 Osaka, SLAC-PUB-8540*, 2000.

[63] BaBar Collaboration. *A measurement of the $B^0\overline{B}^0$ oscillation frequency and determination of flavor-tagging efficiency using semileptonic and hadronic B^0 decays. ICHEP 2000 Osaka, SLAC-PUB-8530*, 2000.

[64] S. Plaszczynski and M. H. Schune. Private communications.

[65] R. Faccini. Private communications, Dec. 2000.

[66] A. Olivas. *Ph.D. thesis in progress*, Dec. 2000. Private communications.

[67] Belle Collaboration. *Measurement of sin($2\phi_1$) at Belle. Submitted to Int. J. Mod. Phys. A, hep-ex/0012028*, 2000.

[68] OPAL Collaboration. *Investigation of CP asymmetry in $B^0 \to J/\psi K^0_S$ decays at LEP. Eur. Phys. J. C5:379-388, hep-ex/9801022*, 1998.

[69] CDF Collaboration. *Measurement of the CP-violation parameter sin(2β) in $B^0 \to J/\psi K^0_S$ decays. Phys. Rev. D61, 072005, hep-ex/9908055*, 2000.

[70] ALEPH Collaboration. *Study of CP asymmetry in $B^0 \to J/\psi K^0_S$ decays in ALEPH. Phys. Lett. B492:259-274, hep-ex/0009058*, 2000.

[71] S. Plaszczynski, M. H. Schune, F. L. Diberder, T. C. Petersen and L. Roos. *Measurement of the mistag fractions with counting methods*, 2001. BAD#151.

[72] J. Charles *et al.* *$B^0_d(t) \to DPP$ time-dependent Dalitz plots, CP-violating angles 2β, $2\beta+\gamma$, and discrete ambiguities. Phys. Lett. B,* **425**, *375*, 1998.

[73] J. Schwinger. *The theory of quantized fields II. Phys. Rev. 713-727*, 1953.

[74] G. L. Lüders. *Ann. Phys. 2*, 1957.

[75] R. L. Walsworth *et al.* *New Clock Comparison Searches for Lorentz and CPT Violation. Procedings of the 3rd International Symposium on Symmetries in Subatomic Physics*, 2000.

16 Abbreviations

Abbreviations are very frequently used in the world of experimental high energy physics, as the field is somewhat technical, and so this thesis contains some as well. They are written in full the first time they occur with the abbreviation in parenthesis, after which only the abbreviation is used. However, abbreviations of accelerators and detectors/experiments are rarely written in full. Below is a list of the abbreviations used, and also non-selfexplainatory words have been included.

BAD	BaBar Analysis Document.
BTEV	B-physics at the TEVatron. Planned experiment at Fermilab.
CDF	Collider Detector at Fermilab. Detector at Tevatron.
CERN	The European Center for Particle Physics.
CKM	Cabibbo, Kobayashi and Maskawa. Inventors of the CKM quark mixing matrix.
CM	Center-of-Mass.
CP	Charge conjugation and Parity. Abstract operators transforming particles.
CPT	Charge conjugation, Parity and Time reversal.
D0	Detector at Tevatron, Fermilab.
DCH	Drift CHamber. Subdetector of BaBar.
DIRC	Detector of Internally Reflected Ĉerenkov light. Subdetector of BaBar.
EMC	ElectroMagnetic Calorimeter. Subdetector of BaBar.
GIM	Glashow, Iliopoulos and Maiani. Inventors of the GIM-mechanism.
HEP	High Energy Physics.
HER	High Energy Ring. The electron storage ring at PEP-II.
HERA-B	B-physics at the HERA-accelerator. Experiment at DESY in Hamburg.
ICHEP	International Conference on High Energy Physics.
IFR	Instrumented Flux Return. Subdetector of BaBar.
KIN	Japanese for 'gold'. Fitting method invented by François Le Diberder.
LEP	Large Electron-Positron collider, CERN, Switzerland.
LER	Low Energy Ring. The positron storage ring at PEP-II.
LHC	Large Hadron Collider. Accelerator in progress at CERN.
LPNHE	Laboratoire de Physique Nucléaire et des Hautes Energies, Paris VI-VII.
MC	Monte Carlo. Technique for numerical integration and estimation.
MNS	Maki, Nakagana, Sakata. Inventors of the MNS lepton mixing matrix.
NBI	Niels Bohr Institute, Copenhagen.
PEP	Positron Electron Project. Storage rings at SLAC.
PID	Particle IDentification.
PMT	Photo Multiplier Tube. Instrument for detecting photons.
QCD	Quantum Chromo Dynamics. Theory of the strong force.
QED	Quantum Electro Dynamics. Theory of the electromagnetic force.
RPC	Resistive Plate Counter. Type of detector used in the IFR.
SLAC	Stanford Linear Accelerator Center, Stanford, USA.
SM	Standard Model.
SUSY	SUper SYmmetry.
SVT	Silicon Vertex Tracker. Subdetector of BaBar.
TagMix	Tagging/Mixing. Expression for the simultanious analysis of the two.
TDR	Technical Design Report.

A The *CPT* theorem

The CPT-theorem states that the requirements for a field theory to be CPT invariant are [73], [74]:

- **Lorentz invariance.**
- **Locality.**
- **Usual connection between spin and statistics.**
- **Hermitian Hamiltonian (in basic theory).**

Of course, decaying particles will not have a hermitian Hamiltonian, since the time operator is not unitary. What is meant by this last statement is the fact that the Hamiltonian governing the physics apart from the decay is hermitian.
By first inspection, all of these requirements seem to be fulfilled for all known theories, but this turns out not to be the case. When reaching the gravity scale, Lorentz invariance does not hold any longer, and some string theories are not local. Thus these theories implicitly predict CPT violation.
Experimental searches have excluded CPT violation to an impressive level, and varies from testing Lorentz invariance and the hermiticity of Hamiltonians to comparing basic quantities. Comparison of m_{K^0} to $m_{\overline{K}^0}$ and of q_{e^-} to q_{e^+} both yield a limit at the level of 10^{-18} [37], while indirect searches looking at time differences in atomic clocks have extended this limit to 10^{-31} [75].

B The GIM mechanism

From the amplitude Eq. (45) of the two dominating mixing diagrams in Fig. 3 (see section 3.1), one can see that it is of the form:

$$\mathcal{M}_{\text{mixing}} = \sum_{i,j=u,c,t} \mathcal{M}_{ij} = \sum_{i=u,c,t} \sum_{j=u,c,t} V_{id}^* V_{ib} V_{jd}^* V_{jb} \; \mathcal{M}(m_i, m_j), \tag{130}$$

that is, the amplitude equals four CKM entries multiplied by an amplitude, which only depends on the mass of the quarks in the loop. If the u, c and t quark masses were equal, then $\mathcal{M}(m_i, m_j) \equiv \overline{\mathcal{M}}$ would be equal for all quarks in the loop, and the diagrams would contribute proportionally to their CKM coupling. Due to the unitarity of V_{CKM}, they would then vanish:

$$\mathcal{M}_{\text{mixing}} = \overline{\mathcal{M}} \sum_{i=u,c,t} V_{id}^* V_{ib} \sum_{j=u,c,t} V_{jd}^* V_{jb} = 0, \quad \text{since} \sum V_{jd}^* V_{jb} = 0. \tag{131}$$

However, the quarks have very different masses, and in order to make this dependency explicit, one uses the unitarity on first and third column:

$$\sum_{i=u,c,t} V_{id}^* V_{ib} = 0 \quad \Rightarrow \quad V_{ud}^* V_{ub} = -(V_{cd}^* V_{cb} + V_{td}^* V_{tb}) \tag{132}$$

to make a *light quark rearrangement*:

$$\begin{aligned}
\mathcal{M}_{\text{mixing}} &= \sum_{i=u,c,t} V_{id}^* V_{ib} \sum_{j=u,c,t} V_{jd}^* V_{jb} \; \mathcal{M}(m_i, m_j) \\
&= \sum_{i=u,c,t} V_{id}^* V_{ib} [V_{ud}^* V_{ub} \; \mathcal{M}(m_i, m_u) + V_{cd}^* V_{cb} \; \mathcal{M}(m_i, m_c) + V_{td}^* V_{tb} \; \mathcal{M}(m_i, m_t)] \\
&= \sum_{i=u,c,t} V_{id}^* V_{ib} [V_{cd}^* V_{cb} (\mathcal{M}(m_i, m_c) - \mathcal{M}(m_i, m_u)) + V_{td}^* V_{tb} (\mathcal{M}(m_i, m_t) - \mathcal{M}(m_i, m_u))] \\
&\equiv \sum_{i=c,t} V_{id}^* V_{ib} [V_{cd}^* V_{cb} \; \overline{\mathcal{M}_c}(m_i) + V_{td}^* V_{tb} \; \overline{\mathcal{M}_t}(m_i)] \\
&\equiv \sum_{i,j=c,t} V_{id}^* V_{ib} V_{jd}^* V_{jb} \; \overline{\mathcal{M}_{ij}},
\end{aligned} \tag{133}$$

with

$$\overline{\mathcal{M}_{ij}} \equiv \overline{\mathcal{M}_i}(m_j) - \overline{\mathcal{M}_i}(m_u) \equiv \mathcal{M}_i(m_j) - \mathcal{M}_i(m_u) - \mathcal{M}_u(m_j) + \mathcal{M}_u(m_u). \tag{134}$$

The four matrix elements in the final sum only depends on differences between quark masses, since all other possible dependencies cancel due to unitarity. Since the (u-type) quark masses are not equal, the contributions do not cancel but are merely suppressed, which is called the GIM-mechanism after its inventors Glashow, Illiopoulos and Maiani (GIM) [6]. From the actual dependence of the matrix elements on mass:

$$S_0(x_q) = x_q \left[\frac{1}{4} + \frac{9}{4} \frac{1}{1-x_q} - \frac{3}{2} \frac{1}{(1-x_q)^2} - \frac{3}{2} \frac{x_q^2 \ln x_q}{(1-x_q)^3} \right], \quad x_q \equiv \frac{m_q^2}{m_W^2}, \tag{135}$$

and the V_{CKM} couplings, one can in principle deduce the relative contribution from u, c and t quarks respectively. The numbers to insert are:

$$x_u \simeq 1.4 \times 10^{-9} \qquad |V_{ud}^* V_{ub}| \simeq 0.0035 \tag{136}$$

$$x_c \simeq 0.016 \qquad |V_{cd}^* V_{cb}| \simeq 0.009 \tag{137}$$

$$x_t \simeq 4.5 \qquad |V_{td}^* V_{tb}| \simeq 0.009 \tag{138}$$

As the numbers reveal, the top quark contribution by far dominates B-mixing, which makes Δm_B a possible probe for $|V_{td}^* V_{tb}|$ and thus $|V_{td}|$[38]. In K-mixing, charm loop contributions are not negligible. This introduces long distance effects, and thus complicates the calculations of Δm_K and thereby the possibility of relating it to CKM matrix elements. In the D-system, it is the d-type quarks that enters the loop diagram, and since they all have masses much lower than m_W, mixing in the D-system is suppressed.

[38] $\Delta m_{B_d}/\Delta m_{B_s}$ is a better probe, since hadronic uncertainties to a large extend cancels out.

C The TrigBin and TwoBin methods

In addition to the OneBin method, more refined counting methods were developed. However, though these methods are more powerful, it is the simplicity of the OneBin method that is its force, and so the following methods have never been put to use.

C.1 TrigBin method

If one wants to refine the OneBin method, the events outside the cuts can be included in the computation. To include them in an efficient manner still using just *one cut*, one can divide events not according to the time, but according to $\cos(\Delta m t)$ for TagMix events and $\sin(\Delta m t)$ for CP events.

This method called *TrigBin* uses the fact that if there were no smearing, $\cos(\Delta m t)$ determines the ratio of unmixed and mixed events, and $\sin(\Delta m t)$ determines the asymmetry. By dividing the events into bins according to their value of $\cos(\Delta m t)$ (TagMix case) and $\sin(\Delta m t)$ (CP case), one exploits that at larger times, t, the number of mixed events exceeds the number of unmixed, and the asymmetry has changed side. In this way one can use the entire range of times, since they are evaluated according to the value of $\cos(\Delta m t)$ and $\sin(\Delta m t)$.

Effectively, this means that the time range is divided into regions composed of several disconnected parts, where in each part the number of events is computed, and the final result is then constructed from the parts. This is what is illustrated in Fig. 48, for (a) TagMix fitting and (b) CP fitting.

In the TagMix case, the number of unmixed and mixed events is compared in each region, whereas in the CP case the number of events with $t > 0$ is compared to the number of events with $t < 0$ in each region. In this way, the TrigBin method is essentially the OneBin method applied to each region, which in turn are determined by the value of $\cos(\Delta m t)$ or $\sin(\Delta m t)$. Once again, one finds the optimal value at which to divide the events by minimizing the statistical uncertainty, and the resulting positions of the bins are also shown in Fig. 48.

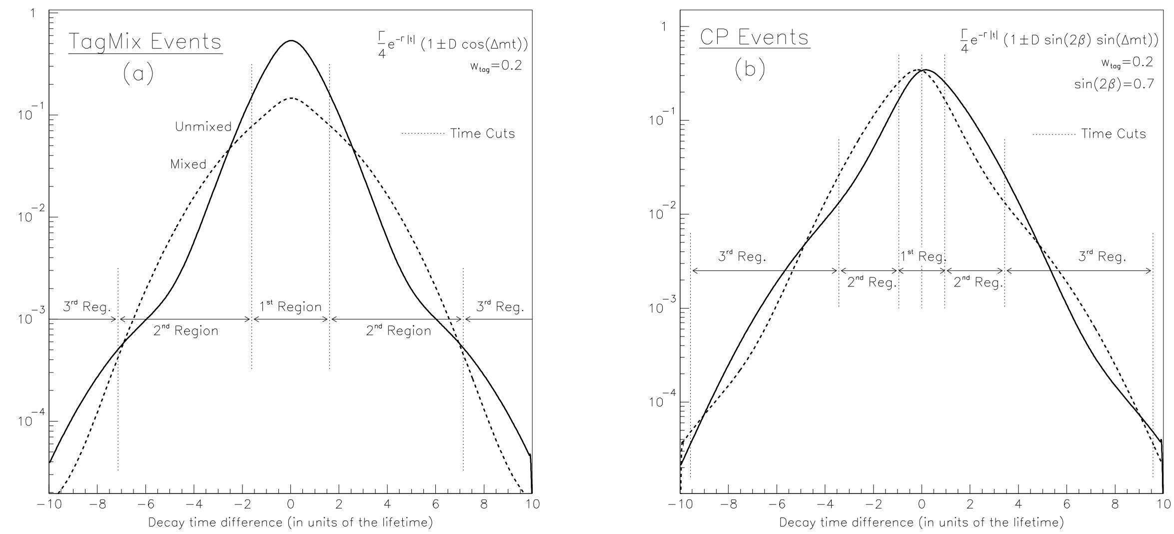


Figure 48: Effective position of bins for TrigBin method, when dividing optimally according to the value of $\cos(\Delta m t)$ (TagMix) and $\sin(\Delta m t)$ (CP). A larger time interval and logarithmic scales are used, to illustrate the idea of cutting on cos and sin.

In the TagMix case, the innermost bin doesn't change position very much compared to OneBin, but now the data outside is used. This contributes in principle, but since it has almost the same number of mixed and unmixed events, it does not have a great impact on the power of the method, as will be seen when discussing results.

In the case of CP fitting, the events with $t \sim 0$ carries only little information, but the statistics are good, and overall they contribute to some extend to the result. Using events with small t in the CP case also turns out to be the best way to include a second set of bins, which will be discussed next.

C.2 TwoBin method

Another way to improve the power of the OneBin method is to include a second set of optimally placed bins. In the TagMix case, these two bins[39] should be placed in the regions where there are more mixed events than unmixed. In the CP case, the outer region beyond the bins from the OneBin method has too little statistics, and so the most information is extracted with two bins on each side of $t = 0$, one next to the other, in effect dividing the bin of the OneBin method.

Once again the statistical uncertainty on the result, which is obtained by combining the result in each set of bins, is minimized, and the result of this is shown in Fig. 49 below.

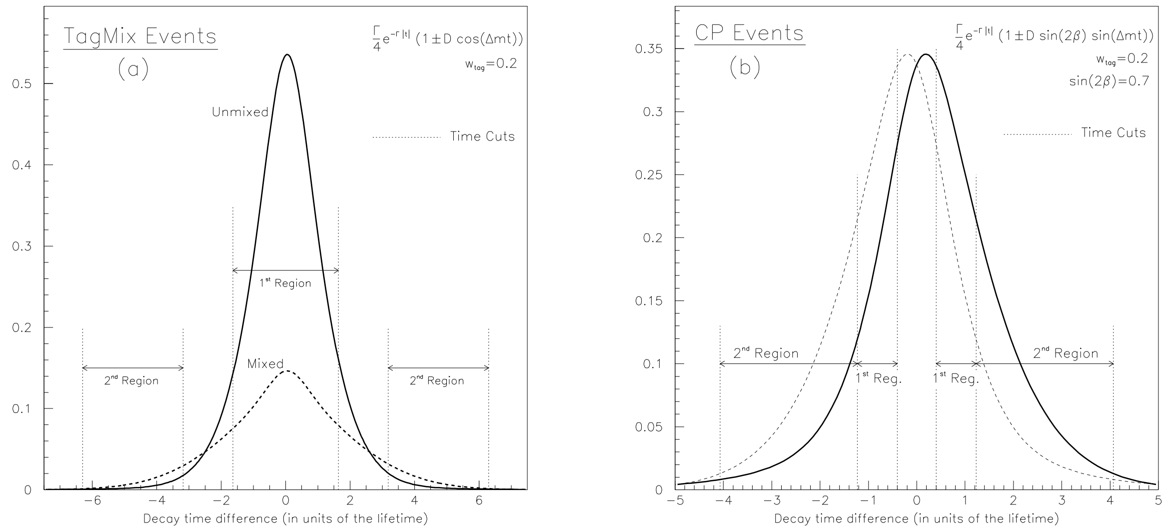


Figure 49: Optimal placing of bins for TwoBin method, again obtained by minimizing the statistical uncertainty.

As can be seen from the figure, the TwoBin method in the TagMix case discards events, which are at a point in time, where they contribute with little or no information, and so the central bin is essentially untouched compared to the OneBin method, whereas one additional bin with two parts has been placed on the sides.

The TwoBin method on CP events is able to include more of the statistically rich central region, resulting in a significant improvement of the methods power.

It should be stressed once again, that the placing of the bins is simply a result of minimizing the statistical uncertainty. As the amount of data increases and systematic uncertainties become important, these should be taken into account. This is not difficult for these simple methods, as the systematic uncertainties can be computed directly.

[39] As the distribution is symmetric, the outer bin is composed of two disconnected parts.

D Validation figures

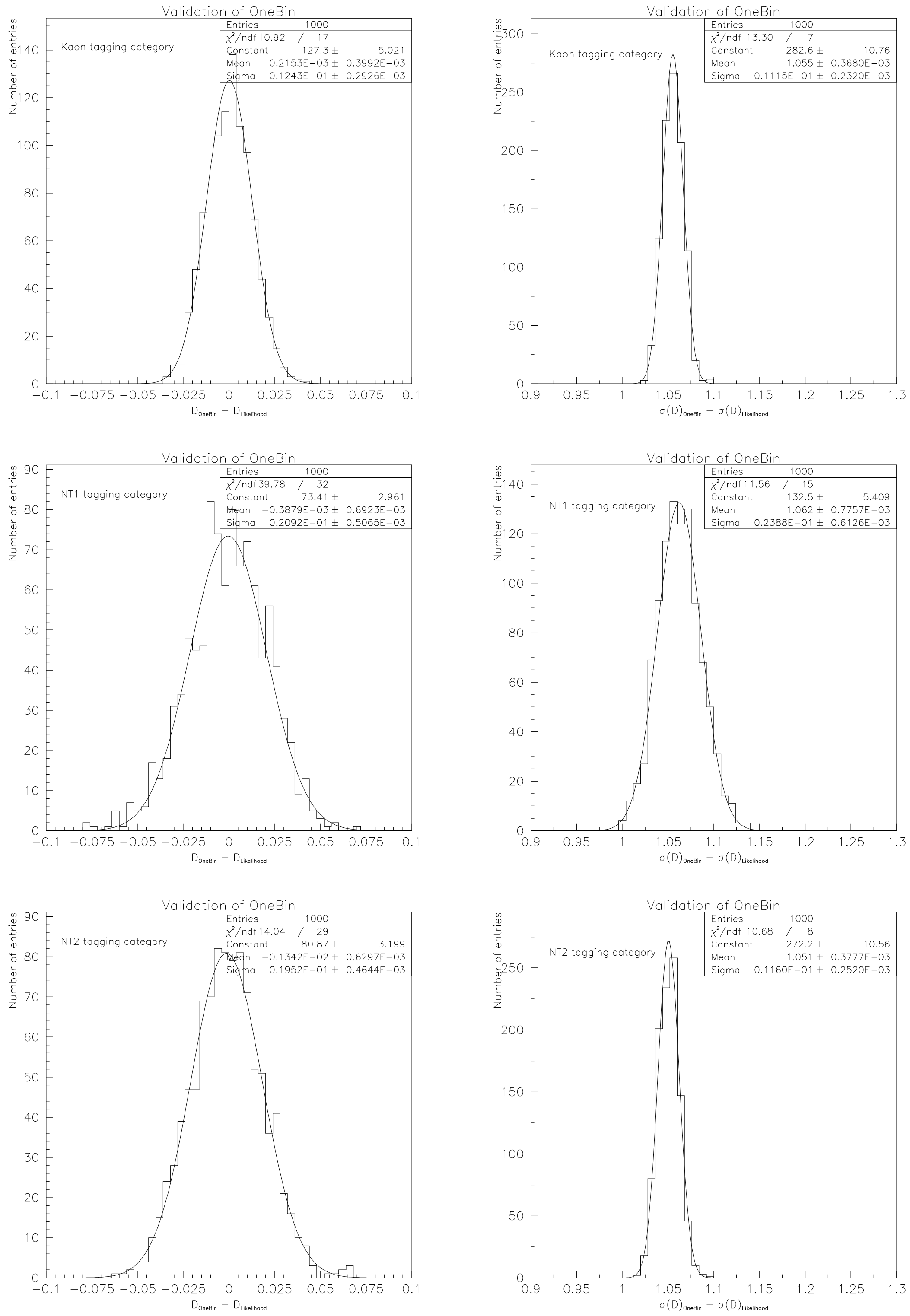


Figure 50: Validation of OneBin method for kaon, NT1 and NT2 tagging categories.