UNIVERSITY OF LJUBLJANA
FACULTY OF MATHEMATICS AND PHYSICS
DEPARTMENT OF PHYSICS

Saša Fratina

MEASUREMENT OF TIME DEPENDENT CP VIOLATION IN $B^0 \rightarrow D^+D^-$ DECAYS

Doctoral Thesis

Advisor: Doc. Dr. Samo Korpar

Ljubljana, 2007

Univerza v Ljubljani
Fakulteta za matematiko in fiziko
Oddelek za fiziko

Saša Fratina

Meritev časovno odvisne kršitve simetrije CP v razpadu nevtralnega mezona $B^0 \rightarrow D^+D^-$

Doktorska disertacija

Mentor: doc. dr. Samo Korpar

Ljubljana, 2007

Acknowledgments

To Samo, for letting me do whatever I wanted.
To Peter, for reminding me what needs to be done.
To Boštjan, for giving me an example of how to do it.
To all three, for being there when things were getting done.

To all who believed in me and made me work some more.
Usually they were one and the same person.

Abstract

We report measurements of the branching fraction and CP violation parameters in $B^0 \to D^+D^-$ decays. The results are based on a data sample that contains 535×10^6 $B\bar{B}$ pairs collected at the $\Upsilon(4S)$ resonance, with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. We obtain $[1.97 \pm 0.20\,(\text{stat}) \pm 0.20\,(\text{syst})\,] \times 10^{-4}$ for the branching fraction of $B^0 \to D^+D^-$. The measured values of the CP violation parameters are: $\mathcal{S} = -1.13 \pm 0.37 \pm 0.09$, $\mathcal{A} = 0.91 \pm 0.23 \pm 0.06$, where the first error is statistical and the second is systematic. We find evidence of CP violation in $B^0 \to D^+D^-$ at the $4.1\,\sigma$ confidence level. While the value of $\mathcal{S}$ is consistent with expectations from other measurements, the value of the parameter $\mathcal{A}$ favors large direct CP violation at the $3.2\,\sigma$ confidence level, in contradiction to Standard Model expectations.

PACS (2007): 13.25.Hw, 11.30.Er
Key words: Standard Model, CP violation, $B^0 \to D^+D^-$ decays, Belle detector

Izvleček

V delu so predstavljeni rezultati meritve razvejitvenega razmerja in kršitve CP v razpadu $B^0 \to D^+D^-$. Analiza je narejena na vzorcu 535×10^6 parov mezonov $B\overline{B}$, izmerjenih z detektorjem Belle na asimetričnem trkalniku e^+e^- KEKB. Izmerimo razvejitveno razmerje $[1.97 \pm 0.20\,(\mathrm{stat}) \pm 0.20\,(\mathrm{syst})\,] \times 10^{-4}$ in parametra kršitve CP $\mathcal{S} = -1.13 \pm 0.37 \pm 0.09$ ter $\mathcal{A} = 0.91 \pm 0.23 \pm 0.06$, kjer je vsakič prva napaka statistična in druga sistematska. Neničelno kršitev CP izmerimo z zanesljivostjo 4.1σ. Neničelno direktno kršitev CP ($\mathcal{A} \neq 0$) izmerimo z zanesljivostjo 3.2σ. Izmerjena vrednost parametra $\mathcal{S}$ se ujema z rezultati drugih neodvisnih meritev, medtem ko vrednost parametra $\mathcal{A}$ kaže na možnost velike direktne kršitve CP, ki je v neskladju z napovedmi Standardnega Modela.

PACS (2007): 13.25.Hw, 11.30.Er
Ključne besede: Standardni model, kršitev CP, razpadi $B^0 \to D^+D^-$, detektor Belle

Contents

Chapter 1

Introduction

The Standard Model (SM) of particle physics is a theory which describes all known elementary particles and their interactions. Fundamental particles are divided into two groups, depending on whether they obey the Pauli exclusion principle or not: half-integer spin fermions do, and therefore two identical fermions can never occupy the same quantum state; on the other hand, there is no such restriction for integer-spin bosons. Fermions are divided into leptons and quarks and form the matter as we know it – quarks are bound into protons and neutrons, which are further bound together to form nuclei. Together with electrons (which themselves are leptons) the nuclei form atoms. Bosons, in contrast, mediate the force between the fermions. They come in three types, which are used to describe strong, weak, and electromagnetic forces. The strong force, mediated by gluons, binds quarks into hadrons. Weak interaction, which is causing the β decay for example, is explained by exchange of weak bosons Z^0, W^+ and W^-. Electromagnetic interaction between the charged particles is mediated by photons.

Fermions can be arranged into three generations, as shown in Table 1.1. For all listed particles there exists an antiparticle (denoted with a bar over the particle sign, such as $\bar{p}$ for anti-proton) of identical mass and opposite quantum numbers, such as charge. The four fermions of each generation behave almost exactly like their counterparts in the other generations; the only difference is a hierarchical increase in mass for the second and third generation. All ordinary matter is made from the first-generation particles. Higher-generation particles are produced by interactions of high-energy cosmic rays with the interstellar medium or the Earth atmosphere and in high-energy physics experiments. They quickly decay into the first-generation particles.

leptons		quarks	
e	ν_e	u	d
μ	ν_μ	c	s
τ	ν_τ	t	b

Table 1.1: Three generations of fermions.

The study of such decays provides insights into the way particles interact and into their fundamental properties. For instance, the Large Electron-Positron collider at CERN tested various SM predictions about the decay of Z bosons and found them confirmed [1]. Still, the SM cannot be a complete theory of fundamental physics, because it lacks at least two properties. Firstly, the model contains 19 free parameters, such as particle masses, which must be determined experimentally (plus another 10 for neutrinos). Secondly, the model does not describe the gravitational interaction. In addition, there are cosmological reasons why the SM is believed to be incomplete. Among those, for example, is the observation that strong interaction treats matter and antimatter in a symmetric way, which suggests that there should be equal amounts of both after the Big Bang, which does not seem to be the case. While the weak interaction in the SM makes a distinction between matter and anti-matter, this effect is too small to explain the observed excess of matter in our universe.

In high energy physics experiments two particles with large kinetic energy are made to collide, converting the available energy into mass according to the Einstein's relation $E = mc^2$. It is by this process that particles of higher generations can be produced. In practice, two oppositely-charged beams of particles are accelerated to almost the speed of light and are stored in an accelerator ring. The beams can then be made to cross in a controlled way at the so-called interaction pont (IP), where a detector is placed in order to detect the resulting decay products. In the past fifty years, there has been a series of successful experiments giving a new insight into the nature of elementary particles, promoting a development of a theory describing this phenomena and testing its predictions. This work describes the analysis of the data collected at one of such experiments, the Belle detector at the High Energy Accelerator Research Organization (KEKB) electron-positron collider.

The main purpose of the Belle experiment is to measure CP violation in B-meson decays. B mesons are particles which contain a quark and an antiquark, one of which is either a b or a $\bar{b}$ quark from the third generation of leptons. CP stands for the product of two transformations: C is the operator of charge conjugation, which transforms a particle into its antiparticle, and P is a parity-reversal operator, which creates the mirror image of a physical system by reflecting all coordinates through the origin. The strong and electromagnetic interactions are invariant under the CP transformation, which means that their laws of physics are unchanged if all particles are replaced with their antiparticles and at the same time the system is replaced with its mirror image. On the other hand, it was shown by an experiment in the kaon system in 1964 that weak interactions violate CP symmetry [2], leading to a Nobel price for James Watson Cronin and Val Logsdon Fitch in 1980. CP violation is incorporated in the SM by a complex phase in the Cabibbo-Kobayashi-Maskawa (CKM) matrix, which contains information on the strength of the interaction between the different types of quarks [3, 4]. It is interesting to note that the CKM matrix is very close to the unit matrix, which means that the CP symmetry is almost conserved. While CP violation (CPV) could, in general, result in the excess of matter over antimatter in our universe, the smallness of the observed CPV is insufficient to account for the observations. Nevertheless, the fact that it is relatively small but non-zero is intriguing, which provides a strong motivation for further studies of its properties. In addition, new sources of CP violation from beyond the SM could result in a difference between the measured values and the SM predictions, providing a research direction for the yet-undiscovered physics. In 2001, soon after the start of the operation of the Belle detector in Tsukuba, Japan, and a similar experiment – Babar at the PEP-II collider in Stanford Linear Accelerator Center (SLAC), USA, CP violation was observed in the B meson system by the two collaborations [5, 6]. This was the first observation of CP violation outside the kaon system. Since then, various measurements have been performed to determine the parameters of the SM with better precision and to search for any deviations from the SM predictions. So far, even if some observables related to the CP violation in the SM have been measured with fairly good accuracy of a few percent [7, 8], no significant deviation from the SM has been observed. This means that CP violation effects due to new physics contributions are either small or already described by the SM. To measure possible small effects we need a considerably larger data sample to decrease the measurement errors. The next-generation experiments, Super-B factories, are being proposed for the purpose [9].

We can learn about the CP symmetry in the B meson system from the study of processes, in which B mesons decay to a CP eigenstate, $B \to f_{CP}$. Speaking in a general way, we can compare the rate at which a B meson decays into a CP eigenstate f_{CP} to the rate at which a B meson decays into the CP conjugate state $\bar{f}_{CP}$, to the rate at which a $\bar{B}$ meson decays into the f_{CP} and to the rate at which a $\bar{B}$ meson decays into the $\bar{f}_{CP}$. Different final states provide complementary information about the system, and only by combining such information from different measurements can we obtain the complete picture of the subject and get accurate results. This thesis presents the measurement of CP violation parameters in $B^0 \to D^+D^-$ decays. The final state D^+D^- is a CP eigenstate, $CP\,|\,D^+D^- >= \,|\,D^+D^- >$. In the case at hand, the difference between the B^0 and $\bar{B}^0$ meson decays is expected to appear only in the time-dependent decay rate, where time refers to the time for which the B meson freely propagates before it decays to f_{CP}. If we are to be able to measure any CP-related property of this decay, we absolutely have to be able to measure two things: the flavor of the B meson decaying to the final state f_{CP} and the time it freely propagates before it decays. Our ability to do so may not be obvious: since both B mesons decay to the same final state, we cannot judge the flavor from its decay products. Moreover, the lifetime of a B^0 meson is about 1.5 ps, which is beyond the precision of any direct time measurement. As we will see, both issues are solved by a special design of the Belle experiment.

The KEKB and PEP-II colliders are referred to as a B-factories, because their main purpose is to produce large numbers of B mesons, i.e. particles, which contain a b quark. The probability of producing a B meson strongly depends on the energy of the electron and positron beams. The probability is largest when the total energy of the colliding e^+e^- is equal to the mass of the $\Upsilon(4S)$ resonance, bound state of a pair of $b\bar{b}$ quarks, which has just enough energy to immediately decay to a pair of B mesons. The two B mesons are not independent of one another since the two b quarks obey the Pauli exclusion principle: when one of them is a b quark, the other has to be an $\bar{b}$. This gives us the possibility of deducing the flavor of the B meson decaying to the CP eigenstate as opposite to the flavor of the other B meson, which can be determined for flavor-specific decays. The problem of the decay time measurement is solved by asymmetric energy of the two beams. The B meson pair is produced in a center-of-mass system (CMS) and is therefore boosted along the direction of the more energetic beam. In the detector rest frame, the B mesons therefore travel for some distance before they decay and by measuring this distance of about $200\,\mu$m we can determine the time between the two decays. As the error on the result gets smaller with a larger data sample, it is important to detect as many B meson pairs as possible. The number of B meson pairs produced is proportional to a collider luminosity, which measures how many particles per unit area and per unit time cross at the IP. Currently, the KEKB holds the world record in luminosity that exceeds 1.7×10^{34} cm^{-2}s^{-1} (17 /nb/s).

We present a measurement of time dependent CP violation in $B^0 \to D^+D^-$ decays based on a data sample that contains 535 ± 7 million $B\bar{B}$ pairs, collected with the Belle detector until the summer of 2006. We will start by an overview of the CP violation in the SM and continue with a description of our experimental apparatus, the Belle detector at the KEKB collider. After these necessary introductions we will move on to our measurement, which is divided into two parts. First, we describe the reconstruction of $B^0 \to D^+D^-$ decays. Second, we show the analysis of the decay time distributions of the reconstructed B mesons. The later involves studies of systematic effects and control samples, as well as a rather elaborate evaluation of the significance of the result. We conclude with a discussion on future prospects of this topic.

Chapter 2

CP Violation in the Standard Model

In the SM, the interactions between the elementary particles of three generations of fermions are described by a gauge field theory with the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The group $SU(3)$ is used to describe Quantum Chromodynamics (QCD), which governs the strong interaction among quarks while $SU(2)_L \times U(1)_Y$ stands for the electroweak gauge group. The weak hypercharge Y is 0, -1, $2/3$, $-1/3$ for neutrinos, leptons, up-type quarks and down-type quarks, respectively. The weak interaction treats left- and right-handed fermions in a fundamentally different way: right-handed components of the leptons and quarks are singlets under the weak $SU(2)_L$ while the left-handed components of leptons transform as doublets. It is important to note that doublets of quark weak eigenstates

$$\begin{pmatrix} u \\ d' \end{pmatrix}_L \quad \begin{pmatrix} c \\ s' \end{pmatrix}_L \quad \begin{pmatrix} t \\ b' \end{pmatrix}_L \tag{2.1}$$

differ slightly from their mass eigenstates, $q' \neq q$. We can write weak eigenstates (d', s', b') in the basis of mass eigenstates (d, s, b). The transformation between the two systems is described by a 3×3 Cabibbo-Kobayashi-Maskawa (CKM) unitary matrix $\mathbf{V}_{\mathrm{CKM}}$ [3, 4].

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \mathbf{V}_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{2.2}$$

Within the SM, CP symmetry is broken by complex phases in the Yukawa couplings of the Higgs field to quarks. When all phases which do not effect the physical observables are removed, a single CP violating parameter remains. In the basis of mass eigenstates, this parameter corresponds to a single irreducible complex phase of the CKM matrix. As a consequence, CPV appears only in the weak charged current interactions mediated by the W boson described by the following term in the interaction Lagrangian.

$$\mathcal{L}^{\mathrm{CC}} = -\frac{g_2}{\sqrt{2}} \begin{pmatrix} \overline{u} & \overline{c} & \overline{t} \end{pmatrix}_L \gamma^\mu \mathbf{V}_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L W^\dagger_\mu + \mathrm{h.c.}\,, \tag{2.3}$$

where g_2 is a coupling constant.

We start with a description of the CKM matrix in Sec. 2.1. Introduction to the CP violation formalism is followed by a section describing the specifics of $B^0 - \overline{B}^0$ mixing. In the last part of this chapter we apply the CP violation formalism to the $B^0 \to D^+D^-$ decay and show how physical observables relate to the SM parameters in this case.

Any reader who is not familiar with the high energy physics (HEP) might consider reading a description of the typical HEP units and a short overview of the mesons which are included in the Appendix A.1.

2.1 Parametrization of the CKM matrix

First, let us just keep in mind that the CKM matrix is just a transformation matrix between the mass and the weak eigenstates. As such, it satisfies the unitary condition $\mathbf{V}\mathbf{V}^\dagger = \mathbf{V}^\dagger\mathbf{V} = \mathbf{I}$. A general complex matrix with elements $V_{ij} = a_{ij}\, \mathrm{e}^{\imath\varphi_{ij}}$ has eighteen real parameters a_{ij} and φ_{ij}. The unitarity condition gives six constraints on the values of a_{ij} and three constraints on phases φ_{ij}, leaving three free parameters a_{ij} and six complex phases φ_{ij}. The freedom to redefine five relative phases between the mass eigenstates phases can be used to remove five of the phases, leaving a single CP violation phase, the Kobayashi-Maskawa phase. All nine matrix elements can therefore be expressed as a function of four independent parameters. Among different parametrizations, the most commonly used is the Wolfenstein parametrization [10]. This choice of parameters makes explicit the experimentally known hierarchy of the matrix elements – the off-diagonal elements being small compared to the diagonal elements, as illustrated in Fig. 2.1. In the Wolfenstein parametrization, the CKM matrix elements are expressed in terms of real parameters λ, A, ρ and η such that $\lambda = |V_{us}| = V_{us}$ and the complex part only appears in the most off-diagonal terms V_{ub} and V_{td}. The remaining elements are not given by a simple analytical expression but rather by a series in terms of $\lambda \simeq 0.22$,

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4)\,. \tag{2.4}$$

Very often the approximation including all terms up to the order of λ^3 is used. For this parametrization a measure for the extent of the CP violation is about $\lambda^6 A^2\eta$, in the case of $\lambda^6 A^2 \eta = 0$ there would be no CP violation.

As already noted, the unitarity of the CKM matrix imposes conditions $\Sigma_k \mathrm{V}_{ki}\mathrm{V}^*_{kj} = \delta_{ij}$ on its elements. This condition can be visualised graphically, with each of the complex terms in the sum represented by a vector in the complex plane. For $i \neq j$, the sum of all three terms (vectors) is zero, therefore the three of them form a triangle in the complex plane. The product of the first and third column is the standard choice for the graphical representation, as $|\mathrm{V}_{11}\mathrm{V}^*_{13}|$, $|\mathrm{V}_{21}\mathrm{V}^*_{23}|$ and $|\mathrm{V}_{31}\mathrm{V}^*_{33}|$ are all of the order of λ^3. This choice of the CKM triangle is shown in Fig. 2.2 (a). For further simplification, all three terms are divided by $V_{cd}V^*_{cb}$, making two of the triangle vertices correspond to the points $(0,0)$, $(1,0)$

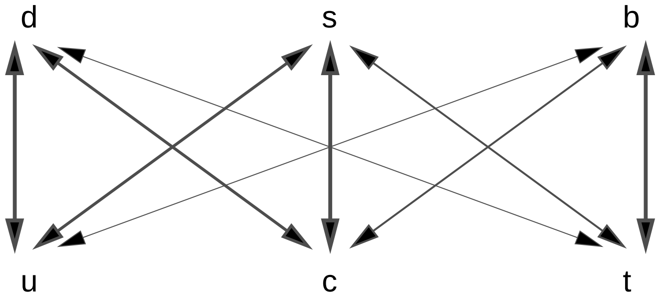


Figure 2.1: The SM allowed transitions between the quarks of different flavors. Quarks in the top row have charge $Q_{d,s,b} = -1/3$ and quarks in the bottom $Q_{u,c,t} = +2/3$. Thicker lines indicate higher probability for the transition. The transition always involves charged current $W^{\pm}$ for $\Delta Q = \pm 1$; there is no transition between the quarks of the same charge, such as $s \rightarrow d$.

and the third to $(\bar{\rho}, \bar{\eta})$, as shown in Fig. 2.2 (b). The angles of this triangle are related to the phase of the quotient of the adjacent vectors;

$$\phi_1 = arg(-V_{cd}V_{cb}^*/V_{td}V_{tb}^*)\,, \tag{2.5}$$

$$\phi_2 = arg(-V_{td}V_{tb}^*/V_{ud}V_{ub}^*)\,, \tag{2.6}$$

$$\phi_3 = arg(-V_{ud}V_{ub}^*/V_{cd}V_{cb}^*)\,. \tag{2.7}$$

(2.8)

Another notation $\alpha = \phi_2$, $\beta = \phi_1$ and $\gamma = \phi_3$ is also used.

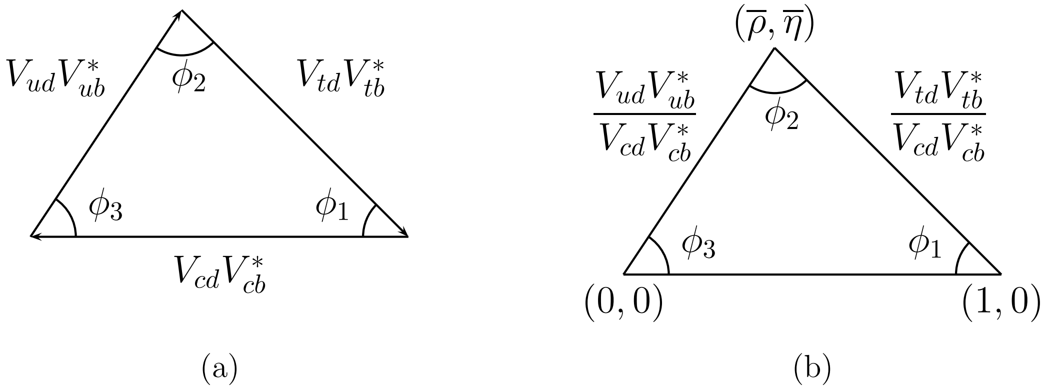


Figure 2.2: The graphical representation of the unitary conditions $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$ (a) and $V_{ud}V_{ub}^*/(V_{cd}V_{cb}^*) + 1 + V_{td}V_{tb}^*/(V_{cd}V_{cb}^*) = 0$ (b).

Different measurements are sensitive to different parameters or combinations of parameters of the CKM matrix. Figure 2.3 shows a list of measurement-based constraints in the above mentioned graphical representation.

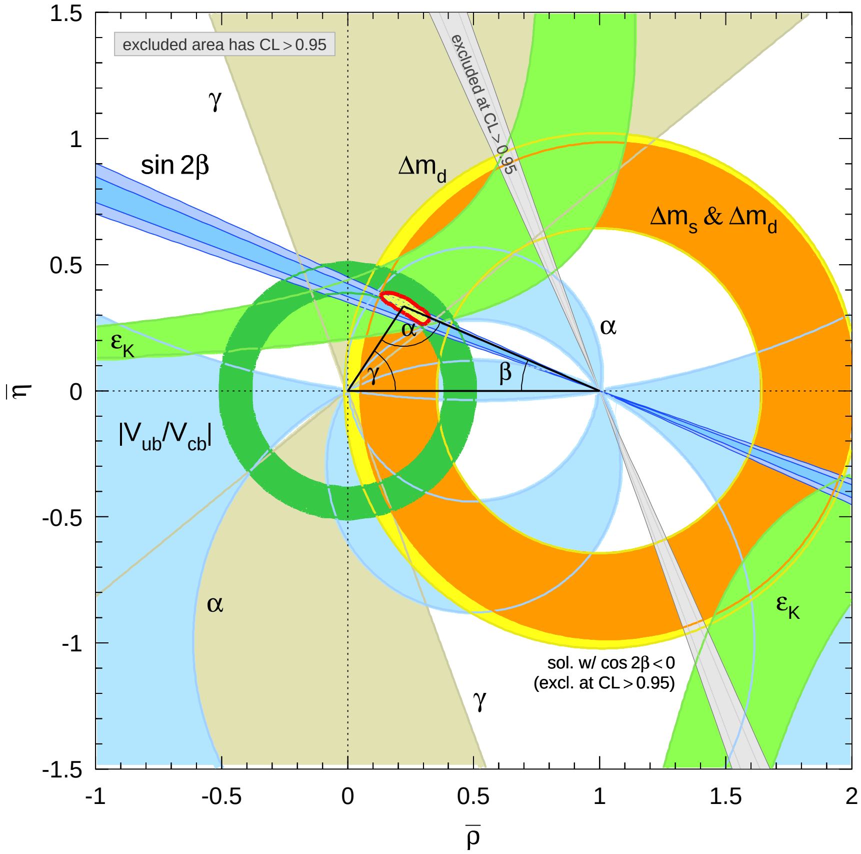


Figure 2.3: The graphical representation in the $(\bar{\rho}, \bar{\eta})$ plane of a several up-to-date measurements [11]. A possible CKM triangle with apex in the preferred region is also shown.

2.2 Formalism and Classification of CP Violation Effects

In this section we present a general formalism for CP violation in the decay of a meson M with an emphasis on the B meson system. Let use define the decay amplitude of M and its CP-conjugate $\overline{M}$ to a final state f and its CP-conjugate $\bar{f}$ as

$$\begin{aligned} A_f \quad &=< f|\mathcal{H}|M > \; , \qquad \bar{A}_f \quad =< f|\mathcal{H}|\overline{M} > \; , \\ A_{\bar{f}} \quad &=< \bar{f}|\mathcal{H}|M > \text{ and } \bar{A}_{\bar{f}} \quad =< \bar{f}|\mathcal{H}|\overline{M} > \; . \end{aligned} \tag{2.9}$$

We define the CP asymmetry as the ratio

$$A \quad = \quad \frac{\Gamma_{\overline{M} \to \bar{f}} - \Gamma_{M \to f}}{\Gamma_{\overline{M} \to \bar{f}} + \Gamma_{M \to f}} \, , \tag{2.10}$$

where Γ is the interaction rate. In a similar way we define the time-dependent CP asymmetry,

$$a(t) = \frac{\mathcal{P}_{\overline{M}\to\bar{f}} - \mathcal{P}_{M\to f}}{\mathcal{P}_{\overline{M}\to\bar{f}} + \mathcal{P}_{M\to f}} . \tag{2.11}$$

Time t is the particle decay time and $\mathcal{P} = \mathrm{d}\Gamma/\mathrm{d}t$ are the corresponding time dependent rates. We distinguish between three types of CP violation in meson decays.

1. CP violation in decay is characterized by

   $$|\bar{A}_{\bar{f}}| \neq |A_f| \tag{2.12}$$

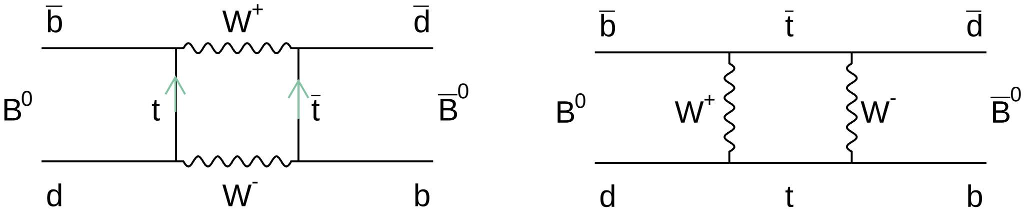
   This is the only possible source of CPV in the charged meson decays.

2. CP violation in mixing, $M^0 \to \overline{M}^0$. This is a time-dependent CP violation as particles moving freely for a longer time have a higher probability of mixing (or oscillating). This is also the only possible source of CPV in the semileptonic decays.

3. The third type of CPV happens when the final state is identical to its CP-conjugate, $f = \bar{f}$. In this case we can not distinguish between two processes, mixing, $M \to \overline{M} \to f$, and direct decay, $M \to f$. According to the principles of quantum mechanics, the CPV in this case is a result of the interference between the complex phases of two indistinguishable processes, CP violation in mixing and decay.

2.3 Mixing in the B meson system

The neutral B^0 meson can change to a $\overline{B}^0$ meson, or vice versa. This process is called mixing. At the quark level, a combination of $\bar{b}d$ quarks is transformed to a $b\bar{d}$, as shown in Fig. 2.4.



(a) (b)

Figure 2.4: Feynman diagrams describing $B^0 - \overline{B}^0$ mixing. Top quark in the box diagram can be replaced by an up or charm quark, but the probabilities for the exchange of c or u quark are smaller by a factor of $(m_{c(u)}/m_t)^2$.

The B^0 meson mixing happens only because the eigenstates of the weak interaction do not coincide with the mass eigenstates. If the decay of the particle is governed by the weak interaction, the process is described in the weak basis. On the other hand, we can use en effective Hamiltonian to describe particle propagation in time as

$$i\frac{d}{dt}|B(t)\rangle = \left(M - i\frac{\Gamma}{2}\right)|B(t)\rangle\,, \tag{2.13}$$

with mass eigenstates $|B_1\rangle$ and $|B_2\rangle$ being the stationary solutions of the equation. The two-by-two Hermitian matrices M and Γ denote mass and decay matrices, respectively. Their diagonal terms are associated with the flavor-conserving transitions, while off-diagonal terms are associated with flavor-changing transitions.

Any B meson state can be written as a superposition of the eigenvectors, either in the weak or mass basis. The mass eigenstates can be expressed in the weak basis as

$$\begin{aligned} |B_1\rangle &= p|B^0\rangle + q|\overline{B}^0\rangle, \\ |B_2\rangle &= p|B^0\rangle - q|\overline{B}^0\rangle, \end{aligned} \tag{2.14}$$

Likewise, the weak interaction states $|B^0\rangle$ and $|\overline{B}^0\rangle$ can be expressed as a function of mass eigenstates

$$\begin{aligned} |B^0(t)\rangle &= \frac{1}{2p}\left(|B_1(t)\rangle + |B_2(t)\rangle\right) \text{ and} \qquad (2.15) \\ |\overline{B}^0(t)\rangle &= \frac{1}{2q}\left(|B_1(t)\rangle - |B_2(t)\rangle\right)\,. \end{aligned}$$

Since $|B_i(t)\rangle$ are solutions of the Eq. 2.13, they can be expressed as $e^{-i\mathrm{H}_i t}|B_i(t=0)\rangle$, which gives

$$\begin{aligned} |B^0(t)\rangle &= g_+|B^0\rangle + \frac{q}{p}\,g_-|\overline{B}^0\rangle \text{ and} \\ |\overline{B}^0(t)\rangle &= \frac{p}{q}\,g_+|B^0\rangle + g_-|\overline{B}^0\rangle\,, \text{ where} \qquad (2.16) \\ g_\pm &= \frac{1}{2}\left(e^{-i\mathrm{H}_1 t} \pm e^{-i\mathrm{H}_2 t}\right) \qquad (2.17) \end{aligned}$$

As the probability for a decay to a final state f is $|\langle f|B(t)\rangle|^2$, we can calculate the time dependent rates $\mathcal{P}_{M\to f}$, $\mathcal{P}_{\overline{M}\to f}$,. . . and calculate the time-dependent asymmetry defined by Eq. 2.11 as follows.

$$\begin{aligned} \mathcal{P}_{M\to f} &= \left|g_+\langle f|B^0\rangle + \frac{q}{p}\,g_-\langle f|\overline{B}^0\rangle\right|^2 \qquad (2.18) \\ \mathcal{P}_{\overline{M}\to f} &= \left|\frac{p}{q}\,g_+\langle f|B^0\rangle + g_-\langle f|\overline{B}^0\rangle\right|^2 \qquad (2.19) \\ a(t) &= \frac{2\mathrm{Im}(\lambda_f)}{1+|\lambda_f|^2}\sin(\Delta m t) - \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}\cos(\Delta m t)\,, \text{ where} \qquad (2.20) \\ \lambda_f &= \frac{q}{p}\frac{\overline{A}_f}{A_f} = |M_{12}|\,e^{-i\Phi_B}\,\frac{\overline{A}_f}{A_f} \qquad (2.21) \end{aligned}$$

and Φ_B is a complex phase due to $B^0 - \overline{B}^0$ mixing. For the Feynman diagrams shown in Fig. 2.4, the later relates to the CKM matrix elements as

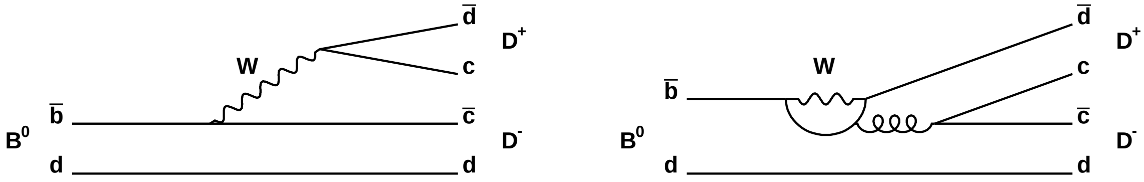
$$e^{-i\Phi_B} \quad = \quad \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \,. \tag{2.22}$$

Since the absolute value of q/p is approximately 1, direct CP-violation requires $|A_f| \neq |\overline{A}_f|$, which could happen if A_f is a sum of at least two decay amplitudes having different weak interaction and strong interaction phases. Indirect CP-violation, on the other hand, is a result of a quantum mechanical interference between the mixing and decay amplitudes.

2.4 CP Violation in $\mathbf{B^0 \to D^+D^-}$ Decays

The final state $f_{CP} = D^+D^-$ is a self-conjugate CP eigenstate, $CP(D^+D^-) = D^+D^-$. CPV of the third type therefore only can occur as a result of the interference between the decay amplitude with and without the B^0 mixing.

The decay $B^0 \to D^+D^-$ is dominated by a tree-level $b \to c\bar{c}d$ transition, shown in Fig. 2.5 (a). The CKM matrix elements associated with each of the W^+ vertices in Fig. 2.5 (a) are V_{cb}^* for the $\bar{b} \to \bar{c}$ transition and V_{cd} for the other vertex. In the Feynman diagram for the $\overline{B}^0 \to D^+D^-$ decay, particles are interchanged with their antiparticles. The corresponding vertices in the Feynman diagram for $\overline{B}^0$ are therefore associated with CKM matrix elements V_{cb} and V_{cd}^*. In the approximation of only the tree level diagram, we follow the Eq. 2.21 and obtain



(a) (b)

Figure 2.5: The tree (a) and the penguin (b) contributions to the $B^0 \to D^+D^-$ decay.

$$\begin{aligned} \lambda_f &= \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \frac{V_{cb} V_{cd}^*}{V_{cb}^* V_{cd}} = \frac{V_{tb}^* V_{td}}{V_{cb}^* V_{cd}} \frac{V_{cb} V_{cd}^*}{V_{tb} V_{td}^*} \\ &= e^{-\imath\phi_1} e^{-\imath\phi_1} = e^{-\imath 2\phi_1} \,, \end{aligned} \tag{2.23}$$

since ϕ_1 is defined as $\phi_1 = arg(-V_{cd}V_{cb}^*/V_{td}V_{tb}^*)$ (Eq. 2.5). It follows that $Im(\lambda_f) =$

$-\sin(2\phi_1)$ and $|\lambda_f| = 1$. The time-dependent decay rates and asymmetry are then

$$\mathcal{P}_{B^0\to D^+D^-} = \frac{e^{-|\Delta t|/\tau}}{4\tau}(1 - 1\,(-\sin(2\phi_1)\sin(\Delta m \Delta t)))\,, \tag{2.24}$$

$$\mathcal{P}_{\overline{B}^0\to D^+D^-} = \frac{e^{-|\Delta t|/\tau}}{4\tau}(1 + 1\,(-\sin(2\phi_1)\sin(\Delta m \Delta t)))\,, \tag{2.25}$$

$$a(t) = -\sin(2\phi_1)\sin(\Delta m \Delta t)\,, \tag{2.26}$$

as shown in Fig. 2.6. The value of $\sin(2\phi_1)$ can therefore be inferred from the Δt distribution of the $B^0 \to D^+D^-$ decays.

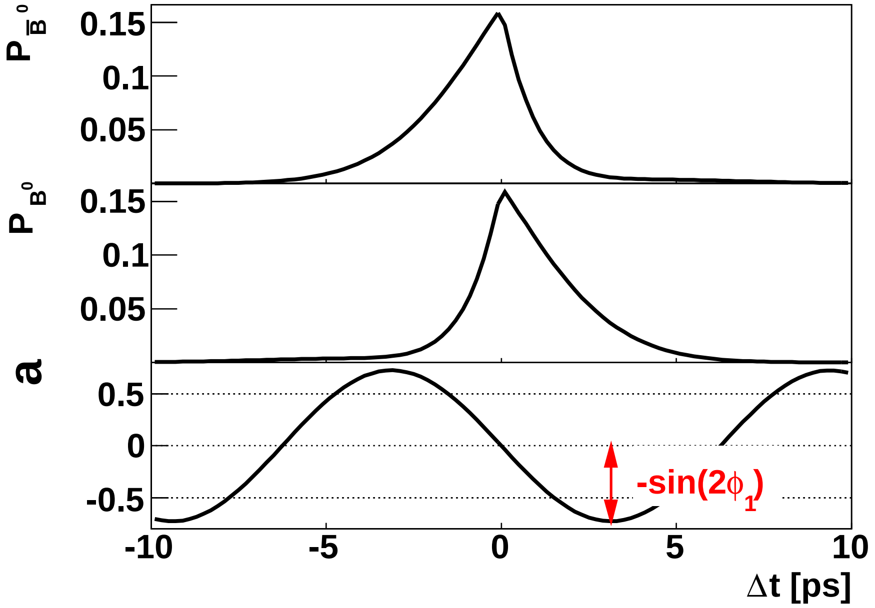


Figure 2.6: The dependencies $\mathcal{P}_{B^0\to D^+D^-}(t)$, $\mathcal{P}_{\overline{B}^0\to D^+D^-}(t)$ and $a(t)$, as defined by Eq. 2.24-2.26.

The second-order gluonic penguin contribution, shown in Fig. 2.5 (b), is expected to change the value of parameter S by less than a few percent and increase the value of A to about 3% [12, 13]. However, particles from physics beyond the SM may give additional contributions within the loop diagrams mediating flavor-changing $b \to d$ transitions. Such contributions may potentially induce large deviations from the SM expectation for time-dependent CP asymmetries. As sin $2\phi_1$ has already been determined with high precision by measurements in charmonium modes [7, 8], the objective of this work is not to measure the value of $\sin(2\phi_1)$ but rather to focus on deviations from expectations in $b \to c\bar{c}d$ transitions. Similar studies have already been carried out for $B^0 \to D^{*\pm}D^{(*)\mp}$ decays, which involve the same quark level weak decay [14, 15, 16, 17].

Chapter 3

The Belle Experiment

The Belle experiment is a collaboration of about four hundred people from over eighty institutes. It is located at KEK, High Energy Accelerator Research Organization center in Tsukuba, Japan. In the following sections we will first make a short introduction to the KEKB particle accelerator, the B meson factory. In the second part we will describe the Belle detector, the experimental apparatus used for the detection of B mesons and other particles produced at the KEKB collider.

3.1 The KEKB Particle Accelerator

The KEKB is an asymmetric electron-positron collider. Two beams of particles are stored in their corresponding rings of 3016 m in circumference. The particles travel in bunches spaced a few meters apart. They intersect at the interaction point (IP), where the Belle detector is placed. Most frequently, particles in the colliding bunches pass each other without interacting. However, approximately once in every 10^5 times, there is some detectable interaction, which we call "the event".

The energies of the electron and positron beams are $E_1 = 8.0\,\mathrm{GeV}$ and $E_2 = 3.5\,\mathrm{GeV}$, respectively. In the center-of-mass system (CMS) the electrons and positrons have equal energies E and opposite momenta $\mp\vec{p}$. Let us denote the z axis as the direction of the positron beam. The electron and positron 4-momenta $p^\mu_{\mathrm{CMS}} = (E, 0, 0, \mp pc)$ transform to the laboratory system as

$$p^\mu_{\mathrm{lab}} = \Lambda\, p^\mu_{\mathrm{CMS}}\,; \quad \Lambda = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix}, \tag{3.1}$$

where Λ is the Lorentz boost transformation matrix. Here we neglect a 22 mrad tilt of the electron beam direction with respect to the positron beam direction in the laboratory system for simplicity reasons. The electron rest energy mc^2 is negligible compared to its total energy. For $E \approx pc$ the above equation for the first (energy) component simplifies to

$E_{1,2} = \gamma E \pm \beta\gamma E$. From this pair of equations we solve for the available energy E in the CMS and Lorentz parameter $\beta\gamma$ as follows.

$$E = \frac{E_1 + E_2}{\gamma} = 2\sqrt{E_1 E_2} = 10.58\,\text{GeV and} \tag{3.2}$$

$$\beta\gamma = \frac{E_1 - E_2}{2\sqrt{E_1 E_2}} = 0.425\,. \tag{3.3}$$

The available energy E is equal to the mass of the $\Upsilon(4S)$ resonance, i.e. a bound state of $b\bar{b}$ quarks, which immediately decays to a pair of $B\overline{B}$ mesons of mass $m_B = 5.28\,\text{GeV}/c^2$. This choice allows for a reasonably large $B\overline{B}$ pair production cross-section. The $\Upsilon(4S)$ resonance is the first of the $\Upsilon(nS)$ resonances above the treshold for the B meson pair production, $2m_B c^2 = 10.56\,\text{GeV}$. Figure 3.1 illustrates the hadron production cross-section at the electron-positron collisions in the 10 GeV energy range. Narrow peaks of $\Upsilon(nS)$ resonances are shown on top of the non-resonant QCD processes involving the production of a pair of lighter quarks, $e^+e^- \to q\bar{q}$, where q stands for u, d, s or c. While the study of the later can give interesting results on its own, these processes are considered a background in the study of B meson decays.

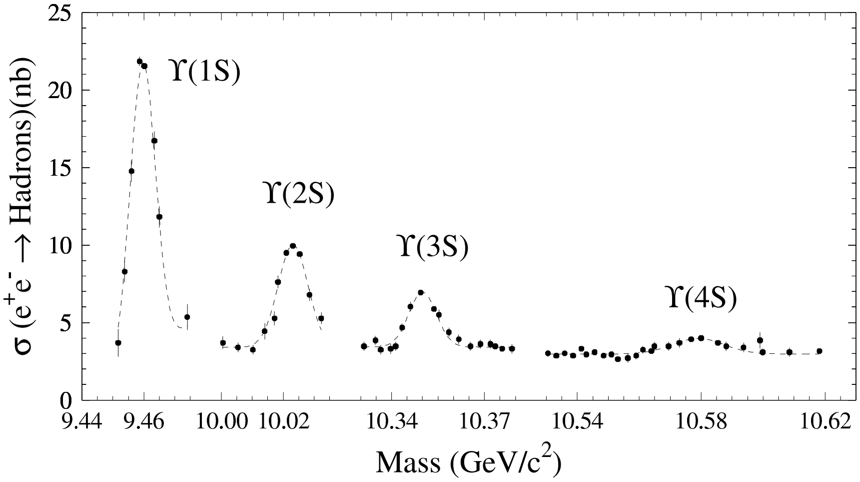


Figure 3.1: The hadron production cross-section as a function of the available energy. Note that the energy axis scale is broken and all of the resonances are in fact quite narrow.

As almost all of the available energy is used for the B meson mass production, B mesons are practically at rest in the CMS. During their lifetime of $\tau_B = 1.53\,\text{ps}$ they travel on average a distance $\beta\gamma\tau_b c = 200\,\mu\text{m}$ in the laboratory system.

The CMS energy requirement (Eq. 3.2) determines only the product of the energy of the two beams E_1E_2 and we have a freedom to choose their relative size. A higher value of $\beta\gamma$ is desirable because it can improve the time resolution, but it comes at a cost – the energy needed to accelerate the two beams, $E_1 + E_2 = 2E\sqrt{1+\beta^2\gamma^2}$, increases at the same time, as well as energy losses and beam instability. In addition, in the case of higher boost particles move more in a forward direction, which increases the level of scattering and its contribution to the vertex measurement error. The design values of E_i are such that the current resolution of the vertex position measurement of about $100\,\mu$m gives a satisfactory accuracy for the time measurement.

In addition to the measurement errors, statistical fluctuations limit the accuracy of our result. As these scale with $\sqrt{N}$, where N is the number of the reconstructed events, collecting a larger data sample is necessary to lower the errors. The interaction rate is proportional to the accelerator luminosity and the interaction cross-section; the later is about 1 nb for the $B\overline{B}$ pair production at the energy of the $\Upsilon(4S)$ resonance. Luminosity is the number of particles per unit area per unit time crossing each other. The KEKB accelerator holds the world record of luminosity of

$$\mathcal{L} = 17.12\,\text{/nb/s (as of November 2006)},$$

exceeding its design value by 70%. In other words, more than ten $B\overline{B}$ pairs are produced each second during the operation. The total number of produced $B\overline{B}$ pairs is proportional to the time-integrated luminosity,

$$N_{B\overline{B}} = \sigma_{B\overline{B}} \int \mathcal{L}\mathrm{d}t$$

and amounts to $492.3\,\mathrm{fb}^{-1}$ at the energy of the $\Upsilon(4S)$ resonance until summer 2006. The total number of $B\overline{B}$ pairs recorded by the Belle detector in this period is 535×10^6 ; 50% are charged B^+B^- and 50% neutral $B^0\overline{B}^0$ pairs. This is currently the world's largest data sample of B mesons. To quantify the $e^+e^- \to q\bar{q}$ continuum background, the KEKB is operating also part of the time at the CMS beam energy of about 60 MeV below the $\Upsilon(4S)$ resonance, collecting an additional $55.5\,\mathrm{fb}^{-1}$ off-resonance time-integrated luminosity.

3.2 The Belle detector

The Belle detector is a large-solid-angle magnetic spectrometer specialized for the detection of particles coming from the electron-positron collisions at the KEKB. It follows the symmetry of the e^+e^- collisions: it is cylindrically symmetric and has a few forward-backward asymmetries in the design due to the overall boost of the incoming particles. We have to be capable of detecting and distinguishing between particles which live long enough to actually reach the detector: charged electrons, muons, pions, kaons and protons and neutral kaons, photons and neutrons. The more massive B and D mesons cannot be detected directly as they decay very close to the IP. Particle detection strongly depends on the type of the particle. Charged particles are easier to detect because they interact electromagnetically with electrons in the detector material. On the other hand, neutral photons and neutrons are detected via an interaction in which a charged electron or proton is produced first. Neutrinos interact only by the weak interaction; too weak for them to be directly observed. Their presence is indicated by a missing energy and momentum in the reconstructed event.

For a detection of a charged particle we have to be able to reconstruct its identity (mass and charge) and trajectory (so that we are able to reconstruct the decay vertex of its mother particle). The mass of a particle could be determined by measuring its energy and momentum using the relation $E^2 = p^2c^2 + m^2c^4$. However, the direct measurement of the particle energy by the hadron calorimeter has a very large error of about 20%, which makes this method impossible. Instead, we use the relation $p = \gamma mv$ and determine the mass of the particle by measuring its momentum and velocity. There are several ways of measuring the particle velocity. One of them is Cherenkov radiation measured by the aerogel Čerenkov counters. Another possibility is to measure the time of flight. Thirdly, one can measure energy per unit distance deposited in the detector. For any of the methods, the problem is not so much to measure the relevant quantity as to determine which one of the five particle hypothesis is most likely to be the case.

The Belle detector consists of seven sub-detectors specializing in the measurements of different quantities, few of which we have already mentioned. The position of sub-detectors with respect to the IP is important as each measurement also makes an impact on the particle being measured. Particle energy, for example, can only be measured once it is fully deposited in the calorimeter. Obviously there is no point in placing the position or particle identification (PID) detector outside the calorimeter as there will be no particle left to measure its properties. The calorimeters are therefore placed in the outer-most part of the detector. The case with other detectors is somewhat less extreme, yet the same issue remains: as a particle passes through any amount of matter, it scatters off its original track. The more matter there is, the more it scatters and more do its position, momentum and energy differ from their values at the IP we are interested in. This imposes the tightest restriction on the vertex detector. As we determine the B meson vertices by extrapolating the tracks of its decay products back to the IP, it is important to measure particle tracks as close to the IP as possible, with as little scattering material between their production vertex and position measurement as is feasible. The particle position is measured with a silicon vertex detector, which is placed at the innermost part of the Belle detector. Next, particle momentum is determined by measuring the radius of its track helix in an external magnetic field. To do this, a greater part of the Belle detector is located inside a superconducting solenoid coil that provides a homogenous magnetic field of 1.5 T. The error on the momentum is smaller if we can detect a longer arc of the particle track. The detector measuring the particle momentum therefore has to be relatively large. In order to minimize the scattering, a gas is used as a filling material. Particle momentum is measured by the central drift chamber, which spans for one meter radially outside the IP. Positively and negatively charged particles have tracks of opposite curvatures. Just outside of the drift chamber are placed aerogel Cherenkov counters, which provide a good separation between pions and kaons in a high momentum region of $pc \approx 1.2 - 3.5\,\text{GeV}$. The particle velocity is determined by measuring the time it takes for a particle to transverse a given distance which is done by a time of flight detector. The greater distance they travel, the longer time they will need to transverse that distance and smaller will be the relative error on their velocity. The time of flight counter is therefore located just outside the Cherenkov counters and before the calorimeters. Since the electron and hadron interaction are fundamentally different, we need two calorimeters. An electromagnetic calorimeter measures the electron and photon energy. It is the only way to detect the photons. Second, hadron calorimeter is used to detect and identify the more penetrating neutral kaons K_L and muons.

The above mentioned detectors and their positions are a somewhat standard set of particle detectors used in the high energy physics experiments. Figure 3.2 shows the Belle detector with different parts marked in different colors. Their relative positions are in accordance with our discussion. The next sections will describe the Belle sub-detectors in more detail. We will start with and devote more time to the silicon vertex detector (SVD), which is crucial for the measurement of the time difference between the decays of two B meson vertices.

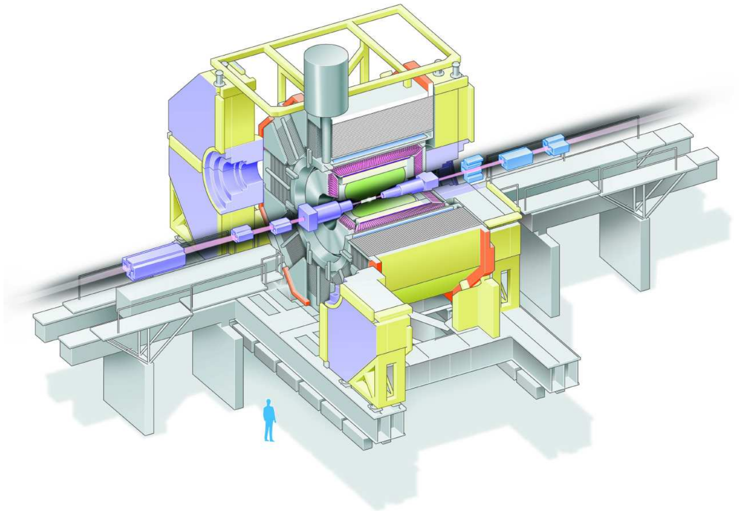


Figure 3.2: The Belle detector with different detector systems marked in different colors: silicon vertex detector in white in the innermost part, then central drift chamber in green, aerogel Cherenkov counter in white, electromagnetic calorimeter in purple and K_L and muon detector in grey color. The time of flight counter is too small to be indicated on this figure.

3.3 The Silicon Vertex Detector

The measurements of time-dependent CP asymmetry parameters require that the resolution of the vertex detector be (at least) better than the average flight distance of B mesons of about 200 μm. The semiconductor silicon detectors are the most accurate particle position measurement detectors available and they meet the above criteria. At Belle, double-sided silicon strip detectors (DSSDs) that provide good resolution are used. When a particle passes through a few hundred micrometers thick semiconductor depletion region it excites the electrons from the valence to the conduction band and thus makes electron-hole pairs, as illustrated in Fig. 3.3. The pairs are separated in an external electric field and collected at the strip electrodes. The position of the signal strip gives us the information about where the particle has passed the detector. The cathode strips on one side of the detector are perpendicular to the anode strips on the other side. This makes possible a simultaneous determination of two track coordinates. The detector is so thin that the third coordinate is simply the middle of the DSSD layer.

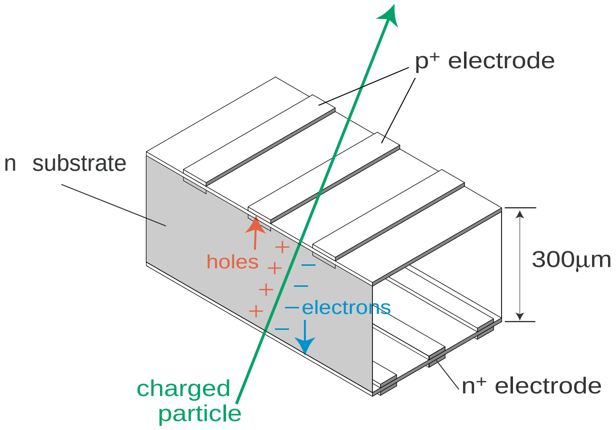


Figure 3.3: A schematic view of the position sensitive semiconductor detector (DSSD).

The DSSD units are organized around the IP in three or four cylindrical layers, as shown in Fig. 3.4. The two measured coordinates approximately correspond to the $r\varphi$ and z coordinates in the cylindrical system. In individual layers, a different number of DSSDs is chosen to cover the polar angle range from $\approx 15^\circ$ to $\approx 150^\circ$. Two inner detector configurations were used. A 2.0 cm radius beam pipe and a 3-layer silicon vertex detector was used for the first sample of 152 $\times 10^6 B\bar{B}$ pairs (SVD1), while a 1.5 cm beam pipe, a 4-layer silicon detector and a small-cell inner drift chamber were used to record the remaining 383 $\times 10^6 B\bar{B}$ pairs (SVD2) [18]. For this detector upgrade, three inner-most layers of CDC were replaced with a small-cell inner drift chamber in order to provide enough space for the larger SVD. Smaller beam pipe gives some additional space for the fourth SVD layer and allows for a smaller degradation of the vertex resolution due to scattering. The main reason for the replacement of the SVD units in 2003 was the radiation damage. As SVD is located closest to the IP, it is hit by a large amount of the synchrotron radiation and beam background. Additionally, the semiconductor material is very sensitive to the radiation damages; any material defects can either trap the signal charge or cause a higher background level, in both cases worsening the signal to noise ratio.

Table 3.1 shows an overview of the SVD parameters for both configurations. The most important parameter is the strip pitch; in the rough approximation the position resolution is of the order of the strip pitch size. For the z side, only every second strip is read out to save the number of read-out channels. The charge on floating strips is shared with

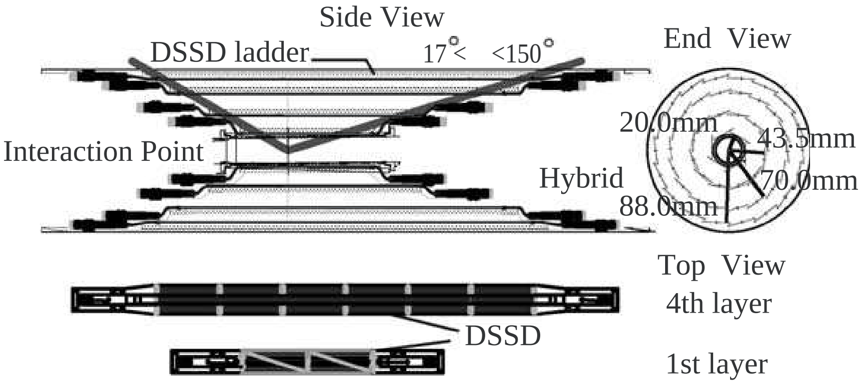


Figure 3.4: A side and front view of the silicon vertex detector.

the neighboring strips via capacitative coupling. Due to a good signal to noise ratio, the resolution is not deteriorated much by this approach [19].

detector configuration	SVD1	SVD2
beam pipe radius [mm]	20	15
number of layers	3	4
Angular coverage (acceptance)	$23° < \vartheta < 140°$	$17° < \vartheta < 150°$
radii of layers 1/2/3(/4) [mm]	30.0/45.5/60.5	20.0/43.5/70.0/88.8
strip pitch for $r\varphi$ [μm]	25 (50 for readout)	50 (65 for layer 4)
strip pitch for z [μm]	84	75 (73 for layer 4)

Table 3.1: Characteristic parameters of Belle SVD.

3.4 Other Belle Detectors

Let us continue with the description of remaining Belle detectors in the order from the IP outwards. The central drift chamber (CDC) is used to determine particle momentum by measuring its track in an external magnetic field [20]. A mixture of 50% helium (He) and 50% ethane (C_2H_6) is used as a filling gas and aluminum for the cathode and anode wires. The wires are aligned in the beam direction, each of them covering a cross-section region of about 5 mm × 5 mm. When a charged particle passes the detector, it ionizes the molecules of gas. Electrons travel towards cathode wire and gain enough energy to further ionize, thus making a sizable signal. The z coordinate is determined by a slight tilt of one set of the wires with respect to the other. The relative error on the transverse momentum ranges from 0.4% to 1.0% for the region from 1 GeV/c to 5 GeV/c, respectively:

$$\sigma_{p_t}/p_t = (0.19\, p_t c/\,\mathrm{GeV} \oplus 0.30/\beta)\,\%\,. \tag{3.4}$$

In the CDC, the size of the signal is also used to measure the particle ionization loss per unit length, $\mathrm{d}E/\mathrm{d}x$, with resolution of about 6% for the $e^+e^- \to e^+e^-$ process electrons. The $\mathrm{d}E/\mathrm{d}x$ provides a K/π separation up to $0.8\,\mathrm{GeV}/c$ and also in the region of relativistic rise above $2.5\,\mathrm{GeV}/c$.

Next is the threshold silica aerogel Cherenkov counters (ACC), shown in Fig. 3.5. The underlining physical principle is that a particle which travels through the matter faster than the speed of light in the matter emits the so-called Cherenkov radiation. The main purpose of ACC is to distinguish between pions and kaons. It works in the threshold mode. Suppose we measure the particle momentum to be $3\,\mathrm{GeV}/c$. When particles of this momenta pass through the aerogel of refractive index $n = 1.01$, the pions of mass $140\,\mathrm{MeV}/c^2$ emit the Cherenkov light while kaons of mass $494\,\mathrm{MeV}/c^2$ do not; hence the name threshold. Aerogels are transparent materials with refractive index in the range of 1.01 to 1.06. This type of detector only works for a limited momentum range, above a certain momentum both pions and kaons emit the Cherenkov photons and below a certain range both do not, leaving no way to distinguish between the two. In the forward direction the particles have in average higher momentum due to the Lorentz boost. A lower refractive index aerogel is used in that range to compensate for the fact. The refractive index of aerogels ranges from 1.010 in the forward to 1.028 in the backward direction to cover the momentum range from 3.5 to $1.2\,\mathrm{GeV}/c$. The ACC system is divided into over a thousand counter modules. Each counter module is equipped with one or two photon multiplier tubes (PMT), sensitive photon detection devices capable of detecting small number of photons of the order of ten (which is a typical number of photons produced in the above case), important to efficiently separate the particles above the Cherenkov threshold where small number of photons are detected from particles bellow the threshold with no photons.

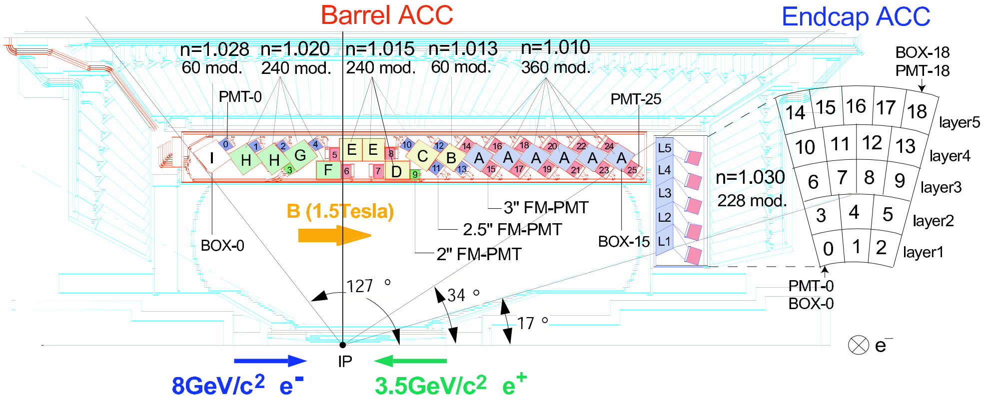


Figure 3.5: Cross-section view of the aerogel Cherenkov counters detector.

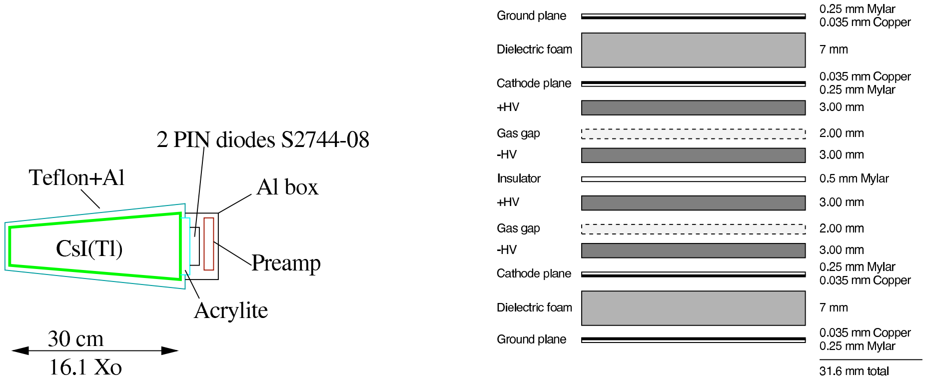
The time of flight (TOF) counter is placed just outside of the ACC, about one meter from the IP. Particles need only about $t \approx l/c = 3\,\mathrm{ns}$ to transverse this distance. Our time resolution requirements are even more demanding: we have to measure the differences in the time of flight between pions and kaons, the later being only about $300\,\mathrm{ps}$ for a $1\,\mathrm{GeV}/c$ momentum particle. The TOF consists of a scintillation rod with attached PMTs to detect the scintillation photons. This provides only the "stop" signal for the time of flight measurement. The "start" signal is determined as the time of the electron-positron bunch collision at the IP. The choice of material is motivated by its very fast time response, allowing for a total time resolution of about 100 ps and over three sigma separation between pions and kaons in the low momentum region below 1 GeV/c. In the higher momentum region, where the PID with the TOF counter gets worse, the ACC serves as the main detector.

In the outermost part of the Belle detector there are two calorimeters measuring the energy of the outgoing particles, an electromagnetic calorimeter (ECL) for the detection of electrons and photons, and K_L and muon detector (KLM) for the detection of the later. Two detectors are needed because the high energy electrons and photons produce the electromagnetic shower which has a significantly different characteristics than the signal left by other particles. First of all, it has a considerably shorter radiation length, distance at which $1 - e^{-1} = 63\%$ of the particle energy is absorbed in the material. If we therefore had the same calorimeter for measuring the energy of electrons and muons, electrons would deposit all of their energy in a small region of detector while muons would loose their energy over a much longer distance. One simply cannot do accurate measurements on both length scales at the same time. Besides, the ECL is the only Belle sub-detector which detects photons so the choice of the active material for the ECL is guided by the requirement for a large photon cross-section, which dictates a different choice of material than it is needed for the K_L and muon detection.

The ECL consists of almost $10,000$ CsI(Tl) crystals, each of a cross-section of about 6 cm × 6 cm and length of 30 cm (corresponding to 16 radiation lengths). The longer side is aligned in the radial direction, so that the total electron or photon energy is ideally deposited in one crystal unit. The light from each crystal is detected by two PIN photodiodes of size 2 cm × 2 cm with a preamplifier, which are glued to the large end face of the crystal, as shown in Fig. 3.6(a). The light-output corresponds to about 5000 photoelectrons per MeV, resulting in the few percent relative error on the measured energy. This allows for a detection of a neutral pion in the $\pi^0 \to \gamma\gamma$ decay mode with a mass resolution of about 5 MeV/c^2.

The Belle KLM detector consists of an alternating sandwich of 4.7 cm thick iron plates separated by a 4.4 cm thick slots with resistive plate counters (RPC) located outside the super-conducting solenoid. The alternating structure, 15 repetitions in total, is used due to the large size of the KLM detector. In this way, the energy deposited in the iron plate in the form of electrons and photons is detected in the RPC before it is dissipated as heat and the distribution of the energy loss along the particle track is also measured. A RPC is a charged particle detector utilizing a constant and uniform electric field produced by two parallel electrode plates made of a 2 mm thick soda-lime glass plates with high resistivity, as illustrated in Fig. 3.6(b). Between the electrodes is a gas mixture of argon, butan and freon. The 4 kV/mm electric field in the gas gap produces a discharge from one electrode to the other upon the passage of an ionizing particle. The iron plates also serve as the flux return for the solenoidal magnetic field.

The last Belle detector, the extreme forward calorimeter (EFC), is a small-angle elec-



(a) ECL (b) RPC

Figure 3.6: The basic building unit of the EFC: CsI(Tl) crystal with attached PIN photodiode (a) and KLM: resistive plate counter (b) detector.

tromagnetic calorimeter designed to extend the Belle angular acceptance to the extreme forward and backward region. The EFC is used to measure the $e^+e^- \to e^+e^-$ scattering and thus monitor the KEKB luminosity. In addition, it can improve the background rejection in rare B decays such as $B \to \tau\nu$ and the reconstruction efficiency for the two-photon physics.

Chapter 4

B meson reconstruction

Our first task is to reconstruct and select from the data sample recorded at the Belle detector only events with the B meson decaying to a pair of charged D mesons, $B^0 \to D^+D^-$. We reconstruct the B and D mesons from their decay products. To achieve a good vertex resolution and signal to background ratio, we use only those D meson decay modes for which the final state decay products (particles that are directly measured by the Belle detector) are charged particles. The modes with notable branching fractions ($\mathcal{B}$) are listed in Table 4.1. Charge conjugated modes are implied throughout this work.

decay	$\mathcal{B}\,[\%]$
$D^+ \to K^-\pi^+\pi^+$	9.2 ± 0.6
$D^+ \to \overline{K}^0\pi^+$	2.82 ± 0.19
$D^+ \to K^+K^-\pi^+$	0.89 ± 0.08
$D^+ \to \overline{K}^0K^+$	0.58 ± 0.06

Table 4.1: Branching fractions for the D meson decays to charged particles [11].

In the case of the $D^+ \to \overline{K}^0\pi^+$ and $D^+ \to \overline{K}^0K^+$ reconstruction channel, the $\overline{K}^0$ meson decays in 50% as K_S and in 50% as K_L, where K_S and K_L are short and long-lived mass eigenstates of the neutral kaon. The K_S is reconstructed in its decay to a pair of charged pions, $K_S \to \pi^+\pi^-$, which happens in $(68.95 \pm 0.14)\,\%$ of all K_S decays. Two of the above modes, $D^+ \to K^+K^-\pi^+$ and $D^+ \to K_SK^+$, were considered as a possibility and rejected due to a higher level of signal-faking background which will be discussed later. Only decays $D^+ \to K^-\pi^+\pi^+$ and $D^+ \to K_S\pi^+$ are therefore used for the reconstruction. Together they amount to about 10% of the D meson decays. A shorter term "$K\pi\pi$" channel is used for the B^0 reconstruction channel when both D mesons are reconstructed in the $D^+ \to K^-\pi^+\pi^+$ decay mode, while "$K_S\pi$" is used when at least one of the D mesons is reconstructed in the $D^+ \to K_S\pi^+$ decay. When terms $K\pi\pi$ and $K_S\pi$ are used for the D meson reconstruction, they refer to the $D^+ \to K^-\pi^+\pi^+$ and $D^+ \to \overline{K}^0\pi^+$ decays, respectively.

The event reconstruction is illustrated in Fig. 4.1. First, the K_S candidates are reconstructed from a pair of opposite-charged tracks. Second, the D mesons are reconstructed from three charged tracks, $K\pi\pi$ candidates, and from the K_S candidate and a charged pion candidate. At the last step, combinations of $D^{\pm}$ mesons are considered as a B^0 candidates. The particle decay vertices are reconstructed in the same order. In this way we avoid smearing of the vertex position due to the finite flight length of D and K_S mesons.

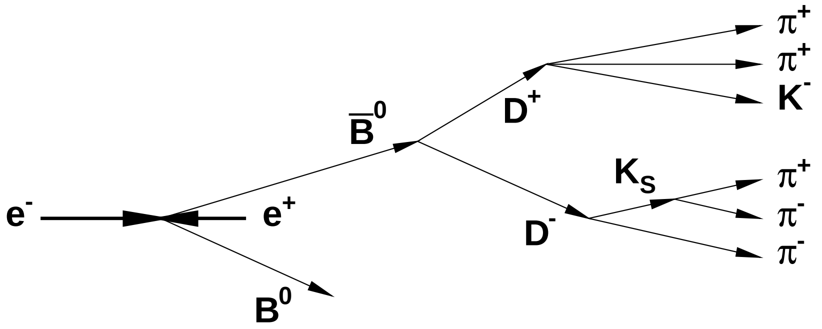


Figure 4.1: A schematic view of a B meson decay chain.

For the purpose of doing the CP analysis on the sample of reconstructed events, we also need to reconstruct the distance between the two B meson vertices Δz and the flavor of the other B meson q, the description of which is included in the next chapter.

Let us estimate how many $B^0 \to D^+D^-$ (signal) events we expect to find in our data sample. We have 535×10^6 $B\overline{B}$ pairs available for the analysis. Half of them are neutral $B^0\overline{B}^0$ ($\times\frac{1}{2}$) and either B^0 or $\overline{B}^0$ can decay to the pair of D mesons D^+D^- ($\times 2$). A previous measurement of the $B^0 \to D^+D^-$ branching fraction reported $\mathcal{B} = (1.9 \pm 0.51(\text{stat.}) \pm 0.30(\text{syst.})) \cdot 10^{-4}$ [21]. Both of the D mesons have to decay in one of our reconstruction modes, the probability for which is 0.1^2. Last of all, we have to reconstruct all six final state pions and kaons. The reconstruction efficiency is about 80% per track, so six tracks reconstruction efficiency is $0.8^6 = 26\%$, if we assume it is independent of the number of tracks. Taking all this into account we estimate

$$N(B^0 \to D^+D^-) \approx 535 \times 10^6 \cdot \frac{1}{2} \cdot 2 \cdot 1.9 \cdot 10^{-4} \cdot (10\%)^2 \cdot 0.8^6 = 265$$

Due to additional selection criteria which are needed to reject wrong combinations of particles (background), event reconstruction efficiency is actually lower, about 12% rather than 26%. Based on this estimation we expect to reconstruct of the order of hundred $B^0 \to D^+D^-$ events, which means our error of the CPV parameters will be statistically dominated – predominantly coming from the statistical fluctuations in the data sample.

In the sample of reconstructed B mesons there are also many B candidates, which are not signal B mesons decaying as $B^0 \to D^+D^-$. We separate signal from background on the basis of two reconstructed kinematic variables, the energy difference

$$\Delta E = E_B^* - E_{\text{beam}}^* \tag{4.1}$$

and the beam-energy-constrained mass

$$M_{\text{bc}} = \sqrt{\left(E^*_{\text{beam}}/c^2\right)^2 - \left(p^*_B/c\right)^2}\,, \tag{4.2}$$

where E^*_B, E^*_{beam}, and p^*_B are the B meson energy, the beam energy, and the B meson momentum in the CM system, respectively. The M_{bc} and ΔE distributions of correctly reconstructed $B^0 \to D^+D^-$ events have a peak with the central values $M_{\text{bc}} = m_B$ and $\Delta E = 0$, smeared due to the detector resolution. We apply different selection criteria to reduce the combinatorial background. In general, tighter selection criteria reduce the level of background, but the signal reconstruction efficiency is decreased at the same time. For the branching fraction measurement it is common to choose such critera that give the largest value of the figure-of-merit (FOM),

$$\text{FOM} \;=\; \frac{S}{\sqrt{S+B}}\,, \tag{4.3}$$

where S and B are the expected number of signal and background events, respectively. The value of $\sqrt{S+B}$ is the level of the Poisson fluctuation in the total number of events, so maximizing the above FOM is equivalent to the minimization of the expected relative error, $\sqrt{S+B}/S$.

In order not to introduce any bias to our analysis and result, we avoid analyzing the data sample until all the reconstruction criteria are decided on. Instead of working with the data sample, we therefore start with a sample of simulated events, referred to as Monte Carlo (MC). The particle decays are simulated by the EvtGen generator [22]. Once the particles are generated, their propagation and the Belle detector response is simulated using the Geant package [23]. The simulated detector signal is then analyzed and particles are reconstructed in the same way as is done for the actual data measurements. The generic MC sample simulates all (known) $e^+e^- \to \ldots$ processes at KEKB, the fraction of each process reflecting our current understanding of its branching ratio. We do a separate simulation of $e^+e^- \to B^0\overline{B}^0$, $e^+e^- \to B^+B^-$, $e^+e^- \to q\bar{q}$ (where q denotes an u, d or s quark), and $e^+e^- \to c\bar{c}$ processes. For a reference to be used later, we name neutral B pair production $e^+e^- \to B^0\overline{B}^0$ as "mixed", charged $e^+e^- \to B^+B^-$ as "charged", $e^+e^- \to u(ds)\bar{u}(\bar{d}\bar{s})$ as "uds" and $e^+e^- \to c\bar{c}$ as "charm". The mixed and charged MC samples are called "$B\overline{B}$" MC while uds and charm are "continuum" MC. The analysis is first tested and all event selection criteria are optimized using the MC sample without looking at the data – this is why we name such an approach a blind analysis.

We do not limit our tests and cross-checks only to the simulated events. After all, the MC only simulates the features that we know of. It is therefore important to test the CP analysis on the real data to ensure that we understand the tagging and vertex reconstruction well. We choose the $B^0 \to D_s^+D^-$ decays to be our control sample. Since there is only one decay amplitude at the tree level and the leading penguin contributions have the same CKM structure as the tree contribution, no sizable CPV is expected for this decay. Nevertheless, for the same reason, we know that the CP parameters for this decay are small and therefore the analysis of the data sample of $B^0 \to D_s^+ D^-$ events serves as a valuable cross-check of our method.

In the first section of this chapter we describe the event selection criteria. We continue with a measurement of the branching fraction, which starts with a determination of the signal yield (the number of reconstructed signal events) and includes also a more detailed background study. In the last section of this chapter we describe the reconstruction of $B^0 \to D_s^+ D^-$ decays, which are used as a control sample for the CP analysis.

4.1 Event Selection Criteria

Event selection criteria are chosen in such a way as to obtain the optimal figure of merit defined by Eq. 4.3. A generic MC sample of $660 \cdot 10^6$ $B\overline{B}$ events and the corresponding number of the continuum MC events is used to determine the expected number of background events N_{bcg}. Apart from the generic MC we generate and simulate also one million $B^0 \to D^+ D^-$ decays (with D mesons decaying as $K\pi\pi$ or $K_S\pi$) to study the properties of the signal events and the expected number of signal events N_{sig}. This sample is referred to as signal MC. The expected number of signal events is rescaled to the full data statistics using previous Belle measurement of the $B^0 \to D^+ D^-$ branching fraction [21].

The selection criterion is essentially a requirement on a value of a particular parameter. For example, the invariant mass of the $K\pi\pi$ combination coming from the $D^+ \to K^-\pi^+\pi^+$ decay is close to the D meson mass, the difference is only due to the detector resolution. On the other hand, random background combinations of three tracks have more or less uniform distribution over a much larger range in the invariant mass. Selecting only the events within a certain range around the D mass therefore reduces the level of background. Exactly which range gives the best performance is determined by optimizing the FOM using the MC sample. For any of the selection criteria optimization, the requirements on all other variables are applied as well. In this way, only the gain due to the particular selection criterion is shown and we avoid the overlap of different criteria. For example, if any of the selection criteria only rejects the events which are already rejected by one of the others, there is no point in having two criteria instead of one. Of course, all criteria can not be optimized at the same time, we repeat the procedure a few times to obtain the best choice.

The fit of the D and B meson vertices is also a way of distinguishing the signal from the background, since the random combinatorial background tracks might not fit to a common vertex. For the same reason, information on the quality of the vertex fit can also be used to suppress the background. We apply such criteria only for the analysis of the Δt distribution, as is discussed in the next section. This criteria are not required for the branching fraction measurement, and neither is a reconstruction of the tag side vertex.

The next paragraphs discuss several event selection criteria. The most selective are the kaon particle identification (PID) criteria and the requirements on the kinematic variables such as the D meson invariant mass and the B meson invariant mass and energy in the CM system.

For any of the requirements we quote the selection criteria efficiency and fake rate. The efficiency is the fraction of tracks satisfying the event selection criteria,

$$\epsilon = \frac{\text{Number of signal tracks passing the selection criteria}}{\text{Number of all signal tracks}} . \tag{4.4}$$

For the calculation of efficiency we only take into account those tracks, which are coming from the signal B meson since we only wish to know by how much is the number of signal events reduced by a particular requirement. Similarly, the signal reconstruction efficiency is the fraction of reconstructed signal events which pass all event selection criteria. On the other hand, fake rate describes how many of the background events are accepted by the selection criteria. It is defined as the fraction of background tracks passing the selection criteria (faking a signal),

$$f = \frac{\text{Number of background tracks passing the selection criteria}}{\text{Number of all background tracks}} . \tag{4.5}$$

4.1.1 Particle Identification Criteria

We start the event reconstruction by selecting pion and kaon tracks from the number of reconstructed charged tracks. Any time we mistake an electron, a muon or a proton for a pion or a kaon, we potentially increase the level of background events. The same happens if we mistake pion for a kaon or vice versa. We therefore use the particle identification criteria to reduce the number of background combinations, the main source of which is the abundance of pions.

To distinguish between pions and kaons we combine information from the CDC, TOF and ACC counters into a kaon likelihood ratio [24],

$$\mathcal{L}_K = \frac{\mathcal{L}(K^\pm)}{\mathcal{L}(K^\pm) + \mathcal{L}(\pi^\pm)} . \tag{4.6}$$

The pion likelihood is $\mathcal{L}_\pi = 1 - \mathcal{L}_K$. Figure 4.2 shows the $\mathcal{L}_K$ distribution for pions and kaons. The kaon tracks have on the average a higher kaon likelihood ratio, the distribution has a peak at $\mathcal{L}_K \approx 1$ while pions have on average lower values of $\mathcal{L}_K$. The peak in the pion distribution around $\mathcal{L}_K \approx 0$ corresponds to a high pion likelihood, $\mathcal{L}_\pi \approx 1$. We select kaons by requiring the value of $\mathcal{L}_K$ to be larger than a certain value, c_K. Following the same argument, we require the pion tracks to have the value of $\mathcal{L}_K$ smaller than a certain value, c_π. Figure 4.3 shows the FOM as a function of c_K and c_π. We select such a value of c_K and c_π that gives maximal FOM, $\mathcal{L}_K$ is required to be higher than $c_K = 0.55$ for kaon and lower than $c_\pi = 0.9$ for pion candidates. The corresponding track PID efficiencies and fake rates are $\epsilon_K = 92\%$, $f_{\pi\to K} = 5\%$, $\epsilon_\pi = 99\%$ and $f_{K\to\pi} = 28\%$, where $f_{\pi\to K}$ is a probability for a pion track to fake (to be mistaken for) a kaon and $f_{K\to\pi}$ is the probability for a kaon track to fake a pion.

The above criteria only help to distinguish between pions and kaons. Electrons, muons and protons are not a major source of background since there are not that abundant

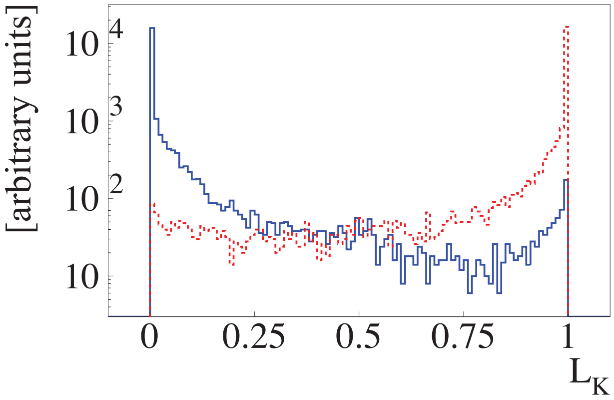
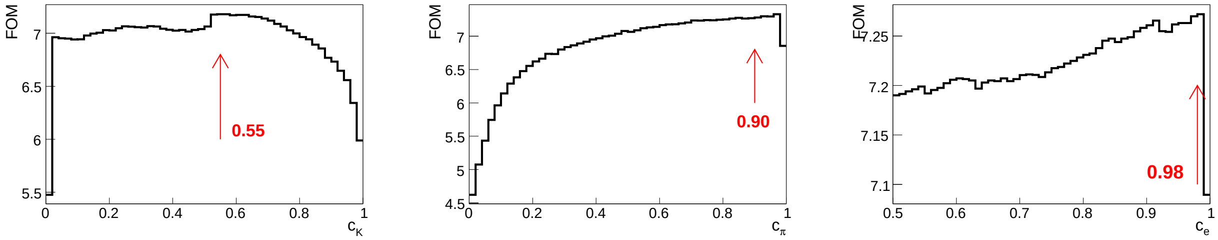


Figure 4.2: The $\mathcal{L}_K$ distribution for the pion (full histogram) and kaon (dashed histogram) tracks for the signal MC $B^0 \to D^+D^-$ events. Note that the scale is logarithmic.



(a) (b) (c)

Figure 4.3: The figure-of-merit $S/\sqrt{S+B}$ as a function of c_K (a), c_π (b) and c_e (c). For each figure the event selection requirement on all the other variables is applied.

among the B (D) meson decay products. We therefore only apply additional criteria to reject electrons. We form electron likelihood as

$$\mathcal{L}_e = \frac{\mathcal{L}(e^\pm)}{\mathcal{L}(e^\pm) + \mathcal{L}(K^\pm) + \mathcal{L}(\pi^\pm)} \tag{4.7}$$

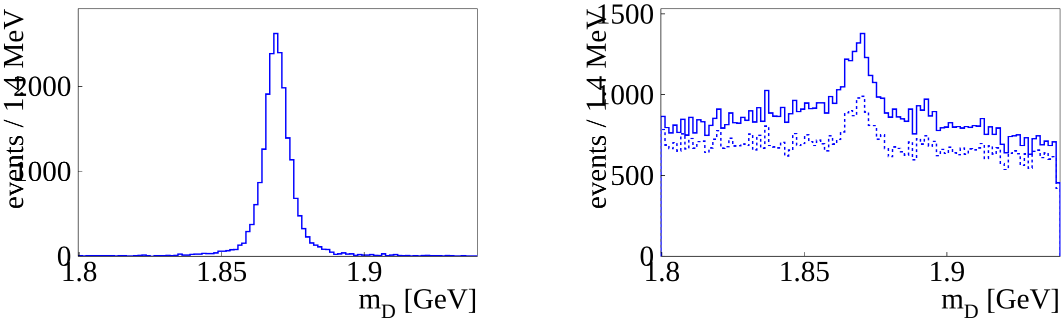
and reject the particles which have a strong indication of being the electrons, with $\mathcal{L}_e > c_e$. We determine the optimal value of c_e to be $c_e = 0.98$, as shown in Fig. 4.3 (c).

4.1.2 Kinematic Variables

There are several kinematic variables which can be used to distinguish between the signal and the background events. All of them are based on the kinematic constraints imposed by the momentum and energy conservation. First of them is the mass of the D meson. If the $K^\mp\pi^\pm\pi^\pm$ ($K_S\pi^\pm$) combination is coming from the $D^\pm \to K^\mp\pi^\pm\pi^\pm$ ($D^\pm \to K_S\pi^\pm$) decay, than the invariant mass $m_{K^\mp\pi^\pm\pi^\pm\,(K_S\pi^\pm)}$ is equal to the D meson mass, smeared due to the energy and momentum measurement resolution of a few MeV (MeV/c). On the other hand, the distribution of the reconstructed D mass for random background combinations is more or less constant in the same range, as shown in Fig. 4.4. However, it should be noted that the background events also have a peak in the D mass distribution, because one of the reconstructed D mesons is often an actual D meson, decaying as $D^+ \to K^-\pi^+\pi^+$ (but not coming from the $B^0 \to D^+D^-$ decay). Using the constraint

$$\left| m_{K^\mp\pi^\pm\pi^\pm\,(K_S\pi^\pm)} - m_{D^\pm} \right| < c_{m_{D^\pm}} \;, \tag{4.8}$$

we can reject a large part of the background events. We determine the value of $c_{D^\pm}$ by optimizing the FOM, as shown in Fig. 4.5. Based on the FOM dependency shown in Fig. 4.5 we require the invariant mass of the $K^\mp\pi^\pm\pi^\pm$ combination to be within $c_{D^\pm} = 10\,\mathrm{MeV}$ $(2.4\,\sigma)$ of the $D^\pm$ nominal mass.



(a) signal (b) background

Figure 4.4: The distribution of the $K^-\pi^+\pi^+$ or $K_S\pi^+$ invariant mass distribution for signal (a) and background (b) combinations (MC). The dashed line in Fig. (b) shows the distribution for events with only one B meson candidate. In this case the combinatorial background shape is not distorted due to the best candidate selection. The fraction of events with more than one B candidate is about 16%. After the D mass criteria is applied, the fraction of such events decreases to 6%.

4.1.3 The Track Impact Parameters

The track impact parameter measures the track's distance from the interaction point (IP). As illustrated in Fig. 4.6, we parametrize the distance in the cylindrical coordinate system, where Δz_{IP} and Δr_{IP} are the distances from the IP to the particle trajectory in the z and $r\varphi$ directions, respectively. Since tracks from the D meson decays come from the IP region, rejection of tracks with large impact parameters can be used to reject the beam background (showers from beam particles scattered by the residual gas or by the interactions with the other beam) and tracks coming from the secondary interactions on a detector material. The optimization of the FOM, as shown in Fig. 4.7, gives the following criteria: $|\Delta r_{\mathrm{IP}}| < 2\,\mathrm{cm}$ and $|\Delta z_{\mathrm{IP}}| < 4\,\mathrm{cm}$. This condition is not applied to the pion tracks which are used to reconstruct the K_S meson, because the K_S is a long-lived particle transversing the distance of the order of a few centimeters.

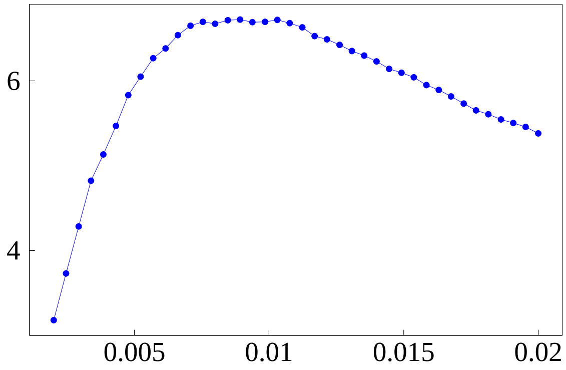
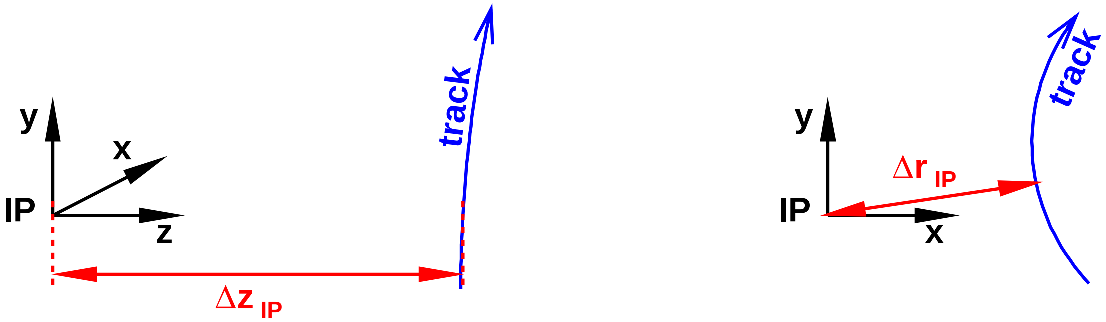


Figure 4.5: The figure-of-merit $S/\sqrt{S+B}$ as a function of $c_{m_{D^\pm}}$.



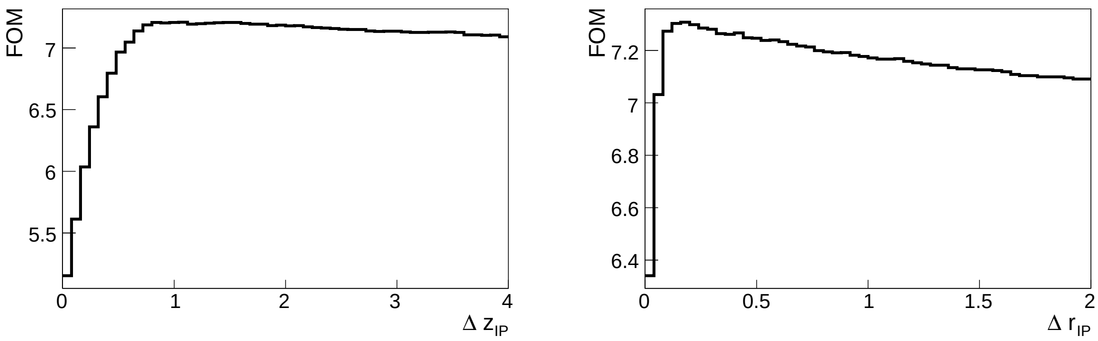
(a) track distance from the IP

(b) front view

Figure 4.6: Track impact parameters Δz_{IP} (a) and Δr_{IP} (b).

4.1.4 The $\mathbf{K_S}$ Reconstruction

The K_S is reconstructed in the $K_S \to \pi^+\pi^-$ decay mode. Its identification takes the advantage of the following characteristics. First, the two pions are coming from a secondary vertex and therefore have a non-zero track impact parameter, as illustrated in Fig. 4.8 (a). Second, the K_S decay vertex is displaced from the IP, $\mathbf{r} = \mathbf{r}_{K_S\text{vtx}} - \mathbf{r}_{\text{IP}} \neq 0$, as shown in Fig. 4.8 (b). Third, since the K_S meson originates from the IP, its momentum is parallel to the displacement vector $\mathbf{r}$, see Fig. 4.8 (b). We therefore use the two pion's impact parameter Δr_{IP}, the K_S vertex displacement $\mathbf{r}$, and the angle α between the momentum direction $\mathbf{p}$ and particle path $\mathbf{r}$ to separate signal from background. Since the spread of the event-by-event z coordinate of the IP is large compared to the spread of the x or y coordinate, for each of the three criteria described above only the transverse component is used to select the K_S candidates. In order to reject background, the angle between the



(a) (b)

Figure 4.7: The FOM for the cut on the parameters $\Delta z_{\rm IP}$ (a) and $\Delta r_{\rm IP}$ (b).

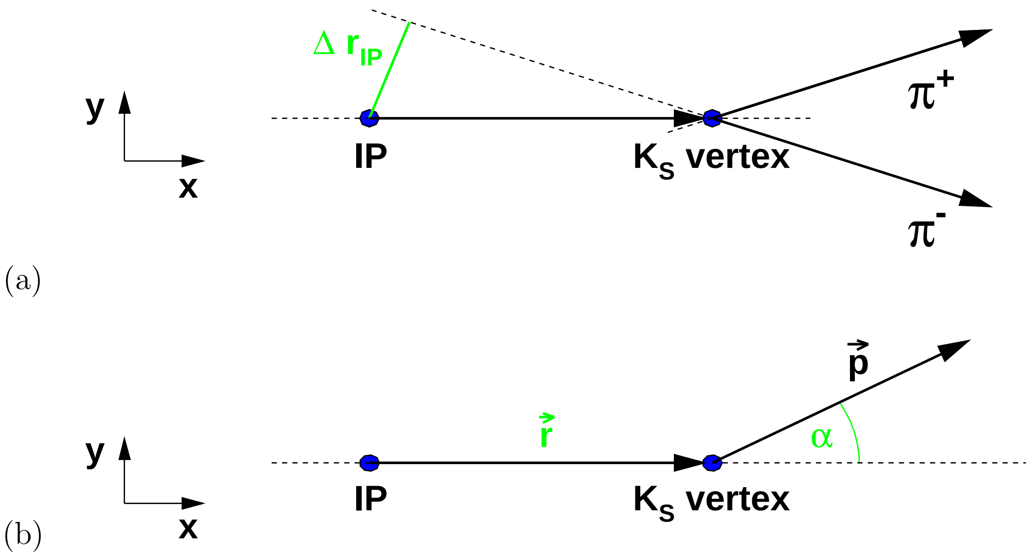


Figure 4.8: For the K_S selection criteria we use constraints on the pion track impact parameter (a), displacement from the IP $\mathbf{r}$ and angle α (b).

reconstructed momentum $\mathbf{p}_\perp = (r_x, r_y, 0)$ and $\mathbf{r}_\perp = (r_x, r_y, 0)$, denoted by α in Fig. 4.8 (b), is required to be smaller than a certain value. Since the momentum and vertex resolution depends on the K_S momentum, the values of the cuts on parameters $\Delta r_{\rm IP}$, $r_\perp$ and α are momentum-dependent. The K_S candidates are divided into three groups depending on their momentum. Particles with $p < 0.5\,{\rm GeV}/c$ are required to have $\Delta r_{\rm IP} > 0.5\,{\rm mm}$ (for both pions) and $\alpha < 0.3\,(17^\circ)$. Next, particles with $0.5\,{\rm GeV}/c < p < 1.5\,{\rm GeV}/c$ are required to have $\Delta r_{\rm IP} > 0.3\,{\rm mm}$, $\alpha < 0.1$ and $r_\perp > 0.8\,{\rm mm}$. The highest-momentum kaons, $p > 1.5\,{\rm GeV}/c$, are required to have $\Delta r_{\rm IP} > 0.2\,{\rm mm}$ (for both pions), $\alpha < 0.03$ and $r_\perp > 2.2\,{\rm mm}$. Additional criteria for the z coordinate of the K_S decay vertex, $|r_z| < 10\,{\rm cm}$, is applied to reject background from the displaced tracks. The above values are standard

Belle selection criteria for the K_S mesons.

4.1.5 Rejection of the Continuum Events

Rejection of the continuum background is based on the event topology. On one hand, B mesons from the $e^+e^- \to B\overline{B}$ events have rest energy mc^2 only slightly smaller than the available energy E and are therefore almost at rest in the CMS, $pc \ll E$. For this reason, the direction of the B meson decay products in the CMS is distributed uniformly in the solid angle and the event topology is more spherical, as illustrated in Fig. 4.9 (a). On the other hand, the two light quarks originating from the $e^+e^- \to q\bar{q}$ process, $q \in \{u, d, s\}$, have very high momentum and boost in the CMS. For that reason, the particle tracks tend to be more jet-like, as shown in Fig. 4.9 (b). The charm production processes, $e^+e^- \to c\bar{c}$ can not be well rejected in this way.

(a) $e^+e^- \to B\overline{B}$

(b) $e^+e^- \to q\bar{q}$; $q \in \{u, d, s\}$

Figure 4.9: Event topology for the $B\overline{B}$ (a) and continuum events (b).

The event topology is described in the basis of Legendre polynomials. We define Super-Fox-Wolfram moments [25]

$$R_l^{SO,OO} = \frac{\Sigma_{i,j}\mathbf{p}_i\mathbf{p}_j P_l(cos\theta_{i,j})}{\Sigma_{i,j}\mathbf{p}_i\mathbf{p}_j}\,, \tag{4.9}$$

where $P_l(x)$ is the l-th Legendre polynomial, p is the particle momentum and $\theta_{i,j}$ is the angle between the momentum of the particles i and j. For the R_l^{SO} terms, the summation is done for $i \in \{\text{signal}\}$, $j \in \{\text{other}\}$ particles and for R_l^{OO} SFW moments, both indices i and j are running over non-signal particle momenta, $i, j \in \{\text{other}\}$. Only moments which give good separation between the signal and background are used: R_3^{OO}, R_2^{OO}, R_4^{OO}, R_2^{SO} and R_4^{SO}. Additional parameters thrust angle and transverse sphericity $S_\perp$ are used to distinguish signal from background. The thrust angle is the angle between the thrust axis of the B_{CP} and B_{tag} final state particles in the CM system,

$$\mathbf{n}_{S(O)} = \frac{\Sigma_i\mathbf{p}_i}{|\Sigma_i\mathbf{p}_i|}\,, \tag{4.10}$$

where $S(O)$ stands for a sum over signal (other) particles, as above. The transverse sphericity is defined as

$$S_\perp = \frac{\Sigma_i\,|\mathbf{p}_{i\,\perp}|}{\Sigma_i\,|\mathbf{p}_i|}\,, \tag{4.11}$$

where $\mathbf{p}_\perp$ is a transverse momentum relative to the thrust axis $\mathbf{n}_S$ and the sum in the numerator does not include particles that do not fall within 45° cone of the B_{CP} thrust axis, $\mathbf{n}_S\mathbf{p}_i/|\mathbf{p}_i| < \cos 45°$. The Fisher linear discriminant is used to calculate an optimal discriminant based on the above seven parameters [26, 27]. We use MC to determinate the discriminant coefficients. The signal MC is used to describe the $B\overline{B}$ events and uds generic MC is used to describe the continuum events. All other event selection criteria are applied before the continuum rejection criteria optimization. The distribution of the composed variable F (for Fisher discriminate) for signal and background events is parameterized and signal likelihood probability is calculated as a function of F.

Besides the event topology, the polar angle between the B meson direction and the z axis in the CM system ϑ can also give some information about the signal probability. The B meson pairs come from a decay of the $\Upsilon(4S)$ resonance of spin $S = 1$, therefore their $\cos\vartheta$ distribution has a $\cos^2\vartheta$ dependence.

Continuum events are suppressed by imposing a criteria on the combined SFW moments based Fisher discriminant F and $\cos\vartheta$ likelihood ratio cF, which is shown in Fig. 4.10. Continuum suppression cut is optimized separately in different r-bins, where r is the flavor reconstruction quality that will be discussed in the next chapter, $r \approx 0$ for events with no tagging information and $r \approx 1$ for unambiguous flavor determination. The optimal criteria is found to be $cF > 0.45$ for $r < 0.25$, $cF > 0.60$ for $0.25 < r < 0.75$ and $cF > 0.1$ for $r > 0.75$.

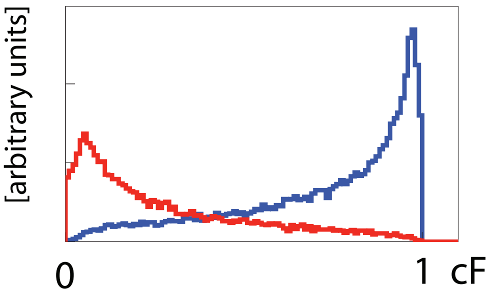


Figure 4.10: The SFW and $\cos\vartheta$ likelihood ratio cF for the signal (in blue) and background (in red) events.

4.1.6 Best Candidate Selection

After applying all of the event selection criteria, 6% of the signal events have more than one B^0 candidate. The candidate with the smallest value of $\chi^2_{D^\pm\ \mathrm{mass}}$, where

$$\chi^2_{D^\pm\ \mathrm{mass}} = \left(\frac{\Delta m_{D^+}}{\sigma_{m_{D^+}}}\right)^2 + \left(\frac{\Delta m_{D^-}}{\sigma_{m_{D^-}}}\right)^2 \tag{4.12}$$

and $\Delta m_{D^\pm} = m_{K\pi\pi/K_S\pi} - m_{D^\pm}$ is a difference between the reconstructed and nominal $D^\pm$ meson mass, is selected as the best candidate. This criteria is successful in 81% of the events with multiple B^0 candidates. In remaining 19% of the events it usually fails due to the cross-feed of the pion or kaon from the other B meson, which can not really be helped unless a tighter PID selection criteria is applied.

4.2 Branching Fraction Measurement

Branching fraction $\mathcal{B}$ for this decay is determined from the total signal yield of both reconstruction modes and averaged reconstruction efficiency. We first describe how we determine the signal yield, and give the result for the signal yield for our data sample. In the next section we show a more detailed background study, which includes a determination of the level of a signal-faking background contribution. We conclude this section with a new result for the branching fraction of these decays, superseding the previous measurements on the subject [21, 28].

4.2.1 Fit of the M_{bc} and ΔE Distributions

The signal yield is determined by an unbinned two-dimensional (2D) maximum likelihood (ML) fit to the M_{bc} and ΔE distribution in the region $5.20\,\mathrm{GeV}/c^2 < M_{\mathrm{bc}} < 5.29\,\mathrm{GeV}/c^2$ and $-0.05\,\mathrm{GeV} < \Delta E < 0.10\,\mathrm{GeV}$. The region of $\Delta E < -50\,\mathrm{MeV}$ is not used due to an additional bump of $B^0 \to D_s^+ D^-$ events around $\Delta E \approx -70\,\mathrm{MeV}$. The M_{bc} and ΔE distributions for signal events are described by a Gaussian function (G) and a sum of two Gaussian functions of the same mean, respectively. On the other hand, the M_{bc} and ΔE distributions for background are described by an ARGUS function [29] and the first order polynomial, respectively.

$$\begin{aligned} w_{\mathrm{sig}}(M_{\mathrm{bc}}, \Delta E) &= \frac{1}{\mathcal{N}_{\mathrm{sig}}}\, G(M_{\mathrm{bc}}|\mu^{M_{\mathrm{bc}}}\sigma^{M_{\mathrm{bc}}}) \cdot && (4.13)\\ &\quad \left(f^{\Delta E} G(\Delta E|\mu^{\Delta E}\sigma_1^{\Delta E}) + (1-f^{\Delta E}) G(\Delta E|\mu^{\Delta E}\sigma_2^{\Delta E})\right) \text{ and} && \\ w_{\mathrm{bcg}}(M_{\mathrm{bc}}, \Delta E) &= \frac{1}{\mathcal{N}_{\mathrm{bcg}}}\, \mathrm{Argus}(M_{\mathrm{bc}}|\ldots) \cdot (1+\alpha\Delta E)\,, && (4.14) \end{aligned}$$

where $\mathcal{N}_{\mathrm{sig}}$ and $\mathcal{N}_{\mathrm{bcg}}$ are normalization constants such that

$$\int_{-0.05\,\mathrm{GeV}}^{0.10\,\mathrm{GeV}} \mathrm{d}(\Delta E) \int_{5.20\,\mathrm{GeV}/c^2}^{5.29\,\mathrm{GeV}/c^2} \mathrm{d}(M_{\mathrm{bc}}) w_{\mathrm{sig\,(bcg)}}(M_{\mathrm{bc}}, \Delta E) = 1$$

and the ARGUS function is defined as

$$\begin{aligned} \mathrm{Argus}(x|\Delta x, E_{\mathrm{beam}}, \beta) &= (x-\Delta x)\,\sqrt{y}\;e^{\beta\,y}\,,\ \text{where} \\ y &= 1-((x-\Delta x)/E_{\mathrm{beam}})^2\,. \end{aligned} \tag{4.15}$$

The event likelihood function is a sum of the signal and background probability density functions,

$$\mathcal{L}_{\mathrm{ev}}(M_{\mathrm{bc}}, \Delta E) = f_{\mathrm{sig}} w_{\mathrm{sig}}(M_{\mathrm{bc}}, \Delta E) + (1 - f_{\mathrm{sig}}) w_{\mathrm{bcg}}(M_{\mathrm{bc}}, \Delta E)\,, \tag{4.16}$$

where f_{sig} is the signal fraction. The parameters describing the M_{bc} and ΔE distributions and f_{sig} are determined by maximizing the likelihood function for all measured events,

$$\mathcal{L} = \Pi_i\,\mathcal{L}_i(M_{\mathrm{bc}}, \Delta E)\,, \tag{4.17}$$

which is equivalent to maximizing

$$\log\mathcal{L} = \Sigma_i \log\mathcal{L}_i \tag{4.18}$$

since logarithm is a strictly monotonic function. It can be observed that any constant factor in the event likelihood function $\mathcal{L}_{\mathrm{ev}}$ results only in an offset in the value of the likelihood function $\log\mathcal{L}$ and does not change the position of the maxima in respect to the value of any of the variables. This is why we can express the event likelihood function $\mathcal{L}_{\mathrm{ev}}$ with the probability density function w rather than the probability distribution function $\mathrm{pdf} = \int \mathrm{d}(\Delta E) \int \mathrm{d}(M_{\mathrm{bc}}) w(M_{\mathrm{bc}}, \Delta E) = w(M_{\mathrm{bc}}, \Delta E)\,\Delta(\Delta E)\Delta(M_{\mathrm{bc}}) = w \cdot \mathrm{const}$. However, for the same reason, also the number of signal and background events is determined only up to a constant factor. To solve this problem, we require the sum of all events as determined from the likelihood function, integral of $\int_{\Delta E}\int_{M_{\mathrm{bc}}} n\,\mathcal{L}_{\mathrm{ev}}$, where n is a constant factor and $\mathcal{L}_{\mathrm{ev}}$ is defined in Eq. 4.16, to be equal to the number of events in the data sample N. Since $\mathcal{L}_{\mathrm{ev}}$ is normalized, it follows that $n = N$.

The signal yield is the number of reconstructed signal events in the M_{bc} and ΔE signal region,

$$N_{\mathrm{sig}} = N f_{\mathrm{sig}} \int_{-0.03\,\mathrm{GeV}}^{0.03\,\mathrm{GeV}} \mathrm{d}(\Delta E) \int_{5.27\,\mathrm{GeV}/c^2}^{5.29\,\mathrm{GeV}/c^2} \mathrm{d}(M_{\mathrm{bc}})\, w_{\mathrm{sig}}(M_{\mathrm{bc}}, \Delta E) = N f_{\mathrm{sig}} I_{\mathrm{sig}}\,, \tag{4.19}$$

where N is the number of events in the data sample and I_{sig} is the above integral; since the integral range is large compared to the widths of the signal Gaussian functions, $I_{\mathrm{sig}} \approx 1$.

In the case of the ML fit, the $n\,\sigma$ confidence intervals for the fitted parameters are the regions for which the change in the value of likelihood function compared to its maximal value is smaller than $n^2/2$, $\mathcal{L}^{\max} - \mathcal{L} < n^2/2$. The $1\,\sigma$ error on the fitted parameter is the change in the parameter at which the value of the ML function is by $1/2$ smaller than its maximal value. This gives an unbiased estimate of the error on the parameters. The covariance matrix $\mathrm{V}_{ij} = \rho_{ij}\sigma_i\sigma_j$ is obtained from the second derivatives of the likelihood function at the maxima, $\mathrm{V}_{ij} = \partial^2\mathcal{L}/\partial x_i \partial x_j$, and ρ_{ij} is the correlation coefficient between the i^{th} and j^{th} parameter. If we neglect the small correlation between the f_{sig} and other fitting parameters, the error on the number of signal events is given by

$$\sigma_{N_{\text{sig}}} = \left(\frac{\sigma_N}{N} \oplus \frac{\sigma_{f_{\text{sig}}}}{f_{\text{sig}}} \oplus \frac{\sigma_I}{I} \right) N_{\text{sig}} \,. \tag{4.20}$$

The uncertainty in the parameters describing the M_{bc} and ΔE probability density functions effects only the value of integral I_{sig}. The change in the signal yield on this account is found to be negligible, $\sigma_I/I < 0.1\%$. The error on the f_{sig} is determined by the fit. The term σ_N/N accounts for the Poisson fluctuations in the total number of events, $\sigma_N/N = 1/\sqrt{N} \approx 2\%$. The error on the signal yield is therefore obtained from a sum of squares of the two non-negligible contributions, σ_N/N and $\sigma_{f_{\text{sig}}}/f_{\text{sig}}$.

Alternatively, one can do an extended unbinned ML fit by maximizing the extended likelihood function

$$\mathcal{L}_{\text{ext}} = \frac{\mathrm{e}^{-\mu_N}\, \mu_N^N}{N!} \, \mathcal{L} \,, \tag{4.21}$$

where the additional parameter μ_N is the expected number of reconstructed events. The additional term describes the Poisson fluctuations of the number of reconstructed events N. Since the new term is uncorrelated with the value of $\mathcal{L}$, the value of $\mathcal{L}_{\text{ext}}$ is maximal when the value of $\mathcal{L}$ and $\mathrm{e}^{-\mu_N}\, \mu_N^N/N!$ are both maximal,

$$\log \mathcal{L}_{\text{ext}} \;=\; -\mu_N + N \, \log \mu_N - \log N! + \mathcal{L} \,. \tag{4.22}$$

The additional terms depend only on μ_N (observed number of events in a data sample N is a constant) and their value is maximal when the derivative with respect to μ_N is zero, giving a constraint

$$-1 + \frac{N}{\mu_N} \;=\; 0 \,, \tag{4.23}$$

or, as expected, $\mu_N = N$. The extended ML fit therefore yields $\mu_N = N$, and the fit result for all other parameters is the same as before. This has been confirmed by performing an extended ML fit (maximizing $\mathcal{L}_{\text{ext}}$) and comparing its results to the ML fit. The same study confirmed that the error on μ_N as determined by the fit is equal to $\sqrt{N}$, as expected. Furthermore, performing the extended ML fit with a different choice of parameters, $\mu_S = f_{\text{sig}} N$ and $\mu_B = (1 - f_{\text{sig}})N$, which directly gives the error on the number of signal events μ_S, resulted in errors on μ_S consistent with those calculated by Eq. 4.20 after performing the (non-extended) ML fit.

A toy MC study was done to confirm that the fit result is unbiased and the errors are neither under- nor over-estimated. Samples of N signal and background events were simulated with event M_{bc} and ΔE distributed according to the Eq. 4.16. To be sure we understand the breakdown of the different error contributions, different simulations were done. First, the number of simulated events was fixed to $N = 2000$ and only f_{sig} was fitted while the other parameters were fxed to the same values as used in the simulation. Second, $N = 2000$ was fixed and all parameters describing the M_{bc} and ΔE distributions were determined by the fit. Then, both cases were repeated with the number of simulated events being distributed according to the Poisson distribution of mean $\mu_N = 2000$. Table 4.2 shows the results of the different simulations.

	fixed N		Poisson distribution	
	$N = 2000$, f_{sig}	$N = 2000$, all	$\mu_N = 2000$, f_{sig}	$\mu_N = 2000$, all
$N_{\mathrm{sig,\ sim}}$	160.2 ± 12.0		160.4 ± 12.8	
$N_{\mathrm{sig,\ fit}}$	160.2 ± 13.8	160.5 ± 14.6	160.4 ± 14.3	160.9 ± 15.1
$\sigma_{N_{\mathrm{sig}},\ \mathrm{fit}}$	13.7 ± 0.4	14.5 ± 0.6	13.8 ± 0.5	14.6 ± 0.6
$\sigma_{N_{\mathrm{sig}}}$	13.7	14.5	$13.8 \oplus 3.6 = 14.3$	$14.6 \oplus 3.6 = 15.0$

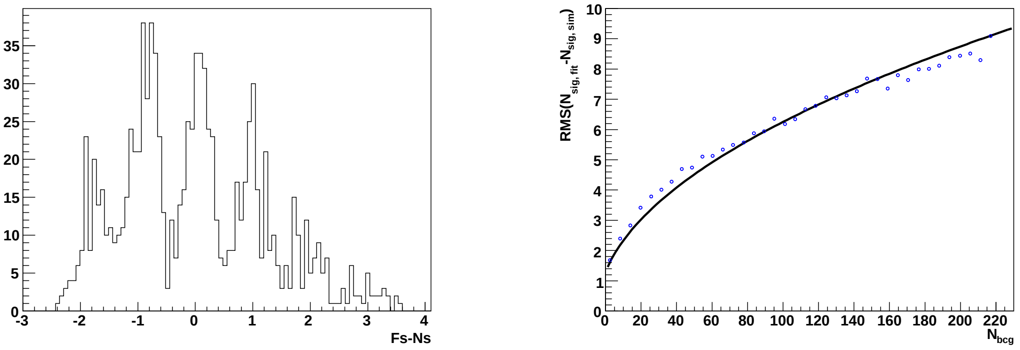
Table 4.2: The number of simulated ($N_{\mathrm{sig,\ sim}}$) and reconstructed ($N_{\mathrm{sig,\ fit}}$) signal events and its error ($\sigma_{N_{\mathrm{sig}}}$) as determined by the fit. The errors shown in the table are root – mean – square (RMS) values of the corresponding distributions. The value of N and $f_{\mathrm{sig}} = 0.0804$ are chosen to be such that the situation is similar to the measured one. Ten thousand samples are simulated for each of the four cases. The average number of simulated signal and background events in the $5.27\,\mathrm{GeV} < M_{\mathrm{bc}}c^2 < 5.29\,\mathrm{GeV}$ and $|\Delta E| < 0.03\,\mathrm{GeV}$ region is $N_S = 160$ and $N_B = 107$, respectively.

There are several things to note. When the number of simulated events in the sample is fixed to $N = 2000$, the error determined by the fit ($\sigma_{N_{\mathrm{sig}},\ \mathrm{fit}}$) is consistent with the spread in the number of signal events (RMS($N_{\mathrm{sig,\ fit}}$), quoted as an error of $N_{\mathrm{sig,\ fit}}$). On the other hand, when N is distributed according to the Poisson distribution, we have to use Eq. 4.20 to obtain the proper errors and then the errors are again consistent, as shown in the Table 4.2. The spread in the estimated errors $\sigma_{N_{\mathrm{sig}},\ \mathrm{fit}}$ is small, therefore they can be well determined by the fit. Since the errors are consistent also in the case when all parameters are determined by the fit, we also confirmed that the effect of the inaccuracy in the other parameters is correctly estimated by the ML fit. By comparing the results of the simulation when only f_{sig} is determined by the fit with those obtained when all of the parameters are varied, we can estimate the effect of the errors on other parameters; it amounts to $14.5 \ominus 13.7 = 4.7$ or only additional 6%. In the later case we also observe a small systematic bias in the number of reconstructed signal events (note that the error on the mean value N_{sig} is its spread divided by a square root of the number of simulated samples, $\sqrt{10000}$). The bias is found to be negligible, $\langle N_{\mathrm{sig,\ fit}} - N_{\mathrm{sig,\ sim}} \rangle = 0.3 - 0.5$ or $0.2 - 0.3\%$. We therefore conclude that our fitting method for a data sample of this size gives an unbiased estimate of the signal yield as well as its error.

It is interesting to note that the spread in $N_{\mathrm{sig,\ fit}}$ (the statistical error) is not much larger than the spread in $N_{\mathrm{sig,\ sim}}$, $\mathrm{RMS}(N_{\mathrm{sig,\ fit}}) - \mathrm{RMS}(N_{\mathrm{sig,\ sim}}) = 6.8$ (for the fixed value of $N = 2000$, similar for Poisson distributed N). One obtains consistent result by calculating $\mathrm{RMS}(N_{\mathrm{sig,\ fit}} - N_{\mathrm{sig,\ sim}}) = 6.5$. This difference is partly due to the fluctuations in the number of background events in the signal region and partly due to the fluctuations in the distribution of the signal events. The relative size of the two contributions is estimated by repeating the above simulation for a fixed average number of signal events $N_S = 160$ and a varied average number of all events, N, in the range from $N = 164$ to $N = 4000$. For this study, the sample size is distributed according to a Poisson distribution and only $f_{\rm sig}$ is determined by the fit. The contribution due to the fluctuations in the number of signal events is estimated by the study in the region $N_B \to 0$, as shown in Fig. 4.11 (a). Clear peaks at the integer values of $N_{\text{sig, fit}} - N_{\text{sig, sim}}$ are due to the fluctuations in the number of events in the tail of the signal distribution (which are mistaken for background) and due to background events in the signal peak region. The RMS of the corresponding distribution is found to be $a_1 = 1.3$ (for a given number of N_S). The contribution due to fluctuations in background is estimated by fitting $\text{RMS}(N_{\text{sig, fit}} - N_{\text{sig, sim}})$ as a function of average background level N_B,

$$\text{RMS}(N_{\text{sig, fit}} - N_{\text{sig, sim}}) = a_1 \oplus \sqrt{a_2 N_B}\,, \text{ with } a_2 = 0.37\,, \tag{4.24}$$

as shown in Fig. 4.11 (b). The value of $a_2 < 1$ compensates for the fact that the signal region is wider than the peak and therefore a smaller number of background events, about 40%, actually causes the fluctuations in the signal yield. This corresponds to a region of about $\pm 2.3\sigma$ in the signal $M_{\rm bc}$ and ΔE distribution. A possible cause for the observed systematic deviations from Eq. 4.24 is that the fluctuations in the narrower signal region are more important than those in the tails, which is being neglected in this simple estimation.



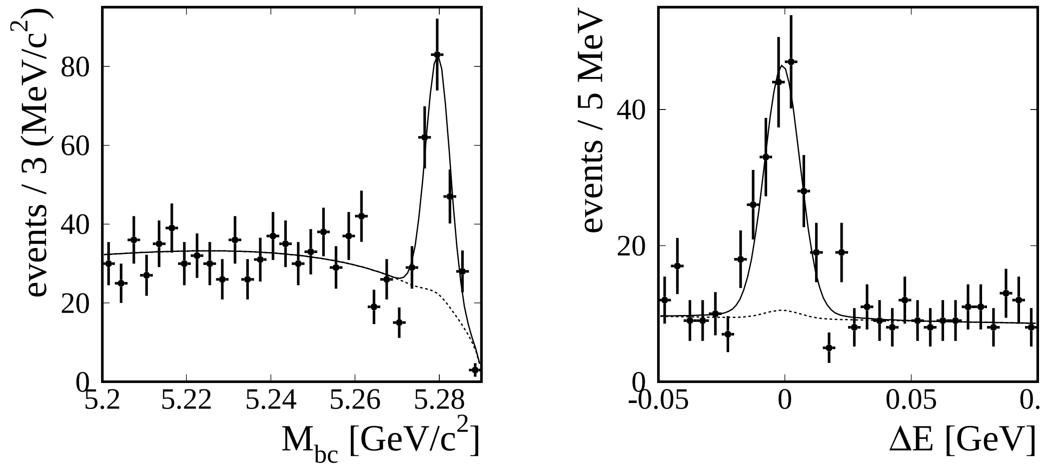
(a) $N = \text{Poisson}(164)$

(b)

Figure 4.11: The distribution of $N_{\text{sig, fit}} - N_{\text{sig, sim}}$ when $N_B \to 0$ (a) and $\text{RMS}(N_{\text{sig, fit}} - N_{\text{sig, sim}})$ as a function of background level N_B (b).

4.2.2 Signal Yield in the Data Sample

Figure 4.12 shows the $M_{\rm bc}$ and ΔE distributions for our data sample. The 2D ML fit described in the previous section yields 150 ± 15 events in the signal peak, where the error is statistical only. The yield for each of the two reconstruction channels separately is shown in Table 4.3. The values of other parameters describing the $M_{\rm bc}$ and ΔE distribution are as follows.



(a) $M_{\rm bc}$, $|\Delta E| < 0.03\,{\rm GeV}$ (b) ΔE, $M_{\rm bc} > 5.27\,{\rm GeV}/c^2$

Figure 4.12: Distributions for the reconstructed events in $M_{\rm bc}c^2$ (a) and ΔE (b). The full (dashed) curves show the projections of the result of the 2D unbinned maximum likelihood fit for all (background) events.

channel	$N_{\rm peak}$	$N_{\rm bcg}$
$K\pi\pi$	124.1±13.6	110.8±2.6
$K_S\pi$	25.7± 5.7	13.8±0.9

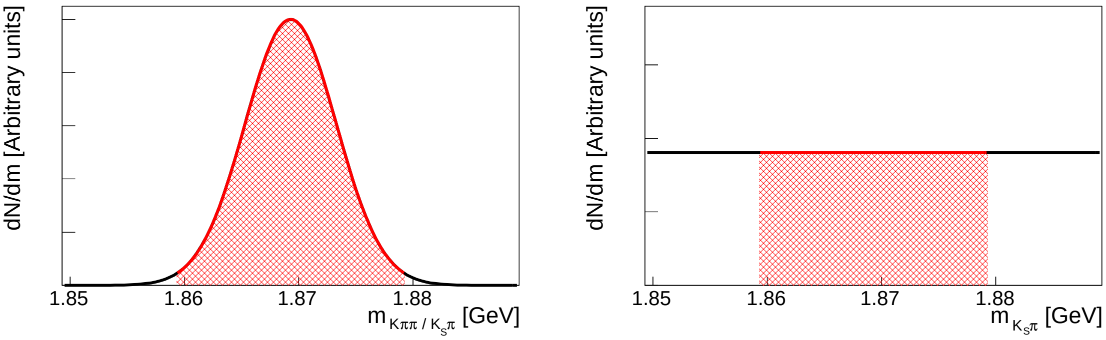
Table 4.3: The number of events in the signal peak $N_{\rm peak}$ and the expected amount of the combinatorial background $N_{\rm bcg}$ in the $5.27\,{\rm GeV}/c^2 < M_{\rm bc} < 5.29\,{\rm GeV}/c^2$ and $|\Delta E| < 0.03\,{\rm GeV}$ region, as extrapolated from the fit.

$$\begin{aligned}
\mu^{M_{\rm bc}} &= (5.2797 \pm 0.0003)\,{\rm GeV} && (4.25)\\
\sigma^{M_{\rm bc}} &= (2.86 \pm 0.22)\,{\rm MeV} \\
\beta &= -20.4 \pm 2.5 \qquad \text{...parameter of the Argus f.} \\
\mu^{\Delta E} &= (-0.97 \pm 0.78)\,{\rm MeV} \\
\sigma_1^{\Delta E} &= (6.98 \pm 0.67)\,{\rm MeV} \\
f^{\Delta E} &= 0.93 \qquad \text{...fixed from MC} \\
\sigma_2^{\Delta E} &= 21\,{\rm MeV} \qquad \text{...fixed from MC} \\
\alpha &= -0.74 \pm 0.48 \qquad \text{...slope of the bcg. } \Delta E \text{ distribution}
\end{aligned}$$

The number of the reconstructed events in the $K\pi\pi$ reconstruction channel for the SVD1 data sample only, 26.2 ± 8.5, is consistent with the previous Belle branching fraction measurement for these decays which reported 28±7 signal events [21]. The small difference can be explained by a few differences between the two analysis in the continuum background rejection and the vertex quality requirement.

4.2.3 Background Study

For a better understanding of the background, we investigate the MC information about the decay chain for all reconstructed MC events in the M_{bc} and ΔE signal region. In this study, we have found a non-negligible number of $B^0 \to D^-\pi^+\overline{K}^0$ decays when (at least) one of the D mesons is reconstructed in its decay to $K_S\pi$. In the case of this recently observed non-resonant decay [30], the invariant mass of the $\pi^+\overline{K}^0$ combination $m_{\pi^+\overline{K}^0}$ is spread over a large kinematic region, as illustrated in Fig. 4.13. In a small fraction of these decays $m_{\pi^+\overline{K}^0}$ lies in the D mass signal region. As a consequence, those events are indistinguishable from the $B^0 \to D^+D^-$ events. It was confirmed that one of the reconstructed D mesons corresponds to the actual D meson while the other one (reconstructed as $K_S\pi$) corresponds to the combination of $\pi^+\overline{K}^0$. Contrary to the combinatorial background there is no random particle mis-association in this case. The reconstructed values of M_{bc} and ΔE for these events are therefore consistent with the B meson mass and zero, respectively. These decays are therefore a source of a peaking background in the M_{bc} and ΔE distributions and potentially fake signal in our fit to the M_{bc} and ΔE distributions.



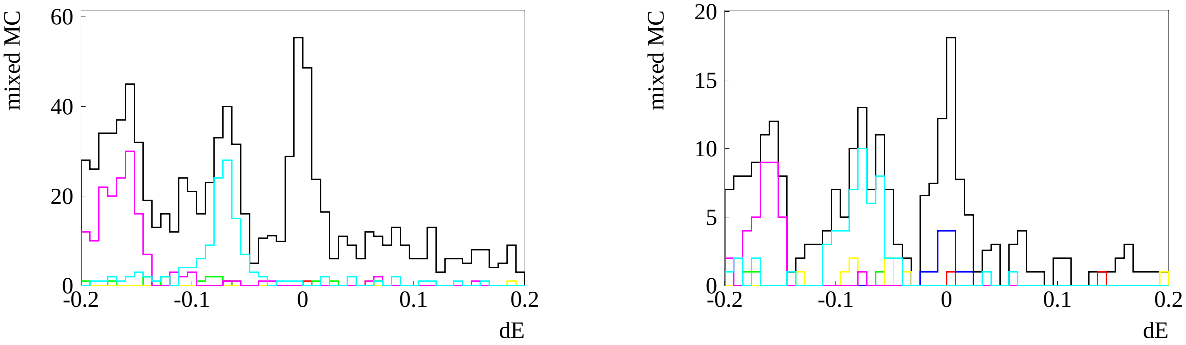
(a) (b)

Figure 4.13: Illustration of the reconstructed D mass distribution for non-resonant $B^0 \to D^-\pi^+\overline{K}^0$ decays. The mass of the $K^-\pi^+\pi^+/K_S\pi^+$ combination for tracks coming from the D^+ meson decay are consistent with the value of the D^+ mass, while the invariant mass of the $K_S\pi^+$ combination from the accompanying $\pi^+\overline{K}^0$ tracks is distributed over a large region. Shaded region indicates the reconstructed D mass signal region.

Following a similar argument, the decay of $B^0 \to D^-\pi^+\overline{K}^{*0}(892)$ with $\overline{K}^{*0}(892) \to K^-\pi^+$ can be faking a signal in the $K\pi\pi$ reconstruction channel. While the amount of these decays is found to be negligible in the case of the MC, it is important to verify this directly form the data sample.

At the beginning of this work, $D^+ \to K^+K^-\pi^+$ and $D^+ \to K_SK^+$ decay modes were also considered for the reconstruction of $D^\pm$ mesons. The same background study in these modes showed large non-resonant background component coming from the $B^0 \to D^-K^+\overline{K}^{*0}(892)$ and $B^0 \to D^-K^+\overline{K}^0$ decays, respectively. As the above background level for these modes is similar to the signal level and the corresponding $D^\pm$ branching fractions are still small compared to the $D^+ \to K^-\pi^+\pi^+$ branching fractions, we decided not to use these reconstruction channels.

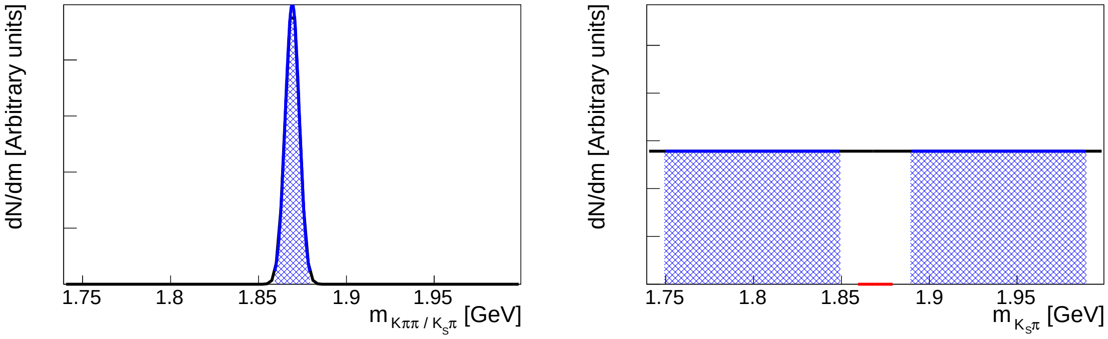
Figure 4.14 shows the ΔE distribution for the mixed ($B^0\overline{B}^0$) MC events. The decays that could cause higher level of background are shown in different colors. Those include $B^0 \to D^{*+}D^-$ decays, with $D^{*+} \to D^+\pi^0$ or $D^{*+} \to D^+\gamma$. The $B^0 \to D^{*+}D^-$ decays with a missing π^0 correspond to the peak in Fig. 4.14 at around -160 MeV and those with a missing γ make another peak outside the region shown in the figure (at $\Delta E < 0.2$ GeV). The last decay being investigated is the $B^0 \to D_s^+D^-$ decay. In the case of $D_s^+ \to K^+K^-\pi^+$ decay, we can mistake a K^+ for a π^+ and reconstruct the corresponding D_s^+ as D^+. As a consequence, there is a bump of the $B^0 \to D_s^+D^-$ events in the region around -70 MeV.



(a) $K\pi\pi$ (b) $K_S\pi$

Figure 4.14: The ΔE distribution for mixed MC events ($B^0 \to \dots$) from the $M_{\mathrm{bc}} > 5.27$ GeV signal region. Different colors indicate different decays, based on the MC simulated information: $B^0 \to D^{*+}D^-$ (purple), $B^0 \to D_s^+D^-$ (light blue), $B^0 \to D^-K^+\overline{K}^{*0}(892)$ (green), $B^0 \to D^-K^+\overline{K}^0$ (yellow), $B^0 \to D^-\pi^+\overline{K}^{*0}(892)$ (red) and $B^0 \to D^-\pi^+\overline{K}^0$ (darker blue).

The amount of the non-resonant background was estimated from the D^+ mass sidebands. The reconstructed invariant mass of one of the two $K^\mp\pi^\pm\pi^\pm/K_S\pi^\pm$ combinations was required to be in the $D^\pm$ mass signal region while the other one was required to be in the $D^\mp$ mass sideband region, as illustrated in Fig. 4.15. The sideband region $20\,\mathrm{MeV} < |m_{K^\mp\pi^\pm\pi^\pm/K_S\pi^\pm} - m_{D^\pm}| < 120\,\mathrm{MeV}$ corresponds to 10-times the width of the signal region. The region close to the signal region was not used since it still contains a fraction of signal events. Sideband was divided into sub-intervals of 20 MeV and the center of each sub-interval was used as a nominal mass of the corresponding D meson to simulate the same treatment of the $K^{\mp}\pi^{\pm}\pi^{\pm}/K_S\pi^{\pm}$ mass constraint and best candidate selection as was used for the $D^{\pm}$ meson candidates in the signal region.



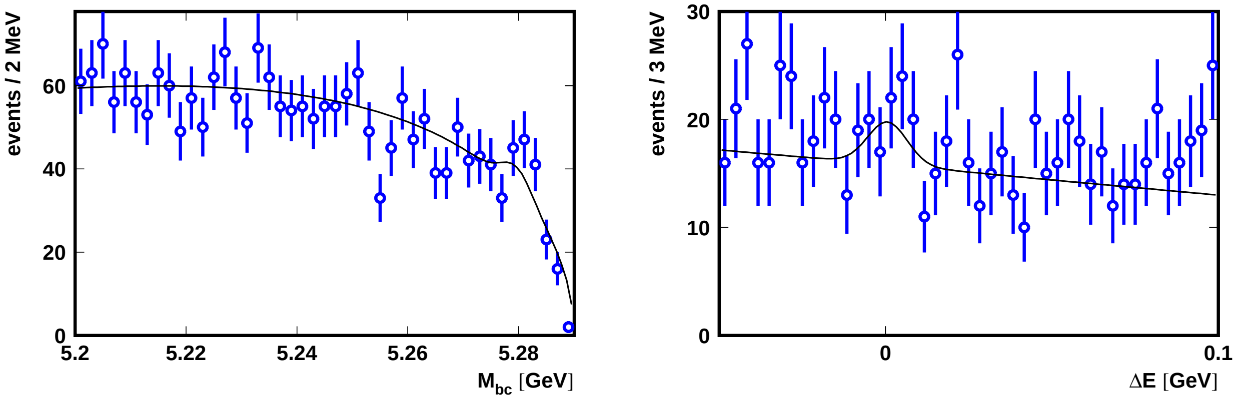
(a) (b)

Figure 4.15: The reconstructed $m_{K\pi\pi/K_S\pi}$ distribution for non-resonant $B^0 \rightarrow D^-\pi^+\overline{K}^0$ decays. The mass of the $K^{\mp}\pi^{\pm}\pi^{\pm}/K_S\pi^{\pm}$ combination for tracks coming from the D meson decay are consistent with the value of the D mass, while the invariant mass of the $K_S\pi^{\pm}$ combination from the accompanying $m_{\pi^+\overline{K}^0}$ is distributed over a large region. Shaded regions indicate the $m_{K\pi\pi/K_S\pi}$ region that is used for the sideband study.

The D mass sideband study was done separately for the $K\pi\pi$ and $K_S\pi$ reconstruction channels. The $M_{\rm bc}$ and ΔE distributions of the reconstructed events are shown in Figs. 4.16 and 4.17, respectively. No prominent peak is observed and the accuracy of the number of events in the peak is limited by a fluctuation in the number of background events. In the case of the $K\pi\pi$ channel, the unbinned ML fit to the distribution gives 20 ± 18 events in the peak. The number of events is consistent with a statistical fluctuation. The rescaled contribution of peaking background events is estimated to be about $N_{\rm nr}(K\pi\pi) = 2 \pm 1.8$ and is therefore found to be negligible. In the case of the $K_S\pi$ reconstruction channel, the rescaled number of peaking background events is estimated to be $N_{\rm nr}(K_S\pi) = 1.4 \pm 0.8$.

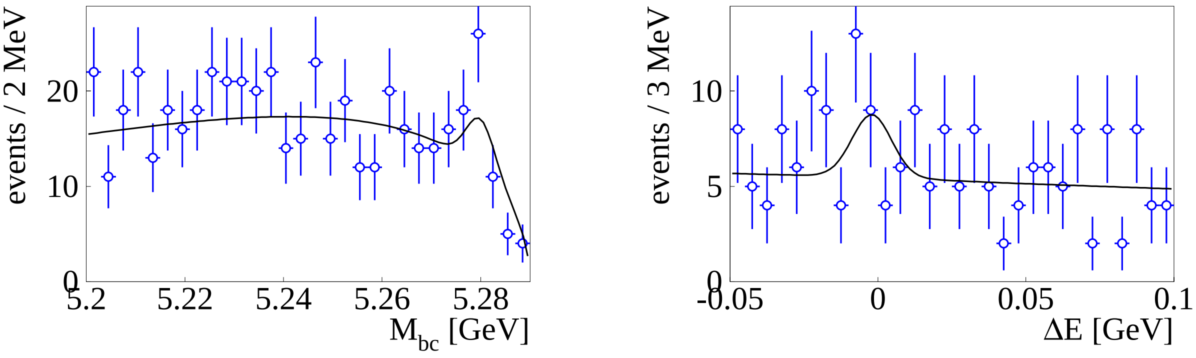
4.2.4 Signal Reconstruction Efficiency

Signal reconstruction efficiency ϵ is a fraction of reconstructed signal events. It only describes what fraction of signal events decaying in the $K\pi\pi$ or $K_S\pi$ decay mode is reconstructed by the reconstruction analysis. It is determined by a signal MC study. Only events that pass all of the event selection criteria and have the values of $M_{\rm bc}$ and ΔE in the signal region are counted as being reconstructed. The resulting error due to about 5-10% wider signal peaks in the $M_{\rm bc}$ and ΔE distributions in the data is estimated by scaling the signal region by the factor $\sigma_{\rm MC}/\sigma_{\rm data}$ so that its size in the units of peak width is the same. The change in efficiency on this account is negligible, 0.1%. Figure 4.18 shows the reconstruction efficiency for different simulated running periods (experiments). Error is MC statistical error only. Reconstruction efficiencies for the SVD1 (up to and including



(a) $M_{\rm bc}$, $|\Delta E| < 0.03\,{\rm GeV}$ (b) ΔE, $M_{\rm bc} > 5.27\,{\rm GeV}/c^2$

Figure 4.16: The $M_{\rm bc}$ (a) and ΔE (b) distributions for the case when both of the $D^{\pm}$ mesons are reconstructed in the $K^{\mp}\pi^{\pm}\pi^{\pm}$ reconstruction channel. The $K\pi\pi$ combination invariant mass of one D meson is required to be in the D mass signal and the other one in the D mass sideband region. The full curve is a projection of the 2D unbinned ML fit result.



(a) $M_{\rm bc}$, $|\Delta E| < 0.03\,{\rm GeV}$ (b) ΔE, $M_{\rm bc} > 5.27\,{\rm GeV}/c^2$

Figure 4.17: The $M_{\rm bc}$ (a) and ΔE (b) distribution for the case when one of the D mesons is reconstructed in the $K_S\pi$ reconstruction channel with $K_S\pi$ invariant mass in the D mass sideband region. The full curve is a projection of the 2D unbinned ML fit result.

experiment 27) and SVD2 (starting with experiment 31) running periods are significantly different, therefore they are treated separately. Efficiencies for the SVD1 and SVD2 samples are determined from the fit to the values for the corresponding experiments. Average efficiency is calculated as

$$\epsilon \;=\; \frac{\mathcal{L}_{\rm SVD1}\;\epsilon_{\rm SVD1} + \mathcal{L}_{\rm SVD2}\;\;\epsilon_{\rm SVD2}}{\mathcal{L}_{\rm SVD1} + \mathcal{L}_{\rm SVD2}}\,. \tag{4.26}$$

The reconstruction efficiencies for the two decay modes are shown in table 4.4.

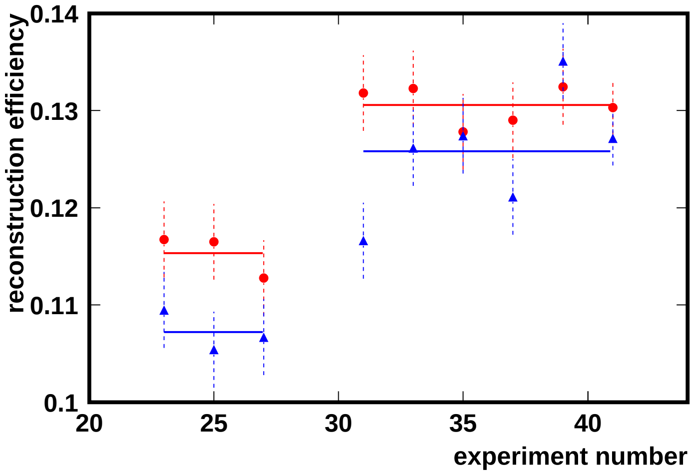


Figure 4.18: Event reconstruction efficiency for a representative number of experiments, determined by MC study. Results for the $K\pi\pi$ (circles) and $K_S\pi$ (triangles) reconstruction channels are shown separately. Lines show fit results.

channel	ϵ_{SVD1} [%]	ϵ_{SVD2} [%]	ϵ [%]
$K\pi\pi$	11.5 ± 0.2	13.1 ± 0.2	12.6 ± 0.2
$K_S\pi$	10.7 ± 0.5	12.6 ± 0.3	12.1 ± 0.4

Table 4.4: Signal reconstruction efficiency for the $K\pi\pi$ and $K_S\pi$ reconstruction channel.

4.2.5 Result for the $\mathrm{B}^0 \to \mathrm{D}^+\mathrm{D}^-$ Branching Fraction

Branching fraction $\mathcal{B}$ for this decay is obtained by summing the signal yield from the two reconstruction channels.

$$\begin{aligned} N_{\mathrm{sig}} &= \Sigma_i \mathcal{B}\, \epsilon_i\, \mathcal{B}(D_i)\, 2\, \frac{1}{2}\, N_{B\overline{B}}\,; \\ \mathcal{B} &= \frac{N_{\mathrm{sig}}}{\epsilon\, \mathcal{B}(D)\, N_{B\overline{B}}}\,, \end{aligned} \tag{4.27}$$

where i stands for a reconstruction channel $K\pi\pi$ or $K_S\pi$ and

$$\epsilon\, \mathcal{B}(D) = \Sigma_i \epsilon_i\, \mathcal{B}(D_i) \tag{4.28}$$

is the average reconstruction efficiency. The world average values for D meson branching fractions are used to calculate $\mathcal{B}(D_i)$ [11],

$$\begin{aligned}
\mathcal{B}(D \to K\pi\pi) &= (9.51 \pm 0.34)\,\% && (4.29)\\
\mathcal{B}(D \to K_S\pi) &= (1.47 \pm 0.06)\,\% && (4.30)\\
\mathcal{B}(K_S \to \pi\pi) &= (69.20 \pm 0.05)\,\% && (4.31)\\
\mathcal{B}(D \to K_S\pi;\, K_S \to \pi\pi) &= \mathcal{B}(D \to K_S\pi)\,\mathcal{B}(K_S \to \pi\pi) = && (4.32)\\
&= (1.017 \pm 0.0423)\,\% && (4.33)\\
\mathcal{B}(D_{K\pi\pi}) &= \mathcal{B}(D \to K\pi\pi)^2 = (0.9044 \pm 0.0647)\,\% && (4.34)\\
\mathcal{B}(D_{K_S\pi}) &= 2\mathcal{B}(D \to K\pi\pi)\mathcal{B}(D \to K_S\pi) + \mathcal{B}(D \to K_S\pi)^2 = && (4.35)\\
&= (0.2038 \pm 0.0150)\,\% && (4.36)
\end{aligned}$$

With the reconstruction efficiencies shown in Table 4.4 we obtain from Eq. 4.28

$$\epsilon\,\mathcal{B}(D) = (0.1386 \pm 0.0664)\,\%\,. \tag{4.37}$$

The number of the reconstructed signal events is the sum of the corresponding signal yields from the Table 4.3, decreased by the number of events in the signal region coming from the non-resonant decays.

$$N_{\mathrm{sig}} = 124.1 + 25.7 - 2.0 - 1.4 = 146.4 \tag{4.38}$$

The resulting $B^0 \to D^+D^-$ branching fraction is then

$$\mathcal{B}(B^0 \to D^+D^-) = (1.97 \pm 0.20 \pm 0.20) \times 10^{-4}\,, \tag{4.39}$$

where the first error is statistical and the second systematic. The result is consistent with previous measurements [21, 28] and has better accuracy. As a cross-check, we calculate the branching fractions separately for the $K\pi\pi$ and $K_S\pi$ modes to see if they yield a consistent result. The results for the branching fractions are indeed consistent, $(2.00 \pm 0.22) \times 10^{-4}$ and $(1.84 \pm 0.44) \times 10^{-4}$ for the $K\pi\pi$ and $K_S\pi$ modes, respectively. The error shown is statistical error only.

The systematic error, shown in Table 4.2.5, is obtained from a quadratic sum of the following contributions. The uncertainty in the D meson branching fractions results in a 5% systematic error. The error in the reconstruction efficiency, based on a MC study, amounts to 2%. The difference in the pion and kaon track reconstruction efficiency between data and MC was estimated using partially reconstructed D^* decays. The errors are added linearly for all six pion and kaon tracks, which yields a 6% uncertainty. The difference in PID efficiency for the simulated and real data is approximately 1% per track, which gives a 6% uncertainty. Smaller contributions come from the uncertainty in the K_S selection efficiency (1%), the number of $B\overline{B}$ events (1.3%) and the number of non-resonant decays (1.5%). The total systematic error of 10% is obtained from the quadratic sum of these uncertainties.

Source	Systematic error [%]
D branching fractions	5.0
event reconstruction efficiency	2.0
track reconstruction efficiency	6.0
K/π PID efficiency	6.0
K_S reconstruction efficiency	1.0
number of $B\overline{B}$ events	1.3
$B^0 \to D^-\overline{K}^0\pi^+$ peaking bcg.	1.5
Total	10.3

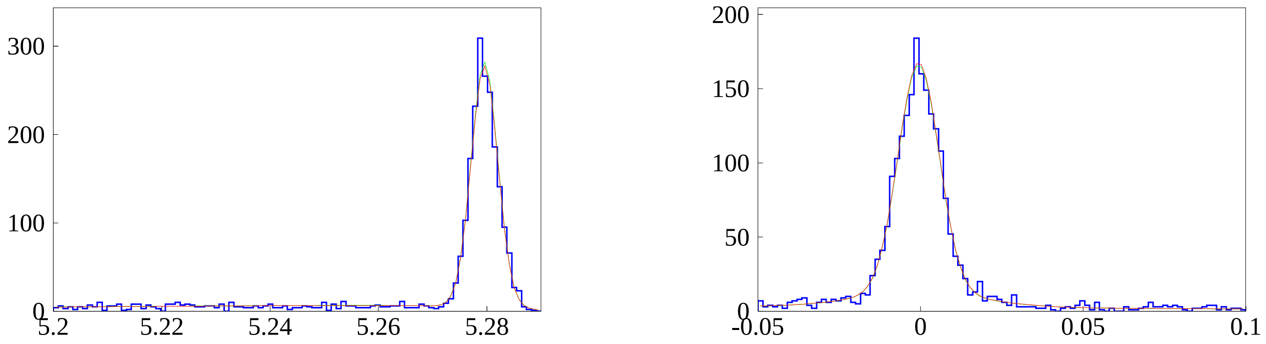
Table 4.5: Contributions to the systematic error on the branching fraction measurement.

4.3 Reconstruction of $\mathbf{B^0 \to D_s^+ D^-}$ Events

The sample of reconstructed $B^0 \to D_s^+ D^-$ events is used to study the resolution of the Δt measurement in data, and as a control sample for the study of $B^0 \to D^+D^-$ decays. Here we give a short description of the reconstruction of these decays.

The $D^\pm_{(s)}$ mesons are reconstructed in their charged particles decay modes, similar to the $B^0 \to D^+D^-$ event reconstruction. For this reason we can expect a similar vertex resolution in both decays. The D^+ mesons are reconstructed from a combination of $K^-\pi^+\pi^+$ or $K_S\pi^+$ tracks, as in the case of $B^0 \to D^+D^-$ reconstruction. The D_s^+ mesons are reconstructed in their $K^+K^-\pi^+$ and K_SK^+ decay modes. The decay $D_s^+ \to K^+K^-\pi^+$ predominantly proceeds via an intermediate resonance, $D_s^+ \to \phi\pi^+$ with $\phi \to K^+K^-$, or $D_s^+ \to \overline{K}^{*0}(892)K^+$ with $\overline{K}^{*0}(892) \to K^-\pi^+$. We therefore require the invariant mass of the $K^+K^-\pi^+$ combination to be consistent with either ϕ or $\overline{K}^{*0}(892)$ intermediate state. Event selection criteria follows previous Belle measurement of the branching fraction for this decay mode [31].

The $M_{\rm bc}$ and ΔE distributions of the reconstructed $B^0 \to D_s^+D^-$ events are parameterized in the same way as was described earlier for the $B^0 \to D^+D^-$ decay. The unbinned 2D ML fit to the $M_{\rm bc}$ and ΔE distribution shown in Fig. 4.19 yields 1896 ± 42 events in the signal peak with signal fraction 0.94 for a data sample of 449×10^6 $B\overline{B}$ events. Again the non-resonant $B^0 \to D^-K^+\overline{K}^{*0}$ and $B^0 \to D^-K^+\overline{K}^0$ decays are a potential source of background peaking in the $M_{\rm bc}$ and ΔE signal region. Similar as before, the amount of non-resonant background is estimated from the D_s^+ sidebands and is found to be 22 ± 2 for the $D_s^+ \to K^+K^-\pi^+$ and negligible for the $D_s^+ \to K_SK^+$ decay mode.



(a) Number of events / 0.9 MeV bin.

(b) Number of events / 1.5 MeV bin.

Figure 4.19: The $M_{\rm bc}$ (a) and ΔE (b) distributions for the data sample of $B^0 \to D_s^+ D^-$ events. The red line is the projection of the result of the 2D unbinned ML fit.

Chapter 5

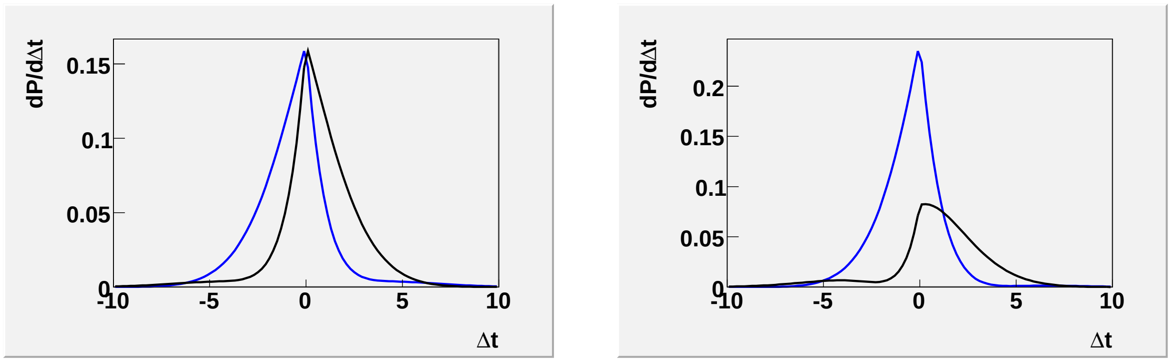
Analysis of the Time Distribution

Once we have a sample of the reconstructed $B^0 \to D^+D^-$ decays, we can continue with the main part of this work, the analysis of their decay time distribution,

$$\mathcal{P}_{\text{sig}}(\Delta t, q) = \frac{e^{-|\Delta t|/\tau}}{4\tau}\left(1 + q\left(\mathcal{S}\sin(\Delta m \Delta t) + \mathcal{A}\cos(\Delta m \Delta t)\right)\right), \qquad (5.1)$$

where $\Delta t = t_{CP} - t_{\text{tag}}$ is the difference between the decay time of the B_{CP} meson (t_{CP}) and the decay time of the B_{tag} meson (t_{tag}). $\mathcal{S}$ and $\mathcal{A}$ are the CPV parameters, $\tau = (1.530 \pm 0.009)$ ps is the B^0 meson lifetime, and $\Delta m = (0.507 \pm 0.005)\,\hbar/\text{ps}$ is the mass difference of the two B mass eigenstates [11]. The flavor q of the B meson decaying to the CP eigenstate is $q = +1(-1)$ for $B_{\text{tag}} = B^0(\overline{B}^0)$.

The decay time probability distributions for different values of $\mathcal{S}$ and $\mathcal{A}$ are shown in Fig. 5.1. In the case of no direct CPV, $\mathcal{A} = 0$, the two distributions for $q = \pm 1$ are symmetric with respect to $\Delta t = 0$, $\mathcal{P}(\Delta t, q) = \mathcal{P}(-\Delta t, -q)$, as shown in Fig. 5.1 (a). In this case, if we did not measure the Δt distribution, we could not make any inference about the value of $\mathcal{S}$. On the other hand, in the case of direct CPV, $\mathcal{A} \neq 0$, the distributions are



(a) $\mathcal{S} = -0.725$ and $\mathcal{A} = 0$ (b) $\mathcal{S} = -0.725$ and $\mathcal{A} = 0.5$

Figure 5.1: The Δt distributions for events tagged as $B_{\text{tag}} = \overline{B}^0$ (in blue) and $B_{\text{tag}} = B^0$ (in black) for different choices of the parameters $\mathcal{S}$ and $\mathcal{A}$.

not only shifted but also have different areas (integrals of Δt), as indicated in Fig. 5.1 (b). Also the time-integrated decay probabilities for events tagged as B^0 or $\overline{B}^0$ are therefore different.

Since the B mesons are almost at rest in the CM system, we use only the vertex z coordinate (in the laboratory system) to determine the value of Δt, $\Delta t = (z_{CP} - z_{\text{tag}})/(\beta\gamma c)$, where $\beta\gamma = 0.425$ is the boost factor. Taking into account the finite detector resolution, the Δt distribution for signal events shown in Fig. 5.1 is smeared by the measurement error of about $100\,\mu\text{m}\,/(\beta\gamma c) = 0.78\,\text{ps}$, as illustrated in Fig. 5.2.

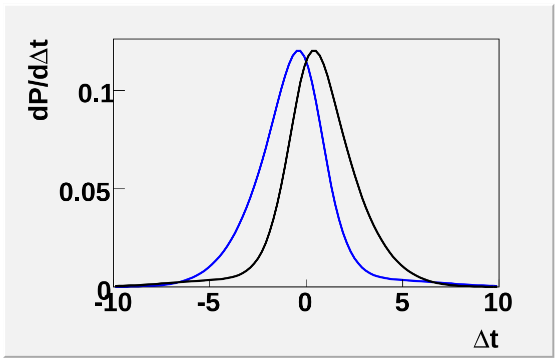


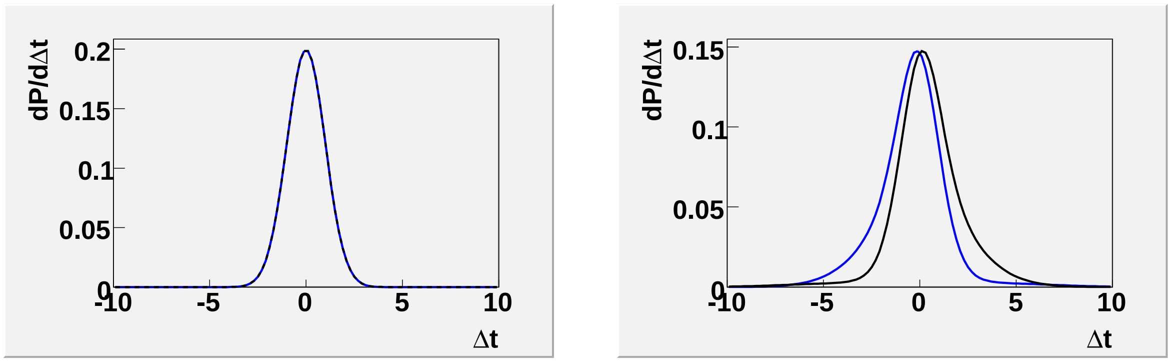
Figure 5.2: The Δt distribution for signal events, smeared by the detector resolution for events tagged as $B_{\text{tag}} = \overline{B}^0$ (in blue) and $B_{\text{tag}} = B^0$ (in black) for $\mathcal{S} = -0.725$ and $\mathcal{A} = 0$.

As we have seen in the previous chapter, our data sample of $B^0 \to D^+D^-$ events is a mixture of signal and background events. The Δt distribution therefore contains also a background component, which has to be described separately. The Δt distribution of background events is approximately a Gaussian function of a width of 1.6 ps, as shown in Fig. 5.3 (a). The Δt distribution for a mixture of signal and background events is shown in Fig. 5.3 (b).

The flavor of the B meson decaying to the CP eigenstate, $B^0 \to D^+D^-$, also needs to be measured. Equation 5.1 describes events for which we have correctly reconstructed the B meson flavor. However, the average flavor reconstruction efficiency is only about 30%. Taking this into account, the Δt distribution of signal events shown in Fig. 5.1 (a) becomes more like the one shown in Fig. 5.4 (a). If we include also the detector resolution and background component, the distributions for events tagged as B^0 or $\overline{B}^0$ are even more similar to one another, as illustrated in Fig. 5.4 (b).

As the accuracy of our measurement is strongly limited by the detector resolution and flavor tagging accuracy, it is important to use all the information we can to measure the CP parameters $\mathcal{S}$ and $\mathcal{A}$. For this purpose, we make an un-binned maximum likelihood fit to the Δt distribution, incorporating into the likelihood function all the information we have about the detector resolution, background Δt distribution and flavor tagging.

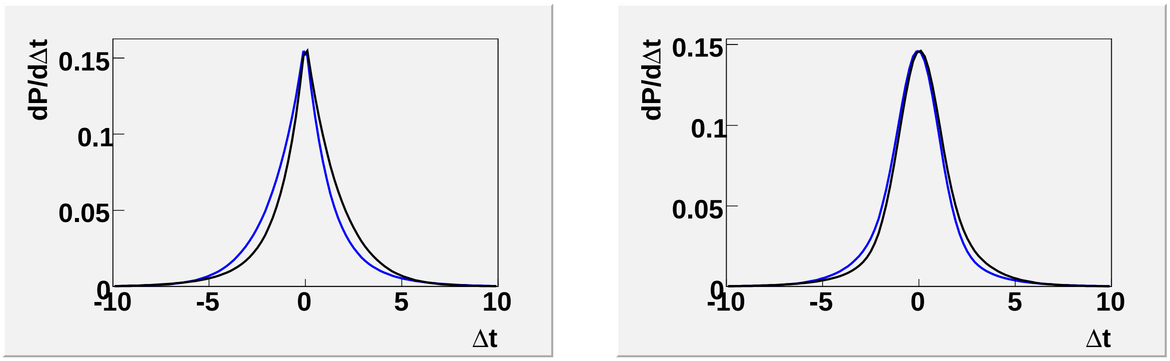
In the following sections we first describe the Δt detector resolution for signal events, measure the Δt distribution for the background events and modify the likelihood function to include the wrong flavor determination. Having thus determined all the parameters of



(a) background events

(b) signal and background

Figure 5.3: The Δt distribution for background (a) and signal and background events, $f_{\text{sig}} = 0.63$ (b). Events tagged as $B_{\text{tag}} = \overline{B}^0$ and $B_{\text{tag}} = B^0$ are shown in blue and black, respectively.



(a)

(b)

Figure 5.4: The Δt distribution for events tagged as $B_{\text{tag}} = \overline{B}^0$ (in blue) and $B_{\text{tag}} = B^0$ (in black) for $\mathcal{S} = -0.725$ and $\mathcal{A} = 0$ if the flavor tagging efficiency is 30%. Fig. (a) shows the Δt distribution for signal events (described by Eq. 5.1 modified to account for the wrong flavor determination) while Fig. (b) shows the distribution for signal and background events with the finite detector resolution.

the event likelihood function, we proceed with the fit of the CP parameters. Before doing the fit for the data sample, we check the consistency of the fit results and estimate the expected errors using the Monte Carlo simulated events. In order not to rely only on the MC, we make another check on a larger data sample of $B^0 \to D_s^+ D^-$ events for which the CP parameters are expected to be zero. At each step, we fit the B^0 lifetime first to test our description of the Δt distribution.

Finally, we perform the CP fit for the data sample of reconstructed $B^0 \to D^+ D^-$ decays. As the later reveals an unexpectedly large direct CP violation, we show a few additional checks to validate our tagging algorithm and the consistency of the fit. Since the result is outside the allowed physical region we do not rely on the errors determined by the fit but instead do a Feldman – Cousins statistical analysis to determine the significance of the result.

5.1 The B Meson Vertex and Flavor Reconstruction

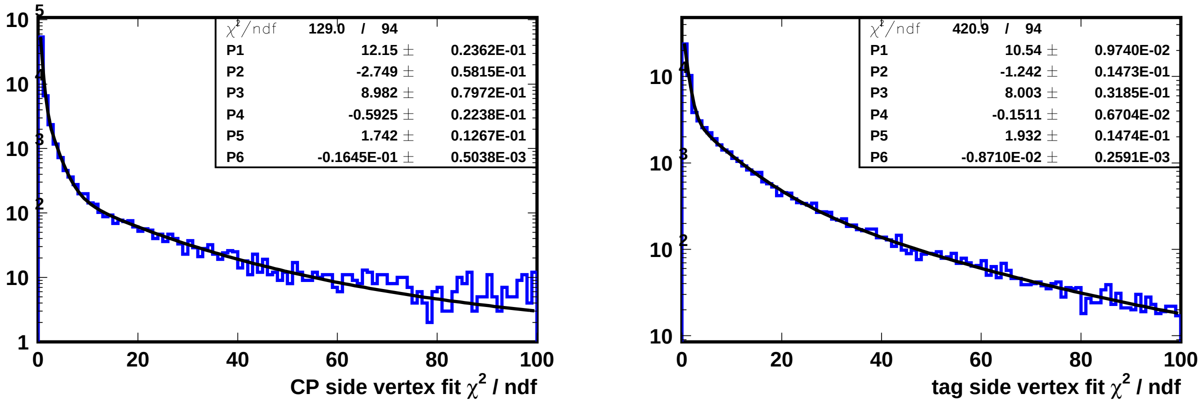
The particle decay vertices are reconstructed in the same order as particles themselves, as illustrated in the previous chapter on the event reconstruction, Fig. 4.1. The K_S vertex is reconstructed from the two charged pion tracks. The $D^{\pm}$ meson vertices are reconstructed either from three charged tracks, $K\pi\pi$ candidates, or from the K_S and π track. In this fit, the $D^{\pm}$ candidate mass is constrained to the D^{+} nominal mass in order to achieve a better M_{bc} and ΔE resolution. The B_{CP} vertex is determined from the two $D^{\pm}$ tracks and the IP information. All remaining charged tracks are associated with the other B meson (B_{tag}) and are used to determine its decay vertex and flavor. In this section we first describe the B_{CP} and B_{tag} vertex reconstruction in more detail and then continue with the description of the B meson tagging algorithm.

5.1.1 The $\mathbf{B_{CP}}$ Vertex Reconstruction

The B_{CP} vertex is determined by constraining the two D meson tracks to a common vertex. The vertex is found by minimizing the $\chi^2 = \Sigma\,\Delta x^T V^{-1} \Delta x$, where $\Delta x = x - x_0$ is the difference between the updated and original parameters describing the particle track, V is their covariance matrix, and the sum in the above expression denotes the summation over all tracks being fitted to a common vertex. The IP profile is described by a three-dimensional Gaussian function with $\sigma_x = 100\,\mu\mathrm{m}$, $\sigma_y = 5\,\mu\mathrm{m}$ and $\sigma_z = 3.3\,\mathrm{mm}$. Since the B mesons come from the IP, an additional term $r^T V_{IP}^{-1} r$, where r is the distance of the vertex from the IP and V_{IP} is the covariance matrix describing the IP region, $V_{IP}^{ij} = \sigma_i \sigma_j \rho_{ij}$ (where σ_i is an error on the i^{th} coordinate and ρ_{ij} is the correlation coefficient between the i^{th} and j^{th} coordinate), is added to the above χ^2 to achieve a better vertex resolution.

The value of χ^2/ndf, where ndf is the number of degrees of freedom, is a measure of the fit quality and is used to reject vertices with a very poor vertex reconstruction quality. For this purpose, we calculate χ^2 without the term related to the IP constraint, $\xi = \chi^2_{\mathrm{w/o\ IP}}/ndf$. Otherwise the decays with vertices close to the IP would automatically carry more weight. Event reconstruction efficiency is reduced by 1.0% due to the $\chi^2/ndf < 100$ criterion. Figure 5.5 shows the ξ distributions for the B_{CP} and B_{tag} vertex.

The most important detector for the accuracy of the vertex position measurement is the silicon vertex detector (SVD). Each particle track can be measured by any of the three (four in the case of SVD2) SVD detector layers. Particle trajectories with more SVD measurements (hits) are determined with a higher accuracy. Ideally, we could require all tracks that are used for the B meson reconstruction to have a minimal number of associated SVD hits. However, due to a limited angular coverage, detector noise and mistakes in the hit association etc., only about 88% of the charged tracks from the D^+D^- final state have at least one $r\varphi$ and two z measurements by the SVD. The usual Belle requirement for each



(a) B_{CP} vertex

(b) $B_{\rm tag}$ vertex

Figure 5.5: Vertex fit χ^2/ndf (without IP) distribution for the (a) CP and (b) tag side B^0 meson. The distribution is described by three exponential terms: $\mathrm{e}^{P1+P2\,x} + \mathrm{e}^{P3+P4\,x} + \mathrm{e}^{P5+P6\,x}$. Fit result for all six fitted parameters is shown in the inset.

charged track to have at least that many SVD hits would therefore cause a significant loss in the reconstruction efficiency. As we have many charged particles in the final state we have a possibility of loosening this criterion.

For this purpose we study the impact of the number of tracks with SVD hits to the D^+ and B^0 meson vertex resolution. The particle track is said to have SVD hits if it has at least one $r\varphi$ and two z measurements by the SVD. The difference is because in the case of SVD1, for any layer the measurement of only one coordinate can be used for tracking (either $r\varphi$ or z – not necessarily both). For SVD2, in any layer only a combination of $r\varphi$ and z coordinate measurements is used, therefore the above condition implies at least two measurements with both, $r\varphi$ and z hits.

We investigate the RMS value of the vertex residual distribution, RMS($z_{\rm rec} - z_{MC}$), where $z_{\rm rec}$ and z_{MC} are the reconstructed and simulated vertex z coordinates. Signal MC sample is used for this study. Table 5.1 compares the D^+ meson vertex resolution for the different situations. It can be noticed that in the case of the $K\pi\pi$ reconstruction channel, the D^+ vertex resolution is significantly worse when only one of the tracks has SVD hits. On the other hand, for the $K_S\pi$ channel, resolutions when only one or two tracks have SVD hits are comparable because in the first case the track with SVD hits is mostly pion coming from the D^+ meson, while in the second the two tracks are typically two pion tracks used for the K_S vertex reconstruction. We conclude that the accuracy of the D^+ vertex measurement is satisfactory when at least two of the final state tracks have enough SVD hits.

For the B^0 meson vertex resolution study, we first compare the average absolute value of the vertex z coordinate residual, $< |z_{\rm rec} - z_{\rm true}| >$, because it is less sensitive to the fluctuations in the tails of the distribution. The results in Table 5.2 show that the two cases when either two or three tracks have enough SVD hits have comparable B meson

number of tracks with SVD hits	1	2	3
RMS [μm] for $K\pi\pi$, SVD1	320	190	130
RMS [μm] for $K_S\pi$, SVD1	270	270	160
RMS [μm] for $K\pi\pi$, SVD2		170	110
RMS [μm] for $K_S\pi$, SVD2	270	280	160

Table 5.1: RMS of the D^+ meson vertex residual distribution $z_{rec} - z_{\text{true}}$ as a function of the number of pion and kaon tracks with at least one $r\varphi$ and two z SVD hits. Values for the $K\pi\pi$ and $K_S\pi$ reconstruction channels are shown separately. The condition $|z_{rec} - z_{\text{true}}| <$ 1 mm is applied to minimize the contribution of outliers. The value for the $K\pi\pi$ channel, SVD2 sample, one track, is not shown as there are not enough D^+ mesons with only one track with SVD hits.

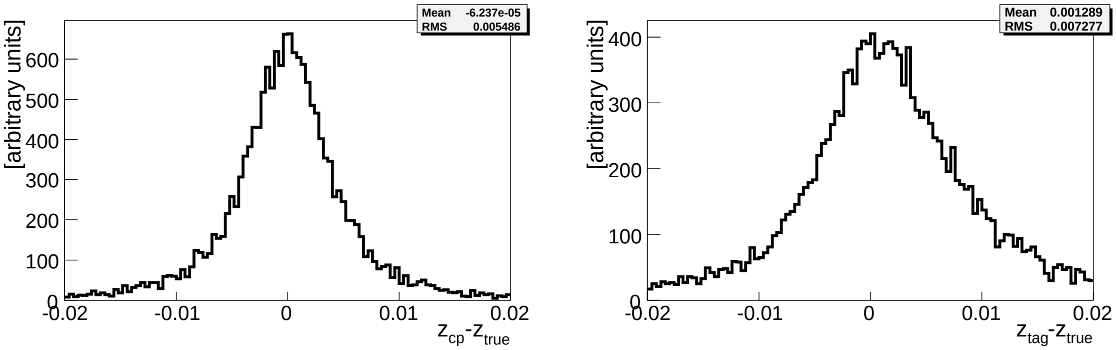
vertex resolution, therefore there is no need to require all three tracks to have SVD hits. In fact, even when only one of the D^+ mesons has at least two tracks with SVD hits, the B^0 vertex resolution is still satisfactory.

number of tracks with SVD hits	3, 3	3, 2	2, 2
$< \lvert z_{\text{rec}} - z_{\text{true}} \rvert >$ [μm] for $K\pi\pi$, SVD1	43	54	74
$< \lvert z_{\text{rec}} - z_{\text{true}} \rvert >$ [μm] for $K_S\pi$, SVD1	43	61	
$< \lvert z_{\text{rec}} - z_{\text{true}} \rvert >$ [μm] for $K\pi\pi$, SVD2	37	45	
$< \lvert z_{\text{rec}} - z_{\text{true}} \rvert >$ [μm] for $K_S\pi$, SVD2	42	43	

Table 5.2: Average error $< |z_{\text{rec}} - z_{\text{true}}| >$ of the B meson vertex residual distribution $z_{\text{rec}} - z_{\text{true}}$ as a function of the number of pion and kaon tracks with SVD hits used in the two D mesons vertex reconstruction. Values for the $K\pi\pi$ and $K_S\pi$ reconstruction channels are shown separately. Condition $|z_{rec} - z_{\text{true}}| < 1$ mm is applied.

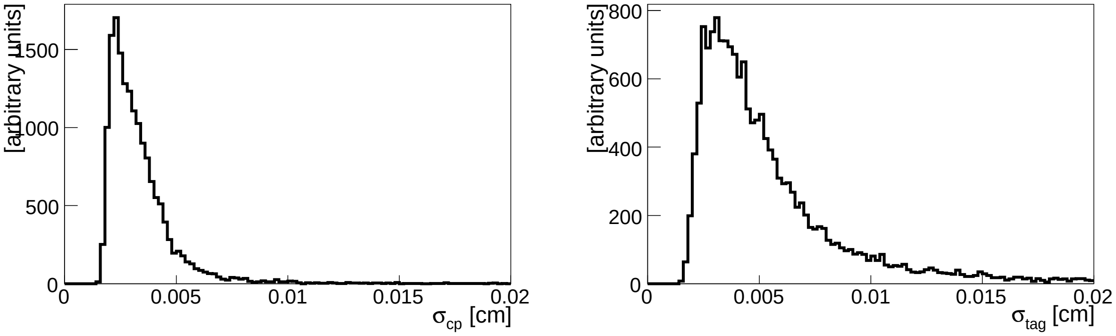
The conclusion of this study is that only one of the D mesons is required to have at least two tracks with enough SVD hits. This requirement causes only 4% loss in the signal reconstruction efficiency. In 87% of the events both D mesons have at least two tracks with enough SVD hits. This fraction is even higher for the SVD2 sample due to the increased number of the SVD layers. Overall B_{CP} vertex reconstruction efficiency is about 90%.

The B_{CP} vertex resolution is about 60 μm, as illustrated in Fig. 5.11. The vertex position error is estimated for each event separately by the vertex fit. The distribution of errors is shown in Fig. 5.7. The residual pull distribution, $(z_{\text{rec}} - z_{MC})/\sigma_z$, is shown in Fig. 5.8. If the residual distribution could be described by a single Gaussian and the error estimates were correct, the pull distribution would be a Gaussian of width $\sigma = 1$. Deviations from this case are described by the vertex resolution function that is described in more detail in the next section.



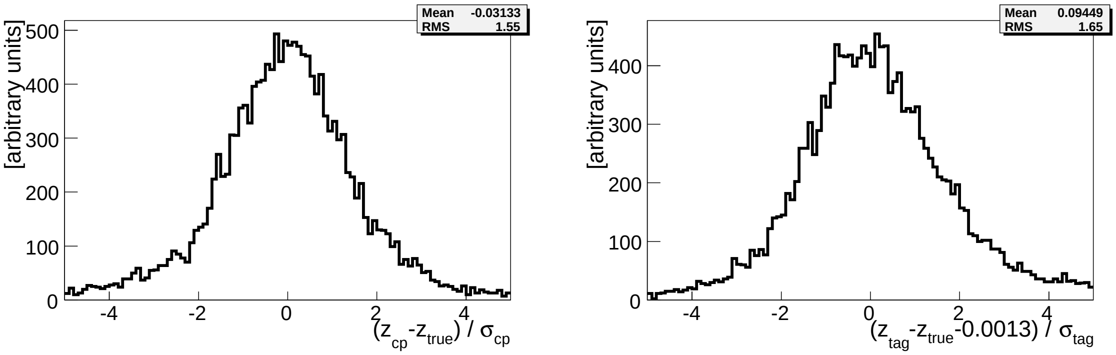
(a) B_{CP} (b) B_{tag}

Figure 5.6: Residual distribution $z_{CP} - z_{\mathrm{true}}$ for correctly reconstructed signal MC events (a). The corresponding distribution for the B_{tag} vertex is shown for a comparison (b). It is shifted and slightly asymmetric due to the effect of the non-primary tracks that will be described later.



(a) B_{CP} (b) B_{tag}

Figure 5.7: The distribution of the B_{CP} (a) and B_{tag} (b) vertex errors σ_z.



(a) B_{CP} (b) B_{tag}

Figure 5.8: The B_{CP} (a) and B_{tag} (b) vertex residual pull distribution, $(z_{\mathrm{rec}} - z_{MC})/\sigma_z$.

5.1.2 The $\mathrm{B}_{\mathrm{tag}}$ Vertex Reconstruction

The B_{tag} decay vertex is measured by fitting all remaining charged tracks (which satisfy a very loose track impact parameter requirement) to a common vertex. The tag side vertex efficiency is 90%. In most of the remaining 10% cases, the vertex is not reconstructed because there are no remaining charged tracks in the event.

Similar as for the B_{CP} vertex, the χ^2/ndf of the tag side vertex fit is also used to reject the events with very poor quality of the vertex reconstruction. The distribution of the B_{tag} vertex ξ is shown in Fig. 5.5 (b). Event reconstruction efficiency is reduced for an additional 1.7% due to the $\chi^2/ndf < 100$ criterion applied to the B_{tag} vertex.

5.1.3 The Flavor Reconstruction

The B meson flavor is determined from the flavor specific decays of the associated B_{tag} meson, such as $B^0 \to K^+X$, where X stands for any accompanying decay products. In this example, a presence of a high-momentum positively charged kaon (containing $\bar{s}$ quark) indicates that the tag-side B meson, B_{tag}, is the B^0 meson (containing $\bar{b}$ quark, which decayed as $\bar{b} \to \bar{c} \to \bar{s}$). Depending on the presence and quality of the flavor-specific signatures, the accuracy of the flavor tagging varies from event to event, depending on the B_{tag} decay process. We divide events into six r-bins based on their tagging quality r. Values of r range from 0 for events with no flavor information to 1 for events with an unambiguous flavor tag, as shown in Table 5.3.

r-bin	1	2	3	4	5	6
range in r	$0-0.25$	$0.25-0.5$	$0.5-0.625$	$0.625-0.75$	$0.75-0.875$	$0.875-1$

Table 5.3: Range in r for the different r-bins.

The Flavor Tagging Quality

The parameter r describing the flavor tagging quality is based on the MC information. The wrong tag fraction w, probability that the measured flavor is wrong, is approximately $(1-r)/2$. For the purpose of the CP analysis, the average w is determined for each of the six r-bins. If the detector response is different for particles and anti-particles, the wrong tag fractions for B^0 and $\overline{B}^0$ mesons can differ as well. We therefore introduce a true-flavor dependent wrong tag fraction,

$$\begin{aligned} w_{B^0} &= P(B_{\text{tag}}^{\text{meas}} = \overline{B}^0 \,|\, B_{\text{tag}}^{\text{true}} = B^0) \\ w_{\overline{B}^0} &= P(B_{\text{tag}}^{\text{meas}} = B^0 \,|\, B_{\text{tag}}^{\text{true}} = \overline{B}^0)\,, \end{aligned} \tag{5.2}$$

where $P(x|y)$ is a probability for x given y. Measured Δt distribution for events tagged as B^0 ($\overline{B}^0$) is then

$$\begin{aligned} \mathcal{P}_{\rm sig}(B_{\rm tag}^{\rm meas} = B^0) &= (1 - w_{B^0})\,\mathcal{P}_{\rm sig}(B_{\rm tag}^{\rm true} = B^0) + w_{\overline{B}^0}\,\mathcal{P}_{\rm sig}(B_{\rm tag}^{\rm true} = \overline{B}^0) \\ \mathcal{P}_{\rm sig}(B_{\rm tag}^{\rm meas} = \overline{B}^0) &= w_{B^0}\,\mathcal{P}_{\rm sig}(B_{\rm tag}^{\rm true} = B^0) + (1 - w_{\overline{B}^0})\,\mathcal{P}_{\rm sig}(B_{\rm tag}^{\rm true} = \overline{B}^0)\,. \end{aligned} \tag{5.3}$$

We write “meas” for the measured and “true” for the actual values to avoid any confusion. If we denote $w = (w_{B^0} + w_{\overline{B}^0})/2$ and $\Delta w = w_{B^0} - w_{\overline{B}^0}$ and substitute the $\mathcal{P}_{\rm sig}$ dependency (Eq. 5.1) into the above equation, we obtain

$$\mathcal{P}_{\rm sig}(q, \Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau}(1 - q\Delta w + q(1 - 2w)(\mathcal{S}\sin(\Delta m\Delta t) + \mathcal{A}\cos(\Delta m\Delta t)))\,, \tag{5.4}$$

where, as introduced earlier, flavor $q = 1$ (-1) for $B_{\rm tag} = B^0$ ($\overline{B}^0$). The values of Δw and w for each of the six bins in the tagging quality parameter r are determined using the following flavor specific B meson decays: $B^0 \to D^{*-}l^+\nu$, $B^0 \to D^{*-}\pi^+$, $B^0 \to D^{*-}\rho^+$ and $B^0 \to D^-\pi^+$ [32]. In this study, the signal B meson is reconstructed in its decay to one of the flavor specific final states listed above, and therefore its flavor is uniquely determined. The flavor of the other B meson is reconstructed by the tagging algorithm and then compared to the flavor determined by the flavor specific decay of the signal B meson. In this way, the wrong tag fractions can be determined. Slightly different values are used for the SVD1 and SVD2 data sample to account for the possibly different reconstruction efficiencies. We show the values of w and Δw for the MC and data sample in Table 5.4 and the coefficient $1 - 2w$ for the different r-bins in Fig. 5.9.

	MC	data
w_1	0.4453 ± 0.0019	0.4658 ± 0.0051
w_2	0.3043 ± 0.0027	0.3257 ± 0.0072
w_3	0.2044 ± 0.0031	0.2251 ± 0.0089
w_4	0.1488 ± 0.0027	0.1610 ± 0.0086
w_5	0.0896 ± 0.0026	0.1028 ± 0.0079
w_6	0.0238 ± 0.0015	0.0203 ± 0.0050
Δw_1	0.0032 ± 0.0028	0.0005 ± 0.0063
Δw_2	-0.0219 ± 0.0042	-0.0196 ± 0.0093
Δw_3	0.0137 ± 0.0048	0.0107 ± 0.0104
Δw_4	0.0028 ± 0.0043	0.0039 ± 0.0098
Δw_5	-0.0139 ± 0.0041	-0.0110 ± 0.0098
Δw_6	0.0024 ± 0.0024	0.0029 ± 0.0059

Table 5.4: The wrong tag fraction parameters w_i and Δw_i for the $i^{\rm th}$ r-bin. The values for the MC and data sample are shown separately. The values for the SVD1 and SVD2 sample are averaged according to the data luminosity.

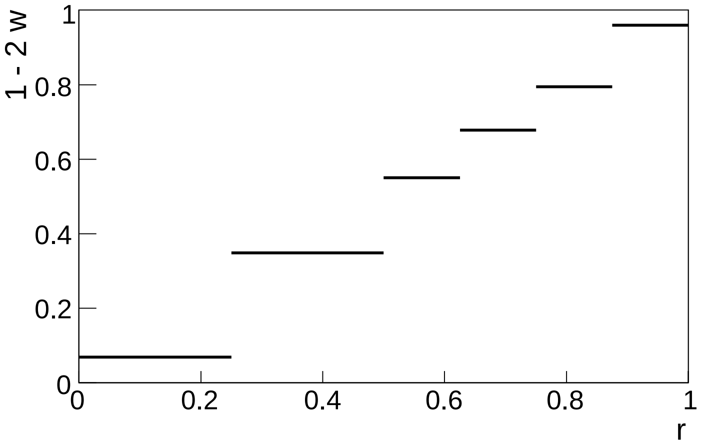


Figure 5.9: The flavor quality coefficient $1 - 2w$ as a function of r (values for the data sample). Average values are used for each of the r-bins.

Flavor Tagging Quality Dependent Signal Fraction

Since the $B\overline{B}$ and continuum events have significantly different distributions in the tagging quality variable r, it is observed that r-bins with a higher average value of r have a higher $N_{\rm sig}/N_{\rm bcg}$ ratio. An r-bin dependent signal fraction $f_{\rm sig} = N_{\rm sig}/(N_{\rm sig} + N_{\rm bcg})$ ratio is therefore used to determine event-by-event signal probability $f_{\rm sig}(M_{\rm bc}, \Delta E, r)$. Due to low statistics only the $1^{\rm st}$ r-bin and the $6^{\rm th}$ r-bin are treated separately. Parameters describing the $M_{\rm bc}$ and ΔE distributions are determined from the fit to all events, their values were already shown in Eq. 4.25. To determine the $N_{\rm bcg}$ in different r-bins, the fit is repeated only for the $f_{\rm sig}$ (with other parameters fixed). While the number of background events is reliably fitted since a large sideband is used, the same approach for the $N_{\rm sig}$is liable to the statistical fluctuations. The number of the signal events in the individual r-bin was therefore determined as the fraction of signal events in that r-bin (as determined by signal MC) times the total number of signal events (as determined from the fit to all r-bins). At this point, we only determine the fraction of events in the signal peak, therefore we do not take into account the non-resonant background component. Instead we quote $f_{\rm peak}$, the fraction of events in the Gaussian $M_{\rm bc}$ and ΔE peaks that can be either signal or non-resonant decays. The result for the r-bin dependent peak fraction, obtained as $f_{\rm peak} = N_{\rm peak}/(N_{\rm peak} + N_{\rm bcg})$, is shown in Table 5.5.

r-bin	1	2 – 5	6
$K\pi\pi$	0.47	0.58	0.74
$K_S\pi$	0.55	0.64	0.78

Table 5.5: The fraction of events in the $M_{\rm bc}$ and ΔE peak, $f_{\rm peak}$, for the $K\pi\pi$ and the $K_S\pi$ reconstruction channels.

5.2 The Δt Likelihood Function

Once we have a sample of reconstructed events, we determine the CP parameters by an un-binned ML fit to their Δt distribution, where the likelihood function is

$$\mathcal{L} = \Pi_i \mathcal{L}_i \,, \tag{5.5}$$

and $\mathcal{L}_i$ is the event likelihood function $\mathcal{L}_{\rm ev}$ for the $i^{\rm th}$ event. To obtain the event likelihood function we modify Eq. 5.1 to include the background contribution and resolution effects. The most significant changes are:

$$\begin{array}{rcll} \mathcal{P}_{\rm sig} & \rightarrow & \mathcal{P}_{\rm sig} \otimes \mathcal{R} & \text{detector resolution,} \quad (5.6) \\ \mathcal{P}_{\rm sig} \otimes \mathcal{R} & \rightarrow & f_{\rm sig}\, \mathcal{P}_{\rm sig} \otimes \mathcal{R} + (1-f_{\rm sig})\mathcal{P}_{\rm bcg} & \text{background and} \quad (5.7) \\ \mathcal{P}_{\rm sig}({\rm d}\Delta t, q) & \rightarrow & \mathcal{P}_{\rm sig}({\rm d}\Delta t, (1-2w)\, q) & \text{flavor tagging.} \quad (5.8) \end{array}$$

The event signal fraction is a function of kinematic variables and flavor tagging quality, $f_{\rm sig} = f_{\rm sig}(M_{\rm bc}, \Delta E, r)$, as described in the previous section.

On the top of these changes we add two minor modifications. First, we find there is a small fraction of events, $f_{\rm ol} \approx 1\%$, with poorly reconstructed vertices. We add an additional wide Gaussian component, $\mathcal{P}_{\rm ol}$, with $\sigma = 20\,{\rm ps}$, to describe such outliers. Second, the non-resonant background cannot be treated in the same way as the combinatoric background. The characteristics of the non-resonant decays, such as $B^0 \rightarrow D^- \overline{K}^0 \pi^+$, are similar as for the $B^0 \rightarrow D^+ D^-$ signal events: the non-resonant event probability is peaking in the $M_{\rm bc}$ and ΔE distributions and they have the exponential B^0 decay Δt distribution. We therefore describe them separately with the term $f_{\rm nr}\, \mathcal{P}_{\rm nr}$, where $f_{\rm nr}$ is proportional to $f_{\rm sig}$, $f_{\rm nr} = a\, f_{\rm sig}$. Taking all this into account, the event likelihood is given by

$$\begin{array}{rcl} \mathcal{L}_{\rm ev} & = & (1-f_{\rm ol})\, \mathcal{P} + f_{\rm ol}\, \mathcal{P}_{\rm ol} \,. \\ \mathcal{P} & = & f_{\rm sig}\, \mathcal{P}_{\rm sig} \otimes \mathcal{R} + f_{\rm nr}\, \mathcal{P}_{\rm nr} \otimes \mathcal{R} + f_{\rm bcg}\, \mathcal{P}_{\rm bcg} \end{array} \tag{5.9}$$

Subscripts sig, nr, bcg and ol refer to signal, non-resonant, background and outlier component, respectively. Each of the terms is described in the next paragraphs. In particular, we show how we obtain the parameters describing the distributions. It should be noted that the event likelihood function does not depend only on the parameters Δt and flavor tagging quality but also on a list of others that are used to parameterize the detector resolution.

All functions describing the distribution of signal and background events are normalized in such a way that

$$\Sigma_q \int_{-T}^{T} \mathcal{P}\,(\Delta t, q) \;=\; 1\,, \tag{5.10}$$

where $T = 70\,{\rm ps}$ limits the Δt fit range and Σ_q denotes a sum over flavor q, $q \in \{-1, 1\}$. A short note on plotting the projection of the fit to the Δt distribution can be found in Appendix, Sec. A.2.

The Kinematic Correction

When we calculate the Δt from the distance between the two B meson vertices Δz as $\Delta t = \beta\gamma\Delta z/c$, we do not take into account that the B mesons have a small, but non-zero momenta in the CMS. The energy of the B meson pair is equal to the beam energy, $2E_B = 2E = 10.58\,\mathrm{GeV}$. Since $E_B = \gamma_B m_B c^2$, the Lorentz factor for the boost from the B meson CMS the beam CMS is

$$\gamma_B = E_B/(m_B c^2) = 1.003795\,, \tag{5.11}$$

$\beta_B\gamma_B = 0.0872$ and $\beta_B = 0.0869$. If we did not take this into account, we would have an additional kinematic smearing of the Δt distribution. The later depends on the angle between the pair of B mesons and the electron beam direction, ϑ. To determine its contribution, we make a transformation from the B meson CM system to the laboratory system, i.e. a boost to the beam CMS followed by a rotation of the z axis by an angle ϑ and by another boost from the beam CMS to the laboratory system. The rotation transformation can be written as

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\vartheta & \sin\vartheta \\ 0 & 0 & -\sin\vartheta & \cos\vartheta \end{pmatrix} \tag{5.12}$$

since the cylindrical symmetry of the system allows us to choose the x direction for a rotation transformation axis.

Let us calculate how much does the fact that $\gamma_B \neq 1$ change the Δt vs. Δz relation. Suppose that the B meson decays at time t. In the B meson rest system, the B decay vertex is $x_0^\mu = (ct, 0, 0, 0)$, which transforms to the laboratory system transforms as

$$x^\mu \;=\; \Lambda\,\mathbf{R}\,\Lambda_B\,x_0^\mu \;= \tag{5.13}$$

$$\begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\vartheta & \sin\vartheta \\ 0 & 0 & -\sin\vartheta & \cos\vartheta \end{pmatrix} \begin{pmatrix} \gamma_B & 0 & 0 & -\beta_B\gamma_B \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta_B\gamma_B & 0 & 0 & \gamma_B \end{pmatrix} \begin{pmatrix} ct \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{5.14}$$

which gives for the decay point z coordinate in the laboratory system

$$z_{\mathrm{lab}} = -\beta\gamma\, ct \;\; \gamma_B \left(1 + \cos\vartheta\,\frac{\beta_B}{\beta}\right)\,. \tag{5.15}$$

This reduces to $z_{\mathrm{lab}} = -\beta\gamma ct$ in the case of $\beta_B = 0$. To calculate the value of Δz we can use the above expression and take into account that the two B mesons move to the opposite directions, $\vartheta_{\mathrm{tag}} = -\vartheta_{CP} = -\vartheta$. We obtain

$$\begin{aligned}
\Delta z &= z_{CP} - z_{\mathrm{tag}} = \\
&= -\beta\gamma\, ct_{CP}\, \gamma_B \left(1 + \cos\vartheta\, \frac{\beta_B}{\beta}\right) + \beta\gamma\, ct_{\mathrm{tag}}\, \gamma_B \left(1 + \cos\vartheta\, \frac{\beta_B}{\beta}\right) = \\
&= -\beta\gamma\, c\Delta t\, \gamma_B \left(1 + \cos\vartheta\, \frac{\beta_B}{\beta}\right)
\end{aligned} \tag{5.16}$$

The relative error on the Δt measurement in the case of kinematic approximation $\Delta z = -\beta\gamma c\Delta t$ is therefore

$$x = \frac{-\Delta z}{\beta\gamma c\Delta t} - 1 = \gamma_B \left(1 + \cos\vartheta\, \frac{\beta_B}{\beta}\right) - 1\,, \tag{5.17}$$

depending on the value of angle ϑ. In the limit two cases, $x = 0.23$ for $\vartheta = 0°$ and $x = 0.0038$ for $\vartheta = 90°$ Since we are able to measure the angle ϑ, we use Eq. 5.17 to describe the Δz versus Δt dependence. In effect, if

$$\begin{aligned}
\mathcal{P}(\Delta t) &= \frac{1}{2\tau}\, \mathrm{e}^{-|\Delta t|/\tau}\,, \text{ then} \\
\mathcal{P}(\Delta z/(\beta\gamma\, c)) &= \frac{1}{2\tau'}\, \mathrm{e}^{-|\Delta t|/\tau'}\,, \text{ where} \\
\tau' &= \gamma_B \left(1 + \cos\vartheta\, \frac{\beta_B}{\beta}\right) \tau
\end{aligned} \tag{5.18}$$

and similar for other terms of $\mathcal{P}_{\mathrm{sig}}$.

5.2.1 The Detector Resolution Function

We divide the detector resolution into three parts,

$$\mathcal{R} = \mathcal{R}_{\mathrm{det}}^{CP} \otimes \mathcal{R}_{\mathrm{det}}^{\mathrm{tag}} \otimes \mathcal{R}_{\mathrm{np}}\,, \tag{5.19}$$

where the sign $\otimes$ denotes the convolution of the corresponding functions. The largest contribution to the Δt measurement error comes from the error on the B_{CP} and B_{tag} vertex position, and is described by the detector resolution $\mathcal{R}_{\mathrm{det}} = \mathcal{R}_{\mathrm{det}}^{\mathrm{CP}} \otimes \mathcal{R}_{\mathrm{det}}^{\mathrm{tag}}$. The third term, $\mathcal{R}_{\mathrm{np}}$, describes additional smearing of the tag-side vertex position measurement due to the tracks originating from the secondary vertices (such as D meson decay vertices), as illustrated in Fig. 5.10. This effect is in principle a part of the B_{tag} vertex resolution function, but is described separately because it is specific for the tag-side vertex only. In this way we can describe the vertex resolution due to the tracking error for the B_{CP} and B_{tag} vertex in the same way.

We discuss each of the terms in the following paragraphs. We start with the parameterization of the B_{CP} and B_{tag} vertex resolution functions, $\mathcal{R}_{\mathrm{det}}^{\mathrm{CP}}$ and $\mathcal{R}_{\mathrm{det}}^{\mathrm{tag}}$. Since all analyses of the Δt distribution of Belle data use common B_{tag} vertex reconstruction, we use common values for the $\mathcal{R}_{\mathrm{det}}^{\mathrm{tag}}$ (that will be referred to as "standard" values). How we obtain

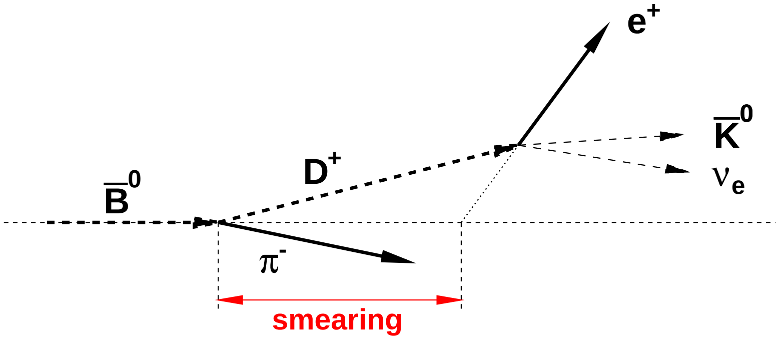


Figure 5.10: The tag-side vertex measurement error due to tracks coming from the non-primary vertices, such as the e^+ track from the $D^+ \to e^+\overline{K}^0\nu_e$ vertex. Since such tracks are assumed to originate from the primary B^0 vertex and the neutral tracks are not used in the vertex fit, an additional error is introduced to the $B_{\rm tag}$ vertex measurement.

these parameters is described in Ref. [33]. On the other hand, $B^0 \to D^+D^-$ decays are reconstructed from many final state charged particle tracks and could have significantly different parameters of the $\mathcal{R}_{\rm det}^{\rm CP}$ distribution. In this work, we therefore only describe how we determine the parameters describing the B_{CP} vertex resolution $\mathcal{R}_{\rm det}^{\rm CP}$ that are obtained specifically only for this analysis. We conclude this section by a description of the $\mathcal{R}_{\rm np}$ function describing the effect of the non-primary tracks.

Parameterization of the $\mathbf{B_{CP}}$ and $\mathbf{B_{\rm tag}}$ Vertex Resolution Function

When we fit the B_{CP} and $B_{\rm tag}$ vertices, we also determine the event-by-event error on the vertex z coordinate, $\sigma_{\rm vtx}$. The first approximation for the vertex resolution function could therefore be a Gaussian of the corresponding width $\sigma_{\rm vtx}$. However, a single Gaussian does not give a satisfactory description of the detector resolution. Therefore we introduce a more elaborate parameterization, with additional scale factors that also correct for possible biases in the values of $\sigma_{\rm vtx}$.

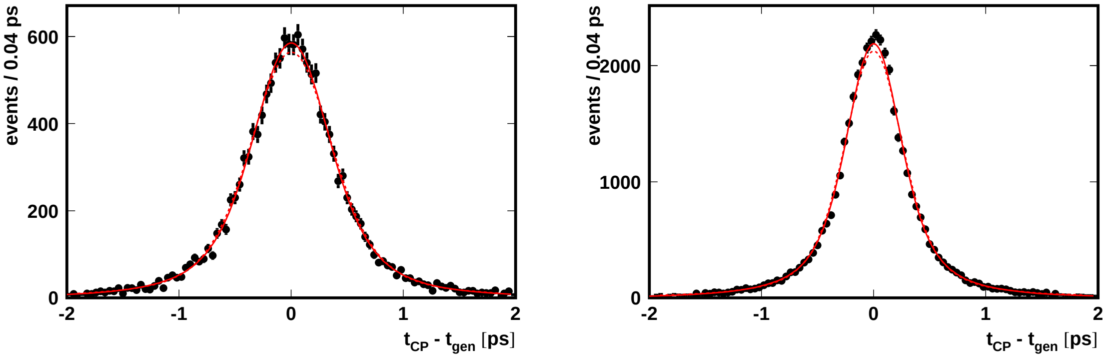
The decay time (i.e. vertex z coordinate) detector resolution for the B_{CP} and $B_{\rm tag}$ vertices are described in the same way, but with different values of the parameters being used for each of the vertices. The functions $\mathcal{R}_{\rm det}^{CP({\rm tag})}$ are each described by two Gaussian functions of mean $\mu = 0$ and widths proportional to the error on the vertex z coordinate $\sigma_{\rm vtx} = \sigma_z/(\beta\gamma c)$, which is determined by the vertex fit.

$$\begin{aligned}
\mathcal{R}_{\rm det}^{CP({\rm tag})} &= f\,\mathcal{G}(\sigma_1) + (1-f)\,\mathcal{G}(\sigma_2) \\
\sigma_1 &= (p_0 + \chi^2/ndf\; p_1)\,\sigma_{\rm vtx} \\
\sigma_2 &= p_2\,\sigma_1
\end{aligned} \tag{5.20}$$

The detector resolution parameters p_0, p_1, f and p_2 can be determined directly from a fit to the residual distribution $t_{CP} - t_{\rm true}$, where $t_{CP} = z_{CP}/(\beta\gamma c)$ and $t_{\rm true} = z_{\rm true}/(\beta\gamma c)$. The z coordinate of the reconstructed B_{CP} vertex is denoted by z_{CP} and its actual (known only for the MC data) value by z_{true}. Another possibility is to fit the Δt distribution which is described by the exponential B^0 decay time distribution and the resolution function $\mathcal{R}$. In the later case, the resolution parameters of $\mathcal{R}^{CP}_{\text{det}}$ are entangled with the parameters describing $\mathcal{R}^{\text{tag}}_{\text{det}}$, and can therefore be determined only with a smaller accuracy.

Fit of the $\mathbf{B_{CP}}$ Vertex Resolution Parameters from the Residual Distribution

First, resolution parameters are determined by a fit to the residual distribution $t_{CP} - t_{\text{true}}$ to check whether the same values can be used for the different reconstruction channels. Figure 5.11 shows the residual distribution and resolution function $\mathcal{R}^{CP}_{\text{det}}$ for the SVD1 and SVD2 signal MC sample. The resolution function parameters are obtained from a fit to the residual distribution in the range of ±2 ps and ±5 ps as shown in Table 5.6. It can be observed that the result is quite sensitive to the fitting range. Furthermore, parameters obtained from the fit to the ±5 ps range do not describe the distribution in the peak well. It can therefore be concluded that the double-Gaussian detector resolution function from Eq. 5.20 is not sufficient for the description of the residual distribution in the full ±5 ps range. In the end this does not present a problem since the error on the $\Delta t = t_{CP} - t_{\text{tag}}$ distribution is described as the sum of the errors for the CP and tag-side B meson vertex and the actual Δt distribution is described well, as we will show in Fig. 5.12 and related paragraphs.



(a) SVD1 (b) SVD2

Figure 5.11: Residual distribution $t_{CP} - t_{\text{true}}$ for correctly reconstructed signal MC events. The full and dashed curves show the detector resolution function with parameters obtained from the fit to the distribution in the range ±2 ps and ±5 ps, respectively.

Table 5.7 shows the fit result for different reconstruction channels separately. Parameters for the SVD2 sample are significantly different from those describing the SVD1 sample, therefore a separate treatment of the two detector layouts is necessary. From the table we also conclude that the resolution parameters for the $K\pi\pi$ and $K_S\pi$ reconstruction channels are consistent, therefore we can use the same set of parameters to describe both.

T [ps]	sample	p_0	p_1	$1-f$	p_2	χ^2/ndf
5	SVD1	1.16 ± 0.02	0.19 ± 0.02	0.07 ± 0.01	3.54 ± 0.17	97/96
2	SVD1	1.07 ± 0.03	0.16 ± 0.02	0.18 ± 0.03	2.24 ± 0.11	96/96
5	SVD2	1.23 ± 0.01	0.21 ± 0.01	0.11 ± 0.01	3.75 ± 0.08	128/96
2	SVD2	1.15 ± 0.01	0.16 ± 0.01	0.19 ± 0.01	2.73 ± 0.06	114/96

Table 5.6: Resolution function parameters obtained by a fit to the $t_{CP} - t_{\text{true}}$ distribution of the signal MC sample in the range [-T, T]. Fit χ^2/ndf is calculated for 100 equidistant bins in the range ± 2 ps.

rec. channel	sample	p_0	p_1	$1-f$	p_2	χ^2/ndf
$K\pi\pi - K\pi\pi$	SVD1	1.07 ± 0.03	0.16 ± 0.02	0.17 ± 0.03	2.25 ± 0.12	93/96
$K_S\pi - K\pi\pi$	SVD1	1.09 ± 0.08	0.13 ± 0.04	0.25 ± 0.08	2.26 ± 0.24	87/96
$K\pi\pi - K\pi\pi$	SVD2	1.16 ± 0.02	0.16 ± 0.01	0.18 ± 0.01	2.74 ± 0.06	119/96
$K_S\pi - K\pi\pi$	SVD2	1.13 ± 0.04	0.17 ± 0.03	0.21 ± 0.03	2.76 ± 0.15	112/96

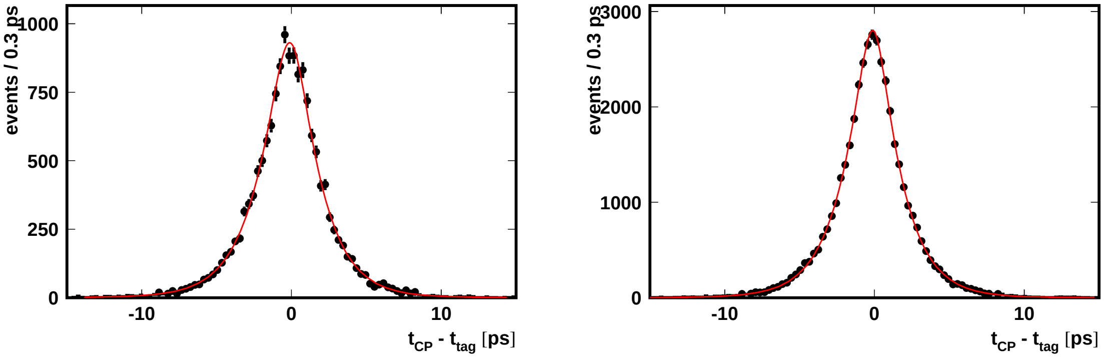
Table 5.7: Resolution function parameters for different decay modes obtained by a fit to the $t_{CP} - t_{\text{true}}$ distribution of the signal MC sample in the range [-2 ps, 2 ps]. Fit χ^2/ndf is calculated for 100 equidistant bins in the fit range.

Fit of the $\mathbf{B_{CP}}$ Vertex Resolution Parameters from the $\mathbf{\Delta t}$ Distribution

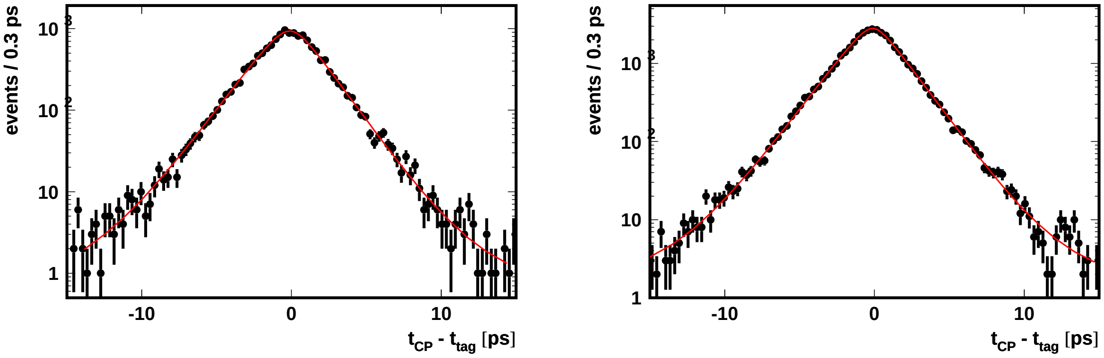
In the second step, the parameters describing the detector resolution $\mathcal{R}_{\text{det}}^{CP}$ are determined from the $\Delta t = t_{CP} - t_{\text{tag}}$ distribution, shown in Fig. 5.12. Parameters describing the B_{tag} vertex resolution, $\mathcal{R}_{\text{det}}^{\text{tag}}$, are fixed to the Belle "default" values [33]. Fit result for different fitting ranges is shown in Table 5.8. The parameters of the outlier component, the fraction and the width of $\mathcal{R}_{\text{ol}}$, are also included in the fit to describe the Δt distribution of the outlier events. The errors on all parameters are larger than those obtained from the fit to the residual distribution because the Δt distribution is smeared also by contributions other than $\mathcal{R}_{\text{det}}^{CP}$,

$$\mathcal{P}_{\text{sig}}(\Delta t) = \frac{e^{-|t|/\tau}}{2\tau} \otimes \mathcal{R}_{\text{det}}^{CP} \otimes \mathcal{R}_{\text{det}}^{\text{tag}} \otimes \mathcal{R}_{\text{kin}} \otimes \mathcal{R}_{\text{np}} \,. \tag{5.21}$$

The value of the parameter p_2 obtained from the $\Delta t = t_{CP} - t_{\text{tag}}$ distribution shown in Table 5.8 is systematically different from the value shown in Table 5.6. As $\sigma_2 = p_2\,\sigma_1$, the Gaussian component of width σ_2 describes the fraction of events that have a residual distribution considerably wider than is their estimated error. Such events are not even within the $[-2\,\text{ps}, 2\,\text{ps}]$ range that was used in the fit to the residual distribution, therefore it is not that surprising that the two fit results differ. The values of other parameters are quite consistent. Since the CPV parameters are determined from a fit to the Δt distribution, we use the parameters which describe the Δt distribution. For the data



(a) SVD1 (b) SVD2



(c) SVD1, logarithmic scale (d) SVD2, logarithmic scale

Figure 5.12: The $\Delta t = t_{CP} - t_{\text{tag}}$ distribution for correctly reconstructed signal MC events. The line shows the detector resolution function with parameters obtained from the fit to the distribution in the range ±70 ps.

T [ps]	sample	p_0	p_1	$1-f$	p_2	χ^2/ndf
70	SVD1	1.07 ± 0.15	0.26 ± 0.05	0.08 ± 0.03	6.07 ± 1.17	113/94
50	SVD1	1.06 ± 0.15	0.26 ± 0.05	0.09 ± 0.04	5.89 ± 1.11	113/94
30	SVD1	1.04 ± 0.15	0.24 ± 0.04	0.09 ± 0.04	5.33 ± 0.96	116/94
70	SVD2	1.19 ± 0.13	0.23 ± 0.03	0.08 ± 0.02	7.46 ± 0.96	87/94
50	SVD2	1.16 ± 0.14	0.24 ± 0.03	0.08 ± 0.02	7.26 ± 0.91	86/94
30	SVD2	1.13 ± 0.15	0.24 ± 0.03	0.08 ± 0.03	6.92 ± 0.89	87/94

Table 5.8: Resolution function parameters for the different decay modes obtained from a fit to the $\Delta t = t_{CP} - t_{\text{tag}}$ distribution in the range $[-T, T]$. A signal MC sample of approximately 15.000 (43.000) correctly reconstructed events for the SVD1 (SVD2) detector layout was used for the purpose. Fit χ^2/ndf is calculated for a 100 equidistant bins in the range ±15 ps.

sample, this is the only way to determine the parameters anyway. The determination of the resolution parameters for data is described later in Sec. 4.3.

Resolution Function Parameters for the Control Sample (MC)

In order not to rely on the MC or on the "standard" resolution parameters that are obtained from the data sample of decays which have a smaller number of tracks in the final state, the resolution parameters for the B_{CP} meson vertex are determined from a fit to the Δt distribution of kinematically similar $B^0 \to D_s^+ D^-$ decays. In order to confirm that the resolution parameters for the $B^0 \to D_s^+ D^-$ and $B^0 \to D^+ D^-$ decays are similar, we first fit the resolution parameters for the $B^0 \to D_s^+ D^-$ events from their $t_{CP} - t_{\rm true}$ residual distribution (shown in Table 5.9) and compare them with the corresponding parameters obtained for the $B^0 \to D^+ D^-$ events, shown in Table 5.6. The values of all parameters are consistent, therefore the same values can be used.

T [ps]	sample	p_0	p_1	$1-f$	p_2	χ^2/ndf
5	SVD1	1.11 ± 0.02	0.20 ± 0.02	0.12 ± 0.01	3.17 ± 0.15	126/96
2	SVD1	1.01 ± 0.04	0.15 ± 0.02	0.26 ± 0.04	2.16 ± 0.10	120/96
5	SVD2	1.26 ± 0.02	0.19 ± 0.01	0.12 ± 0.01	3.78 ± 0.10	91/96
2	SVD2	1.16 ± 0.02	0.15 ± 0.01	0.21 ± 0.02	2.63 ± 0.07	109/96

Table 5.9: The resolution parameters for the $B^0 \to D_s^+ D^-$ sample, obtained by the fit to the residual $t_{CP} - t_{\rm true}$ distribution in the range [-T, T]. Fit χ^2/ndf is calculated for a 100 equidistant bins in the range T.

As the number of reconstructed events is still rather small even in the case of the control sample, the performance of the fit is studied by doing the fit of the resolution function parameters for a generic MC sample that has a similar number of events. The result is shown in Table 5.10. It can be observed that all of the parameters can only be determined with very large errors. The fit is therefore repeated with parameters describing the tail Gaussian (p_2 and f) fixed to the MC values. The corresponding Δt distributions are shown in Fig. 5.13. The improvement in the fit χ^2/ndf is not big enough to justify the additional parameters, therefore only parameters p_0 and p_1 are to be determined by the fit to the data sample.

Resolution Function Parameters for the Data Sample

The resolution parameters for data are obtained from an unbinned ML fit to the Δt distribution of the reconstructed $B^0 \to D_s^+ D^-$ events. The fraction and the width of the outlier component are also floated in this fit. The resolution parameters for the tag-side vertex are fixed to the default values. The result is shown in Table 5.10. Thus obtained values of the resolution parameters for data will be used for the fit of the CP parameters in the data sample. Uncertainty in the resolution function parameters is taken into account in the systematic error study.

sample	# events	p_0	p_1	$1-f$	p_2	χ^2/ndf
$B^0 \to D_s^+ D^-$ generic MC						
SVD1	2448	0.51 ± 0.19	0.09 ± 0.02	0.10 ± 0.06	4.82 ± 1.51	182/94
SVD2	1322	0.83 ± 0.26	0.00 ± 0.02	0.00 ± 0.32	2.47 ± 6.12	186/94
SVD1	2448	0.61 ± 0.23	0.10 ± 0.02			183/96
SVD2	1322	0.75 ± 0.24	0.00 ± 0.02			186/96
$B^0 \to D_s^+ D^-$ data						
SVD1	606	0.20 ± 0.26	0.19 ± 0.06	0.12 ± 0.07	10.29 ± 8.25	131/94
SVD2	1395	0.27 ± 0.08	0.04 ± 0.01	0.46 ± 0.10	12.00 ± 6.87	102/94
SVD1	606	0.54 ± 0.47	0.22 ± 0.07			122/96
SVD2	1395	1.40 ± 0.32	0.22 ± 0.03			99/96

Table 5.10: Resolution parameters obtained by a fit to the $\Delta t = t_{CP} - t_{\rm tag}$ distribution in the range $[-70\,{\rm ps}, 70\,{\rm ps}]$. Fit χ^2/ndf is calculated for 100 equidistant bins in the range of $\pm 15\,{\rm ps}$.

The Effect of the Non-Primary Tracks

We describe the tag-side vertex error due to the non-primary tracks $\mathcal{R}_{\rm np}$ (Fig. 5.10) with a bi-exponential function $E_{\rm pn}(x)$,

$$\mathcal{R}_{\rm np}(\Delta z_{\rm tag}^{\rm np}) = (1-f_{\rm np})\,\delta(\Delta z_{\rm tag}^{\rm np}) + f_{\rm np}\,E_{\rm pn}(\Delta z_{\rm tag}^{\rm np}) \tag{5.22}$$

$$E_{\rm pn}(x) = \begin{cases} f_p \frac{1}{\tau_p}\,{\rm e}^{-x/\tau_p} & x \geq 0 \\ f_n \frac{1}{\tau_n}\,{\rm e}^{x/\tau_n} & x < 0 \end{cases}, \tag{5.23}$$

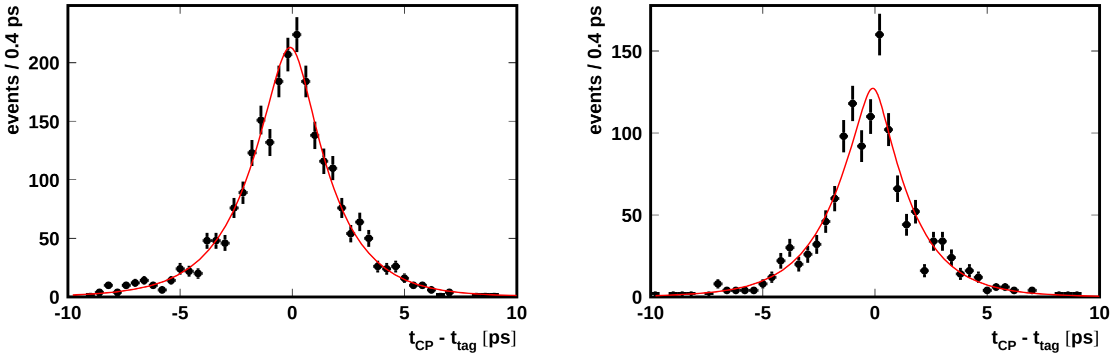
where $\Delta z_{\rm tag}^{\rm np}$ is the error due to the non-primary tracks, $f_{\rm np}$ is the fraction of events with non-primary tracks and $f_{p(n)}$ is the fraction of positive (negative) exponential part such that $f_p + f_n = 1$. In a fraction of events $1 - f_{\rm np}$ there is no additional smearing, as described by the δ-function term $\delta(\Delta z_{\rm tag}^{\rm np})$. We find that the vertex position shift has a linear dependence on both $\sigma_{\rm tag}$ and $\xi_{\rm tag}$ [33]. We therefore parameterize $\tau_{p(n)}$ as

$$\tau_{p(n)} = \tau_{p(n)}^0 + \tau_{p(n)}^1 (s^0 + s^1 \xi_{\rm tag})\,\sigma_{\rm tag}/(\beta\gamma c)\,. \tag{5.24}$$

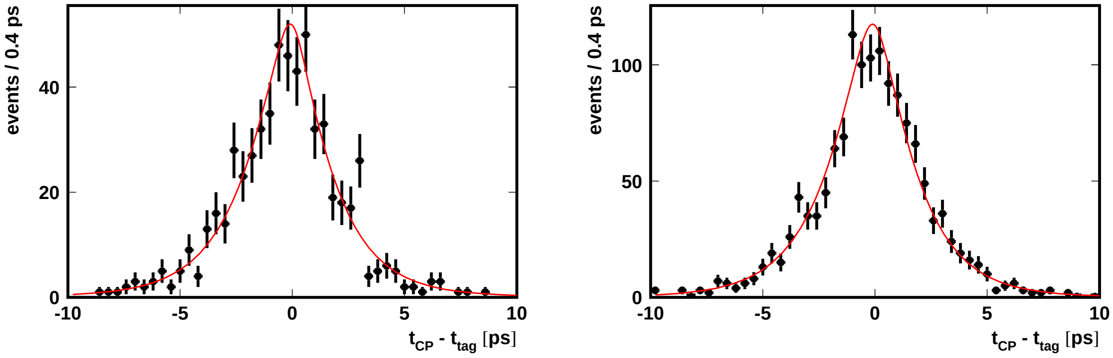
We determine the six parameters f_δ, f_p, τ_p^0, τ_p^1, τ_n^0 and τ_n^1 and the two scale factors s^0 and s^1 from a fit to the $\Delta z_{\rm tag}^{\rm np}$ distributions in MC data.

5.2.2 The Δt Distribution for the Background Events

Based on a MC study, we find that about half of the combinatorial background events are coming from the $B\overline{B}$ decays ($b \to c$ transition), which have an exponential decay Δt



(a) SVD1 generic MC (b) SVD2 generic MC



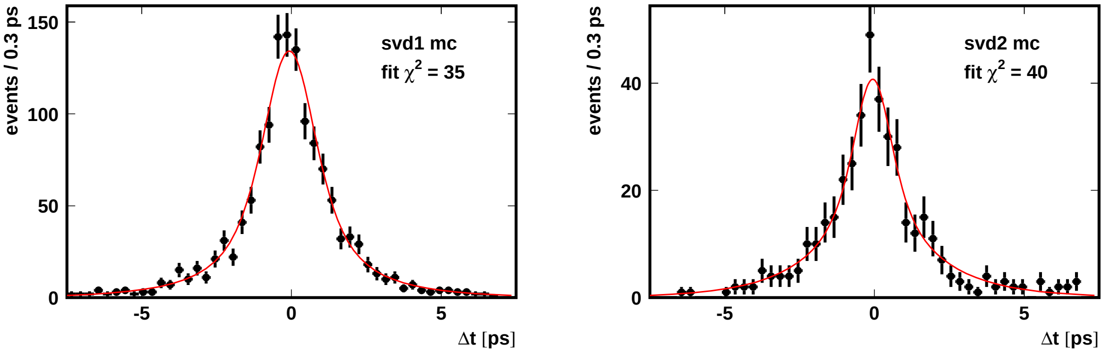
(a) SVD1 data (b) SVD2 data

Figure 5.13: $\Delta t = t_{CP} - t_{\rm tag}$ distribution for the $\overline{B}^0 \to D_s^+ D^-$ events. Line shows the detector resolution function with the parameters obtained from the fit to the distribution.

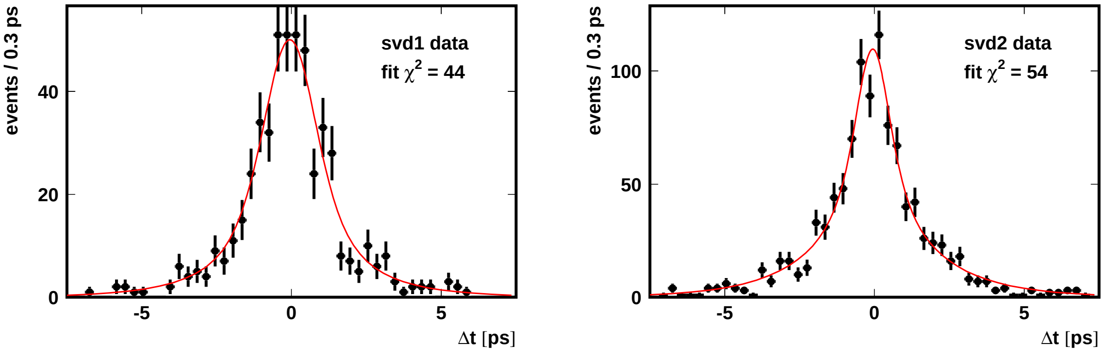
distribution. The other half are continuum $e^+e^- \to q\bar{q}$ ($q \in \{u,d,s,c\}$) events, the Δt distribution for which is a δ-function like. We therefore describe the Δt distribution of the background events by

$$\mathcal{P}_{\rm bcg} = \left((1 - f_\delta)\,\frac{e^{-|\Delta t|/\tau_{\rm bcg}}}{2\tau_{\rm bcg}} + f_\delta\, \delta(\Delta t) \right) \otimes \mathcal{G}(\mu, \sigma_{\rm bcg})\,, \tag{5.25}$$

where for the background resolution function $\mathcal{G}(\mu, \sigma_{\rm bcg})$ we assume a Gaussian function of mean μ and width $\sigma_{\rm bcg}$. Parameters describing the background are determined by a fit to the Δt distribution of events in the $M_{\rm bc} < 5.27\,{\rm GeV}/c^2$ and $\Delta E > 0.06\,{\rm GeV}$ sideband region, which is shown in Fig. 5.14. Separate values are used for the data and MC, SVD1 and SVD2 samples. The fit result is shown in Table 5.11. The negative value of $\mu \approx -0.06 \pm 0.05\,{\rm ps}$ describes the asymmetry of the tag-side vertex resolution function due to the effect of the non-primary tracks described earlier. The possibility of describing the Δt distribution for background events with two Gaussian functions instead of one was also studied. Since adding the second Gaussian did not make any significant improvement in the fit χ^2, it was concluded that a single Gaussian function $\mathcal{G}(\mu, \sigma^{\rm bcg})$ already gives a completely adequate description of the background.



(a) SVD1 generic MC (b) SVD2 generic MC



(c) SVD1 data (d) SVD2 data

Figure 5.14: The Δt distribution for background events from the $M_{\rm bc}$ and ΔE sideband. The χ^2 of the fit is calculated for the Δt range $[-7.5, 7.5]$ ps divided into 50 equidistant bins (46 degrees of freedom).

SVD		# ev.	μ [ps]	$\sigma^{\rm bcg}$ [ps]	f_δ	$\tau_{\rm bcg}$ [ps]	fit χ^2
Fitting range: 70 ps							
1	MC	1369	-0.07 ± 0.05	1.6 ± 0.1	0.66 ± 0.07	1.9 ± 0.3	35
2	MC	404	-0.06 ± 0.09	1.6 ± 0.3	0.36 ± 0.14	1.7 ± 0.3	40
1	data	517	-0.06 ± 0.08	1.6 ± 0.2	0.67 ± 0.16	1.6 ± 0.5	44
2	data	1041	-0.06 ± 0.05	1.4 ± 0.2	0.37 ± 0.07	1.8 ± 0.2	54

Table 5.11: The ML fit result for the parameters describing the Δt distribution of the background events. The last column shows the χ^2 of the fit calculated for the Δt range $[-7.5, 7.5]$ ps divided into 50 equidistant bins (46 degrees of freedom).

The Δt distribution for the non-resonant background is described by an exponential B^0 meson decay Δt distribution smeared by the detector resolution. Since these decays are in fact correctly reconstructed B^0 meson decays, the same detector resolution as for the signal events is used.

5.3 The Monte Carlo Studies

Before actually fitting the CP parameters for the data sample, we have to ascertain the accuracy of the fitting procedure and check for any bias in the fit result or its error as determined by the fit. For this purpose, we perform several MC studies that are described in the next few sections.

5.3.1 The Toy Monte Carlo

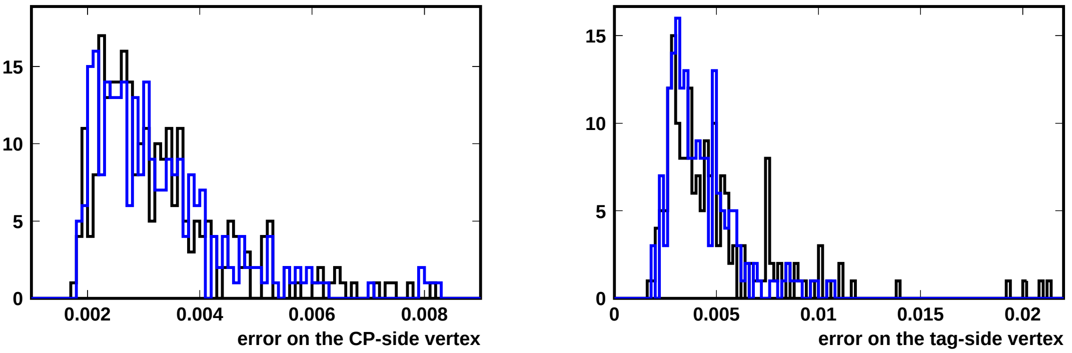
In order to save computer processing power, we simulate only the distributions of all relevant parameters rather than simulating particle decays in the detector and the full detector response. We call such a simulation a toy MC. The distributions of all parameters that are taken into account in the fit of the CP parameters are simulated according to the corresponding distributions in data: vertex z coordinates, their errors and χ^2/ndf for the CP and tag-side B meson vertices, r bins and wrong tag fraction distributions, $N_{\rm sig}/N_{\rm bcg}$ ratio, kinematic variables $M_{\rm bc}$, ΔE and $\cos\theta$. The distributions of the vertex errors and χ^2/ndf for the CP and tag-side vertex are shown in Fig. 5.15. The same resolution function parameters as are used to describe the data are used also for the toy MC study.

Expected Errors on the CP Parameters

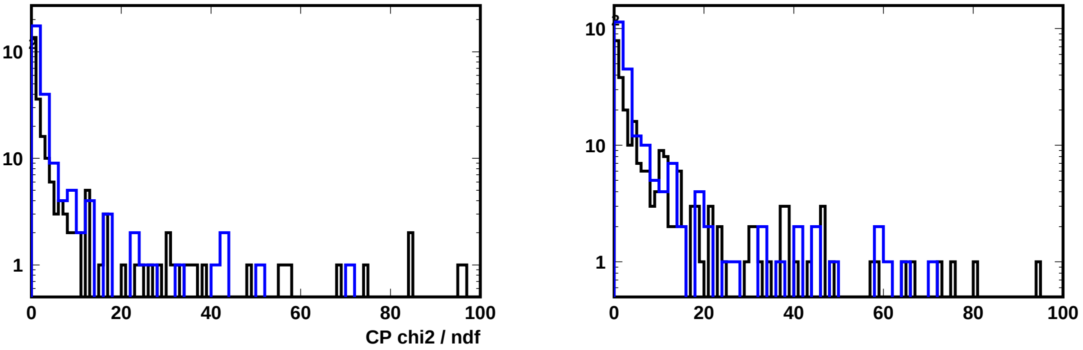
The first goal of the toy MC study was to estimate the errors on the CP parameters for a given number of reconstructed signal and background events. For this purpose, $10,000$ toy MC samples were simulated for values of CP parameters $\mathcal{S}_{\rm sim} = -0.7$ and $\mathcal{A}_{\rm sim} = 0$. Figure 5.16 shows the distribution of the fitted parameters $\mathcal{S}$ and $\mathcal{A}$ and their pull distributions. A fit to the distributions of $\mathcal{S}$ and $\mathcal{A}$ with a Gaussian function results in $\sigma_{\mathcal{S}} = 0.34$ and $\sigma_{\mathcal{A}} = 0.24$, which agrees with their RMS values. A fit to the pull distributions yields $\sigma_{{\rm pull\,in}\,\mathcal{S}} = 1.039 \pm 0.009$ and $\sigma_{{\rm pull\,in}\,\mathcal{A}} = 1.024 \pm 0.008$, which shows that the symmetric errors are only slightly underestimated.

Linearity Test

The second reason for the toy MC study was to test the linearity of the fit response. The toy MC samples were generated for each value of $(\mathcal{S}, \mathcal{A}) \in \{(-0.9, 0), (-0.8, 0), \ldots, (0.9, 0)\}$ and $(\mathcal{S}, \mathcal{A}) \in \{(0, -0.9), (0, -0.8), \ldots, (0, 0.9)\}$. The distribution of the CP fit results for each of the input values is fitted with a Gaussian function, and its mean is compared to the input value. The result is shown in Fig. 5.17. A small systematic shift is observed in the case of the parameter $\mathcal{S}$. It is more prominent for larger absolute values of $\mathcal{S}$. A similar tendency is observed also for the parameter $\mathcal{A}$. Nevertheless, since the observed



(a) CP-side σ_{vtx} (b) tag-side σ_{vtx}



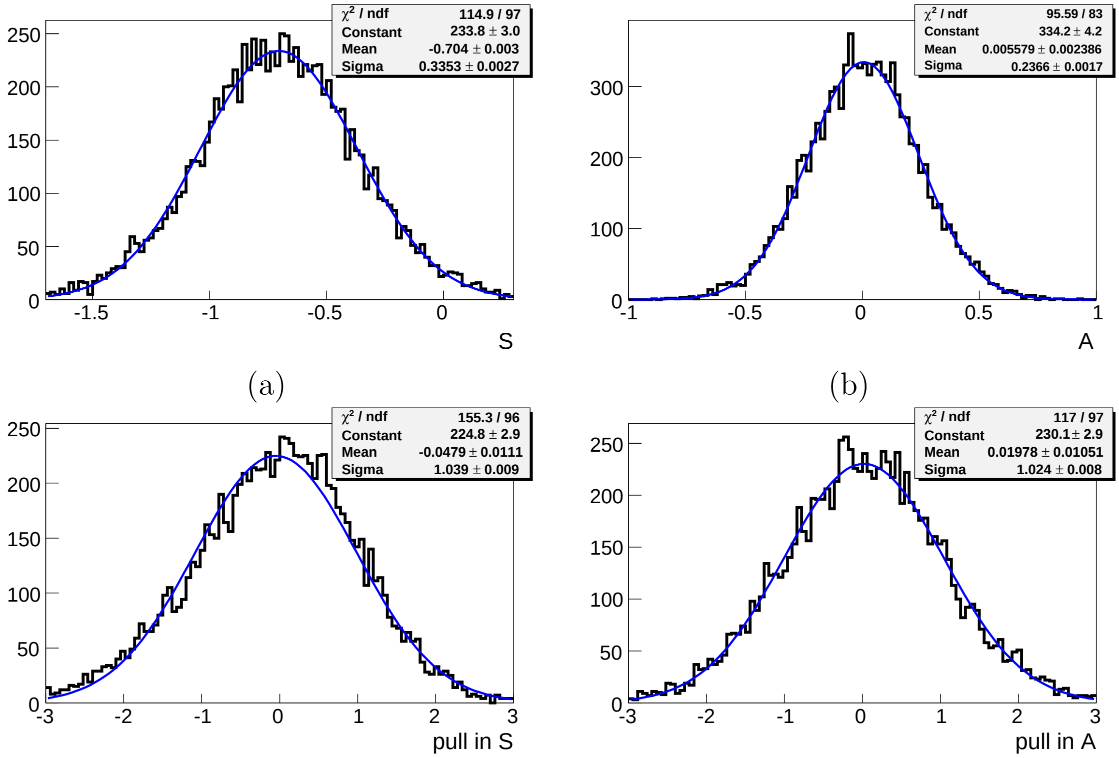
(c) CP-side χ^2/ndf (d) tag-side χ^2/ndf

Figure 5.15: Distributions of the different parameters simulated in the toy MC (blue) compared to the corresponding distributions in data (black).

systematic shift is 20-times smaller than the expected statistical error on CPV parameters, we conclude that this systematic effect is small enough.

5.3.2 Linearity Test for Signal MC

We also test the fit response using a large sample of signal MC events in order to test for other possible sources of the CP fit bias. Samples of 200.000 events were generated for different values of the parameter $\mathcal{S}$ with $A = 0$. The fit for the CP parameters was done for each sample of approximately 3200 (9000) reconstructed events simulated with the SVD1 (SVD2) conditions. The values of resolution parameters for the MC sample are shown in Table 5.14. The result is shown in Fig. 5.18. The difference between the fit result and simulated value was fitted by a constant function to evaluate the significance of a possible systematic shift. A $2\,\sigma$ discrepancy is observed only for the parameter $\mathcal{S}$, SVD2 sample. We therefore conclude that there is no significant systematic error observed by this study.



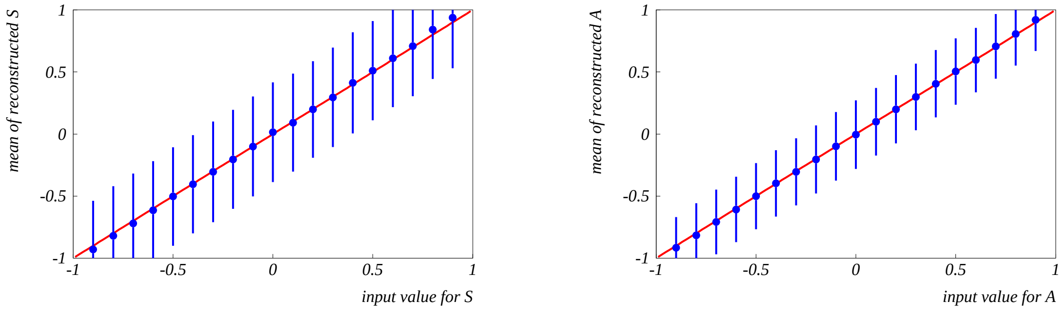
(a) (b)

(c) (d)

Figure 5.16: Distributions of the fit results for 10,000 toy MC samples. Figures (c) and (d) show the pull distribution S/σ_S or A/σ_A, where $\sigma_{S(A)}$ are the symmetric (parabolic) errors determined by the fit.

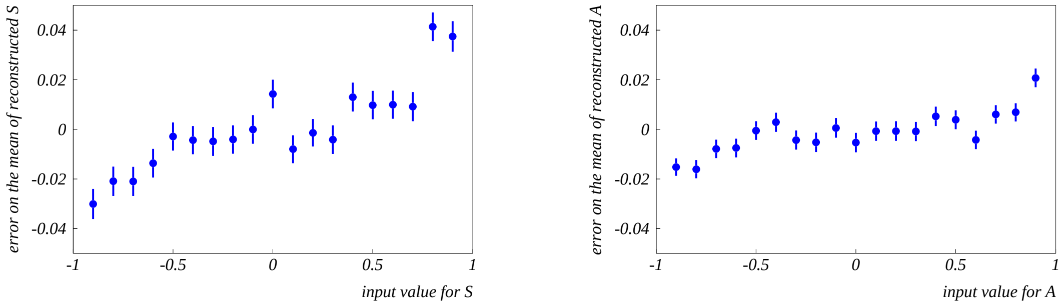
5.3.3 Lifetime and CP Parameters Fit for Generic MC

We also do the lifetime and CP parameters fit for the generic MC sample of signal and background events. The results are shown in Table 5.13 and are consistent with the simulated values of $\tau = 1.536\,\text{ps}$ and $S = A = 0$.



(a) S

(b) A



(c) difference in S

(d) difference in A

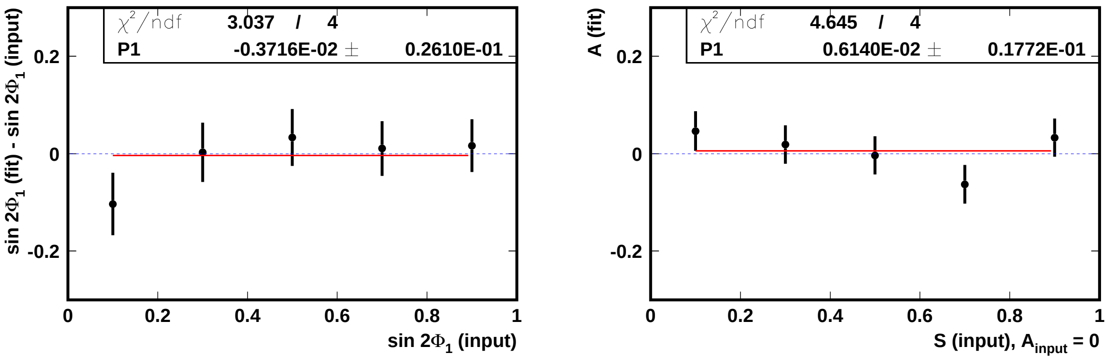
Figure 5.17: A result of the toy MC linearity test for $N_{\text{sig}} = N_{\text{bcg}} = 100$. Errors in (a) and (b) are average errors on the CPV parameters for one toy MC sample. Errors in (c) and (d) are errors on the mean from the fit of the Gaussian function.

parameter	SVD1	SVD2
p_0	1.07	1.19
p_1	0.26	0.23
f	0.08	0.08
p_2	6.07	7.46

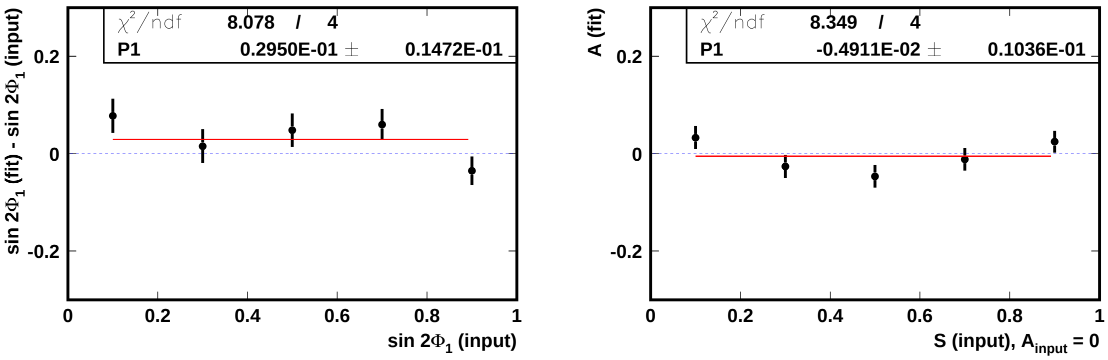
Table 5.12: Values of the physics parameters and resolution function parameters used for the CP fit on generic signal MC sample.

sample	# events	τ [ps]	$\mathcal{S}$	$\mathcal{A}$
SVD1	204	1.662 ± 0.228	0.159 ± 0.322	0.057 ± 0.210
SVD2	67	1.255 ± 0.285	-0.211 ± 0.524	-0.527 ± 0.330
both	271	1.546 ± 0.183	0.045 ± 0.283	-0.104 ± 0.178

Table 5.13: The result of the lifetime and CP fit for the generic MC sample.



(a) $\mathcal{S}$, SVD1 (b) $\mathcal{A}$, SVD1



(c) $\mathcal{S}$, SVD2 (d) $\mathcal{A}$, SVD2

Figure 5.18: Difference between the fit result for a CPV parameter and its input value as a function of the input value for the parameter $\mathcal{S}$.

5.4 Results

After finishing all of the tests for the MC samples we analyzed the data sample. We start by a lifetime and CP parameters fit for $B^0 \to D_s^+ D^-$ sample before we show the result for the $B^0 \to D^+ D^-$ sample. The lifetime fit is done as another consistency check to confirm that we understand the Δt distribution. Table 5.14 shows the values of the B_{CP} vertex resolution parameters that are used to describe the data sample. The Belle "standard" set of parameters is used for the tag-side vertex resolution function and wrong tag fractions [33, 32]. The world average values are used for the B^0 meson mass difference $\Delta m = (0.507 \pm 0.005)/\,\mathrm{s} \cdot \hbar/c^2$ and lifetime $\tau = (1.530 \pm 0.009)\,\mathrm{ps}$ [11].

	data	
	SVD1	SVD2
p_0	$0.54^{+0.44}_{-0.49}$	$1.40^{+0.30}_{-0.33}$
p_1	$0.22^{+0.08}_{-0.06}$	$0.22^{+0.03}_{-0.02}$
$1-f$	0.08	0.08
p_2	6.07	7.46

Table 5.14: Parameters describing the B_{CP} vertex resolution.

5.4.1 The Control Sample Study

The $B^0 \to D_s^+ D^-$ decays proceed via a similar Feynman diagrams as $B^0 \to D^+ D^-$ decays, with a strange antiquark replacing the down antiquark in Fig. 2.5. For that reason, the CKM matrix elements related to the two amplitudes are $V_{cb}V_{cs}$ (tree) and $V_{tb}V_{ts}$, $V_{cb}V_{cs}$ or $V_{ub}V_{us}$ (penguin diagram). In principle, the interference between the tree and the penguin amplitudes could give rise to CP violation for this decay as well. However, the only coefficient of a different weak phase $V_{ub}V_{us}$ is of the order of λ^4 and is therefore small compared to $V_{cb}V_{cs}$. While the terms $V_{tb}V_{ts}$ and $V_{cb}V_{cs}$ are of the same order in λ as the tree contribution, they all have zero complex phase. Since a different weak phase is a condition for a non-zero CP violation, we conclude that the CP violation in this decay mode is negligible and therefore this decay can be used as a cross-check for the $B^0 \to D^+ D^-$ sample. The measured CP parameters for this decay are $\mathcal{S} = -0.064 \pm 0.094$ and $\mathcal{A} = 0.091 \pm 0.060$. The values of both CPV parameters are consistent with zero. The corresponding Δt distribution is shown in Fig. 5.19. As an additional check of our understanding of the Δt distribution we also measure the B^0 lifetime, $\tau = 1.458 \pm 0.049$, and we find it consistent with the world everage [11].

5.4.2 The Lifetime and CP-blind fit for $\mathbf{B^0 \to D^+ D^-}$ Events

Two further tests were performed on the $B^0 \to D^+ D^-$ events, a fit of the lifetime and a so-called CP-blind fit. Result of the lifetime fit for the data sample is consistent with the

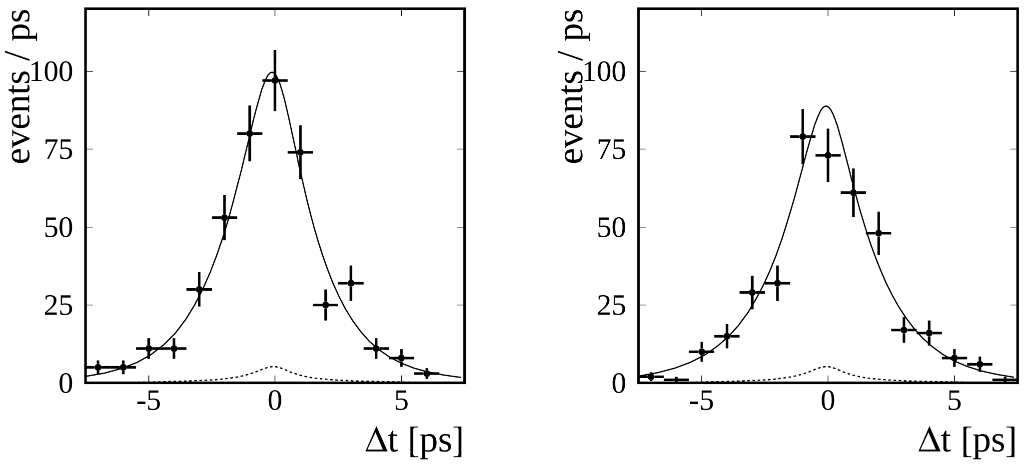


Figure 5.19: The Δt distribution for $B^0 \to D_s^+ D^-$ events with good tagging information ($r > 0.5$) when the tag-side B-meson is reconstructed as B^0 (a) or $\overline{B}^0$ (b). The full and dashed curves show the projection of the fit result for all and background events, respectively.

known B^0 lifetime [11].

$$\begin{aligned} \tau &= (1.56 \pm 0.40)\,\mathrm{ps} \qquad \text{for SVD1 and} \\ \tau &= (1.59 \pm 0.27)\,\mathrm{ps} \qquad \text{for SVD2.} \end{aligned} \tag{5.26}$$

For the CP-blind fit, a random flavor ($q = \pm 1$) is assigned to all events in the data sample. The result of the CP parameters fit for such a sample must of course be consistent with zero and has no information about the actual CP parameters. However, the widths of the $\mathcal{S}$ and $\mathcal{A}$ distributions serve as the error estimates and therefore we can cross-check our previous error estimate done by the toy MC study. The widths of the fit result distributions for 1000 CP-blind data samples are $\mathrm{RMS}_{\mathcal{S}} = 0.37$ and $\mathrm{RMS}_{\mathcal{A}} = 0.22$ and are consistent with the earlier toy MC estimates within 10%.

5.4.3 CP fit for $\mathbf{B^0 \to D^+D^-}$ Data Sample

From an unbinned fit to the measured Δt distribution described by Eq. (5.9), we obtain the CP violation parameters for $B^0 \to D^+ D^-$,

$$\begin{aligned} \mathcal{S} &= -1.13^{+0.30}_{-0.21} \\ \mathcal{A} &= +0.91^{+0.19}_{-0.22}\,, \end{aligned} \tag{5.27}$$

where the quoted error is the statistical error as determined by the fit. The correlation coefficient between the two parameters is 0.038. The Δt distributions are shown in Fig. 5.20.

There are two things to be noted. First, the values of both parameters are outside the physical region of $\mathcal{S}^2 + \mathcal{A}^2 \leq 1$. The signal likelihood function is therefore actually

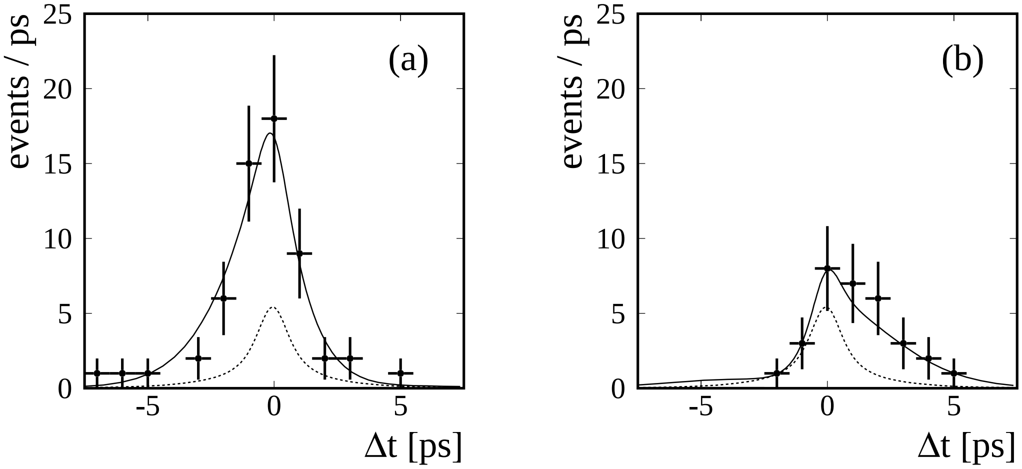


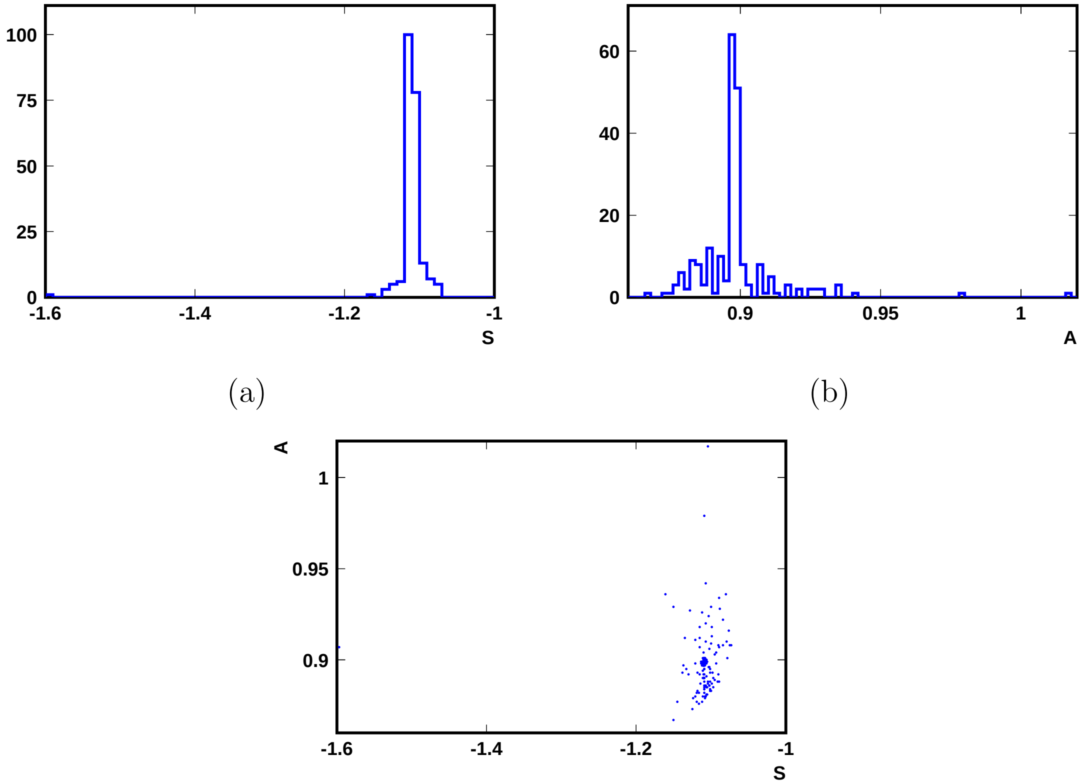
Figure 5.20: The Δt distribution for events with good tagging information ($r > 0.5$) when the tag-side B-meson is reconstructed as B^0 (a) or $\overline{B}^0$ (b). The full and dashed curves show the projection of the fit result and background contribution, respectively.

negative in some region and the errors determined by the fit could be underestimated. To avoid this possibility, we use the toy MC study to determine the statistical errors. Second and perhaps even more intriguing, the result for the parameter $\mathcal{A}$ favors a large direct CP violation in contradiction to the SM expectations. Due to this unexpected result, we perform a few more checks to test the consistency of the fit. We check if there is by any chance one particular event in the data sample that makes an unreasonably large impact on the result. Then we fit the parameter $\mathcal{A}$ only from the signal yield for events tagged as B^0 or $\overline{B}^0$ separately (without the Δt information). To test for a flavor asymmetry in the background events we also do the CP fit for events in the M_{bc} sideband. Last but not least, we use the D meson sidebands as an additional test of the tagging algorithm.

5.4.4 Special event scan

To check if one single event is significantly changing the fit result, the fit is repeated for the data sample with one of the events removed. This is repeated for all events such that each of the events is removed once. The distributions of the fit result for the parameters $\mathcal{S}$ and $\mathcal{A}$ are shown in Fig. 5.21. While there are a few events which make a difference in $\delta\mathcal{S} \approx 0.5$ and $\delta\mathcal{A} \approx -0.1$ there are no events pulling the result outside the physical region – no events in the absence of which the fit result would have a significantly different $\mathcal{S}$ or $\mathcal{A}$.

We repeat this study for a larger number of events N being removed from the data sample. For $N > 1$, we remove N random events from the sample, repeat the fit of the CP parameters and write down the difference in the fit result. We repeat this many times to minimize the statistical fluctuations of the difference. By this study we can estimate how big is the statistical fluctuation due to the removal of a randomly chosen N events. The



(a)

(b)

(c)

Figure 5.21: The fit result for the parameters $\mathcal{S}$ (a), $\mathcal{A}$ (b) and both (c) when one of the events is removed from the data sample. Figures (a) and (b) are projections of (c).

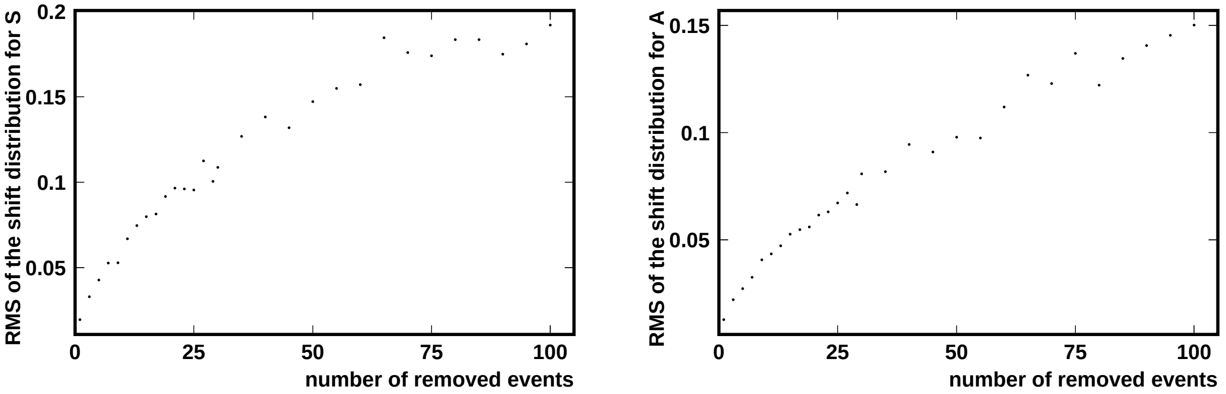
result is shown in Fig. 5.22.

5.4.5 The Time Integrated Fit

A time-integrated direct CP parameter is defined as

$$A = \frac{\Gamma(\overline{B}^0 \to D^+D^-) - \Gamma(B^0 \to D^+D^-)}{\Gamma(\overline{B}^0 \to D^+D^-) + \Gamma(B^0 \to D^+D^-)}, \tag{5.28}$$

where Γ is a time-integrated decay rate. After integrating Eq. 5.4 over the Δt range we obtain the relation between the time-integrated CP parameter A and the CP parameter



(a) parameter $\mathcal{S}$ (b) parameter $\mathcal{A}$

Figure 5.22: The RMS of the distribution of the difference in $\mathcal{S}$ (a) and $\mathcal{A}$ (b) when randomly chosen N events are removed from the sample of reconstructed events as a function of N.

$\mathcal{A}$ describing the Δt distribution as follows.

$$A = \frac{\int \frac{\mathrm{d}P_{\mathrm{sig}}}{\mathrm{d}\Delta t}(\Delta t, q=1)\mathrm{d}\Delta t - \int \frac{\mathrm{d}P_{\mathrm{sig}}}{\mathrm{d}\Delta t}(\Delta t, q=-1)\mathrm{d}\Delta t}{\int \frac{\mathrm{d}P_{\mathrm{sig}}}{\mathrm{d}\Delta t}(\Delta t, q=1)\mathrm{d}\Delta t + \int \frac{\mathrm{d}P_{\mathrm{sig}}}{\mathrm{d}\Delta t}(\Delta t, q=-1)\mathrm{d}\Delta t} \tag{5.29}$$

$$\begin{aligned} A &= \int \frac{e^{-|\Delta t|/\tau}}{2\tau}\left(-\Delta w + (1-2w)(\mathcal{S}\sin(\Delta m \Delta t) + \mathcal{A}\cos(\Delta m \Delta t)\right)\mathrm{d}\Delta t = \\ &= -\Delta w + (1-2w)\,0.626\,\mathcal{A} \end{aligned} \tag{5.30}$$

For $\mathcal{A} \neq 0$, the expected value of the time-integrated parameter A in any r-bin depends on the average wrong tag fraction in that r-bin.

For the time-integrated measurement, an unbinned ML fit to the M_{bc} and ΔE distribution of all reconstructed events is done separately for the events tagged as B^0 and $\overline{B}^0$. The corresponding distributions and fit projections are shown in Fig. 5.23. Table 5.15 shows the number of events in the signal peak and the parameter A calculated from Eq. 5.28. The error on parameter A is determined as

$$\sigma_A = \frac{2N_{B^0}N_{\overline{B}^0}}{(N_{B^0}+N_{\bar{B}^0})^2}\sqrt{\left(\frac{\sigma_{N_{B^0}}}{N_{B^0}}\right)^2 + \left(\frac{\sigma_{N_{\overline{B}^0}}}{N_{\overline{B}^0}}\right)^2}, \tag{5.31}$$

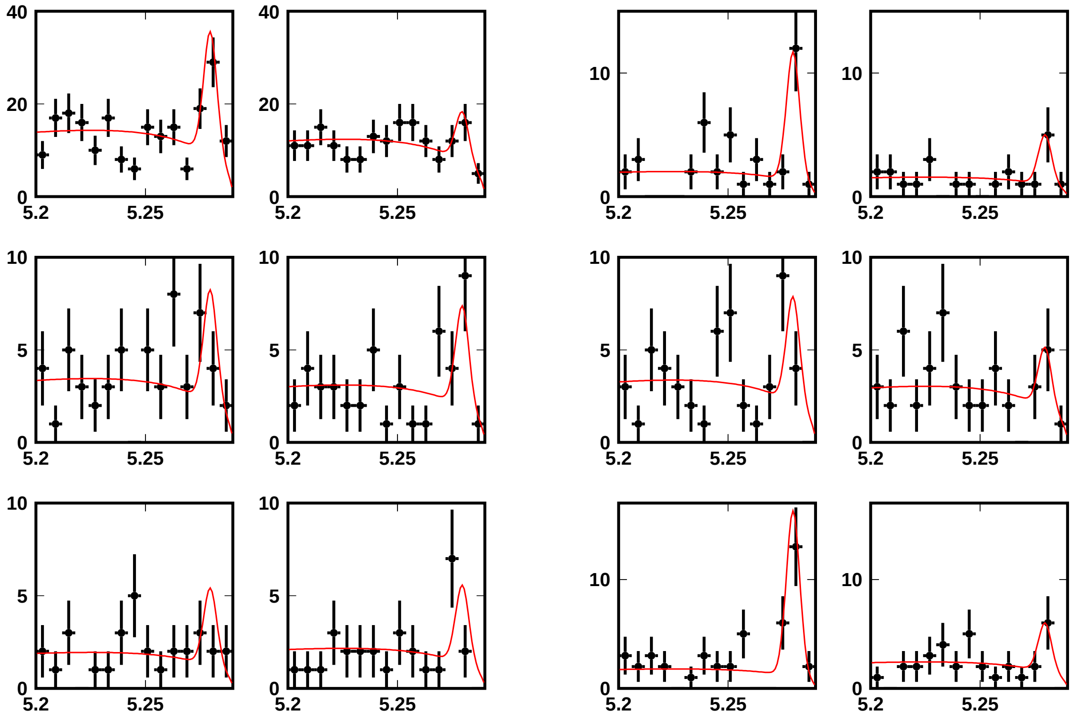
where N_{B^0} and $N_{\overline{B}^0}$ are the number of events tagged as B^0 or $\overline{B}^0$ and σ_{B^0} and $\sigma_{\overline{B}^0}$ the corresponding errors. To determine the parameter $\mathcal{A}$ from the combined information of all six r-bins, a χ^2 minimization fit is done. Here

$$\chi^2(\mathcal{A}) = \Sigma_i\left(\frac{A(\mathcal{A}, i) - A_i}{\sigma(i)}\right)^2, \tag{5.32}$$

where $A(\mathcal{A}, i)$ is the expected raw asymmetry in i^{th} r-bin for a given value of $\mathcal{A}$, calculated from Eq. 5.30, and A_i and σ_i are the measured values shown in Table 5.15. The fit result is

$$\mathcal{A} = 0.86 \pm 0.32\,, \tag{5.33}$$

corresponding to $2.7\,\sigma$.



$f_{\rm tag} = B^0$ $f_{\rm tag} = \overline{B}^0$ $f_{\rm tag} = B^0$ $f_{\rm tag} = \overline{B}^0$

r-bins 1-3 r-bins 4-6.

Figure 5.23: The $M_{\rm bc}$ distribution for events tagged as $B_{\rm tag} = B^0$ in the first and third column and $B_{\rm tag} = \overline{B}^0$ for events in the second and fourth column.

5.4.6 The D Meson Sidebands Study

The $D^{\pm}$ meson mass distributions provide another cross-check of the tagging quality. Since the B^0 meson predominantly decays as $B^0 \to D^- X$ ($\bar{b} \to \bar{c}$ transition), the events with $B_{\rm tag} = \overline{B}^0$ can be expected to have a peak in the D^- mass distribution and no (or small) peak in the D^+ mass distribution. A smaller peak in the D^+ mass distribution can be present due to the $B^0 - \overline{B}^0$ mixing and the best candidate selection. The situation is just the opposite for the events tagged as $B_{\rm tag} = B^0$. Since the background events are often

r-bin	$f_{\rm tag} = B^0$	$f_{tag} = \overline{B}^0$	A	$A(\mathcal{A} = 0.899)$
1	34.4 ± 9.6	13.8 ± 6.4	0.43 ± 0.22	0.04
2	7.8 ± 4.5	7.0 ± 4.3	0.06 ± 0.42	0.22
3	5.4 ± 3.6	5.4 ± 3.8	-0.00 ± 0.48	0.30
4	13.4 ± 5.1	5.2 ± 3.5	0.44 ± 0.31	0.38
5	7.4 ± 4.6	4.2 ± 3.7	0.28 ± 0.49	0.46
6	19.4 ± 6.1	5.8 ± 3.6	0.54 ± 0.25	0.54
3-6	46.5 ± 10.0	20.7 ± 7.3	0.38 ± 0.18	

Table 5.15: Number of events in the signal peak in the different r-bins and different flavors of the tag-side B meson. The parameter A is raw asymmetry calculated from Eq. 5.28. Last column $A(\mathcal{A} = 0.899)$ shows expected A if $\mathcal{A} = 0.899$ (Eq. 5.30).

reconstructed with one correctly reconstructed D meson and the other being a combination of different tracks, $B^0 \to D^- X$ decays are also a part of our background. The $D^\pm$ mass distribution of background events from the $M_{\rm bc}$ and ΔE sideband region can therefore be studied in order to test the tagging qualify of the best r-bin. The corresponding distributions are shown in Fig. 5.24. They are consistent with the above predictions and the levels of background for both flavors are also in a good agreement. We can therefore conclude that this cross-check also does not show any problem.

5.4.7 Systematic Error Study

The systematic errors come from the uncertainties in the different parameters that are used for the event likelihood function. To estimate the different contributions, the fit of the CP parameters is repeated with the corresponding parameters varied by $\pm\sigma_i$, where σ_i is the uncertainty in the parameter. The change in the fit result for the parameters $\mathcal{S}$ and $\mathcal{A}$ is quoted as the systematic error. To obtain the total systematic error, different contributions assumed to be uncorrelated and are added in quadrature, $\sigma^2 = \Sigma_i \sigma_i^2$.

Different contributions to the systematic error are summarized in Table 5.16. The CP fit bias is estimated by the toy MC study. Since it depends on the values of the parameters $\mathcal{S}$ and $\mathcal{A}$, a point inside the physical region of $\mathcal{S}^2 + \mathcal{A}^2 \leq 1$, which is the closest to the measured value, $\mathcal{S} = -0.8$ and $\mathcal{A} = 0.6$, is taken as a reference. The contributions due to the description of the $M_{\rm bc}$ and ΔE distributions (signal fraction) and the B_{CP} vertex resolution function parameters are determined by a toy MC study, as will be explained in the next section. The error due to the inaccuracy in the $B_{\rm tag}$ vertex resolution function parameters is obtained by changing the later by $\pm\sigma$. The error due to the vertex resolution function parameters shown in the table is obtained from the quadratic sum of the B_{CP} and $B_{\rm tag}$ vertex contributions. The effect of the background description is obtained by changing the parameters describing the Δt distribution for background by $\pm\sigma$. The effect of the vertex quality selection criteria is estimated by changing the criteria by a factor of two, i.e. changing the vertex quality requirement of $\xi = \chi^2/ndf < 100$ to $\xi < 50$ and $\xi < 200$ for

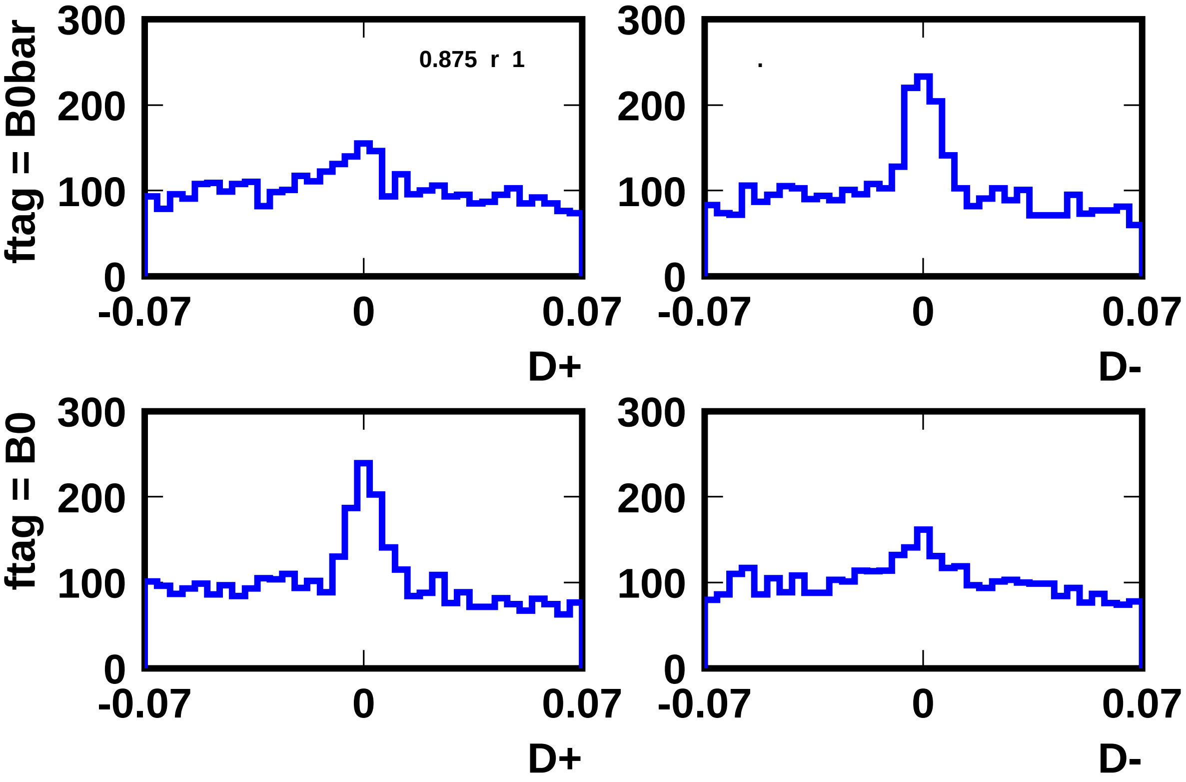


Figure 5.24: D-meson mass distributions for events from the M_{bc} and ΔE sideband. Only events in the best r-bin ($r > 0.875$) are shown.

both, B_{CP} and B_{tag} vertices. As a change in this requirement changes the number of events in the data sample, the change in the result could be only due to a statistical fluctuation. Indeed, the change in $\mathcal{S}$ and $\mathcal{A}$ due to the change in the ξ requirement is consistent with the statistical fluctuation when the same number of randomly chosen events is removed from the data sample, as shown in Fig. 5.22. In order not to overestimate this contribution, the systematic error due to the ξ criteria is therefore estimated by doing the same for the control sample, which is larger and therefore the effect of the statistical fluctuations is smaller. For the tag side interference, the systematic error is assumed to be the same as determined for other Belle analysis [34]. The wrong tag fractions w and Δw are varied by $\pm 1\,\sigma$ for each of the r-bins and the differences in the CP fit result are added quadratures. The effect of the inaccuracy in the B^0 meson lifetime τ and the mass difference Δm is again estimated by changing these parameters by $\pm 1\,\sigma$ [11].

When all contributions are summed up, the systematic errors on $\mathcal{S}$ and $\mathcal{A}$ are 0.09 and 0.06, respectively. It should be noted that the total systematic error is still small compared to the statistical error and the accuracy of our measurement is limited by the statistical fluctuations.

systematic error	$\sigma_{\mathcal{S}}$	$\sigma_{\mathcal{A}}$
CP fit bias	0.06	0.02
signal fraction	0.035	0.015
non-resonant background	0.02	0.03
vertex resolution function parameters	0.04	0.03
background distribution	0.01	0.01
vertex quality selection criteria	0.01	0.01
tag side interference	0.01	0.03
wrong tag fraction	0.017	0.014
values of τ and Δm	0.023	0.007
total	0.09	0.06

Table 5.16: A breakdown of the systematic errors on the parameters $\mathcal{S}$ and $\mathcal{A}$.

5.5 The Feldman – Cousins Statistical Analysis

The CP fit result for the parameter $\mathcal{A}$ indicates an unexpectedly large direct CPV. It is important to determine the confidence levels at which zero CP violation, $\mathcal{S} = \mathcal{A} = 0$, and zero direct CP violation, $\mathcal{A} = 0$, are excluded by our measurement. As the result is outside the physical region of $\mathcal{S}^2 + \mathcal{A}^2 \leq 1$, we have to make a statistical analysis rather than to rely on the errors determined by the CP fit. We determine the confidence level by the Feldman – Cousins (F-C) method [35], as described in this section.

5.5.1 The Confidence Level

The confidence level (CL) states our confidence that the actual value of a parameter being measured is within the quoted confidence interval (CI). If the measured value of the parameter is $\hat{x} \pm \sigma_x$, then the CL for $1\,\sigma$ CI is 68% or, in other words, we believe there is a 68% probability that the parameter is within the range $[\hat{x} - \sigma_x, \hat{x} + \sigma_x]$. We use the hat sign for the measured value ($\hat{x}$) and no sign (x) for the true unknown value of the parameter, to emphasize the difference. If we were rigorous, the value of parameter is a certain number and we therefore can not discuss a probability for the parameter to be within a certain range: either it is within the range or it is not. In this sense, the quote $\hat{x} \pm \sigma_x$ only states that whatever the value of the parameter x is, if we had repeated the measurement many times, 68% of the measured $\hat{x}$ would be within $[x - \sigma_x, x + \sigma_x]$ interval (and not vice versa as stated above).

If we assume a Gaussian distribution of the measurements x_i (which we generally do), there is a unique mapping between the $N\sigma$ CI and the CL, the later being the probability that the measurement falls within a given CI,

$$\mathrm{CL}(N\sigma_x) \quad = \quad \int_{x-N\sigma_x}^{x+N\sigma_x} G(\hat{x}|x, \sigma_x)\, \mathrm{d}\hat{x}\,, \tag{5.34}$$

where $G(\hat{x}|x,\sigma_x)$ denotes a Gaussian distribution of mean x and width σ_x. For that reason, we often quote CL simply in the units of σ, where the relation between the $N\sigma$ CL and the actual CL (probability) is understood, as shown in Table 5.17. In this work, the CL in the units of σ always refers to the CL shown in Table 5.17.

CL [σ]	CL [%]
1	68.3
2	95.4
3	$1-2.7\cdot 10^{-3}$
4	$1-6.3\cdot 10^{-5}$
5	$1-5.7\cdot 10^{-7}$

Table 5.17: Probability coverage (CL [%]) for the $N\sigma$ CL.

In a similar way we can calculate the CL for a simultaneous measurement of two parameters, like CP parameters $\mathcal{S}$ and $\mathcal{A}$.

$$\mathrm{CL}(N\sigma_x, N\sigma_y) \;=\; \int_{x-N\sigma_x}^{x+N\sigma_x}\int_{y-N\sigma_y}^{y+N\sigma_y} G(\hat{x}|x,\sigma_x)\, G(\hat{y}|y,\sigma_y)\,\mathrm{d}\hat{x}\mathrm{d}\hat{y}\,. \tag{5.35}$$

It should be noted that for the two-dimensional distribution, $\mathrm{CL}(1\sigma_x, 1\sigma_y) = 47\%$ (or 39%, if we integrate inside the ellipse defined by $(\hat{x}/\sigma_x)^2 + (\hat{y}/\sigma_y)^2 = 1$), which is smaller than the corresponding integral in one dimension shown by Eq. 5.34.

5.5.2 The Feldman – Cousins Method

The CI and CL can be defined in a more general way than the one given by Eq. 5.34. Instead of quoting $[x-\sigma_x, x+\sigma_x]$ CI, we can define a general $N\sigma$ CI as an interval with the corresponding coverage (CL). The interval can be chosen to be symmetric or not, and can turn out to be unbiased or biased. Biased means that average value of estimates is not equal to the value of the parameter, $\langle\hat{x}\rangle \neq x$. The "standard" choice is to use the interval in $\hat{x}$ that contains the values of $\hat{x}$ for whose the values of PDF are largest,

$$\mathcal{P}(\hat{x}|x)_{\hat{x}\in\mathrm{CI}} \;>\; \mathcal{P}(\hat{x}|x)_{\hat{x}\,\mathrm{outside\,CI}}\,, \tag{5.36}$$

where $\hat{x}$ can be a vector of parameters and $\mathcal{P}(\hat{x}|x)$ is a probability density function (PDF) to measure a value of $\hat{x}$ given the value of the parameter is x. In the simplest case of a Gaussian distribution, $\mathcal{P}(\hat{x}|x) = G(\hat{x}|x,\sigma_x)$, but the CI definition shown by Eq. 5.36 can be used in general for any PDF:

$$\mathrm{CL} \;=\; \int_{\mathrm{CI}} \mathcal{P}(\hat{x}|x)\,\mathrm{d}\hat{x}\,. \tag{5.37}$$

While the condition for the unbiased parameter estimate might seem very natural, it should be noted that as soon as the parameter errors are not symmetric, the standard choice already gives a biased estimate, as will be shown later.

Equation 5.37 defines also the CL determined by the F-C method. The only difference from the standard method is a different choice of the CI. The F-C CI is determined as

$$\left.\frac{\mathcal{P}(\hat{x}|x)}{\mathcal{P}(\hat{x}|x_{\text{best}})}\right|_{\hat{x}\in\text{CI}} > \left.\frac{\mathcal{P}(\hat{x}|x)}{\mathcal{P}(\hat{x}|x_{\text{best}})}\right|_{\hat{x}\text{ outside CI}}, \tag{5.38}$$

where x_{best} is the value of the parameter for which the $\mathcal{P}(\hat{x}|x)$ is maximal. In other words, $\mathcal{P}(\hat{x}|x_{\text{best}})$ is the maximal value of $\mathcal{P}(\hat{x}|x)$ for any value of x. In the F-C approach, only the ordering of the values $\hat{x}$ that are included to the CI (integrating region) is therefore different. The above criteria in a way re-normalizes the PDF. The different choice makes the biggest difference when the measured value is close to or even outside the physical region. As x_{best} is the value of the parameter by definition, it has to be within the physically allowed region ($\mathcal{S}^2 + \mathcal{A}^2 \leq 1$ for the CP parameters measurement). In such a case, the F-C condition moves the CI inside the physical region, as will be illustrated later.

To determine the CL by the F-C method, we only need to know $\mathcal{P}(\hat{x}|x)$ for any value of x; or, specifically for our measurement, $\mathcal{P}(\hat{\mathcal{S}},\hat{\mathcal{A}}|\mathcal{S},\mathcal{A})$ for any value of $(\mathcal{S},\mathcal{A})$. In the following paragraphs we first describe the parameterization of the PDF $\mathcal{P}(\hat{\mathcal{S}},\hat{\mathcal{A}}|\mathcal{S},\mathcal{A})$. Then we determine the parameters describing the PDF for any values of $\mathcal{S}$ and $\mathcal{A}$ using a toy MC study. In this toy MC study, beside the statistical fluctuations we simulate also the systematic uncertainty. Once the PDF is obtained, we continue with the F-C related issues. We describe the "best" parameters and their calculation and compare the F-C and standard CI in more detail. At the end, we show a result of the Feldman – Cousins significance study, i.e. the CL at which the measured values agree with any values of $\mathcal{S}$ and $\mathcal{A}$.

5.5.3 The Probability Density Function

We describe the probability density function, $\mathcal{P}(\hat{\mathcal{S}},\hat{\mathcal{A}}|\mathcal{S},\mathcal{A})$, for a given $\mathcal{S}$ and $\mathcal{A}$ using a bifurcated Gaussian function $\tilde{G}$, where

$$\tilde{G}(x|\mu,\sigma^p,\sigma^n) = \begin{cases} \sqrt{\dfrac{2}{\pi}}\dfrac{1}{\sigma^n+\sigma^p}\exp\left\{-\dfrac{(x-\mu)^2}{2\sigma^{n\,2}}\right\}, & \text{for } x \leq \mu \\ \sqrt{\dfrac{2}{\pi}}\dfrac{1}{\sigma^n+\sigma^p}\exp\left\{-\dfrac{(x-\mu)^2}{2\sigma^{p\,2}}\right\}, & \text{for } x > \mu. \end{cases} \tag{5.39}$$

The normalization of both functions is chosen to be such that $\tilde{G}(x|\mu,\sigma^p,\sigma^n)_{x\downarrow 0} = \tilde{G}(x|\mu,\sigma^p,\sigma^n)_{x\uparrow 0}$. For the independent parameters $\mathcal{S}$ and $\mathcal{A}$ we have

$$\mathcal{P}(\hat{\mathcal{S}},\hat{\mathcal{A}}|\mathcal{S},\mathcal{A}) = \tilde{G}(\hat{\mathcal{S}}|\mu_{\mathcal{S}},\sigma^p_{\mathcal{S}},\sigma^n_{\mathcal{S}})\,\tilde{G}(\hat{\mathcal{A}}|\mu_{\mathcal{A}},\sigma^p_{\mathcal{A}},\sigma^n_{\mathcal{A}}), \tag{5.40}$$

where the parameters of the distribution $\mu_{\mathcal{S}}$, $\sigma^p_{\mathcal{S}}$, $\sigma^n_{\mathcal{S}}$, $\mu_{\mathcal{A}}$ $\sigma^p_{\mathcal{A}}$ and $\sigma^n_{\mathcal{A}}$ depend on the values of $\mathcal{S}$ and $\mathcal{A}$. We neglect a small correlation between the parameters $\mathcal{S}$ and $\mathcal{A}$. We can take the correlation between $\mathcal{S}$ and $\mathcal{A}$ into account by rotating the $(\mathcal{S},\mathcal{A})$ plane by an angle θ, $(\mathcal{S},\mathcal{A}) \to (\mathcal{S}',\mathcal{A}')$, such that

$$\begin{pmatrix} \hat{\mathcal{A}}' - \mu_A \\ \hat{\mathcal{S}}' - \mu_S \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \hat{\mathcal{A}} - \mu_A \\ \hat{\mathcal{S}} - \mu_S \end{pmatrix}. \tag{5.41}$$

The rotation was included to the above PDF and the parameter θ was determined by a toy MC study, which confirmed it is negligible, $\theta \approx 1^\circ$.

It should be noted that the mean value of the distribution $\tilde{\mu}$ described by Eq. 5.39 is different form the parameter μ,

$$\tilde{\mu} \;=\; \mu + \sqrt{\frac{2}{\pi}}\left(\sigma^p - \sigma^n\right). \tag{5.42}$$

For this reason, when we use the standard method to determine the estimate of the parameter $\hat{x}$ (the value of x at which $\mathcal{P}(x)$ is maximal), $\hat{x} = \mu$ and the fit result is biased, $\langle \hat{x} \rangle \neq \tilde{\mu}$, as mentioned earlier. In the case of F-C study, this effect does not introduce a bias since the asymmetry of the distribution is described by the PDF and is properly taken into account.

Parameters of the PDF

The product of bifurcated Gaussians (defined by Eq. 5.39) is described by the parameters μ_A, $\sigma^n_{\mathcal{A}}$, $\sigma^p_{\mathcal{A}}$, μ_S, $\sigma^p_{\mathcal{S}}$ and $\sigma^n_{\mathcal{S}}$. Using $10,000$ toy MC experiments, generated with specific values of $\mathcal{S}$ and $\mathcal{A}$, we get distributions of estimates $\hat{\mathcal{S}}(\mathcal{S},\mathcal{A})$ and $\hat{\mathcal{A}}(\mathcal{S},\mathcal{A})$. Each of these distributions is fitted to $\mathcal{P}(\hat{\mathcal{S}},\hat{\mathcal{A}}|\mathcal{S},\mathcal{A})$, given by Eq. 5.40. From the fits, values of parameters μ_A, $\sigma^n_{\mathcal{A}}$,... are obtained for each of the simulated pairs of values $(\mathcal{S},\mathcal{A})$.

The fit of the parameters is first tested by simulating the $(\hat{\mathcal{S}},\hat{\mathcal{A}})$ distribution and fitting the parameters from the simulated distribution. The first thing that is noticed is that the values of the parameters σ^p and σ^n are anti-correlated. Since it is easier to work with an uncorrelated parameters,

$$\sigma \;=\; \frac{\sigma^p + \sigma^n}{2} \text{ and} \tag{5.43}$$

$$\Delta\sigma \;=\; \sigma^p - \sigma^n \tag{5.44}$$

are used instead of σ^p and σ^n. As the later is strongly correlated with μ according to the Eq. 5.42 (because the mean of the distribution $\tilde{\mu}$ is the one parameter that is well defined), $\tilde{\mu}$ is used instead of μ. The new parameterization is then based on the parameters $\tilde{\mu}_S$, σ_S, $\Delta\sigma_S$, $\tilde{\mu}_A$, σ_A, and $\Delta\sigma_A$. The covariance matrix for this selection of parameters is close to the unit matrix with the off-diagonal terms of the order of a few percent.[1] The errors on the parameters obtained from the fit to the $10,000$ events with $\hat{\mathcal{S}}$ and $\hat{\mathcal{A}}$ distributed according to the bifurcated Gaussian distribution with $\sigma^n_{\mathcal{S}} = \sigma^p_{\mathcal{S}} = 0.3425$, $\sigma^n_{\mathcal{A}} = \sigma^p_{\mathcal{A}} = 0.2350$ and $\theta = \mu_S = \mu_A = 0$ are shown in Table 5.18.

The Dependence of the Parameters on $\mathcal{S}$ and $\mathcal{A}$

To determine the dependence of the parameters $\tilde{\mu}_S$, σ_S, $\Delta\sigma_S$, $\tilde{\mu}_A$, σ_A, and $\Delta\sigma_A$ on the values of $\mathcal{S}$ and $\mathcal{A}$, we make a toy MC study as described in Sec. 5.3.1 for different values

---

[1] The $\sigma_S\Delta\sigma_S$ correlation coefficient is sometimes as large as 15%.

	$\tilde{\mu}$	σ	$\Delta\sigma$
S	0.0034 ± 0.0002	0.00242 ± 0.00002	0.0102 ± 0.0002
A	0.0024 ± 0.0001	0.00166 ± 0.00001	0.0070 ± 0.0001

Table 5.18: The errors on the parameters obtained from the ML fit to the $10,000$ events with $\mathcal{S}$ and $\mathcal{A}$ distributed according to the bifurcated Gaussian distribution with parameters $\sigma^n_{\mathcal{S}} = \sigma^p_{\mathcal{S}} = 0.3425$, $\sigma^n_{\mathcal{A}} = \sigma^p_{\mathcal{A}} = 0.2350$ and $\theta = \mu_S = \mu_A = 0$.

of $\mathcal{S}$ and $\mathcal{A}$ to simulate the CP fit result distribution of $\hat{\mathcal{S}}$ and $\hat{\mathcal{A}}$. A 10,000 toy MC samples were generated and fitted for each of the points $\mathcal{S} \in \{-0.9, 0.8, \ldots, 0.9\}$ and $\mathcal{A} \in \{-0.9, 0.8, \ldots, 0.9\}$ satisfying the condition $\mathcal{S}^2 + \mathcal{A}^2 < 1$. An un-binned ML fit was done for each of the simulated values of $\mathcal{S}$ and $\mathcal{A}$ to determine the values of the parameters for given $\mathcal{S}$ and $\mathcal{A}$. The values of the parameters are shown in Fig. 5.25. Clearly, some of the parameters have a strong dependence on $\mathcal{S}$ and $\mathcal{A}$. The errors on the parameters obtained by this fit are consistent with the estimate shown in Table 5.18.

At the next step, the parameters dependence on $\mathcal{S}$ and $\mathcal{A}$ is described with a 4th order polynomial function.

$$\begin{aligned} f(\mathcal{S},\mathcal{A}) &= c_0 + c_1\,\mathcal{S} + c_2\,\mathcal{A} + c_3\,\mathcal{S}^2 + c_4\,\mathcal{A}^2 + c_5\,\mathcal{S}\,\mathcal{A} + \\ &+ c_6\,\mathcal{S}^3 + c_7\,\mathcal{A}^3 + c_8\,\mathcal{S}^2\,\mathcal{A} + c_9\,\mathcal{S}\,\mathcal{A}^2 + \\ &+ c_{10}\,\mathcal{S}^4 + c_{11}\,\mathcal{A}^4 + c_{12}\,\mathcal{S}^2\,\mathcal{A}^2 + c_{13}\,\mathcal{S}^3\,\mathcal{A} + c_{14}\,\mathcal{S}\,\mathcal{A}^3\,, \end{aligned} \tag{5.45}$$

where f stands for any of the parameters μ_A, $\sigma^n_{\mathcal{A}}$, $\sigma^p_{\mathcal{A}}$, μ_S, $\sigma^p_{\mathcal{S}}$ or $\sigma^n_{\mathcal{S}}$. The coefficients c_i for each of the parameters f are determined by a χ^2 fit,

$$\chi^2 \;=\; \Sigma_j\, \frac{\left(f(\mathcal{S}_j,\mathcal{A}_j) - f^{\mathrm{toy\,MC}}(\mathcal{S}_j,\mathcal{A}_j)\right)^2}{\sigma^{f\,2}_j}\,, \tag{5.46}$$

where χ^2 is summed up for all simulated toy MC samples at different values of $\mathcal{S}$ and $\mathcal{A}$. To avoid using too many unnecessary terms in the polynomial function described by the Eq. 5.46, the fit was first done for all 15 coefficient. The result for all coefficients was evaluated and only the terms with the corresponding coefficients significantly different from zero, $|c_i| > 2\,\sigma_{c_i}$, were used to describe $f(\mathcal{S},\mathcal{A})$ dependence for that parameter. A less tight condition $|c_i| > 1\,\sigma_{c_i}$ gives similar results. The result is shown in Table 5.19. For all parameters f, the removal of those coefficients which have $|c_i| < 2\,\sigma_{c_i}$ does not increase the fit χ^2/ndf and does not cause any shifts in the remaining parameters, therefore they can safely be removed.

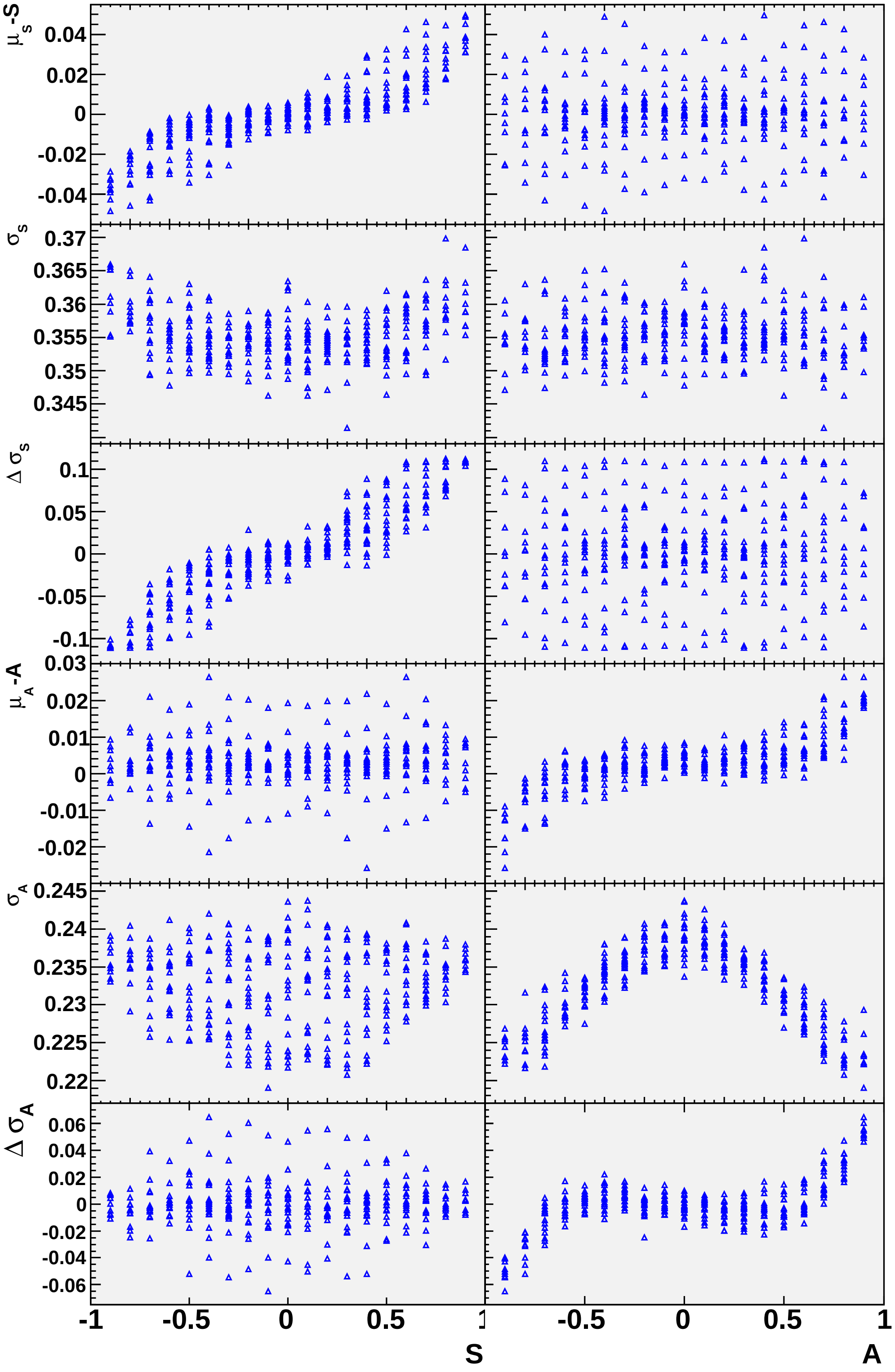


Figure 5.25: Parameters describing $\mathcal{P}(\hat{\mathcal{S}}, \hat{\mathcal{A}}|\mathcal{S}, \mathcal{A})$ as a function of the simulated $\mathcal{S}$ and $\mathcal{A}$.

$[10^{-3}]$	c_0	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}	c_{14}	χ^2/ndf
$\tilde{\mu}_S$	1	989	−0	−3	−10	1	58	1	4	79	3	11	10	−5	−0	1.2
	±1	±1	±1	±4	±4	±4	±2	±2	±2	±2	±4	±4	±7	±6	±6	
	1	989			−9		58		5	79		10	6			1.2
	±0	±1			±3		±2		±2	±2		±4	±4			
σ_S	357	0	−1	−18	−19	1	−1	0	2	−1	25	16	60	−2	−2	1.4
	±0	±1	±1	±3	±3	±3	±1	±1	±2	±2	±3	±3	±5	±4	±4	
	357			−18	−19				1		25	16	60			1.4
	±0			±3	±3				±1		±3	±3	±5			
$\Delta\sigma_S$	2	30	3	−5	−2	9	93	−1	3	154	−4	−2	15	−12	−16	2.1
	±2	±3	±3	±8	±10	±9	±3	±5	±4	±4	±7	±11	±12	±10	±12	
	−1	31					91			154			−1	−1	−3	2.3
	±1	±3					±3			±4			±5	±2	±5	
$\tilde{\mu}_A$	3	−0	990	3	−1	−2	−0	32	31	−4	−4	1	−3	2	1	1.1
	±0	±1	±1	±2	±2	±3	±1	±1	±2	±2	±3	±3	±4	±4	±4	
	3		990	2				32	31	−4	−3					1.1
	±0		±1	±2				±1	±2	±1	±3					
σ_A	240	0	−0	−7	−45	−0	−0	−1	−2	−0	4	28	31	4	−3	1.2
	±0	±1	±1	±2	±2	±2	±1	±1	±1	±1	±2	±2	±3	±3	±3	
	240			−7	−45				−3		4	28	31	1		1.2
	±0			±2	±2				±1		±2	±2	±3	±1		
$\Delta\sigma_A$	−1	−0	−50	0	−3	9	3	126	56	−13	−2	6	26	−16	−16	1.1
	±1	±2	±2	±7	±7	±7	±4	±3	±4	±4	±8	±8	±12	±11	±10	
			−50			8		126	56	−10			13	−15	−16	1.1
			±2			±7		±3	±4	±3			±5	±11	±10	

Table 5.19: The result of the χ^2 fit for the polynomial coefficients from Eq. 5.46 (in units of 10^{-3}), statistical error only. Empty places indicate the coefficients set to zero (the polynomial terms that are not used).

Cross-check of the Toy MC Errors

Since we rely on a toy MC to describe the CP fit result distribution, it is important to confirm that the toy MC really describes our data well. For that purpose, we compare the errors determined by the toy MC simulation to the width of the distribution of the CP-blind fit for the data sample; they are found to be about $5 - 10\,\%$ different. The impact of this difference on the significance of the result was estimated by doing the CL study after re-scaling the widths obtained by the toy MC with the corresponding factors. The result for the CL is about (or less than) 10% higher when compared to the CL that is obtained without re-scaling.

The systematic error

The toy MC studies done so far simulate only the statistical error of the CP parameters being fitted. Since the fit of the CP parameters is done using the identical $M_{\rm bc}$ and ΔE distributions, resolution function parameters,… as are used for the toy MC simulation, the systematic error due to the uncertainty in these parameters is not simulated this way. One of the possible ways to include the systematic error is to make a convolution of the PDF describing the statistical error and another Gaussian function describing the systematic errors. However, due to the asymmetric PDF, this process is demanding on the technical side (and consumes more computer processor time). In addition, the systematic errors can also depend on $\mathcal{S}$ and $\mathcal{A}$. For these reasons, the systematic error simulation is included to the toy MC study by the following approach.

(a) Generate a toy MC sample with the default parameters (describing the $M_{\rm bc}$ and ΔE distributions, resolution function parameters,…).

(b) Fit the CP parameters for the toy MC sample with the same default parameters.

(c) Smear the CP-side resolution parameters used for the fit (parameters are smeared independently according to their covariance matrix, see Appendix A.3) and repeat the fit for the same toy MC sample.

(d) Reset the CP-side resolution parameters to the default values, smear the parameters describing the signal fraction and repeat the fit of the CP parameters.

(e) Set the CP-side resolution parameters and the parameters describing the signal fraction to the values used for (c) and (d), respectively. Repeat the fit.

Steps (a) – (e) are done for each of the toy MC samples. The smearing in (c) and (d) is simulated according to the same error distribution function (covariance matrix) for all samples, but the smearing for each particular sample is random and therefore different from sample to sample. The meaning of the RMS of the distributions $X_{(i)}$ and $X_{(i)} - X_{(\rm b)}$, where X stands for $\mathcal{S}$ or $\mathcal{A}$, is explained in Table 5.20. Only the largest systematic errors are simulated separately as described here. The other contributions are found to be negligible.

$\mathrm{RMS}(X_{(\mathrm{b})})$	The statistical error.
$\mathrm{RMS}(X_{(\mathrm{c})} - X_{(\mathrm{b})})$	The systematic error due to the CP-side resolution parameters.
$\mathrm{RMS}(X_{(\mathrm{d})} - X_{(\mathrm{b})})$	The systematic error due to the signal fraction.
$\mathrm{RMS}(X_{(\mathrm{e})} - X_{(\mathrm{b})})$	Should be consistent with $\mathrm{RMS}(\mathrm{c}-\mathrm{b}) \oplus \mathrm{RMS}(\mathrm{d}-\mathrm{b})$.
$\mathrm{RMS}(X_{(\mathrm{e})})$	The statistical and systematical error, $\mathrm{RMS}(\mathrm{b}) \oplus \mathrm{RMS}(\mathrm{e}-\mathrm{b})$.

Table 5.20: Overview of the meaning of the RMS of the distributions obtained in the steps (b) to (e). The last two serve as a consistency check.

This approach considerably simplifies the PDF description and as a consequence also the CL calculation.

Using this approach, the systematic effects can be studied separately from the statistical fluctuations, since the later cancel out on a sample-by-sample basis. Figures 5.26 and 5.27 show the parameters describing the distributions of $\hat{\mathcal{S}}$ and $\hat{\mathcal{A}}$ due to the different systematic errors only. Their $\mathcal{S}$ and $\mathcal{A}$ dependence is determined in the same way as was done for the statistical error. The systematic errors on $\mathcal{S}$ and $\mathcal{A}$ are the widths of the corresponding distributions, $\sigma_{\mathcal{S}}$ and $\sigma_{\mathcal{A}}$. The dominant dependence of both systematic errors on $\mathcal{S}$ and $\mathcal{A}$ is as follows.

$$\sigma_{\mathcal{S}} = 0.031 + 0.013\mathcal{S}^2 + 0.008\mathcal{S}^4 + 0.008\mathcal{A}^4 + 0.021\mathcal{S}^2\mathcal{A}^2 \tag{5.47}$$

$$\sigma_{\mathcal{A}} = 0.021 + 0.002\mathcal{A}^2 + 0.002\mathcal{S}^4 + 0.009\mathcal{A}^4 + 0.010\mathcal{S}^2\mathcal{A}^2 \tag{5.48}$$

The potential bias due to systematic uncertainties, described by $\tilde{\mu}_{\mathcal{S}}$ and $\tilde{\mu}_{\mathcal{A}}$, is found to be negligible.

$$\tilde{\mu}_{\mathcal{S}} = 0.003\mathcal{S} + 0.001\mathcal{S}\mathcal{A}^2 \tag{5.49}$$

$$\tilde{\mu}_{\mathcal{A}} = 0.002\mathcal{A} \tag{5.50}$$

The cross-check (d) gives consistent results. At the last step, the statistical and systematic error distributions are fitted simultaneously. The comparison of parameters describing statistical error only, systematic error only and both is shown in Table 5.21. We see that the systematic error is negligible compared to the statistical error.

	c_0	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	χ^2/ndf
$\tilde{\mu}_S$		991					58			79				1.2
		±1					±2			±2				
σ_S	359			−17	−19						25	16	61	1.3
	±0			±3	±3						±3	±3	±5	
$\Delta\sigma_S$		47		−3			76			142				2.5
		±3		±0			±3			±4				
$\tilde{\mu}_A$	3		989					34	32					1.2
	±0		±1					±1	±2					
σ_A	241			−4	−44							29	29	1.2
	±0			±1	±2							±2	±3	
$\Delta\sigma_A$	−1		−51		4	−4		130	61	−6				1.1
	±1		±2		±2	±2		±3	±4	±3				

Table 5.21: Fit result for the polynomial coefficients that describe the statistical and the systematical errors. Empty spaces indicate the coefficients that are set to zero.

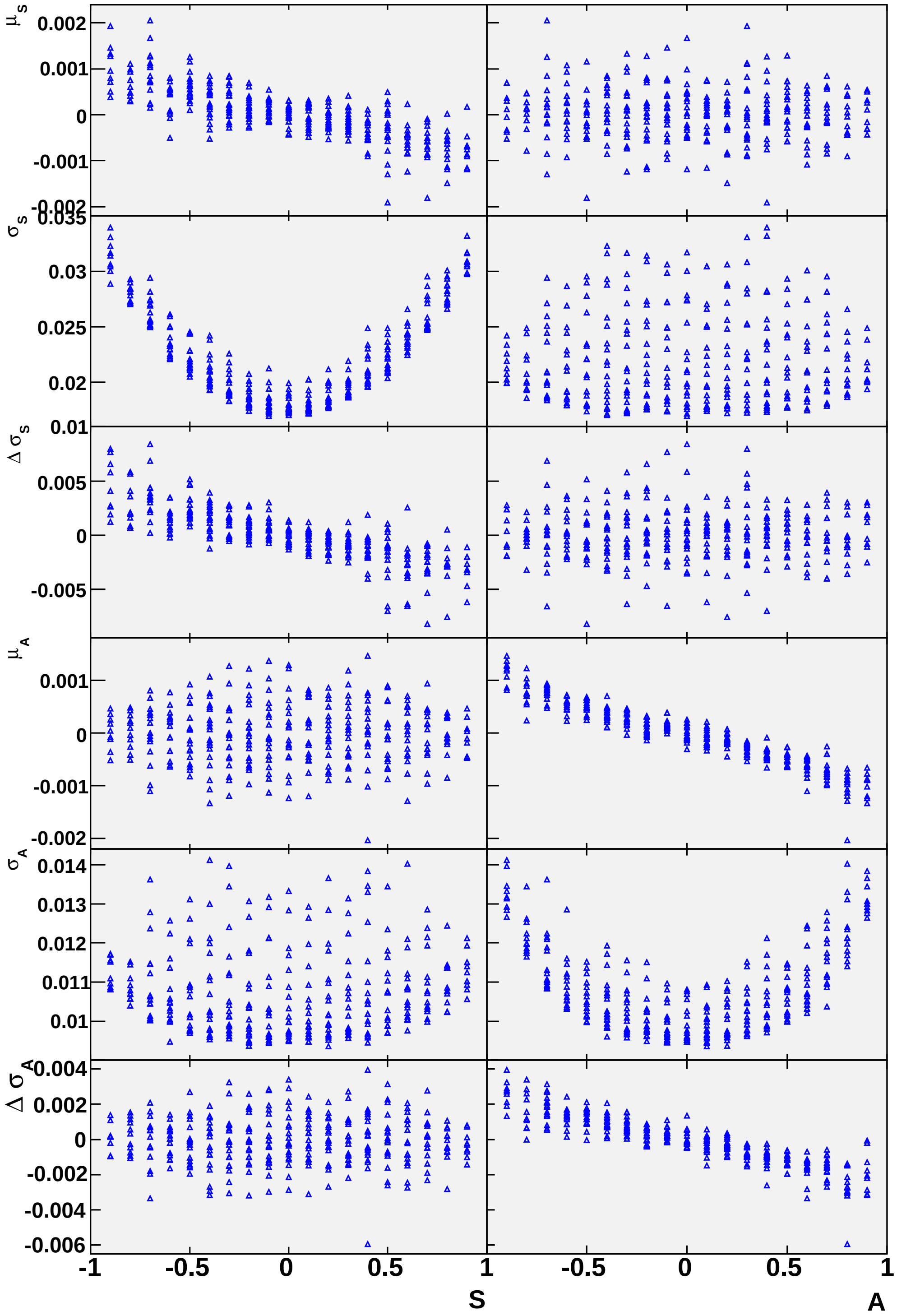


Figure 5.26: Parameters of $\mathcal{P}(\hat{\mathcal{S}}, \hat{\mathcal{A}}|\mathcal{S}, \mathcal{A})$ describing the systematic error due to the CP-side resolution parameters (RMS($X_{(b)} - X_{(a)}$)) as a function of the simulated $\mathcal{S}$ and $\mathcal{A}$.

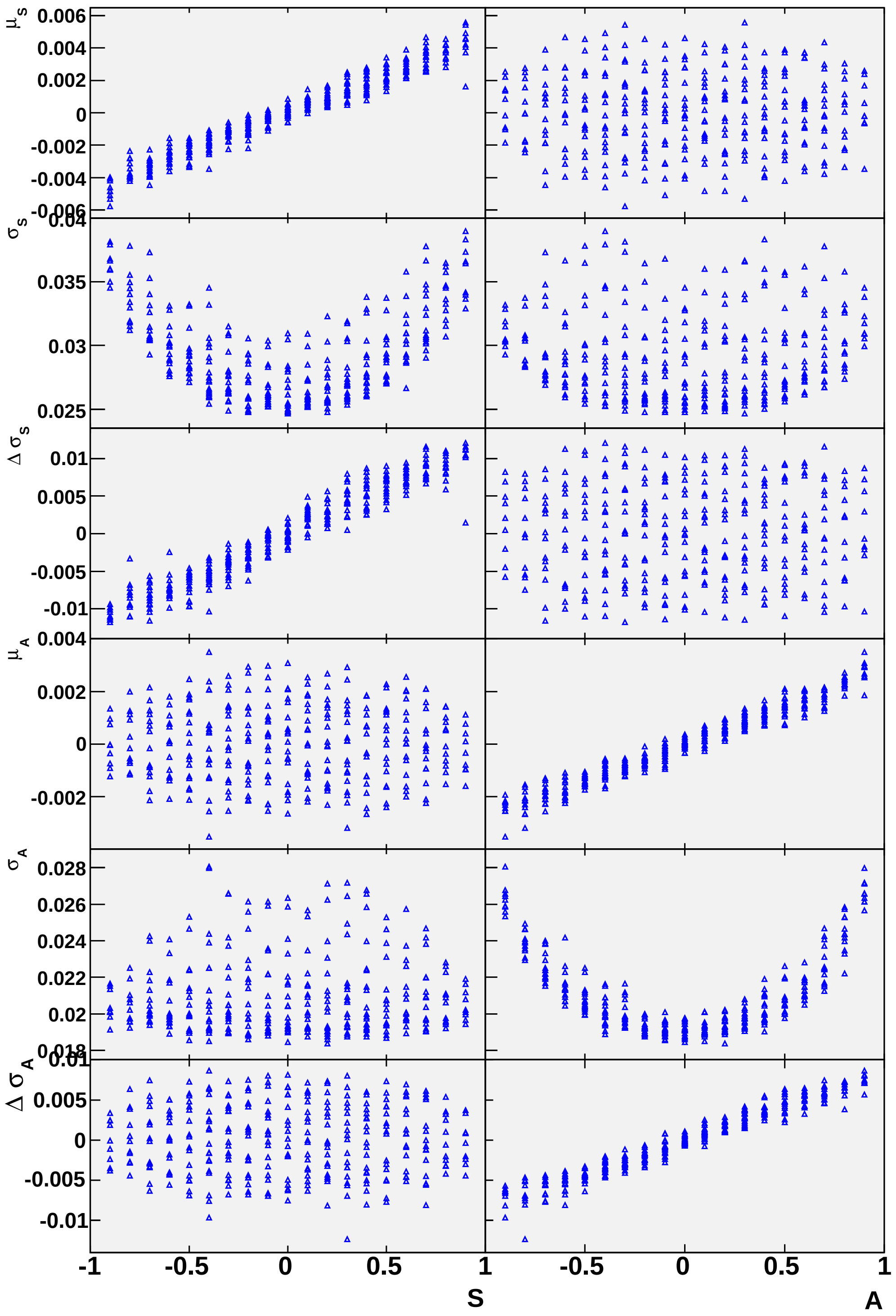


Figure 5.27: Parameters of $\mathcal{P}(\hat{\mathcal{S}}, \hat{\mathcal{A}}|\mathcal{S}, \mathcal{A})$ describing the systematic error due to the uncertainty in the signal fraction ($\mathrm{RMS}(X_{(c)} - X_{(a)})$) as a function of the simulated $\mathcal{S}$ and $\mathcal{A}$.

$\mathcal{P}(\hat{\mathbf{x}}|\mathbf{x}_{\rm best})$

In the Feldman – Cousins approach, a different ordering principle is used, as shown by Eq. 5.38. In the F-C ordering, the "standard" $\mathcal{P}(\hat{x}|x)$ is divided by $\mathcal{P}(\hat{x}|x_{\rm best})$, where the lated is the largest value of $\mathcal{P}$ for any x. Let us illustrate what that means. We will start by a one-dimensional Gaussian distribution,

$$\mathcal{P}(\hat{x}|x) \;=\; G(\hat{x}|\mu(x),\sigma(x)) \;=\; \frac{1}{\sqrt{2\pi}\sigma}\, e^{-(\hat{x}-\mu)^2/(2\sigma^2)}\,, \tag{5.51}$$

where $\mu(x) = x$ and $\sigma(x) = \sigma =$ const. does not depend on the parameter x. Of course, as we are all familiar with this standard case, we could simply write $G(\mu = x, \sigma)$ in the above equation, but instead we write it in the same way as is done later for a more complicated case, just to make a parallel. In this trivial case, the value of x for which $G(\hat{x}|\mu(x),\sigma(x))$ is maximal, is $x_{\rm best} = \hat{x}$ and the maximal value of PDF is $\mathcal{P}(\hat{x}|x_{\rm best}) = G(\hat{x}|x = \hat{x}, \sigma) = 1/(\sqrt{2\pi}\,\sigma)$. The maximal value of $\mathcal{P}(\hat{x}|x_{\rm best})$ is the same for any value of $\hat{x}$, therefore the denominator in the F-C condition (Eq. 5.38) cancels out and the condition is identical to the "standard" case. When the measurement error (width σ) depends on the value of the parameter, this is no longer the case. Since the logarithm is a strictly monotonic function, we can find the maximum of Eq. 5.51 by calculating the value of x for which its derivative is equal to zero,

$$\left.\frac{{\rm d}(\log\mathcal{P}(\hat{x}|x))}{{\rm d}x}\right|_{x=x_{\rm best}} = 0 \tag{5.52}$$

$$-\sigma' - y + \sigma' y^2 = 0\,, \tag{5.53}$$

where $\sigma' = {\rm d}\sigma(x)/{\rm d}x\,|_{x=x_{\rm best}}$ and $y = (\hat{x} - x)/\sigma(x)\,|_{x=x_{\rm best}}$. For $\sigma' \ll 1$ we calculate $y = -\sigma' \neq 0$, which is different from the "standard" case and simplifies to $x_{\rm best} = \hat{x}$ when $\sigma' = 0$. The case when $x_{\rm best} \neq \hat{x}$ happens also in the case of $\mathcal{P}(\hat{\mathcal{S}}, \hat{\mathcal{A}}|\mathcal{S}, \mathcal{A})$, as illustrated in Fig. 5.28. Figure 5.29 shows the difference $x_i^{\rm best} - x_i$ in the case of our F-C study for the CP parameters $\mathcal{S}$ and $\mathcal{A}$. The overall shift in Fig. 5.29 (b) is mainly due to the non-zero values of the parameters $c_0(\tilde{\mu}_A) \neq 0$ and $c_0(\Delta\sigma_A) \neq 0$.

The difference between the F-C and standard method is more pronounced when $\hat{x}$ is outside the physical region. Since the x is a physical parameter, its domain is limited to the physically allowed region. For the Gaussian example, if x is limited to be positive, $x > 0$, then $\mathcal{P}(\hat{x}|x_{\rm best}) = \mathcal{P}(\hat{x}|x_{\rm best} = 0)$ for $\hat{x} < 0$ (if $\sigma = \sigma(x)$, then $x_{\rm best} \approx 0$). When $\hat{x}$ is close to physical region, the two approaches therefore give a significantly different ordering criteria.

Integrating region

The CL calculated by the F-C method is different from the "standard" CL only because a different ordering principle (Eq. 5.38) gives a different integration area, as illustrated in Fig. 5.30.

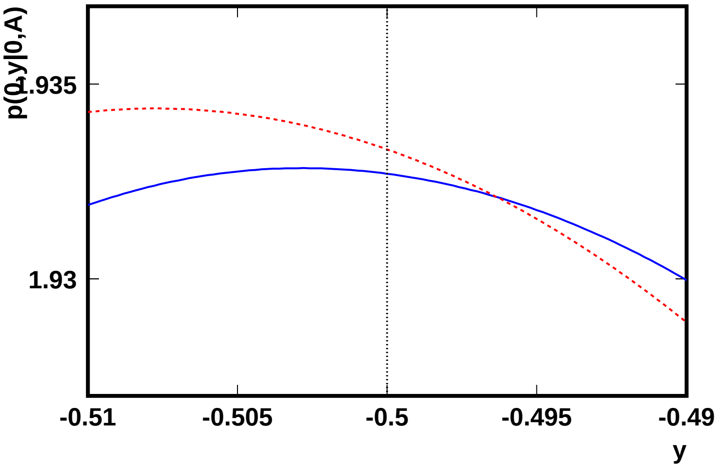
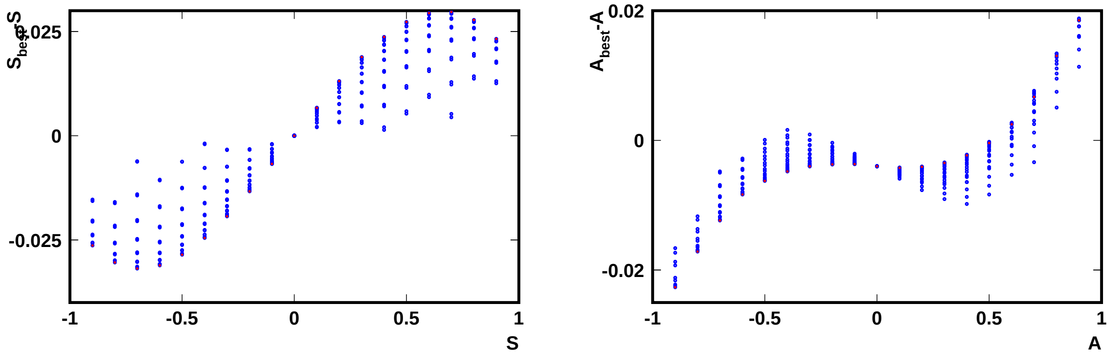


Figure 5.28: $p(0,\hat{\mathcal{A}}|0,\mathcal{A})$ for $\mathcal{A}=-0.500$ (full blue line) and $\mathcal{A}=-0.505$ (dashed red line). It can be observed that, for example, $p(0,\hat{\mathcal{A}}|0,-0.505) > p(0,\hat{\mathcal{A}}|0,-0.500)$ at $\hat{\mathcal{A}}=-0.500$.



(a) $\mathcal{S}_{\text{best}}-\mathcal{S}$ (b) $\mathcal{A}_{\text{best}}-\mathcal{A}$

Figure 5.29: $x_i^{\text{best}}-x_i$ as a function of x_i.

One of the drawbacks of the F-C method is that the integrating region can not be integrated analytically and the numerical integration can be very demanding in terms of the computer processing time, especially for CL ≈ 1. However, in the polar coordinate system (r,φ), the integral of the two-dimensional Gaussian is easily calculated,

$$\int_{\varphi_0}^{\varphi_0+\Delta\varphi}\int_0^R \frac{1}{2\pi}\,e^{-r^2/2}\,\mathrm{d}\varphi\,r\mathrm{d}r \;=\; \frac{\Delta\varphi}{2\pi}\left(1-e^{-R^2/2}\right). \tag{5.54}$$

Since the integrating region is close to an ellipse in any case, replacing the integral with

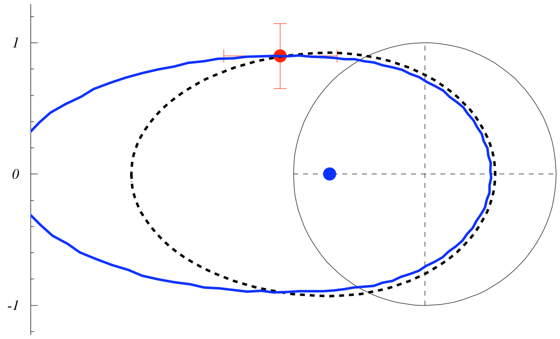


Figure 5.30: The integrating region for the F-C (full line) and "standard" (dashed line) ordering principle for $\mathcal{S} = -0.725$ and $\mathcal{A} = 0$ (shown by the blue point). The red point with errors shows the measured $(\mathcal{S}, \mathcal{A})$. The circle shows the allowed region of $\mathcal{S}^2 + \mathcal{A}^2 \leq 1$.

the following sum

$$\mathrm{CL} = \frac{\Delta\varphi}{2\pi} \Sigma_{i=1}^{N} \left(1 - e^{-R_i^2/2}\right) = \tag{5.55}$$

$$= 1 - \frac{1}{N} \Sigma_{i=1}^{N} e^{-R_i^2/2}\,, \tag{5.56}$$

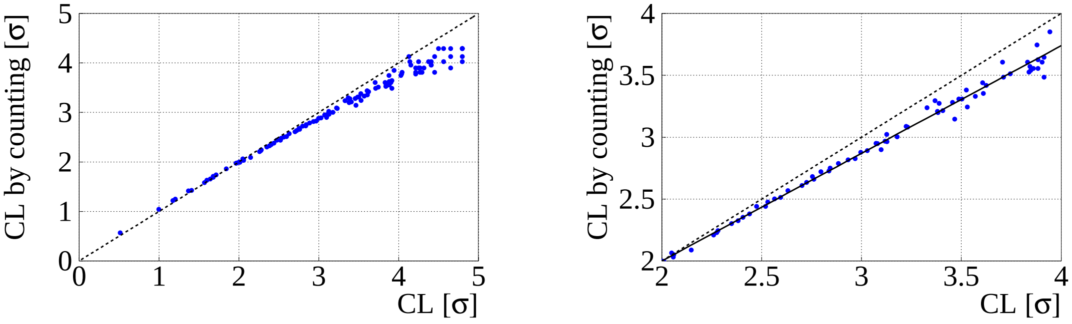
where $R^2 = ((\hat{\mathcal{S}} - \mathcal{S})/\sigma_{\mathcal{S}})^2 + ((\hat{\mathcal{A}} - \mathcal{A})/\sigma_{\mathcal{A}}^2)$, is a good approximation. The widths σ_X in each of the quadrants in $(\mathcal{S}, \mathcal{A})$ are different. We therefore only have to calculate the values of R_i. For the later, numerical method is needed to calculate the maximal PDF, but can be done reasonably fast for a sufficiently large number of intervals N.

Effect of the Wider Than Gaussian Tails

While the description of the $\mathcal{S}$ and $\mathcal{A}$ distribution by the Gaussian function is a reasonable choice, any discrepancy in the tails of the distribution can significantly change the significance of the result when $CL \approx 1$. We therefore compare the CL calculated by integrating the Gaussian function that describes $\mathcal{P}(\hat{x}|x)$ with the CL that we obtain by counting the number of toy MC samples that satisfy the F-C condition 5.38, which we call a counting method. The confidence level obtained by the counting method is then the fraction of the toy MC samples that are within the F-C integrating region and CL in the units of σ is calculated as illustrated in Table 5.17. Figure 5.31 compares the CL obtained by counting with the CL obtained by integration for the values of CP parameters $\hat{S} = -1.1$ and $\hat{A} = 0.9$. The result shows about 4% difference at the $3\,\sigma$ level, which is small enough to be accounted for by re-scaling the CL. We approximate the dependence for CL $> 2\sigma$ by a linear function

$$\mathrm{CL}_{\Sigma} = (2\,\mathrm{CL}_{\int} + 0.6)/2.3 \tag{5.57}$$

that is shown in Fig. 5.31(b).



(a) (b)

Figure 5.31: The CL obtained by counting as a function of the CL obtained by integration for the values $\hat{S} = -1.1$, $\hat{A} = 0.9$. Figure (b) is only an inset of Fig. (a). Full line in Fig. (b) shows dependence described by Eq. 5.31.

5.5.4 Result of the Confidence Level Study

The CL contour plot obtained by the Feldman-Cousins significance study is shown in Fig. 5.32 The case of no CPV, $\mathcal{S} = \mathcal{A} = 0$, is excluded at 4.1σ CL. The case of no direct CPV, $\mathcal{A} = 0$, is excluded at more than 3.2σ CL for any value of the parameter $\mathcal{S}$. In this way we quote the most conservative value and avoid the effect of the uncertainty the value of $\mathcal{S}$ $(-\sin(2\phi_1))$.

The CL for $\mathcal{A} = 0$ is actually not very sensitive to the value of $\mathcal{S}$ as the CL contours are almost parallel to the $A = 0$ axis in that range.

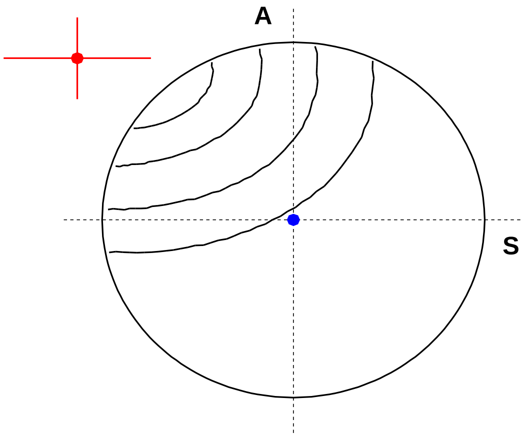


Figure 5.32: The 1−, 2−, 3− and 4−σ level CL contours obtained by the Feldman-Cousins significance study. The red point with error bars is the result of this measurement.

Chapter 6

Discussion

Over the past thirty years, the Standard Model has been extremely successful in explaining results of the various high energy physics experiments. The two experiments Belle and BaBar observed CP violation in the neutral B meson system in 2001. It was hoped that consecutive measurements in the B meson system would show a deviation from the Standard Model predictions. However, all the measurements so far only made it clear that the KM phase is the dominant source of CP violation.

The measurement of CP violation in $B^0 \to D^+D^-$ decays is another possible way of testing the Standard Model predictions. The measurement was performed using a large data sample of 535×10^6 $B\bar{B}$ pairs collected with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. We observe $B^0 \to D^+D^-$ decays with more than $5\,\sigma$ statistical significance and find the branching fraction to be $(1.97 \pm 0.20 \pm 0.20) \times 10^{-4}$. We measure the time dependent CP parameters in $B^0 \to D^+D^-$ decays to be $\mathcal{S} = -1.13 \pm 0.37 \pm 0.09$ and $\mathcal{A} = 0.91 \pm 0.23 \pm 0.06$ and rule out the CP-conserving case of $\mathcal{S} = \mathcal{A} = 0$ at $4.1\,\sigma$ confidence level.

Curiously enough, this measurement indicates a possibly large direct CP violation in $B^0 \to D^+D^-$ decays. We observe evidence for direct CP violation at a 3.2σ confidence level. This is a quite unexpected result as the SM contributions should only result in a small direct CPV term. It would seem that we see some deviation from the SM expectation of $\mathcal{A} \approx 3\%$ [12, 13]. However, while our data sample of reconstructed events does show evidence for direct CP violation at more than $3\,\sigma$ confidence level, it does not reject the SM expectation of $\mathcal{A} \approx 3\%$ at a $3\,\sigma$ CL. In addition, due to large and unknown theoretical uncertainties, the SM contributions could be at a 10% level [36]. Given a large statistical error, the main experimental issue is to reduce the statistical uncertainty in the result. The Belle detector currently records close to 200×10^6 $B\bar{B}$ pairs per year of operation. Since the current measurement already uses a large data sample of 535×10^6 $B\bar{B}$ pairs and the statistical error is only inversely proportional to the square root of the number of events, additional few years of data taking will not drastically change the accuracy of the result. Still, current result being outside the region $\mathcal{S}^2 + A^2 \leq 1$ is obviously caused by a statistical fluctuation and at the current accuracy level, 30% decrease in the measurement error could actually be very helpful in resolving the issue. In order to measure the CP parameters in

$B^0 \to D^+D^-$ decays at a few percent accuracy new technologies are needed to increase the luminosity of the B factory, as proposed for the next generation SuperKEKB accelerator [9].

While we are waiting for the new data to be collected at Belle, we should note there is another B meson factory experiment running at SLAC, Stanford, USA. Previous measurement by the BaBar collaboration already reported a result for the time dependent CP parameters in $B^0 \to D^+D^-$ decays, $\mathcal{S} = -0.29 \pm 0.63\,(\mathrm{stat}) \pm 0.06\,(\mathrm{syst})$ and $\mathcal{A} = -0.11 \pm 0.35\,(\mathrm{stat}) \pm 0.06\,(\mathrm{syst})$ [16]. We note that on a statistically independent data sample of $B\overline{B}$ events they do not observe any direct CPV. Still, their analysis is based on a smaller data sample of 232×10^6 $B\bar{B}$ pairs and therefore has larger statistical errors on both parameters. The comparison of the two results shows that they are consistent at a 2.2σ confidence level.

We do not need to limit our search only to the updated measurements in $B^0 \to D^+D^-$ decays. If the large direct CP violation we observe in this work originates from the physics beyond the SM, similar discrepancies could be observed in other processes involving the same $b \to c\bar{c}d$ transition, such as $B^0 \to D^{*+}D^{(*)-}$ and isospin conjugated $B^+ \to D^+\overline{D}^0$ decays. Previous measurements in these modes did not show any significant deviations from the SM predictions [14, 15, 21, 28]. The number of signal events and the signal to background ratio of these measurements is similar as for the current analysis. To determine whether the disagreement between the SM prediction and our measurement is an indication of new physics or a statistical fluctuation, the statistical uncertainty in the results must be reduced; updating the presented analysis with a larger data sample will show which is the case.

Appendix A

Appendix

A.1 High Energy Physics Basics

A.1.1 Units

The scale of HEP is considerably different from the scale of every day quantities we are used to and on which the international system (SI) of units is based. Additionally, it is more convenient to use the relation $E^2 = p^2c^2 + m^2c^4$ if the same units are used for energy, mass and momentum. For example, a relativistic electron of energy 3 GeV, is therefore said to have a momentum of 3 GeV/c rather than $1.6 \cdot 10^{-27}$ kg m/s. For the sake of simplicity, Planck's constant $\hbar$ and the speed of light c are often omitted. As this thesis is intended for a wider audience we do not follow this convention of $\hbar = c = 1$. Still, the mass of the particle, for example, will be given in units MeV/c^2 or GeV/c^2 rather than the standard SI units of kilogram. Here is a short overview between the SI units and those commonly used in HEP.

- Energy, mass and momentum of a particles are usually measured in GeV (MeV), GeV/c^2 (MeV/c^2) and GeV/c (MeV/c), respectively. To illustrate the different choice of units, we show the mass of electron and proton:

particle	m [MeV/c^2]	m [10^{-31} kg]
electron	0.511	9.11
proton	938	$1.67 \cdot 10^4$

- A decay time is often given in a distance units divided by a speed of light, for example:

particle	τ [μm/c]	τ [ps]
B^0	460	1.50
D^+	312	1.04

A.1.2 Mesons

So far, we have mostly been talking about quarks. Due to the color confinement, in reality we can never detect a single quark - they only come in a bound pair of a quark and an anti-quark, which we call a meson, or in a bound state of three quarks, a barion. When studying the decays of particles it is always useful to keep in mind the underlining quark structure. Table A.1 shows quark structure of various mesons that are relevant for this analysis and are mentioned later on.

	$\bar{d}$	$\bar{u}$	$\bar{s}$	$\bar{c}$	$\bar{b}$	$\bar{t}$
d		π^-	K^0	D^-	B_d^0	
u	π^+		K^+	$\bar{D}^0$	B^+	
s	$\bar{K}^0$	K^-			B_s^0	
c	D^+	D^0				
b	$\overline{B}_d^0$	B^-	$\overline{B}_s^0$			
t						

Table A.1: The quark structure of pions (denoted as π), kaons (K), D and B mesons. Each meson listed consists of the quark shown on the left side of the same row and of the anti-quark shown on the top of the same column.

The lightest mesons which are made of the first generation quarks of type up for u and down for d, pions, were first to be discovered. The neutral pion π^0 is not shown in the table because it is a superposition of $d\bar{d}$ and $u\bar{u}$ states. The more massive mesons contain one quark from the higher generation families and one from the first generation: kaons contain one strange quark s, D mesons have a charm quark c and B mesons contain a bottom quark b. The heaviest top quark t is so massive that it immediately decays and does not make any bound state. The charge of each meson can be deduced from the sum of charges of its quark constituents, $Q_{u,c,t} = +2/3$ and $Q_{d,s,b} = -1/3$. Note that as there is no flavor changing neutral current, the flavor changing transitions (one type of quark transforming to another type of quark) can only occur between the different-charged quarks, as illustrated in Fig. 2.1.

A.2 Displaying the Result of the Fit

When we plot the Δt distribution, we divide the range into bins and show the number of events in the j^{th} bin. To show the fit result on the same figure, we have to calculate the corresponding projection of the probability density function, i.e. the expected number of events in the j^{th} Δt bin, $\mathcal{N}_j$.

$$\mathcal{N}_j = \Sigma_i \int_j \mathcal{L}_i(\Delta t, \dots)\mathrm{d}\Delta t\,, \tag{A.1}$$

where the index i refers to the i^{th} event and index j stands for the j^{th} bin of the Δt distribution. Since we can calculate the value of $\mathcal{N}$ for any value of Δt, we replace a histogram with a continuous projection of the probability density function,

$$\begin{aligned}\mathcal{N}(\Delta t) &= \Sigma_i \int_{\Delta t-\Delta_t/2}^{\Delta t+\Delta_t/2} \mathcal{L}_i(\Delta t', \dots)\mathrm{d}\Delta t' && \text{(A.2)}\\ &\approx \Delta_t\, \Sigma_i\, \mathcal{L}_i(\Delta t, \dots)\,, && \text{(A.3)}\end{aligned}$$

where Δ_t is the Δt bin width, the range of the Δt distribution divided by the number of bins. Since the likelihood function is normalized according to Eq. 5.10, $\Sigma_j \mathcal{N}_j$ is exactly equal to the number of measured events.

A.3 Covariance matrix

For the analysis of the Δt distribution, many different parameters x_i (describing the resolution function, signal fraction, etc) first need to be determined so that they can be used later for the fit of the CP parameters. Due to the uncertainty in the parameters x_i we have systematic errors in the CP parameters. The systematic error due to the uncertainty in is estimated by varying the parameter x_i by $\pm\sigma_{x_i}$. However, this is a proper procedure only when the parameters x_i are uncorrelated. If the parameters are correlated, the correct procedure is as follows.

- Determine the covariance matrix $\mathbf{V}$.
- Diagonalize the covariance matrix $\mathbf{V}$; $\mathbf{V} = \mathbf{U}\mathbf{D}\mathbf{U}^{\mathrm{T}}$, $\mathbf{U}^{\mathrm{T}} = \mathbf{U}^{\mathrm{inv}}$, where $\mathbf{D}$ is a diagonal matrix and $\mathbf{D}_{ii} = \sigma^2_{y_i}$, y_i are the transformed variables Δx_i: $\mathbf{y} = \mathrm{U}^T \Delta\mathbf{x}$.
- Vary each of the y_i by σ_{y_i} (set $\mathbf{y}$ so that $y_j = \sigma_{y_i}$ if i equals j and zero otherwise), calculate $\Delta\mathbf{x} = \mathrm{U}\mathbf{y}$, determine its effect on the CP parameters and sum the contributions in quadratures.

Instead of varying each parameter y_i individually, one can smear all of them according to the independent Gaussian distributions of widths σ_{y_i} and determine the resulting shift in the CP parameters Δx_i. The systematic error is equal to the root mean square (RMS) of the shift distribution.

For both procedures, the error on the x_i (sum in quadratures of $\Delta x_i(y_j)$ in the first and the RMS of the Δx_i distribution in the second case) is consistent with the σ_{x_i} as determined by the un-binned ML fit.

Dodatek B

Povzetek

B.1 Uvod

Standardni Model (SM) opisuje vse znane osnovne delce in njihove interakcije. Osnovne delce delimo v dve skupini, na fermione in bozone. Fermioni imajo polovični spin in ubogajo Paulijevo izključitveno načelo, ki pravi, da dva enaka fermiona ne moreta zasedati istega kvantnega stanja. V nasprotju s fermioni, Paulijevo izklučitveno načelo ne velja za bozone s celoštevilčnim spinom. Vsi osnovni gradniki snovi so fermioni, ki jih nadalje delimo na kvarke in leptone. Kvarki so sestavni delci protonov in nevtronov (skupno poimenovani hadroni), ki se med sabo vežejo v atomska jedra. Jedra skupaj z elektroni (ki jih uvrščamo med leptone) sestavljajo atome. Bozoni po drugi strani opisujejo interakcije; pravimo, da so nosilci sil med osnovnimi gradniki. Poznamo pet vrst bozonov, ki opisujejo močno, šibko in elektromagnetno interakcijo. To so gluoni, Z^0, W^+, W^- in fotoni. Močno silo, ki je odgovorna za vezavo kvarkov, opisujejo gluoni. Šibko interakcijo, zaradi katere npr. poteka razpad β, opisujejo šibki bozoni Z^0, W^+ in W^-. Elektromagnetna interakcija je opisana z izmenjavo fotonov.

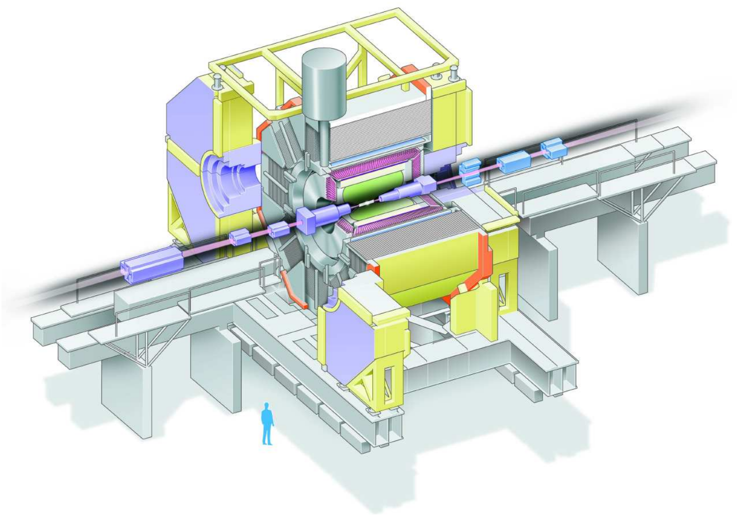
Mase fermionov kažejo hierarhično strukturo. Na osnovi njihovih mas fermione delimo v tri generacije, kot je prikazano v Tabeli B.1. Vsakemu naštetemu delcu pripada anti-delec enake mase in nasprotnih kvantnih števil (nasprotnega naboja,...), ki ga označimo s črtico, npr. $\bar{p}$ za anti-proton. Navadna snov okoli nas je sestavljena iz delcev prve generacije. Delci druge in tretje generacije lahko nastanejo pri interakciji visokoenergijskih kozmičnih žarkov s snovjo v vesolju ali z zemeljsko atmosfero ter pri eksperimentih fizike visokih

leptoni		kvarki	
e	ν_e	u	d
μ	ν_μ	c	s
τ	ν_τ	t	b

Tabela B.1: Tri generacije fermionov.

energij. Z merjenjem prehodov med delci različnih generacij študiramo lastnosti interakcij med njimi. V tem delu so predstavljeni rezultati meritev, opravljenih z detektorjem Belle, ki je specializiran za proučevanje lastnosti kvarka b.

Detektor Belle, prikazan na sliki B.1, meri lastnosti dogodkov, ki se zgodijo ob trkih elektronov energij 8 GeV in pozitronov energij 3.5 GeV na trkalniku KEKB v Tsukubi na Japonskem. Težiščna energija obeh žarkov ustreza masi resonance $\Upsilon(4S)$. Resonanca $\Upsilon(4S)$ je prvo vezano stanje kvarkov $b\bar{b}$, ki je nad energijskim pragom za razpad v par mezonov $B\overline{B}$. Pri tej energiji imamo znaten presek za nastanek para mezonov $B\overline{B}$ in to je tudi razlog, da KEKB včasih imenujemo kar tovarna mezonov B. Druga pomembna lastnost žarkov je njuna asimetrična energija. Nastali mezoni B se zato v laboratorijskem sistemu gibljejo v smeri bolj energijskega žarka. To nam omogoča posredno meritev časa iz razdalje, ki jo mezona preletita v laboratorijskem sistemu. Čas med razpadoma obeh mezonov izmerimo iz razdalje med mestoma razpadov obeh mezonov B v smeri z, $\Delta t = (z_{CP} - z_{\text{tag}})/(\beta\gamma c)$, kjer je $\beta\gamma = 0.425$ Lorentzov faktor. Zaradi teh razlogov je naš eksperiment specializiran prav za meritev časovno odvisnih razpadov mezonov B, kot je razpad $B^0 \to D^+D^-$.



Slika B.1: Detektor Belle z različnimi podsistemi označenimi v različnih barvah: silicijev detektor mesta razpada v najbolj notranjem delu je označen z belo. Sledi z zeleno označen ionizacijski detektor, katerega poglavitni namen je rekonstrukcija poti delcev. Na njegovi zunanji strani so pragovni števec sevanja Čerenkova v beli, elektromagnetni kalorimeter v vijolični in K_L in muonski detektor v sivi barvi. Merilec časa peleta je premajhen, da bi bil označen na sliki, leži pa tik za števcem Čerenkova.

V tem povzetku bom najprej na kratko predstavila teoretično podlago za razpad $B^0 \to D^+D^-$ ter našo motivacijo. Nadaljevala bom z opisom rekonstrukcije razpadov $B^0 \to D^+D^-$ ter pokazala, koliko teh dogodkov rekonstruiramo iz vzorca podatkov, ki ga imamo na razpolago za to meritev, to je 535×10^6 parov $B\overline{B}$. V drugem delu bom na kratko opisala analizo porazdelitve po razpadnem času in predstavila rezultate meritve. V zaključku bom izmerjene rezultate primerjala z drugimi, sorodnimi meritvami.

B.2 Standardni Model

V Standardnem Modelu (SM) je simetrija CP kršena samo pri šibki interakciji, ki jo opisuje naslednji člen interakcijskega Lagrangiana.

$$\mathcal{L}^{\mathrm{CC}} = -\frac{g_2}{\sqrt{2}} \begin{pmatrix} \overline{u} & \overline{c} & \overline{t} \end{pmatrix}_L \gamma^\mu \mathbf{V} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L W_\mu^\dagger + \mathrm{h.c.}\,, \tag{B.1}$$

kjer je g_2 sklopitvena konstanta in $\mathbf{V}$ matrika Cabibbo-Kobayashi-Maskawa (CKM) [3, 4].

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \mathbf{V} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{B.2}$$

Splošna kompleksna matrika je določena z osemnajstimi parametri, $V_{ij} = a_{ij} e^{\imath\varphi_{ij}}$. Matrika CKM je unitarna, $\mathbf{V}^\dagger\mathbf{V} = \mathbf{V}\mathbf{V}^\dagger = \mathbf{I}$, zato osemnajst parametrov med sabo ni neodvisnih. Unitarni pogoj določa šest zvez med velikostmi elementov a_{ij}, in tri zveze med kompleksnimi fazami φ_{ij}. Neodvisno od unitarnega pogoja lahko poljubno definiramo pet relativnih faz med kvarkovskimi stanji. Zaradi teh omejitev so samo štirje parametri CKM matrike prosti parametri; to so tri absolutne vrednosti a_{ij} in ena kompleksna faza φ_{ij}.

Unitarne pogoje lahko predstavimo tudi grafično, tako, da vsakega izmed členov enačbe opišemo z vektorjem v kompleksni ravnini. Običajna izbira je produkt tretje vrstice adjungirane matrike in prvega stolpca, $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$. Pri tej izbiri so vsi trije členi enačbe podobnih velikosti, kot je prikazano na sliki B.2 (a). Za poenostavitev enačbo delimo z $V_{cd}V_{cb}^*$, v tem primeru drugo ogljišče trikotnika ustreza koordinati $(1, 0)$, kot je prikazano na sliki B.2 (b). Edina parametra tega trikotnika sta tako koordinati vrha trikotnika, ki ga označimo z $(\bar{\rho}, \bar{\eta})$. Koti trikotnika ustrezajo razliki faz priležnih stranic,

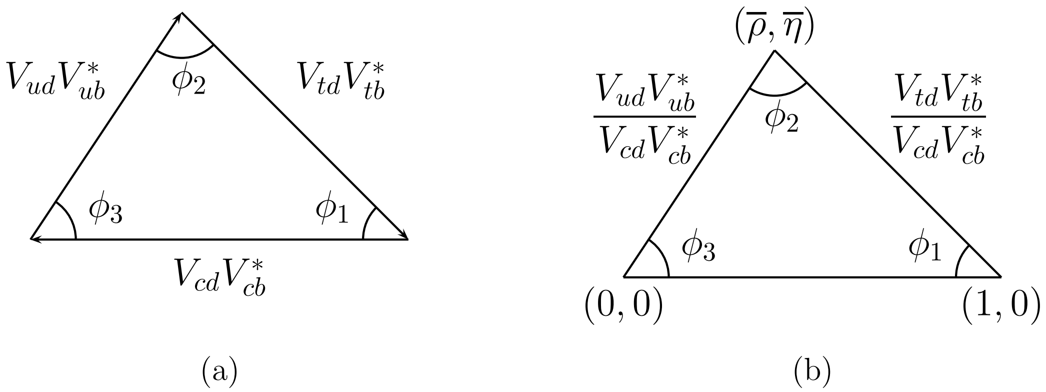
$$\phi_1 = arg(-V_{cd}V_{cb}^*/V_{td}V_{tb}^*)\,, \tag{B.3}$$

$$\phi_2 = arg(-V_{td}V_{tb}^*/V_{ud}V_{ub}^*)\,, \tag{B.4}$$

$$\phi_3 = arg(-V_{ud}V_{ub}^*/V_{cd}V_{cb}^*)\,. \tag{B.5}$$

V uporabi je tudi drugačna izbira oznak kotov trikotnika: $\alpha = \phi_2$, $\beta = \phi_1$ in $\gamma = \phi_3$.

Pri razpadu $B^0 \to D^+D^-$ lahko v končno stanje D^+D^- razpade mezon B^0 ali $\overline{B}^0$. Zato je kršitev CP v tem razpadu posledica interference med amplitudama dveh procesov, med razpadom z mešanjem ($B^0 \to \overline{B}^0 \to D^+D^-$) in brez njega ($B^0 \to D^+D^-$). Mešanje je proces, pri katerem se nevtralni mezon spremeni v svoj anti-delec, v našem primeru $B^0 \to \overline{B}^0$. Na kvarkovskem nivoju se par kvarkov $\bar{b}d$ spremeni v $b\bar{d}$, kot je prikazano na sliki B.3. Verjetnost za razpad mezona B^0 ($\overline{B}^0$) pri takem procesu je odvisna od časa, v



Slika B.2: Grafična reprezentacija unitarnega pogoja $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$ (a) in $V_{ud}V_{ub}^*/(V_{cd}V_{cb}^*) + 1 + V_{td}V_{tb}^*/(V_{cd}V_{cb}^*) = 0$ (b).

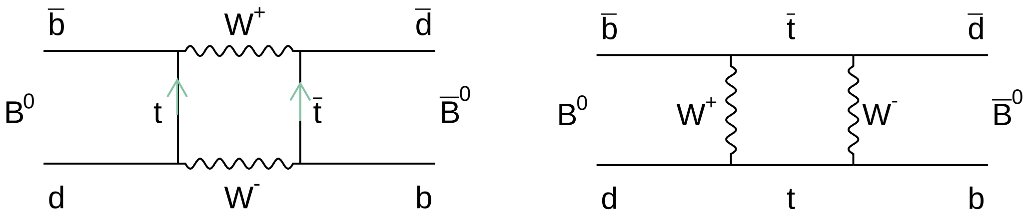
katerem mezon B prosto potuje, preden razpade. Pri eksperimentu Belle pari mezonov B nastanejo koherentno. Zato je, vse dokler prvi mezon ne razpade, vedno eden od njiju B^0 in drugi $\overline{B}^0$. Čas, od katerega je odvisna verjetnost za razpad, je zato čas med razpadom mezona B_{CP} in B_{tag}, $\Delta t = t_{CP} - t_{\mathrm{tag}}$. Časovno odvisna verjetnost za razpad je

$$\mathcal{P}_{\mathrm{sig}}\left(\Delta t, q\right) \;=\; \frac{e^{-|\Delta t|/\tau}}{4\tau}\left(1 + q\left(\mathcal{S}\sin(\Delta m \Delta t) + \mathcal{A}\cos(\Delta m \Delta t)\right)\right), \tag{B.6}$$

kjer $q = +1(-1)$ označuje dogodek, pri katerem je $B_{\mathrm{tag}} = B^0(\overline{B}^0)$. Drugi parametri porazdelitve so razpadni čas mezona B^0, $\tau = (1.530 \pm 0.009)\,\mathrm{ps}$, razlika med masami dveh masnih lastnih stanj mezonov B, $\Delta m = (0.507 \pm 0.005)\,\hbar/\mathrm{ps}$ [11], in parametra kršitve CP $\mathcal{S}$ in $\mathcal{A}$.

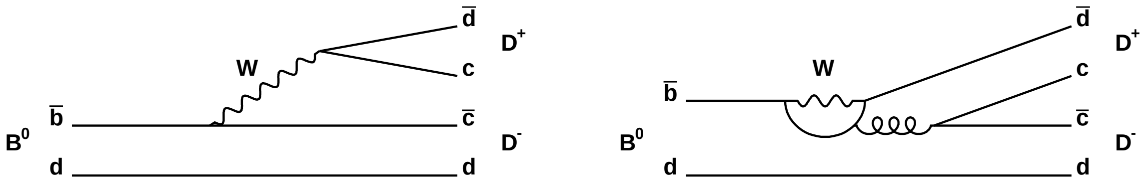
Razpadno amplitudo za $B^0 \to D^+D^-$ dominira drevesni diagram $b \to c\bar{c}d$, ki je prikazan na sliki B.4 (a). Opisujejo ga matrični elementi V_{cb}^* ($\bar{b} \to \bar{c}$) in V_{cd} ($d \to c$). V enakem razpadu mezona $\overline{B}^0$ so kvarki zamenjani z ustreznimi anti-kvarki, zato ta razpad opisujejo kompleksno konjugirani matrični elementi. Matrični elementi, ki opisujejo mešanje, so $V_{tb}^*V_{td}$. Izkaže se, da je parameter kršitve CP, ki opisuje ta razpad, povezan z razmerjem amplitud za razpad in mešanje, $V_{cd}V_{cb}^*/V_{td}V_{tb}^*$, oz. s kotom ϕ_1 trikotnika CKM (enačba B.3). Če upoštevamo samo diagram v drevesnem redu, je parameter $\mathcal{S} = -\sin(2\phi_1)$ in $\mathcal{A} = 0$.

V okviru SM pričakujemo, da je vpliv pingvinskega diagrama, ki je prikazan na sliki B.4 (b), na vrednosti parametrov $\mathcal{S}$ in $\mathcal{A}$ velikostnega reda nekaj procentov [12, 13]. Napoved SM za parametra kršitve CP je zato $\mathcal{S} \approx -\sin(2\phi_1)$ in $\mathcal{A} \approx 0.03$. Meritev parametra $\mathcal{S}$ je zato tudi meritev kota ϕ_1. Vendar pa ima naš vzorec rekonstruiranih dogodkov omejeno statistiko, zato se natančnost naše meritve ne more primerjati z visoko natančnostjo meritve $-\sin(2\phi_1)$ v razpadih s čarmonijem v končnem stanju, ki dosega relativno napako nekaj procentov [7, 8]. Zato namen tega dela ni sama meritev parametra ϕ_1, pač pa neodvisna meritev parametrov $\mathcal{S}$ in $\mathcal{A}$. Le – te lahko primerjamo s pričakovanimi vrednostmi in na ta



(a) (b)

Slika B.3: Opis mešanja $B^0 - \overline{B}^0$ s Feynmanovim diagramom. Kvark t v diagramu lahko zamenjata kvarka u ali c, vendar je verjetnost za tak proces manjša za faktor $(m_{c(u)}/m_t)^2$.



(a) (b)

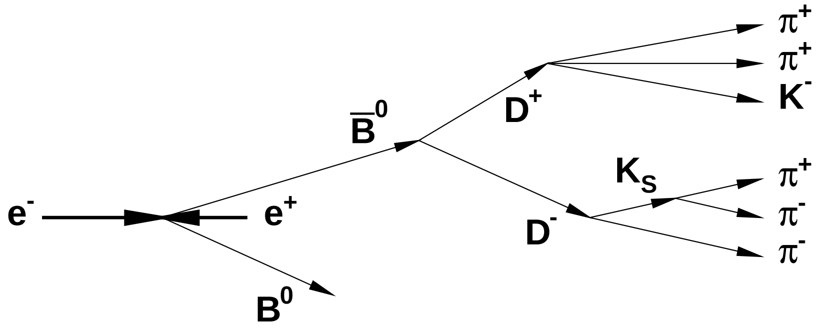
Slika B.4: Drevesni (a) in pingvinski prispevek (b) k razpadu $B^0 \to D^+D^-$.

način testiramo morebiten vpliv procesov, ki jih SM ne opisuje in zato niso upoštevani v njegovi napovedi. Taki procesi lahko bistveno vplivajo predvsem na pingvinske diagrame, ki imajo zanko. Podobne študije so že bile opravljene v sorodnih procesih $B^0 \to D^{*\pm}D^{(*)\mp}$ [14, 15, 16, 17].

B.3 Rekonstrukcija razpadov $\mathbf{B^0 \to D^+D^-}$

Naša prva naloga je izmed dogodkov, izmerjenih z detektorjem Belle, izbrati samo tiste, pri katerih mezon B razpade v dva nabita mezona D, $B^0 \to D^+D^-$. Mezone D^+ rekonstruiramo iz njihovih razpadnih produktov. Pri tem uporabimo le razpadne verige z znatnim razvejitvenim razmerjem ($\mathcal{B}$), in sicer samo take, pri katerih so v končnem stanju izključno nabiti delci. Tem pogojem ustrezata razpada $D^+ \to K^-\pi^+\pi^+$ in $D^+ \to \overline{K}^0\pi^+$, kjer $\overline{K}^0$ razpade kot $K_S \to \pi^+\pi^-$, kot je ilustrirano na sliki B.5. V primeru, ko sta oba mezona D rekonstruirana kot $D^+ \to K^-\pi^+\pi^+$, uporabljamo za razpad krajšo oznako "$K\pi\pi$". V primeru, ko je vsaj en od dveh mezonov D rekonstruiran kot $D^+ \to K_S\pi^+$, uporabljamo za razpad krajšo oznako "$K_S\pi$". V pričujočem delu so nabojno konjugirani kanali vedno vključeni, razen, kadar so obravnavani ločeno.

Med rekonstruiranimi mezoni B so tudi taki, pri katerih mezon B ni razpadel v par nabitih mezonov D. Signalne razpade $B^0 \to D^+D^-$ ločimo od ozadja na osnovi dveh



Slika B.5: Shamatska slika rekonstrukcije razpadov $B^0 \to D^+D^-$.

kinematskih spremenljivk, razlike energij

$$\Delta E = E_B^* - E_{\rm beam}^* \tag{B.7}$$

in invariantne mase mezona B, ki jo rekonstruiramo z upoštevanjem zakona o ohranitvi četverca gibalne količine,

$$M_{\rm bc} = \sqrt{(E_{\rm beam}^*/c^2)^2 - (p_B^*/c)^2}\,, \tag{B.8}$$

kjer so E_B^*, $E_{\rm beam}^*$, in p_B^* energija mezona B, energija žarkov, in gibalna količina mezona B. Porazdelitvi $M_{\rm bc}$ in ΔE signalnih dogodkov imata vrh pri vrednostih $M_{\rm bc} = m_B = 5.28\,{\rm GeV}/c^2$ in $\Delta E = 0$, razmazan zaradi detektorske resolucije. Porazdelitvi sta dobro opisani z Gaussovima funkcijama širin $\sigma_{\Delta E} = 3\,{\rm MeV}$ in $\sigma_{M_{\rm bc}} = 7\,{\rm MeV}/c^2$. Število dogodkov ozadja zmanjšamo s pomočjo vrste selekcijskih kriterijev. Med njimi sta najpomembnejša kaonska identifikacija in rekonstuirana invariantna masa mezonov D. Vsakič izberemo tak kriterij, za katerega je $N_{\rm sig}/\sqrt{N_{\rm sig}+N_{\rm bcg}}$ največji, kar ustreza čim manjši relativni statistični napaki na razvejitvenem razmerju, $\sqrt{N_{\rm sig}+N_{\rm bcg}}/N_{\rm sig}$.

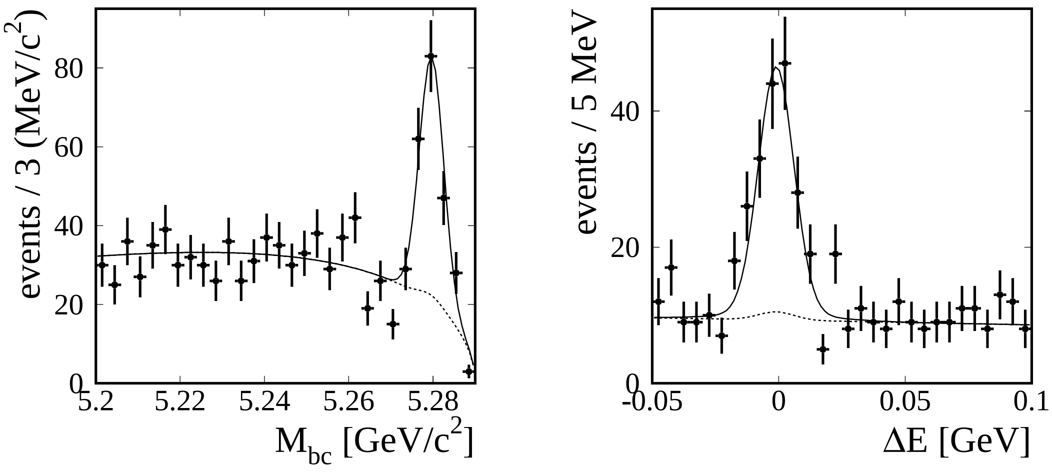
Slika B.6 prikazuje porazdelitev $M_{\rm bc}$ in ΔE rekonstruiranih mezonov B. Število signalnih dogodkov določimo z maksimiziranjem cenilke verjetnosti (maximum likelihood, ML)

$$\mathcal{L} = \Pi_i\, \mathcal{L}_i(M_{\rm bc}, \Delta E)\,, \tag{B.9}$$

kjer $\mathcal{L}_i(M_{\rm bc}, \Delta E)$ opisuje porazdelitev $M_{\rm bc}$ in ΔE rekonstruiranih dogodkov. Število rekonstruiranih dogodkov tipa $B^0 \to D^+D^-$ v signalnem območju $M_{\rm bc} > 5.27\,{\rm GeV}/c^2$ in $|\Delta E| < 0.03\,{\rm GeV}$ je 150 ± 15, kjer ± 15 označuje statistično napako meritve. Manjši dodatek k ozadju v signalnem območju (slika B.6, črtkana črta) opisuje delež neresonantnih razpadov $B^0 \to D^-\pi^+\overline{K}^0$ in $B^0 \to D^-\pi^+\overline{K}^{*0}(892)$ v vzorcu rekonstruiranih mezonov B. Število teh razpadov, $N_{\rm nr}$, ocenimo iz analize dogodkov, ki imajo rekonstruirano vrednost mase enega D mezona zunaj signalnega območja. Izmerimo $N_{\rm nr} = 3.4 \pm 2.0$.

Ob poznavanju razvejitvenih razmerij za $D^+ \to \ldots$ in izkoristka rekonstrukcije lahko izračunamo razvejitveno razmerje za razpad $B^0 \to D^+D^-$,

$$\mathcal{B}(B^0 \to D^+D^-) \;=\; (1.97 \pm 0.20 \pm 0.20)\,\times 10^{-4}\,. \tag{B.10}$$



(a) $M_{\rm bc}$, $|\Delta E| < 0.03\,{\rm GeV}$ (b) ΔE, $M_{\rm bc} > 5.27\,{\rm GeV}/c^2$

Slika B.6: Porazdelitev $M_{\rm bc}c^2$ (a) in ΔE (b). Polna (črtkana) črta prikazuje projekcijo cenilke $\mathcal{L}_i(M_{\rm bc}, \Delta E)$ za vse dogodke (ozadje).

Prva napaka označuje statistično in druga sistematsko napako. Sistematska napaka je izračunana iz vsote kvadratov posameznih prispevkov, največji med njimi so nenatančnosti v razvejitvenih razmerjih mezonov D (5%), izkoristku rekonstrukcije (6%) in identifikaciji delcev (6%). Rezultat je konsistenten s prejšnjimi meritvami [21, 28] in jih presega v natančnosti.

B.4 Analiza porazdelitve razpadnega časa

V drugem delu analize iz izmerjene porazdelitve Δt določimo parametra $\mathcal{S}$ in $\mathcal{A}$ kršitve CP. Ker je gibalna količina mezonov B v težiščnem sistemu (center-of-mass system, CMS) majhna, $pc \ll m_B c^2$, uporabimo za določitev vrednosti Δt samo meritev koordinate mesta razpada v smeri gibanja težiščnega sistema. Približno velja, da je $\Delta t \approx (z_{CP} - z_{\rm tag})/(\beta\gamma c)$, kjer je $\beta\gamma = 0.425$ Lorentzov faktor. Pri meritvi parametrov $\mathcal{S}$ in $\mathcal{A}$ uporabimo natančnejšo zvezo $\Delta t(\Delta z)$, ki upošteva tudi gibanje para mezonov B v CMS.

Za meritev kršitve CP moramo izmeriti tudi okus mezona B (označen s q v enačbi B.6). Pri tem si pomagamo z analizo rekonstruiranih razpadnih produktov, ki ostanejo po rekonstrukciji mezona B_{CP}. Okus lahko izmerimo, ker je vrsta razpadov mezonov B okusno značilnih, kot npr. razpad $B^0 \to K^+X$, kjer X predstavlja preostale razpadne produkte. V tem primeru je prisotnost visokoenergijskega pozitivno nabitega kaona (ki vsebuje kvark $\bar{s}$) indikacija, da je mezon $B_{\rm tag}$ razpadel kot B^0 (ki vsebuje kvark $\bar{b}$, ki razpade kot $\bar{b} \to \bar{c} \to \bar{s}$). Kvaliteta (natančnost) določitve okusa, ki jo označimo z r, je odvisna od prisotnosti in kvalitete okusno občutljivih razpadnih produktov. Glede na vrednost r razdelimo dogodke v šest razdelkov, ki so prikazani v tabeli B.2. Vrednost $r \approx 0$ označuje dogodke, za katere nimamo praktično nobene informacije o okusu, medtem ko $r \approx 1$ označuje dogodke, za katere je okus skoraj enolično določen. Za vsakega od razdelkov v r izmerimo verjetnost, da je okus napačno izmerjen, $w \approx (1-r)/2$ (tabela B.2). V splošnem razlikujemo med vrednostmi w za primere, ko je $B_{\rm tag} = B^0$ in $B_{\rm tag} = \overline{B}^0$ in na ta način upoštevamo morebitne razlike v detektorskem izkoristku za rekonstrukcijo delcev in antidelcev.

r razdelek	1	2	3	4	5	6
r območje	$0-0.25$	$0.25-0.5$	$0.5-0.625$	$0.625-0.75$	$0.75-0.875$	$0.875-1$
w	0.47	0.33	0.23	0.16	0.10	0.02

Tabela B.2: Razdelki v r in povprečna verjetnost za napačno izmerjen okus w.

Parametre kršitve CP določimo z nebiniranim ML prilagajanjem porazdelitve Δt, kjer je verjetnostna cenilka

$$\mathcal{L} = \Pi_i \mathcal{L}_i \tag{B.11}$$

in $\mathcal{L}_i$ verjetnostna porazdelitev za i-ti dogodek. Verjetnostno porazdelitev $\mathcal{L}_i$ opišemo z Δt odvisnostjo za signalne dogodke (enačba B.6), ki jo popravimo tako, da opisuje tudi ozadje in detektorsko ločljivost. Najpomembnejše spremembe so

$$\begin{aligned}
\mathcal{P}_{\rm sig} &\to \mathcal{P}_{\rm sig} \otimes \mathcal{R} && \text{detektorska ločljivost,} && \text{(B.12)}\\
\mathcal{P}_{\rm sig} \otimes \mathcal{R} &\to f_{\rm sig}\mathcal{P}_{\rm sig} \otimes \mathcal{R} + (1-f_{\rm sig})\mathcal{P}_{\rm bcg} && \text{ozadje in} && \text{(B.13)}\\
\mathcal{P}_{\rm sig}(\mathrm{d}\Delta t, q) &\to \mathcal{P}_{\rm sig}(\mathrm{d}\Delta t, (1-2w)\, q) && \text{meritev okusa.} && \text{(B.14)}
\end{aligned}$$

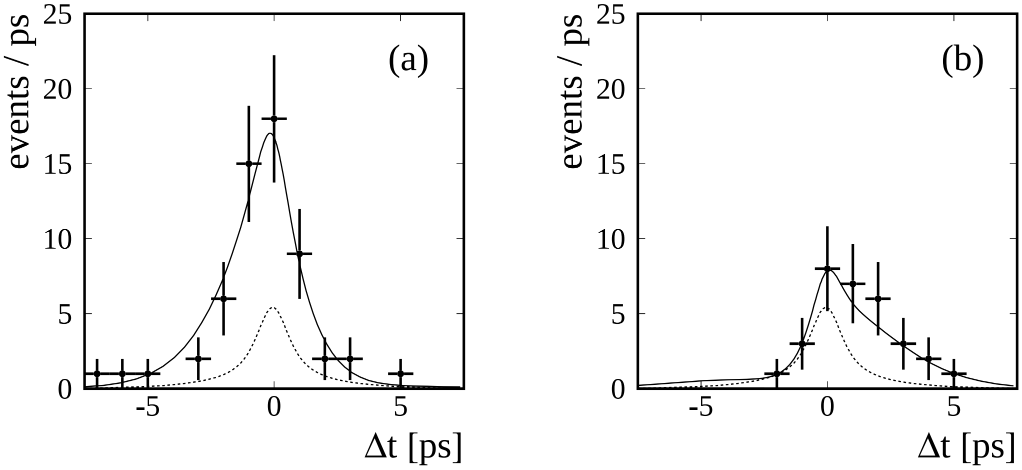
Verjetnost, da dogodek ustreza signalnemu razpadu $B^0 \to D^+D^-$, $f_{\rm sig}$, izračunamo na osnovi vrednosti kinematskih parametrov $M_{\rm bc}$ in ΔE ter kvalitete določanja okusa r, $f_{\rm sig} = f_{\rm sig}(M_{\rm bc}, \Delta E, r)$. Majhna deleža neresonantnih razpadov ter dogodkov z veliko večjo napako od izmerjene opišemo ločeno.

Preden izmerimo vrednosti parametrov kršitve CP, naredimo vrsto testov našega postopka. Med njimi so razne Monte Carlo simulacije, s katerimi preverimo nepristranskost metode. Pomemben test je tudi neodvisna meritev parametrov $\mathcal{S}$ in $\mathcal{A}$ v kontrolnem vzorcu razpadov $B^0 \to D_s^+D^-$, pri katerih je pričakovana vrednost obeh parametrov majhna, $\mathcal{S} \approx 0$ in $\mathcal{A} \approx 0$. Vzorec razpadov $B^0 \to D_s^+D^-$ vsebuje okoli 2000 dogodkov, med katerimi je le okoli 5% dogodkov ozadja. Večji vzorec nam omogoča meritev z manjšo statistično napako in torej bolj natančen test metode.

Slika B.7 prikazuje porazdelitev Δt za dogodke z dobro izmerjenim okusom. Rezultat meritve parametrov kršitve CP je

$$\begin{aligned}
\mathcal{S} &= -1.13 \pm 0.37 \pm 0.09\\
\mathcal{A} &= +0.91 \pm 0.23 \pm 0.06\,,
\end{aligned} \tag{B.15}$$

kjer je prva napaka statistična in druga sistematska. Največji prispevki k sistematski napaki so nenatančno določeni delež signalnih dogodkov (0.035 za $\mathcal{S}$ in 0.015 za $\mathcal{A}$), parametri resolucijske funkcije (0.04 za $\mathcal{S}$ in 0.03 za $\mathcal{A}$) in manjša pristranskost metode zaradi majhnega vzorca rekonstruiranih dogodkov (0.06 za $\mathcal{S}$ in 0.02 za $\mathcal{A}$).

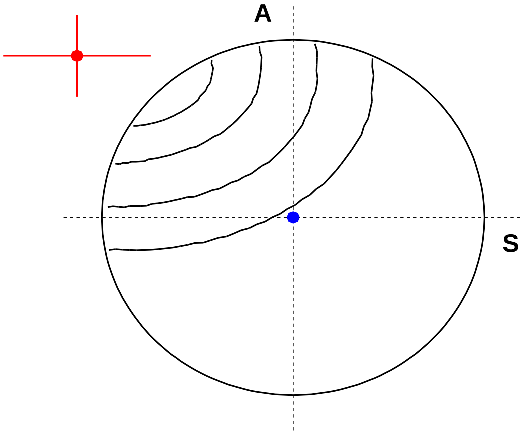


Slika B.7: Porazdelitev Δt za dogodke z dobro izmerjenim okusom ($r > 0.5$); ločeno za dogodke z izmerjenim okusom $B_{\rm tag} = B^0$ (a) in $B_{\rm tag} = \overline{B}^0$ (b). Polni in črtkani črti prikazujeta projekcijo verjetnostne cenilke $\mathcal{L}$ za vse dogodke (polna) in ozadje (črtkana črta).

Pri tem rezultatu je potrebno biti pozoren na dve lastnosti. Najprej je potrebno poudariti, da so vrednosti parametrov zunaj fizikalnega območja $\mathcal{S}^2 + \mathcal{A}^2 \leq 1$. Napake na parametrih, določene z metodo ML, so zato lahko podcenjene. Zato da se izognemo napaki pri ocenjanju napake parametrov, napako raje določimo na osnovi Monte Carlo simulacije. Naslednji, morda bolj zanimiv rezultat je, da je vrednost parametra $\mathcal{A}$ v neskladju z napovedmi SM. Pomembno je, da določimo s kolikšno zanesljivostjo (confidence level, CL) naš rezultat ovrže vrednost $\mathcal{A} = 0$. Ker je izmerjen rezultat zunaj fizikalnega območja, v ta namen uporabimo pristop, ki sta ga vpeljala G. J. Feldman in R. D. Cousins (metoda F-C) [35]. Rezultati te statistične analize so prikazani na sliki B.8. Slika prikazuje 1σ, 2σ, 3σ in 4σ CL krivulje, kjer je pomen $N\sigma$ CL predstavljen v tabeli B.3. Neničelna kršitev CP v razpadu $B^0 \to D^+D^-$ je potrjena s 4.1σ CL. Neničelna direktna kršitev CP, $\mathcal{A} \neq 0$, je potrjena z več kot 3.2σ CL za katerokoli vrednost $\mathcal{S}$.

CL [σ]	CL [%]
1	68.3
2	95.4
3	$1 - 2.7 \cdot 10^{-3}$
4	$1 - 6.3 \cdot 10^{-5}$
5	$1 - 5.7 \cdot 10^{-7}$

Tabela B.3: Pokritje (CL [%]) za $N\sigma$ CL.



Slika B.8: Krivulje označujejo 1σ, 2σ, 3σ in 4σ CL, izračunan s Feldman-Cousins statistično analizo.

B.5 Zaključek

V tem delu sem izmerila razvejitveno razmerje in parametra kršitve CP za razpad $B^0 \to D^+D^-$. Rezultati analize so

$$\mathcal{B}(B^0 \to D^+D^-) \;=\; (1.97 \pm 0.20 \pm 0.20) \times 10^{-4} \tag{B.16}$$

za razvejitveno razmerje in

$$\begin{aligned} \mathcal{S} &= -1.13 \pm 0.37 \pm 0.09 \\ \mathcal{A} &= +0.91 \pm 0.23 \pm 0.06\,, \end{aligned} \tag{B.17}$$

za parametra kršitve CP. Zanimivo je, da je rezultat za parameter $\mathcal{A}$ v neskladju s pričakovano vrednostjo $\mathcal{A} \approx 0$. Naša meritev izloči ničelno direktno kršitev CP z zanesljivostjo 3.2σ. Če je razlog za razkorak med izmerjeno in pričakovano vrednostjo pojav, ki ga SM ne opisuje, bi pričakovali podobna razhajanja v procesih, pri katerih gre za isti prohod med kvarki, $b \to c\bar{c}d$. Dosedajšnje študije tovrstnih razpadov, to so $B^0 \to D^{*+}D^{(*)-}$ in izospinsko konjugiran razpad $B^+ \to D^+\overline{D}^0$, niso pokazale nobenih omembe vrednih razhajanj [14, 15, 21, 28]. Parametra $\mathcal{S}$ in $\mathcal{A}$ v razpadu $B^0 \to D^+D^-$ je izmerila tudi kolaboracija Babar. Njihov razultat se sklada z majhno direktno kršitvijo CP, $\mathcal{S} = -0.29 \pm 0.63\,(\mathrm{stat}) \pm 0.06\,(\mathrm{syst})$ in $\mathcal{A} = -0.11 \pm 0.35\,(\mathrm{stat}) \pm 0.06\,(\mathrm{syst})$, vendar imajo večjo napako [16]. Meritvi sta konsistentni z zanesljivostjo 2.2σ. V luči teh primerjav je povsem verjetno, da je naš provokativen rezultat v veliki meri posledica statistične fluktuacije. S časom, ko bodo na voljo večji vzorci podatkov [9], bomo lahko ponovili to meritev z večjo natančnostjo. Takrat se bo izkazalo ali se vrednost parametra $\mathcal{A}$ v resnici znatno razlikuje od nič ali gre morda le za statistično fluktuacijo.

Bibliography

[1] G. Alexander et al. Electroweak parameters of the z0 resonance and the standard model: the lep collaborations. *Phys. Lett.*, B276:247–253, 1992.

[2] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2 pi decay of the k(2)0 meson. *Phys. Rev. Lett.*, 13:138–140, 1964.

[3] N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–532, 1963.

[4] Makoto Kobayashi and Toshihide Maskawa. Cp violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[5] K. Abe et al. Observation of large cp violation in the neutral b meson system. *Phys. Rev. Lett.*, 87:091802, 2001.

[6] B. Aubert et al. Observation of cp violation in the b^0 meson system. *Phys. Rev. Lett.*, 87:091801, 2001.

[7] K. Sumisawa et al. Measurement of time-dependent cp-violating asymmetries in b0 → k0(s) k0(s) k0(s) decay. *Phys. Rev. Lett.*, 95:061801, 2005.

[8] B. Aubert et al. Branching fraction and cp asymmetries of $b^0 \to k_s^0 k_s^0 k_s^0$. *Phys. Rev. Lett.*, 95:011801, 2005.

[9] M. T. Cheng et al. Letter of intent for a study of cp violation in b meson decays. *KEK report*, 2004. KEK-94-2.

[10] Lincoln Wolfenstein. Parametrization of the kobayashi-maskawa matrix. *Phys. Rev. Lett.*, 51:1945, 1983.

[11] W. M. Yao et al. Review of particle physics. *J. Phys.*, G33:1–1232, 2006.

[12] Michael Gronau. Cp violation in neutral b decays to cp eigenstates. *Phys. Rev. Lett.*, 63:1451, 1989.

[13] Zhi-zhong Xing. Cp violation in b/d → d+ d-, d*+ d-, d+ d*- and d*+ d*- decays. *Phys. Rev.*, D61:014010, 2000.

[14] T. Aushev et al. Search for cp violation in the decay b0 → d*+- d-+. *Phys. Rev. Lett.*, 93:201802, 2004.

[15] H. Miyake et al. Branching fraction, polarization and cp-violating asymmetries in b0 → d*+ d*- decays. *Phys. Lett.*, B618:34–42, 2005.

[16] B. Aubert et al. Measurement of time-dependent cp asymmetries in $b^0 \to d^{(*)\pm} d^{\mp}$ decays. *Phys. Rev. Lett.*, 95:131802, 2005.

[17] B. Aubert et al. Measurement of time-dependent cp asymmetries and the cp-odd fraction in the decay $b^0 \to d^{*+} d^{*-}$. *Phys. Rev. Lett.*, 95:151804, 2005.

[18] Y. Ushiroda. Belle silicon vertex detectors. *Nucl. Instrum. Meth.*, A511:6–10, 2003.

[19] P. Collins. Silicon detectors. *Lecture course at X ICFA school*, page 37, 2003.

[20] A. Abashian et al. The belle detector. *Nucl. Instrum. Meth.*, A479:117–232, 2002.

[21] G. Majumder et al. Evidence for b0 → d+ d- and observation of b- → d0 d- and b- → d0 d*- decays. *Phys. Rev. Lett.*, 95:041803, 2005.

[22] D. J. Lange. The evtgen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001.

[23] R. Brun et al. Geant 3.21. *CERN Report*, DD/EE/84-1, 1984.

[24] E. Nakano. Belle pid. *Nucl. Instrum. Meth.*, A494:402–408, 2002.

[25] K. Abe et al. Measurement of branching fractions for b → pi pi, k pi and k k decays. *Phys. Rev. Lett.*, 87:101801, 2001.

[26] Ronald Aylmer Fisher. The use of multiple measurements in taxonomic problems. *Annals Eugen.*, 7:179–188, 1936.

[27] K. Abe et al. Evidence for b0 → pi0 pi0. *Phys. Rev. Lett.*, 91:261801, 2003.

[28] B. Aubert et al. Measurement of branching fractions and cp-violating charge asymmetries for b meson decays to d(*) anti-d(*), and implications for the ckm angle gamma. *Phys. Rev.*, D73:112004, 2006.

[29] H. Albrecht et al. Search for hadronic b → u decays. *Phys. Lett.*, B241:278, 1990.

[30] A. Drutskoy et al. Observation of b → d(*) k- k(*)0 decays. *Phys. Lett.*, B542:171–182, 2002.

[31] A. Zupanc et al. Improved measurement of b0bar –> ds-d+ and search for b0bar –> ds+ds- at belle, 2007.

[32] H. Kakuno et al. Neutral b flavor tagging for the measurement of mixing- induced cp violation at belle. *Nucl. Instrum. Meth.*, A533:516–531, 2004.

[33] H. Tajima et al. Proper-time resolution function for measurement of time evolution of b mesons at the kek b-factory. *Nucl. Instrum. Meth.*, A533:370–386, 2004.

[34] Owen Long, Max Baak, Robert N. Cahn, and David Kirkby. Impact of tag-side interference on time dependent cp asymmetry measurements using coherent b0 anti-b0 pairs. *Phys. Rev.*, D68:034010, 2003.

[35] Gary J. Feldman and Robert D. Cousins. A unified approach to the classical statistical analysis of small signals. *Phys. Rev.*, D57:3873–3889, 1998.

[36] Yuval Grossman and Mihir P. Worah. Cp asymmetries in b decays with new physics in decay amplitudes. *Phys. Lett.*, B395:241–249, 1997.

Izjava

Spodaj podpisana Saša Fratina izjavljam, da sem v disertaciji predstavila rezultate lastnega znanstvenoraziskovalnega dela.

Saša Fratina

Ljubljana, 30.4.2007