Measurement of the CP-odd Fraction and Time-Dependent CP Violation in the Decay $B^0 \to D^{*+}D^{*-}$ at Belle

THÈSE N^{O} 4178 (2008)

PRÉSENTÉE LE 19 SEPTEMBRE 2008

À LA FACULTÉ DES SCIENCES DE BASE
LABORATOIRE DE PHYSIQUE DES HAUTES ÉNERGIES
PROGRAMME DOCTORAL EN PHYSIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Kim VERVINK

Licentiaat in de Natuurkunde, Universiteit Antwerpen, Belgique
et de nationalité belge

acceptée sur proposition du jury:

Prof. G. Gremaud, président du jury
Prof. O. Schneider, directeur de thèse
Prof. M. Calvi, rapporteur
Prof. M. Hazumi, rapporteur
Prof. U. Langenegger, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2008

Abstract

We present a study of the CP asymmetry in the $B^0 \to D^{*+}D^{*-}$ decay, using data produced by the electron-positron asymmetric-energy accelerator KEKB (Tsukuba, Japan) running at the $\Upsilon(4S)$ resonance. We use the full data sample (as of summer 2008) of $657\times10^6 B\overline{B}$ events, which corresponds to an integrated luminosity of $605\,\mathrm{fb}^{-1}$ recorded with the Belle detector.

From the distributions of the proper-time intervals between a neutral B meson decay to the $D^{*+}D^{*-}$ final state and another flavor-tagged B^0 or $\overline{B^0}$ decay, we measure the CP-violation parameters $\mathcal{A}$ and $\mathcal{S}$. $\mathcal{A}$ is a measure of the direct CP violation in the decay amplitudes while a non-zero $\mathcal{S}$ value indicates CP violation in the interference between the decay and the mixing amplitudes.

At the quark level the $B^0 \to D^{*+}D^{*-}$ decay is a $b \to c\bar{c}d$ transition, where the tree amplitude is Cabbibo-suppressed. The contribution of penguin amplitudes in this decay is estimated to be at the percent level [1]. When ignoring penguin corrections, the SM prediction for the CP parameters are $\mathcal{A}_{D^{*+}D^{*-}} = 0$ and $\mathcal{S}_{D^{*+}D^{*-}} = -\eta_{D^{*+}D^{*-}} \sin 2\beta$, where $\eta_{D^{*+}D^{*-}}$ is the CP eigenvalue of $D^{*+}D^{*-}$ and $\beta = \arg[-V_{cd}V_{cb}^*]/[V_{td}V_{tb}^*]$ is an observable phase related to the coefficients V_{ij} of the Cabibbo-Kobayashi-Maskawa quark-mixing matrix. Any large measured deviation with respect to prediction can be a sign of New Physics.

As the D^* particle is a vector-meson three different partial waves contribute to the final states: the S and D correspond to the CP-even state, while the P wave corresponds to a CP-odd state. To avoid too large a dilution of the CP asymmetry, the two CP components should be separated on a statistical basis. We measure the CP-odd fraction by performing a time-integrated angular analysis in the transversity basis and obtain:

$$R_\perp \;=\; 0.125 \pm 0.043(\mathrm{stat}) \pm 0.023(\mathrm{syst}),$$

which is consistent with the prediction from theory as well as the previous measurements.

We have taken into account the CP-odd dilution and the effect of incorrect flavor assignment to model the decay rates. The proper-time difference of the two B mesons is obtained using the reconstructed D-meson trajectories and the beam spot profile. A "null" measurement is performed by fitting the CP asymmetry of the control sample, $B^0 \to D^{(*)+}D_s^{(*)-}$ which gives $\mathcal{A} = -0.02 \pm 0.03(\mathrm{stat})$ and $\mathcal{S} = -0.07 \pm 0.04(\mathrm{stat})$ and is consistent with no CP violation. The lifetime fit to the $B^0 \to D^{*+}D^{*-}$ is used to verify the precision of the vertex resolution function. The measured lifetime is consistent with the world average value.

Finally, the CP study of the $B^0 \to D^{*+}D^{*-}$ decay gives an evidence of CP violation with a $3.5\,\sigma$ significance:

$$\begin{aligned} \mathcal{S}'_{D^{*+}D^{*-}} &= -0.96 \pm 0.25(\mathrm{stat})^{+0.12}_{-0.16}(\mathrm{syst}), \\ \mathcal{A}_{D^{*+}D^{*-}} &= +0.15 \pm 0.13(\mathrm{stat}) \pm 0.04(\mathrm{syst}). \end{aligned}$$

where $\mathcal{S}'_{D^{*+}D^{*-}} = \frac{1}{\eta}\mathcal{S}_{D^{*+}D^{*-}}$. This result is consistent with the SM prediction for tree decays.

Keywords: high-energy physics, KEK, Belle, Standard Model, New Physics, B physics, CP violation, angular analysis.

Résumé

Nous présentons l'étude de l'asymétrie CP dans la desintégration $B^0 \to D^{*+}D^{*-}$, à partir des données obtenues lors des collisions electron-positron fournies par l'accélérateur asymétrique KEKB à Tsukuba au Japon, à l'énergie de la résonance $\Upsilon(4S)$. Nous avons analysé l'ensemble complet des données enregistrées à l'aide du détecteur BELLE jusqu'à l'été 2008, constitué de 657×10^6 événements $B\overline{B}$ correspondant à une luminosité intégrée de $605\,\mathrm{fb}^{-1}$.

A partir des distributions d'intervalles de temps propre entre la désintégration d'un méson B neutre en $D^{*+}D^{*-}$ et celle de l'autre méson identifié comme un B^0 ou un $\overline{B^0}$, nous avons pu mesurer les paramètres $\mathcal{A}$ et $\mathcal{S}$ de violation CP pour ce canal. $\mathcal{A}$ est une mesure de la violation de CP directe, alors qu'une valeur non-nulle de $\mathcal{S}$ indique une violation de CP causée par l'interférence entre les amplitudes de désintégration et celui de mélange des B neutres.

Au niveau des quarks, la désintégration $B^0 \to D^{*+}D^{*-}$ est une transition $b \to c\bar{c}d$, pour laquelle l'amplitude du diagramme de Feynman au premier ordre (*tree diagram*) est réduite par le mécanisme CKM. La contribution des diagrammes de second ordre (*penguin diagram*) est estimée quant à elle à quelques pourcents [1]. Les prédictions du Modèle Standard pour les paramètres CP, en ignorant les corrections des diagrammes de second ordre, sont de $\mathcal{A}_{D^{*+}D^{*-}} = 0$ et $\mathcal{S}_{D^{*+}D^{*-}} = -\eta_{D^{*+}D^{*-}} \sin 2\beta$, où $\eta_{D^{*+}D^{*-}}$ est la valeur propre de CP de $D^{*+}D^{*-}$ et $\beta = \arg[-V_{cd}V_{cb}^*]/[V_{td}V_{tb}^*]$ est une phase observable liée aux éléments V_{ij} de la matrice de Cabibbo-Kobayashi-Maskawa décrivant le mélange des quarks. Des valeurs mesurées incompatibles avec les prédictions du Modèle Standard seraient un signe de "nouvelle physique".

Comme la particule D^* est un méson-vecteur, trois ondes partielles différentes contribuent à l'état final : les ondes S et D correspondent à un état de CP pair, alors que l'onde D correspond à un état de CP impair. Pour éviter une trop grande dilution de l'asymétrie CP, les deux composantes doivent être séparées sur une base statistique. Nous avons mesuré la fraction d'état propre de CP impair en effectuant une analyse angulaire intégrée sur le temps, dans la base de transversité, et nous avons obtenu :

$$R_\perp \;=\; 0.125 \pm 0.043(\mathrm{stat}) \pm 0.023(\mathrm{syst}),$$

ce qui est en accord avec les prédictions du Modèle Standard, ainsi qu'avec les résultats obtenus avec des mesures antérieures.

Nous avons pris en compte la dilution par les états impairs de CP ainsi que les effets dus à une mauvaise détermination de la saveur pour modéliser les taux de désintégration. La différence de temps propre entre les deux mésons B a été reconstruite à partir des trajectoires des mésons D et du profil du point d'interaction. Nous avons effectué une vérification mesurant l'asymétrie CP d'un échantillon de contrôle $B^0 \to D^{(*)+}D_s^{(*)-}$, pour lequel nous obtenons $\mathcal{A} = -0.02 \pm 0.03(\mathrm{stat})$ et $\mathcal{S} = -0.07 \pm 0.04(\mathrm{stat})$, ce qui est

compatible avec une asymétrie nulle. L'ajustement du temps de vie dans la désintégration $B^0 \to D^{*+}D^{*-}$ a été utilisé pour vérifier la précision de la fonction de résolution spatiale sur les points de desintégration. Le temps de vie mesuré est compatible avec la valeur moyenne mondiale.

Finalement, l'étude des paramètres CP dans la désintégration $B^0 \to D^{*+}D^{*-}$ donne une indication de violation CP avec une "significance" de $3.5\,\sigma$:

$$\begin{aligned} \mathcal{S}'_{D^{*+}D^{*-}} &= -0.96 \pm 0.25(\mathrm{stat})^{+0.12}_{-0.16}(\mathrm{syst}), \\ \mathcal{A}_{D^{*+}D^{*-}} &= +0.15 \pm 0.13(\mathrm{stat}) \pm 0.04(\mathrm{syst}). \end{aligned}$$

où $\mathcal{S}'_{D^{*+}D^{*-}} = \frac{1}{\eta}\mathcal{S}_{D^{*+}D^{*-}}$. Ces résultats sont compatibles avec les prédictions du Modèle Standard pour les diagrammes du premier ordre.

Mots-clés : physique des hautes énergies, KEK, Belle, Modèle Standard, Nouvelle Physique, physique de B, violation de CP, analyse angulaire.

Acknowledgments

This page is by far the one I've spent most time on writing and editing, as there are so many people who I wanted to sincerely thank for their contribution throughout these four years.

First of all I want to thank Olivier Schneider, my thesis director, who was a great boss all along. He gave me enough space and freedom to work in an independent way on subjects which interested me, while always being present when you need valuable advice. Thank you for correcting in depth (believe me) this thesis and for your patient explanations of particle physics over the years. He also made sure that I could work in collaboration with a post-doc who is an expert in CP violation, Belle software and deadlines: Tagir Aushev. Thank you very much, Tagir, for the discussions we had on how to make sense of the new figures and plots, for pushing me to get results, for motivating me, for getting me ready to present at conferences and for being as enthusiastic as I was when we saw the "CP wiggle" for the first time (Feb. 22nd at midnight!).

Furthermore I want to thank Aurelio Bay, Olivier Schneider, Erika and Esther for the smooth operation of this lab and for giving me the opportunity to attend different conferences, physics schools and workshops all around the world. Thanks go to Cedric for always patiently helping me out with any computer or printer problem.

The LPHE lab consists of a large group of fantastic colleagues and I believe for a lot of us, the 10 o'clock coffee breaks are a great motivation to get to work in the morning. Thank you all very much for the great atmosphere here. Mathias, Aurelie, Jeremie, Shirit, Marco, Nicolas ... it was great to spend these four years in your presence. Thank you for the discussions on physics, the good atmosphere in as well as outside the lab, for teaching together and for being each others allies in this sometimes frustrating world of research.

During my first year in Lausanne I had the chance to work in close contact with the EPFL technicians. M. Hertig, Alain, Raymond Bernard and Rodolphe always created a joyful atmosphere, merci les gars! During that time Matt Needham, Helge Voss and Jeroen Van Hunen guided me through the wonderful world of glue, screws, high voltage, XML, pointers and Panoramix. Thank you all for your help and dedication. Also the other members of the Silicon Tracker team (Olaf, Uli, Matt, Pablo ...) and the Nikhef crowd (Marcel, Jacopo...) created such an energy and team spirit that motivated us time after time to overcome the toughest hardware and software problems we encountered.

I am also indebted to my colleagues in Belle: Hazumi-san, Tagir, Sakai-san, Tom, Karim ... for their valuable and diplomatic comments during our many physics meetings. For the social life at Tsukuba Pablo, Tagir, Galina, Jean, Nicolas, Olivier, Jan, Sam... definately contributed to the quality of my stay there. Thank you for the fun trips, parties and dinners we had.

Finally I want to express my greatest gratitude to my family who always supported me in all my decisions, for their enthusiasm and curiosity about my work while making

sure that there is a safe port I can return to. Also Eileen and Peter Brown deserve a big thanks. Eileen, reading over this manuscript and fishing out the Germanic translations was a great help! And then there is one last person I want to thank, Andrew. Where do I start... thanks for the humongous support you gave me, for visiting me in Japan, for listening enthusiastically to my stories of work and for so much more.

for Walter and Roger,
who piqued my curiousity.

Contents

Introduction

PHYSICISTS have always been intrigued by symmetries in nature and their relation to conservation laws. For example, space and time symmetries are responsible for momentum and energy conservation. However it is often more exciting to discover that a symmetry of nature is broken. The work presented in this dissertation is a study of the symmetry between matter and anti-matter called CP symmetry. When this symmetry is broken it means that matter and anti-matter have a different behavior. It is the violation of this symmetry during the creation of the universe which makes it possible for large amounts of matter to exist nowadays [1]. Around $10^{-35}\,\mathrm{s}$ after the Big Bang the universe was in a very hot and energetic phase and cosmologists believe that equal numbers of matter particles and antimatter particles would have been created. When a particle interacts with its antiparticle they annihilate. At this early stage, when the universe was still very small, particles and antiparticles would be continuously interacting with each other. The annihilations and recreations from the large amounts of energy, would however always take place in equal amounts. As the universe cooled down and expanded further the particles would start decaying instead of interacting. It is at this moment that CP violation must have taken place. One particle (baryon) out of $\sim 10^{10}$ did not annihilate and survived. The universe which we observe today is just the little surplus of matter which was left after the Big Bang.

Particle physicists have been able to reproduce some of the decays which took place at the early age of the universe. By colliding particles with high energy they can study the interactions which take place after the collision, which resembles the reaction of the early stage of the universe. By doing this they observed in 1964 the CP-forbidden decay of a neutral K_L meson which decayed into two charged pions. In 1976 Weinberg, Salam and Glasgow were able to formulate a Quantum Mechanical theory which describes the particles and the reactions taking place at this subatomic level. They describe three fundamental forces: the strong, the electro-magnetic and the weak force as well as the elementary particles of which everything is made: leptons and quarks. This theory, called the Standard Model is based on the experiment evidence and theoretical calculations. Also CP violation is incorporated through a complex phase in the quark transition matrix of the weak interaction. The model foresaw the existence of neutral weak currents and when these were discovered at CERN, Weinberg, Salam and Glasgow received the Nobel Prize in physics.

However there are still some problems and issues in the Standard Model. It bothers physicists that the fourth fundamental force, gravity, is not taken into account or that there is a large number of independent parameters and that there exist a hierarchy problem. But more important for us, the CP violation explained by the Standard Model can only

[1] Note that matter and antimatter are just arbitrary names, we could also say that everything nowadays exists of antimatter and all the matter is gone.

explain the survival of one out of $\sim 10^{20}$ baryons instead of one out of $\sim 10^{10}$. Particle physicists are now convinced that there must be a theory beyond the Standard Model. The experimental results which do not conform with Standard Model predictions are now referred to as "New Physics".

We can look for hints of New Physics by measuring accurately the parameters of the CP-violation theory of the Standard Model. Besides the kaon system we now know that CP violation also takes place in the B system, where we expect large asymmetries and where theoretical predictions can be done in a clean or even in a model-independent way. Two large experiments are built to study B physics, the Belle experiment at the KEKB accelerator in Japan and the BaBar experiment at the PEP II collider in California. They are referred to as the B factories due to the large number of $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ events they have produced. In this work we have studied the CP asymmetry between the $B^0 \to D^{*+}D^{*-}$ and the $\overline{B^0} \to D^{*-}D^{*+}$ using the Belle detector. This decay can be produced through a so-called penguin diagram which contains an internal loop and is therefore an excellent place for new exotic particles to show up in a virtual manner. A measured asymmetry which has a large deviation with respect to the Standard Model prediction can be a sign of New Physics.

In the first chapter of this document the Standard Model and the theory of CP violation is explained, as well as the method to extract the CP-violating parameters. We give an overview of the current experimental status to compare with the Standard Model predictions.

In the second chapter the apparatus is described. This includes both the KEKB accelerator producing the e^+e^- collisions and the Belle detector installed around the interaction point. We end this chapter by describing some of the analysis tools which are used to obtain the information necessary to perform a precise study.

The next four chapters will describe in detail our analysis. First the $B^0 \to D^{*+}D^{*-}$ events are filtered out of a data sample of more than 600 million $B\overline{B}$ decays. This is done by imposing selection criteria which disfavor background events. Once we have obtained a satisfying pure subset of the data sample we can study the properties of our CP-violating decay mode. The current results on the branching fraction of $B^0 \to D^{*+}D^{*-}$ are largely dominated by the systematical uncertainty, contrary to the other measurements presented in this work. We have thus not aimed to perform a measurement of the branching fraction.

The $D^{*+}D^{*-}$ state can decay through three different partial waves, which have different CP properties. We therefore use an angular analysis in Chapter 4 to disentangle these states before performing the final asymmetry study.

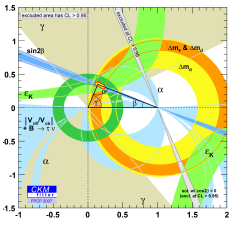
In Chapter 5 we examine the proper-time difference of the two B mesons in the event and study the detector resolution model which describes effect on the reconstructed vertices. Next a measurement of the lifetime of the B meson is performed to compare the result with the precise world average in order to test the accuracy of the resolution model we used.

Finally the CP asymmetry between the $B^0 \to D^{*+}D^{*-}$ and $\overline{B^0} \to D^{*-}D^{*+}$ is studied. We present our final result in Chapter 6, together with various checks including a "null" measurement with a control sample of $B^0 \to D_s^{(*)+}D^{(*)-}$ decays where no CP violation is expected. The significance and the systematic uncertainty of the result are discussed as well.

For completeness, Chapter 7 gives an overview of a cross-check analysis which was performed on a smaller data sample to compare with earlier results of this decay published by Belle.

Chapter 1

Theoretical Motivation



We introduce in this chapter the CP operator and its symmetry violation. The Standard Model of particle physics is described as well and CP violation is interpreted within this model, which gives rise to the CKM picture and the unitarity triangle. We start by studying CP violation in the general neutral meson-antimeson system but then work towards a more concrete decay which is studied in this work. An overview of the latest experimental results is given as well.

1.1 Symmetries

1.1.1 Introduction

Physicists have always been interested in investigating symmetries in nature. Noether's theorem [2] published in 1918, shows that symmetries lie at the basis of any conservation law. For example, the conservation of energy is a consequence of the invariance of physical laws under a time shift. The idea behind the study presented in this thesis work is the study of the violation of a discrete symmetry called CP. This can be studied by comparing the properties of decays of a fundamental particle such as its lifetime, mass, decay products etc with their CP conjugated state. The CP operator is the product of the parity and the charge-conjugation operator. CP is also closely related to the T operator which reverses time. The different operators have the following characteristics.

Parity operator

The parity operator reflects the space. Parity holds in particle physics when the mirror image of a process is indistinguishable from a real process. For example:

$$(\pi^+ \rightarrow \mu_R^+ + \nu_L) \xrightarrow{P} (\pi^+ \rightarrow \mu_L^+ + \nu_R).$$

Parity conservation implies that nature makes no distinction between right- and left handed rotations or between opposite sides of a particle. Thus two particles rotating in opposite direction μ_L and μ_R are identical if parity conservation holds. Parity violation was discovered by Chien-Shiung Wu [3] in 1957, using nuclear β decays. The emitted electron

decays preferably in the direction of the nuclear spin. In the parity-conjugated process the emitted electron will decay to the opposite direction. The results showed that one of the two directions is favored which violates the parity symmetry. Until then it was believed that parity conservation was one of the fundamental geometric conservation laws in nature.

Charge-conjugation operator

The charge-conjugation operator transforms a particle into its anti-particle and vice versa. For example:

$$(\pi^+ \to \mu_R^+ + \nu_L) \xrightarrow{C} (\pi^- \to \mu_R^- + \overline{\nu}_L).$$

Charge and Parity operator

The right-handed neutrinos (ν_R) in the example of parity exchange, or left-handed anti-neutrinos ($\overline{\nu}_L$) in the example of charge exchange, have never been observed in nature. But the combination of a parity and a charged transformation process have been observed:

$$(\pi^+ \to \mu_R^+ + \nu_L) \xrightarrow{CP} (\pi^- \to \mu_L^- + \overline{\nu}_R).$$

Also in the β decays, when applying both the P and C operator, everything becomes symmetric again. Historically CP symmetry was proposed to restore order after the discovery of P violation in the β system. CP conservation means that exchanging all particles with their antiparticles is equivalent to taking the mirror image of all processes.

C, P, T, CP, CT, PT and CPT symmetries are conserved in the strong and electromagnetic interactions. CPT is a conserved fundamental symmetry. Therefore the lifetime and mass of a particle and antiparticle must be equal. No experimental evidence has been able to disagree with this. However, a CP violation has been observed in the weak decays of neutral K or B mesons.

In this work CP violation in weak interactions in the B sector is studied by looking at differences between a decay of a B particle and its anti-particle.

1.1.2 Importance of CP violation on a cosmological scale

We believe that our universe today is made out of matter, as there has been no observation of antinuclei in cosmic rays or γ rays from $p\overline{p}$ annihilation in space. If symmetries would have held throughout the creation of the universe, the pure energy state that the big bang created would have condensed to equal amounts of matter and antimatter. But as matter and antimatter annihilate to photons, we would be left with only photons and no more matter or antimatter. This is definitely not the case. A force must have governed which violated the symmetry between matter and antimatter. The measured ratio of baryons [1] in the universe with respect to photons from annihilation processes (cosmic microwave background radiation); this is of the order of 10^{-10}. Which means that one out of 10^{10}

[1] a baryon is a hadron composed of three quarks, see Section 1.2

baryons did not annihilate. According to Sakharov [4](1967) the baryon inequality can be caused if the following three conditions are satisfied:

- there exists a transformation which violates the baryon number conservation.
- the C and CP symmetries need to be violated, such that the partial decay widths are different.
- the CP violation has to take place outside of thermal equilibrium, such that no reversing reaction can install the initial state.

or if CPT is violated. There is no experimental evidence so far of CPT violation. Also baryon number violation has not been observed experimentally. As will be seen in Section 1.2, CP violation is incorporated in the theoretical model of particle physics (the Standard Model) and has experimentally been observed in weak decays. The asymmetry can be produced outside thermal equilibrium. It seems thus that CP violation could be a good candidate to explain the matter antimatter asymmetry in the universe. However both the experimental observation and the theoretical prediction from the Standard Model can just explain the existence of one galaxy but not the entire universe. This is one of the reasons why physics beyond the Standard Model is expected.

1.2 The Standard Model of particle physics

The Standard Model (SM), formulated between 1970 and 1973, is a theory that describes the elementary particles and the fundamental forces which act on them. The theory is consistent with both quantum mechanics and special relativity. In this section we will give a brief introduction to the Standard Model. For more detailed explanations, the reader is referred to [5]. The three fundamental forces described in the SM are the strong, electromagnetic and weak force. Gravity however falls outside this model. Still, because all experimental results on these three forces so far have been consistent with the predictions of the Standard Model, it is a very successful theory. Some issues left which the Standard Model fails to accommodate are discussed in Section 1.2.2.

1.2.1 Brief overview of the Standard Model

The Standard Model is based on a gauge theory of $\mathrm{SU(3)_C \otimes SU(2)_L \otimes U(1)_Y}$ groups. The $SU(3)_C$ group describes the **strong interaction** by Quantum Chromo Dynamics (QCD). The strong force is governed by massless gluons as mediator particles and acts on the color charge of particles. A gluon itself carries two colors and as there exist three colors in total (red, blue and green) eight different gluons can exist with different color combinations (ignoring the color singlet). The $SU(2)_L$ and U_Y fields form together a unified theory of the electroweak interactions. The **electro-magnetic interactions** are described by Quantum Electro Dynamics (QED) and mediated through massless photons, acting on electrical charge. The **weak interactions** exist through charged current interactions mediated by the $W^{\pm}$ boson, while neutral current interactions take place through the Z^0 boson. The weak interaction is the weakest force of the three and acts on the flavor of quarks and leptons.

Next we will describe the different particles in the Standard Model and their classification. The elementary particles of nature all have spin $1/2$ and are thus fermions.

Amongst the fermions two different classes are made to distinguish quarks from leptons. Both quarks and leptons exist in three so-called generations, each corresponding to a different mass scale. For quarks a generation consists of an up-type quark and a down-type quark. Up-type quarks have electrical charge $+2/3$ and they can be the up quark (u), charm quark (c) or top quark(t). The down-type quarks have electrical charge $-1/3$ and are respectively the down quark (d), strange quark (s) and bottom quark (b). Table 1.1 summarizes the different quarks and their corresponding masses. For leptons the same type of classification can be made, according to their mass. Each doublet is made from a charged lepton and a neutral neutrino as shown in the right side of Table 1.1.

Gen.	Quark	Mass [$\mathrm{MeV/c^2}$]	Charge[e]	Lepton	Mass [$\mathrm{MeV/c^2}$]	Charge[e]
1	u	$1.5\,\mathrm{to}\,3.0$	$2/3$	e	0.511	-1
1	d	$3.0\,\mathrm{to}\,7.0$	$-1/3$	ν_e	$< 2 \times 10^{-6}$	0
2	c	1250 ± 90	$2/3$	μ	105.7	-1
2	s	95 ± 25	$-1/3$	ν_μ	< 0.19	0
3	t	$(174 \pm 3) \times 10^3$	$2/3$	τ	1777.0	-1
3	b	4200 ± 70	$-1/3$	ν_τ	< 18.2	0

Table 1.1: Elementary fermions of the Standard Model. The values are taken from [6].

Using these fundamental bricks larger particles can be constructed. Quarks don't exist in an isolated state but combine in groups of two or three to form a real physical particle, called a hadron. Elements containing two quarks, more precisely, one quark and one anti-quark, are called mesons, such as the K^0 meson which exists as a $d\bar{s}$ combination or the B^0 meson ($b\bar{d}$) which will play an important role in this work. Particles containing three quarks are called baryons, like the proton which contains two up quarks and one down quark.

The unification of the electro-magnetic and the weak theory to the electroweak theory is described by the $U(1)_Q$ group. The cost of this unification is the introduction of the Higgs particle, the only particle in the Standard Model which has not yet been observed in experiments. The presence of a Higgs field causes a spontaneous symmetry breaking which introduces masses to the $W^\pm$ and Z^0 bosons. The masses of fermions are also created through the spontaneous symmetry breaking, as they interact through the Yukawa coupling with the Higgs field. This coupling however is not diagonal, meaning that the mass eigenstates are not the same as the eigenstates in the weak interaction. The link between these two sets of eigenstates for the quarks will be described in Section 1.4.

1.2.2 Open issues

All experimental results are consistent with the Standard Model. The Standard Model even predicted the existence of particles which were only discovered later, for example the top quark, which was discovered in 1995 [7] [8]. Still one particle needs to be discovered, the Higgs particle. Besides this there are still some open issues, which make physicists believe that there should be a more complete and general theory behind the Standard Model. The main problems that are not explained are:

- The large number (19) of independent parameters.

- The hierarchy problem: the scale at which the parameters are measured experimentally is very different from the scale at which the Lagrangian operates. This can be due to renormalization problems, which are hard to calculate as the quantum corrections are divergent. Still the large scale difference is a disturbing feature in the Standard Model.
- It is also not clear from the SM theory why there are three generations of fermions corresponding to different mass regions.
- In order to explain the imbalance of matter and antimatter in nature CP violation larger than that of the Standard Model is needed.
- The Standard Model can only explain the survival of one baryon out of 10^{20} while the observed quantity of the matter in the universe, requires that one baryon out of 10^{10} survived instead of having been annihilated.
- There is no explanation in the Standard Model on the presence of dark matter and dark energy.
- Gravity is not explained by the Standard Model.

Several extensions beyond the Standard Model have been proposed by theorists, such as the Super Symmetry model (SUSY) or models which suggest the existence of additional dimensions of space, but none of them have yet been proved experimentally.

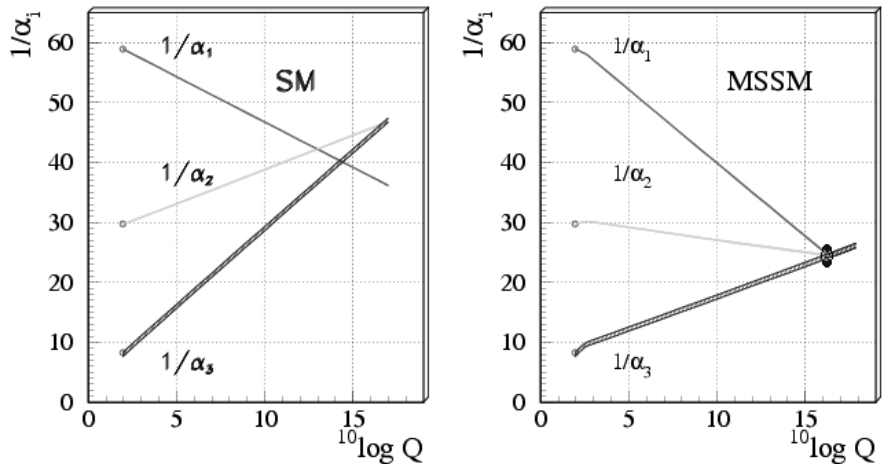


Figure 1.1: Running coupling constants of the three forces converging to a single point at a high energy scale (figure from [6]).

The coupling constants of the three forces seem to converge to a single value as energy increases ($\approx 10^{16}\,\mathrm{GeV}$), leading to the idea of a Grand Unified Theory (GUT), which unifies the three forces at higher energies. However the point where the coupling constants of the three forces come together is not exactly at one point. This problem can be solved by SUSY, which relates each fermion (boson) of the Standard Model with a new super-symmetric boson (fermion). The Super Symmetry model also solves the hierarchy problem.

The problem with GUT however is that it predicts the existence of a mediator particle which would cause the proton to decay, something that so far has never been observed.

1.3 General formalism of CP violation

In the following sections we describe the basic formalism of CP violation, first by studying the neutral particle-antiparticle system and its time-dependent evolution, from a quantum mechanics point of view. The mixing, decay mechanism and the possible appearances of CP violation will be discussed in the next part and then applied to the more specific case i.e. the neutral B_d system: the relevant system in this thesis. Finally the specific properties of CP violation in the $B^0 \to D^{*+}D^{*-}$ system will be shown.

1.3.1 The neutral meson-antimeson system

Consider a meson $|P\rangle$ and its antimeson $|\overline{P}\rangle$ at rest. Both states are eigenstates of the strong and electromagnetic interactions such that:

$$(H_s + H_{em})|P\rangle = m|P\rangle \qquad (H_s + H_{em})|\overline{P}\rangle = \overline{m}|\overline{P}\rangle.$$

where m and $\overline{m}$ are the masses of $|P\rangle$ and $|\bar{P}\rangle$ respectively. H_{em} and H_s represent the Hamiltonians that govern the electromagnetic and strong interactions. The flavor of each state is well defined:

$$F|P\rangle = +|P\rangle \qquad F|\overline{P}\rangle = -|\overline{P}\rangle.$$

H_s and H_{em} are assumed to be invariant under CPT, which makes $m = \overline{m}$. Let both the meson and the antimeson be connected through the CP operator by the arbitrary CP phase θ_{CP}.

$$CP|P\rangle = e^{i\theta_{CP}}|\overline{P}\rangle \quad \overset{CP^2=\mathbb{1}}{\Longrightarrow} \quad CP|\overline{P}\rangle = e^{-i\theta_{CP}}|P\rangle.$$

When the T operator acts on a particle at rest, it only gives rise to an extra phase, but doesn't change the state of the particle:

$$T|P\rangle = e^{i\theta_T}|P\rangle \qquad T|\overline{P}\rangle = e^{i\bar{\theta}_T}|\overline{P}\rangle$$

and $2\theta_{CP} = \bar{\theta}_T - \theta_T$ when $CPT|P\rangle = TCP|P\rangle$. Let $|f\rangle$ be any eigenstate of the strong and electromagnetic interactions with eigenvalue E_f,

$$(H_s + H_{em})|f\rangle \;=\; E_f|f\rangle,$$

that is accessible as final state in the weak decay of $|P\rangle$ or $|\overline{P}\rangle$. Assume that only the strong and electromagnetic interactions are present at $t = 0$. $|P\rangle$ and $|\bar{P}\rangle$ are considered stable under these forces. At $t > 0$, the weak interaction comes into play. This will allow both states to mix and eventually decay into a final state $|f\rangle$. A state $|\psi\rangle$ of this system can be written in a general way as a linear combination of all possible final states $|f\rangle$ and initial states $|P\rangle$ and $|\bar{P}\rangle$:

$$|\psi(t)\rangle \;=\; a(t)|P\rangle + b(t)|\bar{P}\rangle + \sum_f c_f(t)|f(t)\rangle, \tag{1.1}$$

where $a(t)$, $b(t)$ and $\sum_f c_f(t)$ are the corresponding probability amplitudes. Because we have no decay products present at $t = 0$:

$$\begin{aligned} t = 0: &\quad |a(0)|^2 + |b(0)|^2 = 1 \quad |c_f(0)|^2 = 0 \\ &\text{and}: \\ t > 0: &\quad |a(t)|^2 + |b(t)|^2 \leq 1 \quad |c_f(t)|^2 \geq 0. \end{aligned}$$

The general state $|\psi\rangle$ can be found by solving the Schrödinger equation:

$$-i\frac{d}{dt}|\psi\rangle = (H_s + H_{em} + H_w)|\psi\rangle,$$

where H_w is the Hamiltonian function of the weak interaction. Assuming that the weak interactions between the final states are small compared to the strong and electromagnetic interactions, one can ignore the term:

$$\langle f|H_w|f'\rangle$$

and use the Wigner-Weiskopf approximation [9]. Terms with a second order H_w component as well as terms containing $\langle P|H_w|f\rangle$ can be ignored because we assume weak interactions between the final state and initial state are small. In this way the term containing the c_f factor will disappear from the equations as it contains a negligible H^2_{eff} term, where H_{eff} denotes the effective Hamiltonian. For the detailed calculations we refer to [10]. The equation describing the meson-antimeson system in the rest frame is now simplified to a two-state solution:

$$|\psi(t)\rangle \;=\; a(t)|P\rangle + b(t)|\bar{P}\rangle. \tag{1.2}$$

The time-evolution of the probabilities $a(t)$ and $b(t)$ becomes:

$$i\frac{d}{dt}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = H_{\text{eff}}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \Lambda\begin{pmatrix} a(t) \\ b(t) \end{pmatrix}, \tag{1.3}$$

where

$$\Lambda = M - \frac{i}{2}\Gamma = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}. \tag{1.4}$$

The 2X2 matrix M is called the mass matrix. Its elements are:

$$M_{11} \;=\; m + \langle P|H_w|P\rangle + \sum_f \mathcal{P}\Big(\frac{\langle P|H_w|f\rangle\langle f|H_w|P\rangle}{m - E_f}\Big), \tag{1.5}$$

$$M_{22} \;=\; m + \langle \overline{P}|H_w|\overline{P}\rangle + \sum_f \mathcal{P}\Big(\frac{\langle \overline{P}|H_w|f\rangle\langle f|H_w|\overline{P}\rangle}{m - E_f}\Big), \tag{1.6}$$

where m is the mass of $|P\rangle$ and $|\bar{P}\rangle$ and $\mathcal{P}$ denotes the principal part. The elements on the diagonal of the decay matrix, Γ, are:

$$\begin{aligned} \Gamma_{11} \;&=\; 2\pi\sum_f |\langle P|H_w|f\rangle|^2\delta(m_0 - E_f), \\ \Gamma_{22} \;&=\; 2\pi\sum_f |\langle \overline{P}|H_w|f\rangle|^2\delta(m_0 - E_f), \end{aligned} \tag{1.7}$$

The off-diagonal elements are given by:

$$\begin{aligned} M_{12} \;&=\; \langle P|H_w|\overline{P}\rangle + \sum_f \mathcal{P}\Big(\frac{\langle P|H_w|f\rangle\langle f|H_w|\overline{P}\rangle}{m - E_f}\Big), \\ \Gamma_{12} \;&=\; 2\pi\sum_f \langle P|H_w|f\rangle\langle f|H_w|\overline{P}\rangle\delta(m_0 - E_f). \end{aligned} \tag{1.8}$$

the factor $\delta(m_0 - E_f)$ in the decay matrix elements ensures energy conservation. The decay states $|f\rangle$ represent therefore real physical states. In the mass matrix equation the sum is taken over all possible $|f\rangle$ states, real or virtual.

If H_{eff} is Hermitian and invariant under T, CP and CPT then the following conditions must be met:

$$\begin{array}{ll} \text{if } CPT \text{ is conserved then :} & \Lambda_{11} = \Lambda_{22} \\ \text{if } CP \text{ is conserved then :} & \Lambda_{11} = \Lambda_{22} \text{ and } |\Lambda_{12}| = |\Lambda_{21}| \\ \text{if } T \text{ is conserved then :} & |\Lambda_{12}| = |\Lambda_{21}| \end{array}$$

As we assume CPT conservation we can define the elements on the diagonal as follows: $M_{11} = M_{22} = M,\ \Lambda_{11} = \Lambda_{22} = \Lambda,\ \Gamma_{11} = \Gamma_{22} = \Gamma$.

The eigenvalues of H_{eff} are obtained by solving Eq. (1.3) and using the Ansatz

$$a(t) \quad = \quad C_+ e^{-i\lambda_+ t} + C_- e^{-i\lambda_- t}, \tag{1.9}$$

where $\lambda_\pm$ are the eigenvalues of Λ which are obtained from the equation:

$$\lambda_\pm^2 - 2\lambda_\pm \Lambda - \left(\Lambda_{21}\Lambda_{12} - \Lambda^2\right) = 0. \tag{1.10}$$

such that the eigenvalues of Λ are:

$$\lambda_\pm = \Lambda \pm \sqrt{\Lambda_{21}\Lambda_{12}}. \tag{1.11}$$

Consequently we can find:

$$b(t) \quad = \quad \frac{q}{p}(C_+ e^{-i\lambda_+ t} - C_- e^{-i\lambda_- t}), \tag{1.12}$$

with

$$\frac{q}{p} = -\sqrt{\frac{\Lambda_{21}}{\Lambda_{12}}}. \tag{1.13}$$

and $|q|^2 + |p|^2 = 1$. The eigenstates of Λ are defined as $|P_\pm\rangle$, which are fixed by the diagonalizing of the effective Hamiltonian:

$$H_{\text{eff}}|P_\pm\rangle = \lambda_\pm |P_\pm\rangle. \tag{1.14}$$

$|P_\pm\rangle$ are now physical eigenstates which can be related back to $|P\rangle$ and $|\bar{P}\rangle$ by:

$$|P_\pm\rangle = |p|\big(|P\rangle \pm \frac{q}{p}|\bar{P}\rangle\big). \tag{1.15}$$

The mass and decay width of $|P_\pm\rangle$ are related to its eigenvalue:

$$\lambda_\pm = m_\pm - \frac{i}{2}\Gamma_\pm, \tag{1.16}$$

where

$$m_\pm = \Re\lambda_\pm \tag{1.17}$$

$$\Gamma_\pm = -2\Im\lambda_\pm. \tag{1.18}$$

The solution of (1.9) and (1.12) for an initial state ($t = 0$) which is purely $|P\rangle$ is

$$a(0) = C_+ + C_- = 1, \tag{1.19}$$

$$b(0) = \frac{q}{p}(C_+ - C_-) = 0, \tag{1.20}$$

which means that $C_+ = C_- = 1/2$.
For the sake of visibility let us define

$$h_\pm(t) = \frac{1}{2}\big(e^{-i\lambda_+ t} \pm e^{-i\lambda_- t}\big). \tag{1.21}$$

The time evolution of an initially pure $|P\rangle$ state can be written as:

$$|P(t)\rangle = h_+(t)|P\rangle + \frac{q}{p}h_-(t)|\bar{P}\rangle \tag{1.22}$$

$$= \frac{1}{2|p|}\Big(e^{-i\lambda_+ t}|P_+\rangle + e^{-i\lambda_- t}|P_-\rangle\Big). \tag{1.23}$$

$|P(t)\rangle$ is now a superposition of the original $|P\rangle$ and $|\bar{P}\rangle$ state. An oscillation between the two neutral states takes place, with a time-dependent probability amplitude of being in the $|P\rangle$ or $|\bar{P}\rangle$ state equal to $h_+(t)$ or $\frac{q}{p}h_-(t)$ respectively. If the initial state would have been $|\bar{P}\rangle$ then $a(0) = C_+ + C_- = 0$, $b(0) = \frac{q}{p}(C_+ - C_-) = 1$ and $C_+ = -C_- = 1/2$. The time-dependent evolution of the antiparticle would then have been:

$$|\overline{P}(t)\rangle = \frac{p}{q}h_-(t)|P\rangle + h_+(t)|\bar{P}\rangle \tag{1.24}$$

$$= \frac{p}{2|p|q}\Big(e^{-i\lambda_+ t}|P_+\rangle - e^{-i\lambda_- t}|P_-\rangle\Big). \tag{1.25}$$

1.3.2 CP violation in the oscillation between particle and antiparticle

So far we developed a general solution for the neutral particle-antiparticle system, which gave rise to an oscillation under the weak force. Next we will show how at this level, before the decaying process occurs, a possible CP violation can take place.

The probability of obtaining a state $|P\rangle$ at time t when the initial state was $|P\rangle$ is given by:

$$|\langle P|P(t)\rangle|^2 = |h_+(t)|^2 = \frac{1}{4}\big(e^{-\Gamma_+ t} + e^{-\Gamma_- t} + 2e^{-\overline{\Gamma}t}\cos(\Delta m t)\big). \tag{1.26}$$

with $\Delta m = m_- - m_+$ and $\overline{\Gamma} = \frac{\Gamma_+ + \Gamma_-}{2}$. Oppositely the probability for an initial particle $|P\rangle$ to oscillate to a $|\bar{P}\rangle$ is:

$$|\langle \overline{P}|P(t)\rangle|^2 = |\frac{q}{p}h_-(t)|^2 = \frac{|q|^2}{4|p|^2}(e^{-\Gamma_+ t} + e^{-\Gamma_- t} - 2e^{-\overline{\Gamma}t}\cos(\Delta m t)). \tag{1.27}$$

Remember that both $|P\rangle$ and $|\bar{P}\rangle$ are flavor eigenstates and the CP conjugate of each other. They are however not physical states with a well defined mass or decay width. The probability of finding the physical state $P_\pm$ at time t, when starting off with a pure $|P\rangle$ state is:

$$|\langle P_\pm|P(t)\rangle|^2 = \frac{1 + |\frac{q}{p}|^2}{4}e^{-\overline{\Gamma}t}. \tag{1.28}$$

In order to conserve the CP symmetry during the oscillation the following equality needs to be satisfied:

$$|\langle\overline{P}|P(t)\rangle|^2 = |\langle P|\overline{P}(t)\rangle|^2 \tag{1.29}$$

or

$$\left|\frac{q}{p}h_-(t)\right|^2 = \left|\frac{p}{q}h_-(t)\right|^2 \Longrightarrow \left|\frac{q}{p}\right|^2 = \left|\frac{p}{q}\right|^2 \Longrightarrow \left|\frac{q}{p}\right| = 1. \tag{1.30}$$

We can conclude that in order to have CP symmetry conservation in the oscillation of a neutral particle-antiparticle system, the following condition needs to be met:

$$|p| = |q|. \tag{1.31}$$

The violation of this condition is sometimes called indirect CP-violation.

1.3.3 CP violation in the decay

Imagine now a particle $|P\rangle$ decaying to a final state $|f\rangle$ and an antiparticle $|\bar{P}\rangle$ decaying to $|\overline{f}\rangle$. The two decay amplitudes corresponding to this process are:

$$A_f = \langle f|H_{\text{eff}}|P\rangle, \tag{1.32}$$
$$\overline{A}_{\overline{f}} = \langle\overline{f}|H_{\text{eff}}|\overline{P}\rangle. \tag{1.33}$$

The two conditions to be satisfied in order to conserve CP symmetry in the decay are:

$$|A_f| = |\overline{A}_{\overline{f}}|, \tag{1.34}$$
$$|A_{\overline{f}}| = |\overline{A}_f|. \tag{1.35}$$

If the final state is a CP eigenstate, so that $|f\rangle = |\overline{f}\rangle$, the above equations simplify to:

$$|A_f| = |\overline{A}_f|. \tag{1.36}$$

The violation of this condition is called direct CP-violation.

1.3.4 CP violation in the interplay between oscillation and decay

When combining the oscillation and decay mechanism, a third source of CP violation can occur due to the interplay. This process occurs when the final state is accessible both from the particle and antiparticle. This is always the case when the final state is a CP eigenstate. But also processes like $B^0 \to D^{\pm}\pi^{\mp}$ have a final state accessible by both the particle B^0 and $\overline{B^0}$. These processes are schematically represented in Figure 1.2. The different probability amplitudes of the oscillation and decay mechanisms are added on the figure. The total probability amplitude of obtaining a final state f when the initial state is $|P\rangle$ is

$$P_f(t) \propto A_f h_+(t) + \overline{A}_f \frac{q}{p} h_-(t). \tag{1.37}$$

Similarly, the total probability amplitude to get the final state f at time t from an initial state $|\bar{P}\rangle$ at time $t = 0$ is:

$$\overline{P}_f(t) \propto \overline{A}_f h_+(t) + A_f \frac{p}{q} h_-(t). \tag{1.38}$$

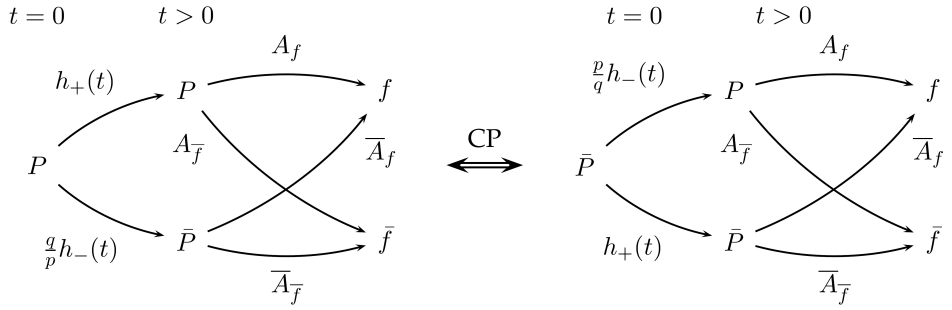


Figure 1.2: Oscillation and decay interference for an initial particle $|P\rangle$ (left) and the CP conjugated process (right).

The parameter that can be measured is the time-dependent decay-rate $\Gamma_f(t) = |P_f(t)|^2$. CP violation occurs when $\Gamma_f(t) \neq \overline{\Gamma}_f(t)$. For the two types of processes described above the time-dependent decay-rate becomes:

$$\Gamma(t) = |A_f|^2|h_+(t)|^2 + |\overline{A}_f|^2\Big|\frac{q}{p}\Big|^2|h_-(t)|^2 + 2\Re\big(A_f^*\overline{A}_f\frac{q}{p}h_+^*(t)h_-(t)\big), \quad (1.39)$$

$$\overline{\Gamma}(t) = |\overline{A}_f|^2|h_+(t)|^2 + |A_f|^2\Big|\frac{p}{q}\Big|^2|h_-(t)|^2 + 2\Re\big(\overline{A}_f^*A_f\frac{p}{q}h_+^*(t)h_-(t)\big). \quad (1.40)$$

The first term in both equations describes the decay mechanism. When

$$|A_f|^2 \neq |\overline{A}_f|^2, \quad (1.41)$$

the CP symmetry is broken in the decay amplitude. The second term in Eq. (1.39) and (1.40) comes from the oscillation between the particle and antiparticle before decaying. Here CP violation in the oscillation can occur when

$$\Big|\frac{q}{p}\Big| \neq 1, \quad (1.42)$$

as said before.

The third term describes the interplay between the decay and the oscillation. Even when there is no CP violation in the oscillation or the decay, there can still be a difference in the total decay rate, $\Gamma(t) \neq \overline{\Gamma}(t)$. To see this the third term is further developed.

$$\begin{aligned}\Re\big(A_f^*\overline{A}_{\overline{f}}\frac{q}{p}h_+^*(t)h_-(t)\big) &= \Re\big(A_f^*\overline{A}_{\overline{f}}\frac{q}{p}\big)\Re\big(h_+^*(t)h_-(t)\big) \\ &- Im\big(A_f^*\overline{A}_{\overline{f}}\frac{q}{p}\big)Im\big(h_+^*(t)h_-(t)\big), \quad (1.43)\\ \Re\big(\overline{A}_{\overline{f}}^*A_f\frac{p}{q}h_+^*(t)h_-(t)\big) &= \Re\big(A_f^*\overline{A}_{\overline{f}}\frac{q^*}{p}\big)\Re\big(h_+^*(t)h_-(t)\big) \\ &+ Im\big(A_f^*\overline{A}_{\overline{f}}\frac{p}{q}\big)Im\big(h_+^*(t)h_-(t)\big). \quad (1.44)\end{aligned}$$

It can be seen from these equations that even when there is no direct CP violation $Im(A_f^*\overline{A}_{\overline{f}}\frac{q}{p})$ can be different from zero, which implies a difference in the decay rates between the CP conjugated processes. This third type of CP violation is often referred to as mixing-induced CP violation.

1.4 CP violation in the Standard Model

So far we have described the CP violation mechanism from a quantum mechanical point of view. The relevant force for CP violation is the weak interaction, which can occur through a neutral current, with a Z boson or through the charged current, with a $W^{\pm}$ mediator. It is in the latter case that CP violation can occur.

The weak charged current can change an up-type (left) quark (U) to a down-type (left) quark (D). These transitions are governed by nine coupling constants V, with different strengths. The corresponding Feynman diagram as well as its CP conjugate are shown in Figure 1.3. It is the complex nature of the coupling V_{UD} that gives rise to CP violation.

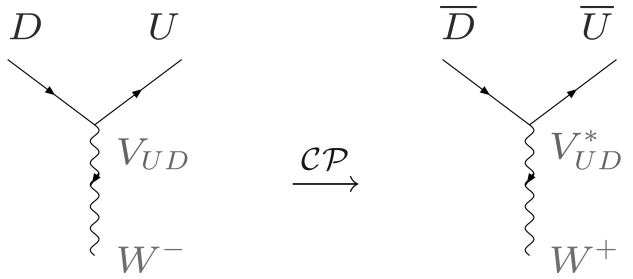


Figure 1.3: Charged-current quark-level interaction process D → U W^- and its CP conjugate in the Standard Model.

The Lagrangian describing these charged-current interactions can be written as

$$\mathcal{L} = \frac{-g}{\sqrt{2}}(V_{ij}\overline{U}_{Li}\gamma^{\mu}D_{Lj}W^{\dagger}_{\mu} + V^{*}_{ij}\overline{D}_{Li}\gamma^{\mu}U_{Lj}W_{\mu}), \tag{1.45}$$

and the CP-conjugated Lagrangian is:

$$\mathcal{L}_{CP} = \frac{-g}{\sqrt{2}}(V_{ij}\overline{D}_{Li}\gamma^{\mu}U_{Lj}W_{\mu} + V^{*}_{ij}\overline{U}_{Li}\gamma^{\mu}D_{Lj}W^{\dagger}_{\mu}). \tag{1.46}$$

In order to have CP conservation $\mathcal{L}$ needs to be equal to $\mathcal{L}_{CP}$, or $V^{*}_{ij} = V_{ij}$.

To generate CP violation there needs to be at least one phase in the V matrix [11]. A general unitary matrix has N^2 independent parameters. We have the freedom to rephase $2N-1$ quark fields which makes the number of independent parameters of an $N \times N$ unitary quark-mixing matrix:

$$N^2 - (2N-1) = \underbrace{\frac{1}{2}N(N-1)}_{\text{Euler angles}} + \underbrace{\frac{1}{2}(N-1)(N-2)}_{\text{Complex phases}} = (N-1)^2.$$

In order to accommodate CP violation in the Standard Model at least three generations of quarks are needed, as this will give one complex phase in the V matrix. A more detailed calculation of the quark-field rephrasing can be found in [10].

1.4.1 CKM matrix and unitarity triangle

The 2×2 quark-mixing matrix originally proposed by Cabibbo was extended in 1973 by Kobayashi and Maskawa to a 3×3 matrix known today as the Cabibbo-Kobayashi-Maskawa (CKM) matrix [12, 13].

$$V_{\rm CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{1.47}$$

This matrix is usually parametrized in a way to incorporate the constraints of unitary. This can be done by rephasing invariance, i.e. the possibility to change the overall phase of any row or any column of the CKM matrix, without changing the physics contained in that matrix. The CKM matrix can be parametrized by three rotation angles and one phase, as explained in the previous paragraph. However the most commonly used parametrization is the Wolfenstein parametrization [14]. Experimental results showed that the elements of the matrix could be grouped according to their size. Wolfenstein realized this by noticing that most bottom quarks decay into a charm quark which means that $|V_{cb}| \gg |V_{ub}|$. He then defined the parameter $\lambda = V_{us}$ and rewrote the CKM matrix using λ as an expansion parameter, while holding approximately the unitarity:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \approx \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \tag{1.48}$$

The value for λ is small, $\lambda \approx 0.22$. Experiment showed that $|V_{cb}| \approx |V_{us}|^2$ which makes $A \approx 1$. And because $\frac{|V_{ub}|}{|V_{cb}|} \approx \lambda/2$, ρ and η should be smaller than one. In this parametrization the unitarity relations are satisfied up to order λ^3. The higher-order terms of the CKM matrix up to $\mathcal{O}(\lambda^5)$ are:

$$\begin{pmatrix} -\frac{1}{8}\lambda^4 & 0 & 0 \\ A^2\lambda^5(\frac{1}{2} - \rho - i\eta) & -\frac{1}{8}\lambda^4(1+4A^2) & 0 \\ \frac{1}{2}A\lambda^5(\rho + i\eta) & A\lambda^4(\frac{1}{2} - \rho - i\eta) & -\frac{1}{2}A^2\lambda^4 \end{pmatrix}. \tag{1.49}$$

While λ and A are relatively well known: $\lambda = 0.2205 \pm 0.0018$ and $A = 0.824 \pm 0.075$, large uncertainties remain on the ρ and η values. It is the term $i\eta$ that gives rise to the complex phase and induces CP violation when different from 0.

The unitarity of the CKM matrix allows us to write nine independent equations of the forms $\sum_k = V_{ik}V_{jk}^* = \delta_{ij}$ and $\sum_k = V_{ik}^*V_{jk} = \delta_{ij}$. Six of these nine relationships express the normalization of the columns and rows of the matrix, while six others represent the orthogonality and are written explicitly as follows:

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0\,, \tag{1.50a}$$
$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0\,, \tag{1.50b}$$
$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0\,, \tag{1.50c}$$
$$V_{ud}^*V_{cd} + V_{us}^*V_{cs} + V_{ub}^*V_{cb} = 0\,, \tag{1.50d}$$
$$V_{cd}^*V_{td} + V_{cs}^*V_{ts} + V_{cb}^*V_{tb} = 0\,, \tag{1.50e}$$
$$V_{ud}^*V_{td} + V_{us}^*V_{ts} + V_{ub}^*V_{tb} = 0\,. \tag{1.50f}$$

Each of these equations can be represented as a triangle in the complex plane. The triangle used most often corresponds to: $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$ (Eq. (1.50c)) and

is referred to as "the unitarity triangle". In terms of the Wolfenstein parametrization this equation in leading order is given by:

$$[(\rho + i\eta) + (-1) + (1 - \rho - i\eta)]A\lambda^3 + \mathcal{O}(\lambda^4) = 0. \tag{1.51}$$

We introduce the generalized Wolfenstein's parameters:

$$\overline{\rho} = \rho\Big(1 - \frac{1}{2}\lambda^2\Big), \quad \overline{\eta} = \eta\Big(1 - \frac{1}{2}\lambda^2\Big), \tag{1.52}$$

so we write:

$$[(\overline{\rho} + i\overline{\eta}) + (-1) + (1 - \overline{\rho} - i\overline{\eta})]A\lambda^3 + \mathcal{O}(\lambda^4) = 0. \tag{1.53}$$

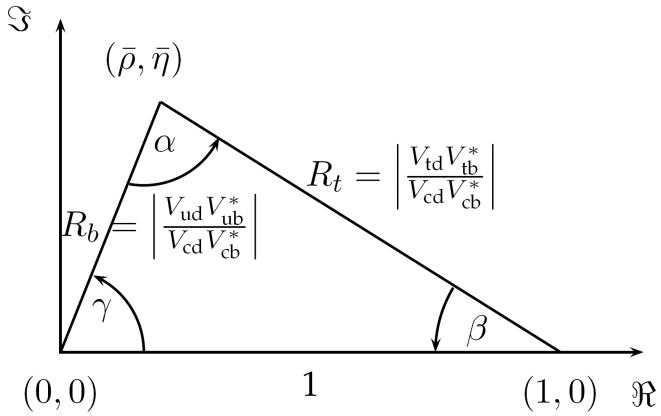


Figure 1.4: Rescaled unitarity triangle in the $(\overline{\rho}, \overline{\eta})$ plane.

In Figure 1.4 the rescaled unitarity triangle is shown in the complex plane, which can be obtained by dividing Eq. (1.50c) by $V_{cd}V_{cb}^*$, so that one of the sides is aligned on the real axis with a length of one. The lengths of the other two sides are:

$$R_b = \left|\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right| = (1 - \frac{\lambda^2}{2})\frac{1}{\lambda}\left|\frac{V_{ub}}{V_{cb}}\right|, \tag{1.54}$$

$$R_t = \left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right| = \frac{1}{\lambda}\left|\frac{V_{td}}{V_{cb}}\right|. \tag{1.55}$$

The angles of the unitarity triangle are defined as follows:

$$\alpha = \phi_2 = \arg\Big(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\Big), \tag{1.56}$$

$$\beta = \phi_1 = \arg\Big(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\Big), \tag{1.57}$$

$$\gamma = \phi_3 = \arg\Big(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\Big). \tag{1.58}$$

where the $\phi_{1,2,3}$ notation is mostly used in Asia while the European and American experiments mainly work with the α, β, γ notation.

The sum of the three angles yields π, which means that the angles are linearly dependent on each other. Using the Wolfenstein parametrization it becomes clear that these phases induce CP violation as they depend on the parameter $i\eta$:

$$\left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right| e^{-i\beta} = R_t e^{-i\beta} \approx 1 - \overline{\rho} - i\overline{\eta}, \tag{1.59}$$

$$|\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}| e^{-i\gamma} = R_t e^{-i\gamma} \approx \overline{\rho} - i\overline{\eta}. \tag{1.60}$$

Finally some useful relationships between the angles of the triangle and the CKM matrix:

$$\beta \approx -\arg(V_{td}), \tag{1.61}$$

$$\gamma \approx -\arg(V_{ub}). \tag{1.62}$$

As said before, the main experimental uncertainties are not on λ or on A but on the ρ and η parameters. Measuring more precisely the angles of the unitarity triangle is the main goal of CP violation experiments. By trying to over constrain the unitarity triangle one could, possibly, find inconsistencies which would mean a sign of the existence of physics beyond the Standard Model. In the next section the latest experimental results on the angles and sides of the unitarity triangle are summarized.

1.4.2 Experimental results of the Unitarity Triangle

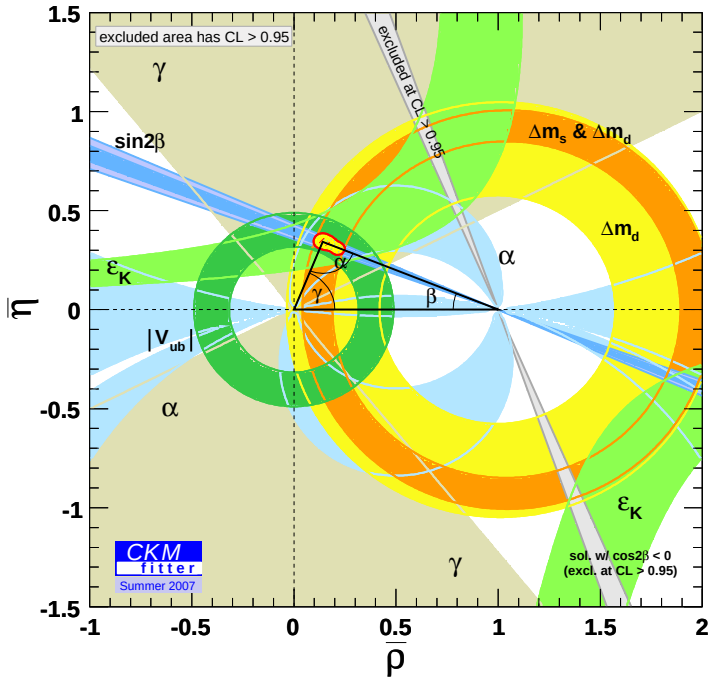


Figure 1.5: CKM unitarity triangle with the experimental results as of summer 2007.

Many experiments have measured the elements of the CKM matrix. The CKM fitter collaboration [15] collects the measurements and displays them on the $(\overline{\rho}, \overline{\eta})$ plane, as shown in Figure 1.5. The main measurements which constrain the triangle are briefly summarized below.

ϵ_K:
This was the first measurement of CP violation, performed in the kaon system. This parameter is defined by $\epsilon_K = \frac{A(K_L \to \pi\pi)_{I=0}}{A(K_S \to \pi\pi)_{I=0}}$ and it was measured by observing the CP-violating decay $K_L \to \pi\pi$, which happened at a level of 0.1%. The most precise result yields $\epsilon_K \approx 2 \times 10^{-3}$ and is shown in the figure by the light green "boomerang" belt.

α angle:
The angle at the top of the triangle is experimentally challenging to measure as penguin diagrams enter the scene. This angle is theoretically determined by $-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}$ and can experimentally be measured by looking at the CP asymmetries in $b \to u\overline{u}d$ transition, of which the best known decay is $B^0 \to \pi^+\pi^-$. Penguin contributions however contaminate the measurement of α. Solutions have been proposed, for example, by constraining the penguin contributions using isospin relations [16]. Measurements of α produce four allowed regions in the $\overline{\rho}, \overline{\eta}$ plane, shown as blue arc-shaped areas.

β angle:
This angle currently constrains most the undetermined top of the triangle. It is extensively measured in the B-factories at SLAC (BaBar) and KEK (Belle). The phase β, or better $\sin 2\beta$, arises from the phase which is induced from the interference of the mixing between the B^0 and $\overline{B^0}$ particles. The most famous "golden channel" is $B^0 \to J/\psi K_S$. Also the analysis described in this work aims to measure this angle. Because $\sin 2\beta$ is measured instead of β, there are four possibilities for β, two are shown by the lightblue lines and two more by the white-grey lines.

γ angle:
The γ angle is defined by $-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}$ and remains one of the least known parameters of the CKM matrix. The angle can be determined, for example, by measuring $B^+ \to K^+\overline{D^0}$ simultaneously with $B^+ \to K^+D^0$. The first decay is color-allowed and its decay amplitude is proportional to $|V_{cb}|$. The second decay is color-suppressed, the decay amplitude is proportional to $|V_{ub}|e^{i\gamma}$ and thus γ can be determined by : $\frac{A(B^+ \to K^+D^0)}{A(B^+ \to K^+\overline{D^0})} = e^{2i\gamma}$, which gives a theoretically very clean measurement. In practice however it is more challenging as the color-suppressed decay amplitude is very small. The picture above shows how the γ measurement constrains the top of the triangle to the gray wedges coming from the bottom left of the triangle.

Left side, R_b:
The length R_b is determined theoretically by $\frac{|V_{ub}|}{|V_{cb}|}$. By measuring decay rates of $b \to u$ and $b \to c$ transitions, the corresponding CKM amplitude can be obtained. The determination of V_{cb} can be done through inclusive and exclusive semi-leptonic decays, for example $B^0 \to D^+ l\nu_l$. The theoretical calculations of the decay amplitude however contain hadronic form factors which are challenging to calculate, as will be explained in

Section 1.6. Still the experimental results of inclusive and exclusive decays agree [11]:

$$|V_{cb}| = (42.0 \pm 0.7) \times 10^{-3}. \tag{1.63}$$

Both experimentally and theoretically the extraction of V_{ub} is a lot more difficult. The results of the inclusive and exclusive measurements have a 1σ discrepancy:

$$\begin{aligned} \text{exclusive measurement}: \quad |V_{ub}| &= (3.8 \pm 0.6) \times 10^{-3}, \\ \text{inclusive measurement}: \quad |V_{ub}| &= (4.4 \pm 0.3) \times 10^{-3}. \end{aligned}$$

Combining the $|V_{cb}|$ with both $|V_{ub}|$ measurements allows us to extract a value for R_b:

$$\begin{aligned} \text{exclusive measurement}: \quad R_b &= 0.39 \pm 0.06, \\ \text{inclusive measurement}: \quad R_b &= 0.45 \pm 0.03. \end{aligned}$$

which still causes one of the main "tensions" in the CKM picture and is therefore not included in Figure 1.5.

Right side, R_t:
This length is determined by $\frac{|V_{td}|}{|V_{cb}|}$, where the experimental results of V_{cb} have already been described above. $|V_{td}|$ can be measured through the mass-difference in the $B^0 - \overline{B^0}$ system: $\Delta M_d \propto |V_{td}V_{tb}^*|^2$. We can further improve this result by simultaneously measuring $\Delta M_s \propto |V_{ts}V_{tb}^*|^2$, where ΔM_s is the mass difference between the two weak-eigenstates. The ratio yields:

$$\frac{\Delta M_s}{\Delta M_d} \propto \frac{M_{B_s}}{M_{B_d}} \xi^2 \frac{|V_{ts}|^2}{|V_{td}|^2}, \tag{1.64}$$

where ξ^2 has to be obtained from challenging lattice QCD calculations [17]. .

1.5 CP violation in the B^0 system

As mentioned before CP violation has only been measured in the kaon system and in the B system. As the measurement of CP violation in this thesis is made in the B^0 system, we will apply the theory described in the previous section to the case where the initial particle is a B^0 or a B^0.

1.5.1 B^0 - $\overline{B^0}$ mixing

The flavor eigenstates of the B^0 system are exactly B^0 and $\overline{B^0}$ equivalent to $|P\rangle$ and $|\bar{P}\rangle$ in the previous equations. The mixing between the two particles is caused by the off-diagonal terms of the mass and decay matrix. In Eq. (1.7) we specified that the final states need to be real, which excludes almost all effects from New Physics to the decay width Γ_{12}. The off-diagonal mass elements, however, are summed over all virtual and real states, meaning that intermediate states need to be counted for as well, which leaves more possibilities for New Physics to appear. The lowest order quark transitions in the B^0 system take place through box diagrams, shown in Figure 1.6.

In the box a virtual u, c or t (anti-)quark can occur, with a probability depending on its mass and its corresponding coupling constants, which makes having a virtual top quark in

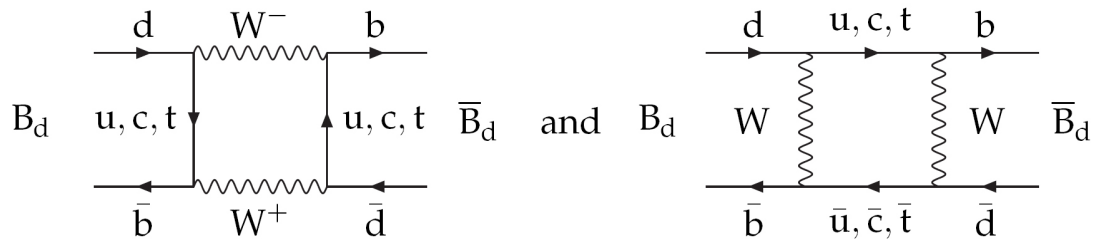


Figure 1.6: B^0 mixing diagrams.

the box the most probable. With the use of the formulas described in Section 1.3.1 we can calculate the terms of the mass and decay matrix in the B^0 system. The exact calculation of the element M_{12} is rather challenging due to hadronic effects [10]

$$M_{12} = \frac{G_F^2 m_W^2}{4\pi^2} \langle B^0 | H_{\text{eff}} | \overline{B^0} \rangle, \tag{1.65}$$

but when ignoring any QCD corrections, the relation can be simplified to:

$$M_{12} \propto (V_{td}^* V_{tb})^2 e^{i(\pi - \theta_{CP})}. \tag{1.66}$$

As M_{12} is a complex number it can be rewritten as:

$$M_{12} = e^{i\phi_M} |M_{12}|, \tag{1.67}$$

where the phase ϕ_M is:

$$\phi_M = 2\arg(V_{td}^* V_{tb}) + \pi - \theta_{CP}, \tag{1.68}$$

$$\phi_M \overset{SM}{=} 2\beta + \pi - \theta_{CP}, \tag{1.69}$$

and where β is one of the angles of the unitarity triangle. Without losing generality, let's define $\theta_{CP} = \pi$ so that the phase of the off-diagonal element of the mass matrix for the B^0 system gets simplified to:

$$\phi_M = \arg(M_{12}) = 2\arg(V_{td}^* V_{tb}) \overset{SM}{=} 2\beta. \tag{1.70}$$

Again, this equality only holds when hadronic effects are ignored.

The off-diagonal element of the decay matrix, Γ_{12}, is given by

$$\Gamma_{12} = \sum_f \langle f | T | B^0 \rangle \langle f | T | \overline{B^0} \rangle, \tag{1.71}$$

where the physical states f have to be accessible to both B^0 and $\overline{B^0}$. The value of Γ_{12} is dominated by the mass available in the decays of both B^0 mesons, i.e, $\approx m_b^2$ and because $M_{12} \propto m_t^2$ [18] one finds that:

$$\frac{|\Gamma_{12}|}{|M_{12}|} \sim \frac{m_b^2}{m_t^2} \sim 4 \times 10^{-3}. \tag{1.72}$$

and thus $|\Gamma_{12}| \ll |M_{12}|$ in the B^0 system. The CP parameter for the oscillation becomes:

$$\frac{q}{p} = -\sqrt{\frac{\Lambda_{21}}{\Lambda_{12}}} \overset{|\Gamma_{12}| \ll |M_{12}|}{\approx} -\left(\frac{M_{12}^*}{|M_{12}|}\right). \tag{1.73}$$

This means that $|q| \approx |p|$ and that CP violation in the $B^0 - \overline{B^0}$ oscillation is very small. This however does not mean that the oscillation period itself is small. $Im(\frac{\Gamma_{12}}{M_{12}})$ is of the order of $\mathcal{O}(10^{-4})$ [10], we can therefore say that:

$$\frac{q}{p} \approx e^{-i(2\beta - \phi_{\mathrm{CP}})}. \tag{1.74}$$

The weak eigenstates with definite masses and decay widths of the B^0 system are:

$$|B_\pm\rangle = p|B^0\rangle \pm q|\overline{B^0}\rangle. \tag{1.75}$$

The corresponding masses are: $m_\pm = M \mp |M_{12}|$. The two B weak-eigenstates are therefore referred to as the heavy and light particle:

$$|B_-\rangle = |B_H\rangle \quad |B_+\rangle = |B_L\rangle.$$

The average mass and decay width of the B^0 system as well as the mass and width difference between the B_H and B_L particle are given by the following relationships:

$$\overline{M} = \frac{M_H + M_L}{2} = M_{11}, \qquad \overline{\Gamma} = \frac{\Gamma_H + \Gamma_L}{2} = \Gamma_{11}, \tag{1.76}$$

$$\Delta m = M_H - M_L > 0, \quad \Delta\Gamma = \Gamma_L - \Gamma_H. \tag{1.77}$$

The experimental results are [19]:

$$\begin{aligned} \Delta m &= 0.507 \pm 0.004\,\mathrm{ps}^{-1}, \\ \tau = 1/\overline{\Gamma} &= 1.530 \pm 0.009\,\mathrm{ps}, \\ \frac{\Delta\Gamma}{\overline{\Gamma}} &= \left(40.9^{+8.9}_{-9.9}\right) \times 10^{-4}, \end{aligned}$$

where the latter is obtained using the first two experimental measurements and theoretical calculations (see [19]). As can be seen from these results, the difference in the two masses is very small. Directly measuring the two masses is experimentally still too challenging, but the mass difference can be obtained by measuring interference effects. The decay width difference in the B_d system is estimated using experimental results and theoretical calculations. This difference is $\mathcal{O}(10^{-15})$ and is therefore neglected in further calculations. As we assume $\Delta\Gamma = 0$, we will in the next chapters refer to $\overline{\Gamma}$ just as Γ or $\frac{1}{\tau_{B^0}}$.

Like in the previous section we can now determine the time-dependent decay rates for an initially ($t = 0$) pure B^0 or $\overline{B^0}$ respectively:

$$|B^0\rangle = h_+(t)|B^0\rangle + \frac{q}{p}h_-(t)|\overline{B^0}\rangle \tag{1.78}$$

$$= \frac{1}{2}\sqrt{1 + \left|\frac{q}{p}\right|^2}\Big(e^{-i\lambda_H t}|B_H\rangle + e^{-i\lambda_L t}|B_L\rangle\Big), \tag{1.79}$$

$$|\overline{B^0}\rangle = \frac{p}{q}h_-(t)|B^0\rangle + h_+(t)|\overline{B^0}\rangle \tag{1.80}$$

$$= \frac{p}{2|p|q}\Big(e^{-i\lambda_L t}|B_L\rangle - e^{-i\lambda_H t}|B_H\rangle\Big), \tag{1.81}$$

where $h_\pm = \frac{1}{2}(e^{-i\lambda_+ t} \pm e^{-i\lambda_- t})$ and $\lambda_\pm = M_{\frac{L}{H}} - \frac{i}{2}\Gamma_{\frac{L}{H}}$. With the parameters defined in this section we can simplify the following relationships:

$$\begin{aligned} h_\pm(t) &= \pm e^{i\overline{M}t} e^{-\overline{\Gamma}t/2}\big[\cosh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} + \sinh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2}\big] \\ &\approx \pm e^{i\overline{M}t} e^{-\overline{\Gamma}t/2}\cos\frac{\Delta m t}{2}, \end{aligned} \tag{1.82}$$

$$\begin{aligned} |h_\pm|^2 &= \frac{e^{-\overline{\Gamma}t}}{2}\big[\cosh\frac{\Delta\Gamma t}{2} \pm \cos\Delta m t\big] \\ &\approx \frac{e^{-\overline{\Gamma}t}}{2}\big[1 \pm \cos\Delta m t\big], \end{aligned} \tag{1.83}$$

$$\begin{aligned} h_+^*(t)h_-(t) &= \frac{e^{-\overline{\Gamma}t}}{2}\big[-\sinh\frac{\Delta\Gamma t}{2} + i\sin\Delta m t\big] \\ &\approx \frac{e^{-\overline{\Gamma}t}}{2} i\sin\Delta m t. \end{aligned} \tag{1.84}$$

where we have applied the $\Delta\Gamma = 0$ approximation. These relationships will be used in the next section to determine the time-dependent decay rates.

1.5.2 B^0 decay to a CP eigenstate.

When time-dependent decay rates for a B system are studied, a separation is made between four types of decays depending on the final state.

- the final state is a unique CP eigenstate, for example $J/\psi K_S$.
- the final state is a mixture of CP eigenstates, like $D^{*+}D^{*-}$.
- flavor specific final states like $l^+\nu X^-$.
- flavor non-specific final states like $D^-\pi^+$ and $D^+\pi^-$.

In this analysis CP violation in $B \to D^{*+}D^{*-}$ is studied. The final state, $D^{*+}D^{*-}$, is a mixed CP eigenstate and thus falls under the second type. We will apply now the general theory seen in the previous sections to the case where the final state is a CP eigenstate. The process is as follows, we can start with a B^0 state which can decay to the CP final state or oscillate to the $\overline{B^0}$ which in its turn can decay to the same CP final state. Figure 1.7 shows a schematic view of the oscillation and decay process. There are now only two (instead of four) decay amplitudes to be defined as $f = \overline{f}$:

$$\begin{aligned} A_f &= A_{\overline{f}} = \langle f_{\rm CP}|H_{\rm eff}|B^0\rangle, \\ \overline{A}_f &= \overline{A}_{\overline{f}} = \langle f_{\rm CP}|H_{\rm eff}|\overline{B^0}\rangle. \end{aligned}$$

Because $|f\rangle$ is a CP eigenstate we have:

$$|\overline{f}\rangle = (CP)|f\rangle = \eta|f\rangle, \tag{1.85}$$

where η is the CP eigenstate of f. We can write the decay amplitude as a product of its weak phase and an effective operator $\mathcal{O}$ [20] such that:

$$\overline{A}_f = e^{i\phi_D}\langle f|\mathcal{O}|\overline{B^0}\rangle \quad A_f = e^{-i\phi_D}\langle f|\mathcal{O}^\dagger|B^0\rangle, \tag{1.86}$$

$$\Rightarrow \overline{A}_f = \eta e^{i(\phi_D - \theta_{CP})}\langle f|\mathcal{O}^\dagger|B^0\rangle = \eta e^{i(2\phi_D - \theta_{CP})}A_f, \tag{1.87}$$

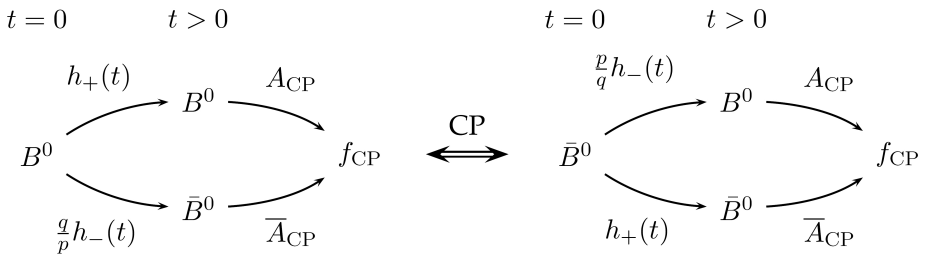


Figure 1.7: Oscillation and decay interference for an initial B^0 to a CP eigenstate (left) and the CP conjugated process (right).

such that the ratio of the amplitudes is given by:

$$\frac{\overline{A}_f}{A_f} = \eta e^{i(2\phi_D - \theta_{CP})}. \tag{1.88}$$

At time $t = 0$ we start with a B^0 particle that decays to the CP eigenstate at $t > 0$. The decay rate can be written as:

$$\begin{aligned} R_f &= |\langle f|H_{\rm eff}|B^0(t)\rangle|^2 \\ &= \left|h_+(t)\langle f|H_{\rm eff}|B^0\rangle + \frac{q}{p}h_-(t)\langle f|H_{\rm eff}|\overline{B^0}\rangle\right|^2 \\ &= N_f|A_f|^2\Big\{|h_+(t)|^2 + \left|\frac{q}{p}\right|^2\left|\frac{\overline{A}_f}{A_f}\right|^2|h_-|^2 + 2\Re\Big[\frac{q}{p}\frac{\overline{A}_f}{A_f}h_+^*(t)h_-(t)\Big]\Big\}. \end{aligned} \tag{1.89}$$

where N_f is a normalization factor. For the CP conjugated decay we would have:

$$\begin{aligned} \overline{R}_f &= |\langle f|H_{\rm eff}|\overline{B^0}(t)\rangle|^2 \\ &= \left|\frac{p}{q}h_-(t)\langle f|H_{\rm eff}|B^0\rangle + h_+(t)\langle f|H_{\rm eff}|\overline{B^0}\rangle\right|^2 \\ &= N_f|A_f|^2\left|\frac{p}{q}\right|^2\Big\{|h_-(t)|^2 + \left|\frac{q}{p}\right|^2\left|\frac{\overline{A}_f}{A_f}\right|^2|h_+|^2 + 2\Re\Big[\frac{q}{p}\frac{\overline{A}_f}{A_f}h_+(t)h_-^*(t)\Big]\Big\}. \end{aligned} \tag{1.90}$$

The key observables for the study of CP violation can be given by the complex quantity:

$$\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f}. \tag{1.91}$$

When the modulus of λ_f is not equal to unity, CP violation occurred in the oscillation or in the decay:

$$|\lambda_f| = \left|\frac{q}{p}\right|\left|\frac{\overline{A}_f}{A_f}\right| \approx \left|\frac{\overline{A}_f}{A_f}\right|. \tag{1.92}$$

where the last approximation holds in the $B^0 - \overline{B^0}$ system. When the imaginary part of λ_f differs from zero, CP violation in the interference between the oscillation and the decay

takes place. Using the equations (1.84) and setting $\left|\frac{q}{p}\right| = 1$, we can simplify the previous equations to:

$$R_f = N_f|A_f|^2\frac{e^{-\overline{\Gamma}t}}{2}\Big\{1+|\lambda_f|^2+(1-|\lambda_f|^2)\cos\Delta mt - 2\Im(\lambda_f)\sin\Delta mt\Big\}, \quad (1.93)$$

$$\overline{R}_f = N_f|A_f|^2\frac{e^{-\overline{\Gamma}t}}{2}\Big\{1+|\lambda_f|^2-(1-|\lambda_f|^2)\cos\Delta mt + 2\Im(\lambda_f)\sin\Delta mt\Big\}. \quad (1.94)$$

Thus, by measuring the decay rates of an initially pure B^0 or $\overline{B^0}$ state we can measure CP violation, as the last and second-to-last terms in the previous equation have an opposite sign. We can therefore write:

$$\begin{aligned}\overset{(-)}{R}_f &\propto \frac{e^{-\overline{\Gamma}t}}{2}\Big\{1\pm\Big(\frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}\Big)\cos\Delta mt \pm \frac{-2\Im(\lambda_f)}{1+|\lambda_f|^2}\sin\Delta mt\Big\},\\ &\propto \frac{e^{-\overline{\Gamma}t}}{2}\Big\{1\mp(\mathcal{A}_{CP}^{\rm dir}\cos\Delta mt + \mathcal{S}_{CP}^{\rm mix-ind}\sin\Delta mt)\Big\}. \end{aligned} \quad (1.95)$$

where the $+$ sign in the top equation ($-$ sign in the bottom equation) should be used when the initial state is B^0 and the $-$ sign in the top equation ($+$ sign in the bottom equation) when the initial state is $\overline{B^0}$. The two new variables introduced are the two independent CP violating parameters, defined as:

$$\mathcal{A}_{CP}^{\rm dir} \equiv \frac{|\lambda_f|^2-1}{|\lambda_f|^2+1} \quad \mathcal{S}_{CP}^{\rm mix-ind} \equiv \frac{2\Im\lambda_f}{|\lambda_f|^2+1}, \quad (1.96)$$

There is actually still a third CP parameter which has not entered in our formulas due to the $\Delta\Gamma = 0$ approximation. This variable is:

$$\mathcal{A}_{\Delta\Gamma} \equiv \frac{2\Re\lambda_f}{|\lambda_f|^2+1}, \quad (1.97)$$

while we assume $\mathcal{A}_{\Delta\Gamma} = 0$. The three parameters are related in the following way:

$$|\mathcal{A}_{\Delta\Gamma}|^2 + |\mathcal{A}_{CP}^{\rm dir}|^2 + |\mathcal{S}_{CP}^{\rm mix-ind}|^2 = 1. \quad (1.98)$$

We finally define the time-dependent CP asymmetry as:

$$\begin{aligned}\mathcal{A}_{CP}(t) &= \frac{\Gamma[\overline{B^0}(t)\to f] - \Gamma[B^0(t)\to f]}{\Gamma[\overline{B^0}(t)\to f] + \Gamma[B^0(t)\to f]}\\ &= \frac{\overline{R}_f - R_f}{\overline{R}_f + R_f}. \end{aligned} \quad (1.99)$$

when applying the equations for the decay rates we obtain:

$$\mathcal{A}_{CP}(t) = \mathcal{A}_{CP}^{\rm dir}\cos(\Delta mt) + \mathcal{S}_{CP}^{\rm mix-ind}\sin(\Delta mt), \quad (1.100)$$

which is correct up to the approximation $|p| = |q|$ and $\Delta\Gamma = 0$. We can conclude that in the neutral B^0 system we can measure the direct and mixing-induced CP parameters from the time-dependent CP asymmetry for decays to a CP eigenstate.

1.6 Hadronic effects

So far we have ignored the strong interactions but as the decay which is studied in this analysis, $B \rightarrow D^{*+}D^{*-}$, has a hadronic final state, these complex strong interactions between quarks will come into play. Figure 1.8 shows an artistic example of what lies behind the clean weak interactions, once gluons enter the scene.

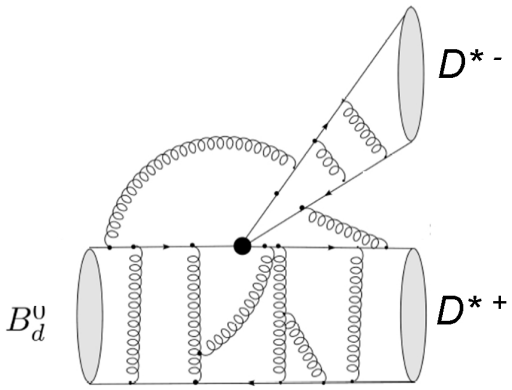


Figure 1.8: Illustration of possible strong interactions in a weak decay.

To understand the complex effects of the strong interaction on the CP measurement in $B \rightarrow D^{*+}D^{*-}$, we will first look at a simpler example, with less hadronic effects $B \rightarrow l\nu_l$ (see Figure 1.9). The transition amplitude for this leptonic decay is:

$$T_{fi} = \frac{g^2}{8} V_{ub} [\overline{u}_l \gamma^\alpha (1-\gamma_5) v_\nu] \big[\frac{g_{\alpha\beta}}{k^2 - M_W^2}\big] \langle 0|\overline{u}\gamma^\beta(1-\gamma_5)b|B^-\rangle, \tag{1.101}$$

where the factor $[\overline{u}_l\gamma^\alpha(1-\gamma_5)v_\nu]$ contains the Dirac spinors and $\langle 0|\overline{u}\gamma^\beta(1-\gamma_5)b|B^-\rangle$ represents the hadronic matrix element. The transferred momentum k of the W boson is equal to M_B^2. The mass of the B meson is roughly 40 times smaller than the mass of the W, which allows us to simplify the above equation using:

$$k^2 = M_B^2 \ll M_W^2, \tag{1.102}$$

$$\big[\frac{g_{\alpha\beta}}{k^2 - M_W^2}\big] \longrightarrow \big[\frac{g_{\alpha\beta}}{-M_W^2}\big] \equiv \big(\frac{8G_F}{\sqrt{2}g^2}\big) g_{\alpha\beta}, \tag{1.103}$$

where the momentum dependence is neglected and G_F is the Fermi constant. The transition amplitude becomes:

$$T_{fi} = \big(\frac{8G_F}{\sqrt{2}}\big) V_{ub} [\overline{u}_l \gamma^\alpha (1-\gamma_5) v_\nu] \langle 0|\overline{u}\gamma^\beta(1-\gamma_5)b|B^-\rangle, \tag{1.104}$$

and the W boson is "integrated out" of the equation. The only hadronic part left is the hadronic matrix element, $\langle 0|\overline{u}\gamma_\alpha(1-\gamma_5)b|B^-\rangle$. As B^- is a pseudo-scalar meson we can write that:

$$\langle 0|\overline{u}\gamma_\alpha b|B^-\rangle = 0, \tag{1.105}$$

$$\langle 0|\overline{u}\gamma_\alpha\gamma_5)b|B^-\rangle = i f_B q_\alpha, \tag{1.106}$$

where f_B is the B-meson decay constant and q is the four-momentum of the B^-. f_B is an important hadronic parameter which can be calculated from QCD. From this example it can be understood that for a fully hadronic decay a lot more parameters will be needed to describe the hadronic behaviors. For non-leptonic decays theorists use the low-energy effective Hamiltonian, which are expressed using the Operator Product Expansion (OPE). The transition amplitude is written as:

$$\langle f|H_{\text{eff}}|i\rangle = \Big(\frac{G_F}{\sqrt{2}}\Big)V_{CKM}\sum_k C_k(\mu)\langle f|Q_k(\mu)|i\rangle. \tag{1.107}$$

V_{CKM} represent the CKM factor and μ is the renormalization scale. The operator product expansion allows separation of the short-distance from the long-distance contributions. The factors C_k govern the short-distance physics. They are called the Wilson coefficients and can be calculated in a perturbative manner.

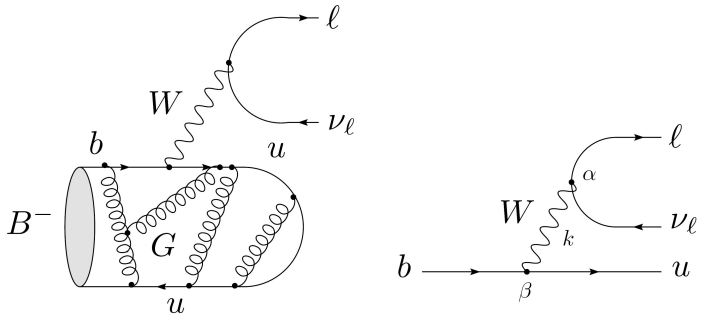


Figure 1.9: Possible strong interactions in $B \to l\nu_l$ decays.

The hadronic matrix element $\langle f|Q_k(\mu)|i\rangle$ represents the long-distance physics, and is a non-perturbative term, which makes calculating it very difficult so that large uncertainties remain. The Q_k are local operators, which are generated though the electroweak interaction and QCD and govern "effectively" the considered decay.

Finally the hadronic matrix element, $\langle f|O_k(\mu)|i\rangle$ can be factorized, as was shown for the leptonic case, but now using a decay constant and a form factor. These calculations bear still large theoretical uncertainties, as no non-perturbative method is established yet. So far data have shown large non-factorisable corrections, which means that we cannot ignore these effects.

1.7 The $B^0 \to D^{*+}D^{*-}$ decay

Now that we have the tools from the theory of the weak and strong interactions, we can apply them on the $B^0 \to D^{*+}D^{*-}$ decay. The lowest-order Feynman diagrams are shown in Figure 1.10. The tree diagram on the left is governed by a V_{cb}^* and a V_{cd} transition. The loop diagram on the right is called a penguin diagram and the quark q in the loop can be an up, charm or top quark. The corresponding CKM factor is $V_{qb}^*V_{qd}$ where $q = u, c$ or t. We will now calculate the time-dependent decay rates $R_{D^{*+}D^{*-}}(t)$ and $\overline{R}_{D^{*+}D^{*-}}(t)$ in more detail. The total decay amplitude from the tree (A_T) and penguin (A_P) diagrams, can be written as:

$$A_{D^{*+}D^{*-}}(t) = \lambda_c^{(d)}(A_T^c + A_P^c) + \lambda_u^{(d)}A_P^u + \lambda_t^{(d)}A_P^t, \tag{1.108}$$

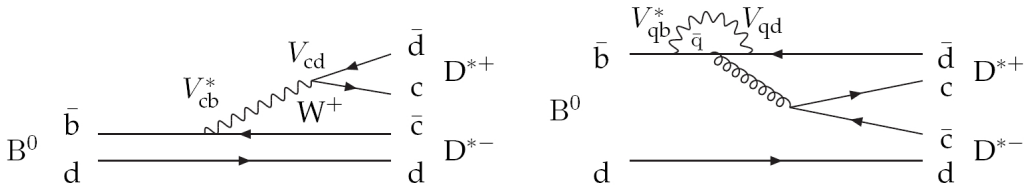


Figure 1.10: Lowest-order Feynman diagrams for the $B^0 \to D^{*+}D^{*-}$ decay. Left: the tree diagram. Right: the penguin diagrams.

with $\lambda_j^{(d)} = V_{jd}V_{jb}^*$. Using the CKM unitarity constraint from Eq. (1.50c)

$$\lambda_u^{(d)} + \lambda_c^{(d)} + \lambda_t^{(d)} = 0, \tag{1.109}$$

we can eliminate $\lambda_t^{(d)}$ and obtain:

$$A_{D^{*+}D^{*-}}(t) = \lambda_c^{(d)}(A_T^c + A_P^c - A_P^t)\left[1 + \frac{\lambda_u^{(d)}}{\lambda_c^{(d)}}\frac{A_P^u - A_P^t}{A_T^c + A_P^c - A_P^t}\right]. \tag{1.110}$$

Let us define $ae^{i\theta}$ as:

$$ae^{i\theta} = \frac{A_P^u - A_P^t}{A_T^c + A_P^c - A_P^t}. \tag{1.111}$$

Using:

$$\frac{\lambda_u^{(d)}}{\lambda_c^{(d)}} = \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} = \frac{(1-\lambda^2/2)A\lambda^3(\rho+i\eta)}{-\lambda A\lambda^2} = -(1-\lambda^2/2)\sqrt{\rho^2+\eta^2}e^{i\gamma}, \tag{1.112}$$

the decay amplitude can be rewritten as:

$$A_{D^{*+}D^{*-}}(t) = -\lambda^3 A(A_T^c + A_P^c - A_P^t)[1 - (1-\lambda^2/2)ae^{i\theta}\sqrt{\rho^2+\eta^2}e^{i\gamma}]. \tag{1.113}$$

With the same reasoning the CP-conjugated decay amplitude can be found:

$$\overline{A}_{D^{*+}D^{*-}}(t) = \lambda_c^{(d)*}(\overline{A}_T^c + \overline{A}_P^c - \overline{A}_P^t)\left[1 + \frac{\lambda_u^{(d)*}}{\lambda_c^{(d)*}}\frac{\overline{A}_P^u - \overline{A}_P^t}{\overline{A}_T^c + \overline{A}_P^c - \overline{A}_P^t}\right]. \tag{1.114}$$

Remember from Eq. (1.88): $\overline{A}_f = \eta A_f e^{-i\theta_{CP}}$. The phase ϕ_D is left as it is included in the CKM factors λ. This allows us to write:

$$\begin{aligned}\overline{A}_{D^{*+}D^{*-}}(t) &= \lambda_c^{(d)*}\eta e^{-i\theta_{CP}}(A_T^c + A_P^c - A_P^t)\left[1 + \frac{\lambda_u^{(d)*}}{\lambda_c^{(d)*}}\frac{\eta e^{-i\theta_{CP}}(A_P^u - A_P^t)}{\eta e^{-i\theta_{CP}}(A_T^c + A_P^c - A_P^t)}\right] \\ &= \eta e^{-i\theta_{CP}}\lambda_c^{(d)*}(A_T^c + A_P^c - A_P^t)\left[1 + \frac{\lambda_u^{(d)*}}{\lambda_c^{(d)*}}ae^{i\theta}\right].\end{aligned} \tag{1.115}$$

Using:

$$\frac{\lambda_u^{(d)*}}{\lambda_c^{(d)*}} = \frac{V_{ud}^*V_{ub}}{V_{cd}^*V_{cb}} = \frac{(1-\lambda^2/2)A\lambda^3(\rho-i\eta)}{-\lambda A\lambda^2} = -(1-\lambda^2/2)\sqrt{\rho^2+\eta^2}e^{-i\gamma}, \tag{1.116}$$

such that:

$$\overline{A}_{D^{*+}D^{*-}}(t) = -\eta e^{-i\theta_{CP}}\lambda^3 A(A_T^c + A_P^c - A_P^t)[1-(1-\lambda^2/2)\sqrt{\rho^2+\eta^2}e^{-i\gamma}ae^{-i\theta}] \quad (1.117)$$

The factor introduced in the previous section is:

$$\lambda_f = \frac{q}{p}\frac{\overline{A}}{A}, \quad (1.118)$$

with $\frac{q}{p} = e^{-i(2\beta-\theta_{CP})}$ if we assume no CP violation in the mixing and β is an angle of the unitary triangle. We then have:

$$\begin{aligned}
\lambda_{D^{*+}D^{*-}} &= e^{-i(2\beta-\theta_{CP})}\frac{\overline{A}_{D^{*+}D^{*-}}}{A_{D^{*+}D^{*-}}} \\
&= \eta e^{-i(2\beta-\theta_{CP})}e^{-i\theta_{CP}}\frac{[1-(1-\lambda^2/2)\sqrt{\rho^2+\eta^2}e^{-i\gamma}ae^{i\theta}]}{[1-(1-\lambda^2/2)\sqrt{\rho^2+\eta^2}e^{+i\gamma}ae^{i\theta}]} \\
&= \eta e^{-i2\beta}\frac{1-(1-\lambda^2/2)ae^{i\theta}\sqrt{\rho^2+\eta^2}e^{-i\gamma}}{1-(1-\lambda^2/2)ae^{i\theta}\sqrt{\rho^2+\eta^2}e^{+i\gamma}}. \quad (1.119)
\end{aligned}$$

Until this point no approximations have been made (besides the Wolfenstein parametrization which is correct to the order of $O(10^{-3})$). A clean measurement of the angle β would be possible if the fraction in the last equation would become close to one, which can happen for example when $ae^{i\theta}$ is small. To calculate this term the difficult QCD effects need to be taken into account. However the Standard Model predicts that the fraction of penguin diagrams with respect to tree diagrams ("P/T" ratio) is only of the order of a few percent [21, 1]. In the limit where $ae^{i\theta}$ is negligible the CP-violating parameter becomes:

$$\lambda_{D^{*+}D^{*-}} = \eta e^{-i2\beta}. \quad (1.120)$$

Therefore in the absence of extra weak phases coming from the penguin decays the Standard Model prediction is:

$$\mathcal{A}_{CP}^{\rm dir} \equiv \frac{|\lambda|^2-1}{|\lambda|^2+1} = 0, \quad (1.121)$$

$$\mathcal{S}_{CP}^{\rm mix-ind} \equiv \frac{2\Im\lambda}{|\lambda|^2+1} = -\eta\sin 2\beta, \quad (1.122)$$

and thus no direct-CP violation is expected in the $B \to D^{*+}D^{*-}$ decays while the indirect-CP violating parameter offers a clean measurement of $\sin 2\beta$. The time-dependent decay rates of an initially pure B^0 state which decays to $D^{*+}D^{*-}$ becomes:

$$R_{D^{*+}D^{*-}}(t) \propto \frac{e^{-\overline{\Gamma}t}}{2}(1+\eta\sin 2\beta\sin\Delta mt) \quad (1.123)$$

and for an initially pure $\overline{B^0}$:

$$\overline{R}_{D^{*+}D^{*-}}(t) \propto \frac{e^{-\overline{\Gamma}t}}{2}(1-\eta\sin 2\beta\sin\Delta mt). \quad (1.124)$$

The time-dependent CP-asymmetry becomes:

$$\mathcal{A}_{CP} = -\eta\sin 2\beta\sin(\Delta mt) \quad (1.125)$$

The reasoning in this section is very similar to that of the "golden channel" $B^0 \to J/\psi K_S$, where the final state is a CP eigenstate as well. In the $J/\psi K_S$ analysis one of the CKM amplitudes is dominant, which will make the hadronic phases and the other CKM phases cancel out. The $B \to D^{*+}D^{*-}$ decay is not referred to as the "golden channel" for two reasons from the theoretical side (there are also more challenges to face from the experimental side namely the signal reconstruction). First of all the dominating amplitude in $B \to D^{*+}D^{*-}$, $V_{cd}V_{cb}^*$ is of the order of λ^3, while the $B^0 \to J/\psi K_S$ channel is color suppressed but its mixing parameters, $V_{cs}V_{cb}^*$ are of the order of λ^2 in the tree diagram, which results in a larger branching fraction. The second reason is that the approximation we made, $ae^{i\phi} = 0$, ignores the few percent of penguin diagrams that are predicted. The accuracy of the approximations made in $B^0 \to J/\psi K_S$ are only of the order of $O(10^{-3})$.

If we assume that the Standard Model penguin diagrams are negligible then the measurement of $\sin 2\beta$ is an excellent probe for New Physics. The loops diagrams are prime candidates to host New Physics particles. As the tree diagram is already CKM suppressed (compared to for example $J/\psi K_S$), any effect on the penguin diagrams can be potentially significant, which makes this decay a very good probe for searches of physics beyond the Standard Model. However also in the tree diagram new physics can enter and affect the decay amplitudes. The expressions will become more complicated when extra particles can enter the scene. When this complication is ignored while the penguin contributions are present, a shift in the measured $\sin 2\beta$ will be observed. The size of this shift depends on the "P/T" ratio as well as the relative weak phases which enter due to penguin diagrams [22] or new physics effects in the decay amplitudes. The latter can even give rise to a shift of β of $\Delta\beta = 0.6$ in models with enhanced chromomagnetic dipole operators or in Super Symmetry models without R-parity [23]. Also the theoretical prediction of $|\lambda_{D^{*+}D^{*-}}|$ would be different from one such that direct CP violation can exist. Large uncertainties would enter in the theoretical equations due to the low-energy hadronic effects, which will translate into a large uncertainty for $\sin 2\beta$. These theoretical uncertainties could be constrained by comparing data from different channels to the hadronic models. The proper-time differences could be measured simultaneously for the $B^0 \to D^{*+}D^{*-}$ channel and the $B_s^0 \to D_s^{*+}D_s^{*-}$ channel. These two decays are related through the U-spin symmetry [24] such that the CP violating effect should be similar in both decays. However the Belle experiment cannot probe the $B_s^0 \to D_s^{*+}D_s^{*-}$ decays, but the new generation experiment, LHCb will have access to both the B_s and the B^0 decay such that a simultaneous measurement is possible.

To summarize, the $B \to D^{*+}D^{*-}$ allows us to extract $\sin 2\beta$, the amplitude of the mixing-induced CP violation, from measuring the difference in decay rates between an initial B^0 particle and an initial $\overline{B^0}$ particle. There is no direct CP violation expected. These predictions only hold if the penguin diagrams are negligible with respect to the tree diagrams. However if this assumption is incorrect, due to incorrect Standard Model predictions or more importantly due to New Physics entering in the loops, the calculations need to incorporate hadronic effects as well as effects from the new phases. The hadronic effects are theoretically very challenging to compute and have still large uncertainties.

1.8 Determining the CP eigenvalue

In the previous equations the CP eigenvalue of $D^{*+}D^{*-}$ has been symbolized by η. We will see in this section that η can be ± 1 and how this dilutes the measurement of $\sin 2\beta$.

The D^{*+} and D^{*-} are vector mesons, with spin ($J^P = 1^-$) while the B meson is a spin

0 particle. Due to the conservation of the total spin, the final state can exist with different relative orbital momenta. Three angular states exist: an S, P or D wave corresponding to the orbital angular momentum $l = 0, 1$ or 2. The CP eigenvalue of $D^{*+}D^{*-}$ is given by the following relationship:

$$CP|D^{*+}D^{*-}\rangle = (-1)^l|D^{*+}D^{*-}\rangle, \tag{1.126}$$

which means that it can either be $+1$ (CP even) when $l = 0, 2$ or -1 (CP odd) when $l = 1$. The final state is therefore an admixture of CP even and CP odd states. As η has an opposite sign for the two CP states, it will "dilute" the overall CP asymmetry. The dilution factor $D < 1$, coming from the cancellation of the two different CP components enters in the asymmetry rate in the following way:

$$\begin{aligned} a(t) &= \frac{\overline{R}_{D^{*+}D^{*-}} - R_{D^{*+}D^{*-}}}{\overline{R}_{D^{*+}D^{*-}} + R_{D^{*+}D^{*-}}} \\ &= D \sin 2\beta \sin \Delta m t, \end{aligned} \tag{1.127}$$

where the dilution factor is given by:

$$D = \frac{\Gamma_{CP\,\mathrm{ev}} - \Gamma_{CP\,\mathrm{odd}}}{\Gamma_{\mathrm{tot}}}. \tag{1.128}$$

When CP odd and CP even states can be disentangled, a more precise measurement of the CP violating parameters is possible. To extract the different angular momenta, we can look at the direction of the decay products of $D^{*+}D^{*-}$: the D mesons and slow pions. The direction of their momenta gives us insight into their partial wave distribution.

This formalism can be developed in the helicity framework, where three amplitudes, H_0, H_+ and H_- are defined corresponding to the helicity of the two D^* mesons. The helicity is defined as $\lambda = \vec{s}\cdot\vec{p}/(|\vec{s}||\vec{p}|)$ and holds the value $\lambda = -1, 0, +1$. The helicity framework is popular for theoretical calculations as it gives a straightforward procedure to determine the longitudinal rate. However in practice a different approach is used by working with the transversity basis. The transversity formalism offers a convenient method to determine the CP-odd fraction in the data sample because each decay amplitude contributes to only one CP eigenstate. A_0 and $A_\parallel$ are CP even states while $A_\perp$ is the CP-odd state. The following relationships show how the transverse amplitudes are constructed:

$$A_0 = H_0, \quad A_\parallel = \frac{1}{\sqrt{2}}(H_+ + H_-), \quad A_\perp = \frac{1}{\sqrt{2}}(H_+ - H_-). \tag{1.129}$$

At time $t = 0$ there is no CP violation possible, we have:

$$A_0 = \overline{A}_0, \quad A_\parallel = \overline{A_\parallel}, \quad A_\perp = \overline{A_\perp}. \tag{1.130}$$

While time elapses CP-even and CP-odd states evolve. The three polarization states are normalized such that:

$$\Gamma(t) \propto |A_0|^2 + |A_\parallel|^2 + |A_\perp|^2. \tag{1.131}$$

Let us define the final states which are in a CP even state as $D^{*+}D^{*-}{}_{\mathrm{ev}}$ and the CP odd state as $D^{*+}D^{*-}{}_{\mathrm{odd}}$. The decay amplitudes described by Eq. (1.94) now become:

$$R_{D^{*+}D^{*-}_{ev}}(t) = N_{D^{*+}D^{*-}_{ev}}|A_{D^{*+}D^{*-}_{ev}}|^2 e^{-\Gamma t}\{1 - \sin 2\beta \sin \Delta m t\}, \tag{1.132}$$

$$R_{D^{*+}D^{*-}_{odd}}(t) = N_{D^{*+}D^{*-}_{odd}}|A_{D^{*+}D^{*-}_{odd}}|^2 e^{-\Gamma t}\{1 + \sin 2\beta \sin \Delta m t\}, \tag{1.133}$$

and for the situation where the initial particle was $\overline{B^0}$ we get:

$$\overline{R}_{D^{*+}D^{*-}_{ev}}(t) = N_{D^{*+}D^{*-}_{ev}}|A_{D^{*+}D^{*-}_{ev}}|^2 e^{-\Gamma t}\{1+\sin 2\beta \sin \Delta m t\}, \quad (1.134)$$

$$\overline{R}_{D^{*+}D^{*-}_{odd}}(t) = N_{D^{*+}D^{*-}_{odd}}|A_{D^{*+}D^{*-}_{odd}}|^2 e^{-\Gamma t}\{1-\sin 2\beta \sin \Delta m t\}. \quad (1.135)$$

Note that the only difference between the even and odd equations is the sign multiplying $\sin 2\beta$.

Finally an overview of the different relationships between transversity amplitudes, helicity amplitudes and partial wave decomposition is given in Table 1.2 and illustrated in Figure 1.11.

Wave	Transversity	Helicity	CP
S	$\frac{1}{\sqrt{3}}(\sqrt{2}\mathrm{A}_{\parallel} - \mathrm{A}_0)$	$\frac{1}{\sqrt{3}}(H_+ - H_0 + H_-)$	even
P	$\mathrm{A}_{\perp}$	$\frac{1}{\sqrt{2}}(H_+ - H_-)$	odd
D	$\frac{1}{\sqrt{3}}(\mathrm{A}_{\parallel} + \sqrt{2}\mathrm{A}_0)$	$\frac{1}{\sqrt{6}}(H_+ + 2H_0 + H_-)$	even

Table 1.2: Relationships between the wave decomposition, the transversity and the helicity framework.

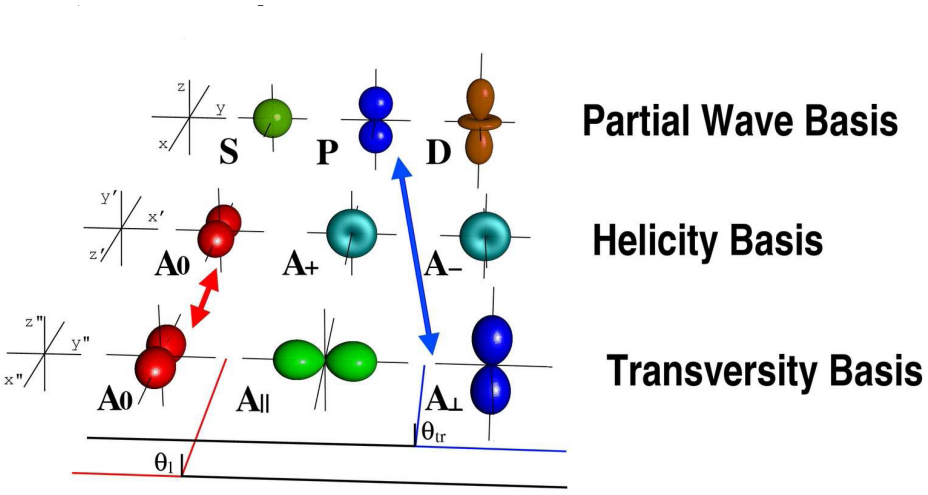


Figure 1.11: The three possible angular states of the $B^0 \to D^{*+}D^{*-}$ decay are presented on the x axis and visualized in each of the three bases (top: partial wave basis, middle: helicity basis, bottom: transversity basis). The partial wave basis shows the angular distribution of the D^{*+}, the two others show the angular distribution of the slow π^+.

1.8.1 Transversity basis and transversity angles

In the transversity basis the direction of the momenta of the daughter particles of $D^{*+}D^{*-}$ is used to define three independent physical angles. The transversity amplitudes are defined by the direction of the polarization of the two D^* mesons in a way that:

- The x axis is defined along the D^{*+} momentum in the B^0 meson rest frame.

- The xy plane is constructed with this x-axis and the momenta of the D^{*-} daughters seen in the D^{*+} rest frame, which defines the y axis. The sign of the y axis is such that the projection of the slow pion of the D^{*-} on the y axis is positive.

- The z axis is such that the basis becomes a right-handed Cartesian system.

An illustration is shown in Figure 1.12.

In this basis the decay amplitudes are defined as spin projections for one vector particle parallel or perpendicular to the plane of the decay of the other. The direction of the different polarization amplitudes is thus given by:

$$\begin{array}{llll} A_0 & \hat{\epsilon}_{D^{*-}} \parallel \hat{x} & \hat{\epsilon}_{D^{*+}} \parallel \hat{x} \\ A_\parallel & \hat{\epsilon}_{D^{*-}} \parallel \hat{y} & \hat{\epsilon}_{D^{*+}} \parallel \hat{y} \\ A_\perp & \hat{\epsilon}_{D^{*-}} \parallel \hat{y} & \hat{\epsilon}_{D^{*+}} \parallel \hat{z}, \end{array}$$

where $\hat{epsilon}$ is the polarization vector of the D^{*+} or D^{*-}.

There are three independent angles relevant to this analysis, θ_{tr}, θ_1 and ϕ_{tr}, which can be defined in the transversity basis: θ_{tr}, θ_1 and ϕ_{tr}.

- θ_{tr} is defined as the angle between the z axis and the momentum vector of the slow pion from the D^{*+} decay in the D^{*+} rest frame.

- θ_1 is the angle between the x axis and the momentum of the slow pion from the D^{*-} in the D^{*-} rest-frame.

- The angle ϕ_{tr} is the angle between the x-axis and the momentum vector of the slow pion from the D^{*+} projected onto the xy plane in the D^{*+} rest frame.

1.8.2 Extraction of the angular state

The three different polarization states can be distinguished due to their different $\cos\theta_{tr}$ and $\cos\theta_1$ probability distributions. The angular dependence of the decay can be written as follows [25]

$$\begin{aligned} \frac{1}{\Gamma}\frac{d^3\Gamma(B\to D^{*+}D^{*-})}{d\cos\theta_{tr}d\cos\theta_1 d\phi_{tr}} &= \frac{9}{16\pi}\frac{1}{|A_0|^2+|A_\perp|^2+|A_\parallel|^2} \\ &\big\{2\cos^2\theta_1\sin^2\theta_{tr}\cos^2\phi_{tr}|A_0|^2 \\ &+\sin^2\theta_1\cos^2\theta_{tr}|A_\perp|^2 \\ &+\sin^2\theta_1\sin^2\theta_{tr}\sin^2\phi_{tr}|A_\parallel|^2 \\ &-\sin^2\theta_1\sin 2\theta_{tr}\sin\phi_{tr}\,\mathrm{Im}(\mathrm{A}_\parallel^*\mathrm{A}_\perp) \\ &-\frac{1}{\sqrt{2}}\sin 2\theta_1\sin^2\theta_{tr}\sin 2\phi_{tr}\,\mathrm{Re}(\mathrm{A}_0^*\mathrm{A}_\parallel) \\ &+\frac{1}{\sqrt{2}}\sin 2\theta_1\sin 2\theta_{tr}\cos\phi_{tr}\,\mathrm{Im}(\mathrm{A}_0^*\mathrm{A}_\perp)\big\}. \end{aligned} \tag{1.136}$$

The CP-conjugated equation is similar to this one but with the A terms replaced with $\overline{A}$. When integrating over ϕ_{tr}, the interferences between different polarization states disappear:

$$\frac{1}{\Gamma}\frac{d^2\Gamma(\overset{(-)}{B}\to D^{*+}D^{*-})}{d\cos\theta_{tr}d\cos\theta_1} = \frac{9}{64}\sum_{i=0,\parallel,\perp} |\overset{(-)}{A_i}|^2 H_i(\cos\theta_{tr},\cos\theta_1). \tag{1.137}$$

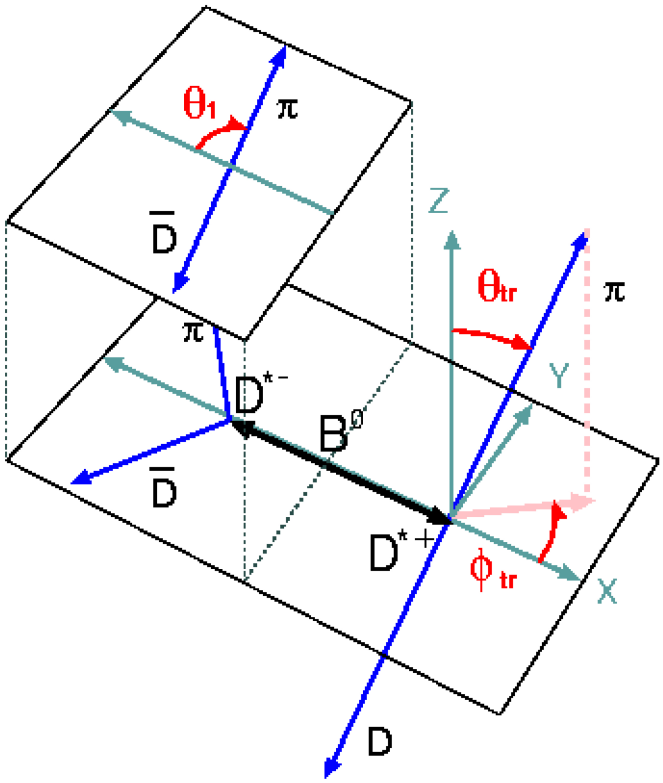


Figure 1.12: Transversity basis and angles for the $B^0 \to D^{*+}D^{*-}$.

where the angular terms are

$$\begin{aligned} H_0(\cos\theta_{tr},\cos\theta_1) &= 2(1-\cos 2\theta_{tr})(1+\cos 2\theta_1), \\ H_{\|}(\cos\theta tr,\cos\theta_1) &= (1-\cos 2\theta_{tr})(1-\cos 2\theta_1), \\ H_{\perp}(\cos\theta_{tr},\cos\theta_1) &= 2(1+\cos 2\theta_{tr})(1-\cos 2\theta_1). \end{aligned} \tag{1.138}$$

From these equations it is seen that the three helicity states can be disentangled, by the different shape of the $\cos\theta_{tr}$ and $\cos\theta_1$ distribution depending on the transversity. However to obtain the dilution factor for the CP fit, D, it is sufficient to obtain the fraction of the CP-odd component defined as:

$$R_{\perp} = \frac{|A_{\perp}|^2}{|A_0|^2+|A_{\perp}|^2+|A_{\|}|^2}. \tag{1.139}$$

By integrating equation (1.137) over θ_1, one obtains an equation where the CP-odd component can be extracted by a one-parameter fit over θ_{tr}:

$$\frac{1}{\Gamma}\frac{d\Gamma(B\to D^{*+}D^{*-})}{d\cos\theta_{tr}} = \frac{3}{4}(1-R_{\perp})\sin^2\theta_{tr}+\frac{3}{2}R_{\perp}\cos^2\theta_{tr}. \tag{1.140}$$

Thus using the transversity basis allows us to define a variable $\cos\theta_{tr}$ of which its distribution reveals the CP-odd fraction in the data sample. The theoretical prediction of the CP odd fraction in the data sample is 11% [1].

Chapter 2

The Belle experiment at KEKB

We introduce in this chapter the Belle experiment at the asymmetric-energy accelerator KEKB in Tsukuba, Japan. The special features of the detector are presented as well as the data-acquisition. Finally some analysis tools such as the flavor tagging, luminosity measurement and Monte Carlo techniques are developed.

THE practical aspects of measuring mixing and CP-violation phenomena will be explained in this chapter. The challenges involved in the experimental setup can be divided into three groups: obtaining optimal initial conditions for the $B^0\overline{B^0}$ state, measuring the time-evolution process and detecting the final states. The latter involves determining the flavor of the B-mesons which can be obtained with a "tagging" technique (see Section 2.7.4). The $B^0\overline{B^0}$ pair is produced through the $\Upsilon(4S)$ resonance. The asymmetric accelerator accounts for a boost of the system which allows us to measure the proper time difference between the neutral B-mesons.

2.1 The $\Upsilon(4S)$ production

The first observation of the $\Upsilon(1S)$ meson was made in 1977 by the CFS Collaboration [26]. Experiments at the CESR and DORIS machines confirmed the existence of this bound $b\bar{b}$ state with $J^{PC} = 1^{--}$ and measured even higher excited states. The e^+e^- hadronic cross-section was further explored as a function of the center-of-mass (CM) energy, in particular by the CLEO experiment. The result is shown in Figure 2.1; the first three peaks correspond to the $\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$ excited states and are very narrow. Their measured width is merely dominated by the energy resolution of the detector.

The $\Upsilon(4S)$ state has an energy just $20\,\mathrm{MeV}$ above the threshold for $B\overline{B}$ production and has a $10.58\,\mathrm{GeV/c^2}$ mass. At higher energies the $\Upsilon(5S)$ shows up. The $\Upsilon(4S)$ decays more than 96% of the time into a $B\overline{B}$ state, which is either a $B^0\overline{B}^0$ $(49.1 \pm 0.7\%)$ or a B^+B^- $(50.9 \pm 0.7\%)$ pair. The B^+B^- couple does not mix but can be used to test direct CP violation. Due to this almost exclusive decay to neutral and charged $B\overline{B}$ decays, the $\Upsilon(4S)$ makes a very practical state for CP-violation studies in the B-system. Many

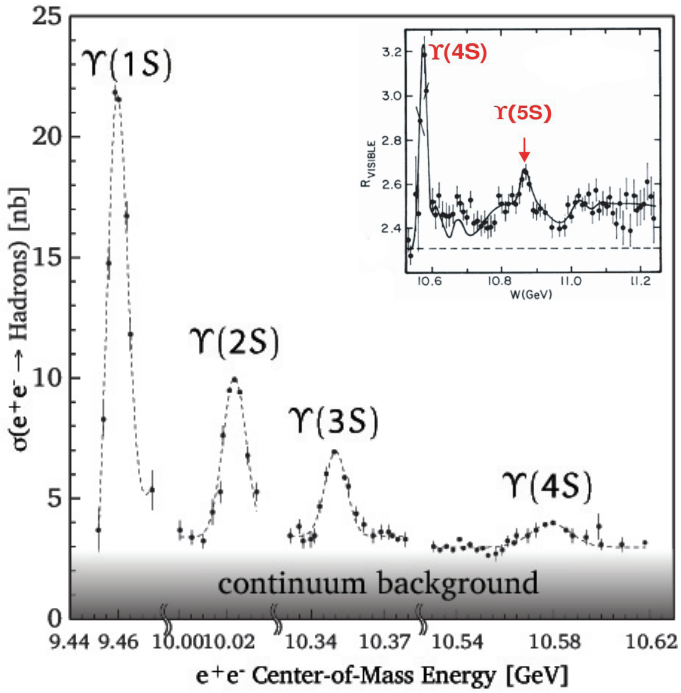


Figure 2.1: e^+e^- hadronic cross-section as a function of the CM energy, measured at CLEO and CUSB, illustrating the different Υ resonances. The gray band symbolizes the $q\bar{q}$ continuum background. The insert on the top right shows the cross-section for higher CM energies, where the $\Upsilon(5S)$ resonance appears. The figures are taken from [27] and [28].

experiments dedicated to CP violation have used this excited state, amongst those CLEO and the two B-factories Belle and BaBar; the first two experiments are still running. BaBar stopped taking data in the spring of this year, 2008.

Figure 2.1 shows also that the e^+e^- annihilation does not only create the excited $\Upsilon(4S)$ state around $10.58\,\mathrm{GeV}$ but also non-resonant contributions which are referred to as "continuum". It consists of lighter pairs of the u, c, d or s quarks. The continuum production cross-section accounts for about three quarters of the total electron-positron cross-section at the $\Upsilon(4S)$ resonance. In Chapter 3 we will explain how selection criteria are used to discriminate these continuum events from the $B\overline{B}$ decays.

2.2 The KEKB asymmetric-energy accelerator

Since 1980 the experiments ARGUS, using the accelerator DORIS-II at DESY (Germany) and CLEO, using the accelerator CESR at Cornell (United States) have produced many $B^0\overline{B}^0$ and B^+B^- pairs. They discovered the mixing between the B^0 and anti-B^0 mesons and were able to measure the absolute values of elements of the CKM matrix.

These two experiments were not able, however, to measure any CP violation in the B system, therefore two new experiments were built: Belle at KEK (Japan) and BaBar at SLAC (United States), using the KEK-B and PEP-II accelerator facilities. These accelerators have electron-positron beams with an asymmetric energy so that the $\Upsilon(4S)$ is produced in a boosted system. The B mesons will thus have a measurable decay length.

The construction of the KEKB collider took place from 1994 to 1998. The two beams are stored in a 3 km long ring and are fed by a linear accelerator, Linac, as shown in Figure 2.2. The "High Energy Ring " (HER) contains the accelerated electrons while the "Low Energy Ring" (LER) stores the positrons. A detailed description of the KEKB accelerator can be found in [29]. The center-of-mass energy ($\sqrt{s}$) of the two beams corresponds to the mass of the $\Upsilon(4S)$ particle:

$$\sqrt{s} = 2\sqrt{E_{\mathrm{HER}}E_{\mathrm{LER}}} = 10.58\,\mathrm{GeV}.$$

In order to accommodate for the boosted production of the $\Upsilon(4S)$, the energy of the LER beam is $3.5\,\mathrm{GeV}$ while the HER beam's energy is $8.0\,\mathrm{GeV}$. The created boost is:

$$\beta\gamma = \frac{E_{\mathrm{HER}} - E_{\mathrm{LER}}}{\sqrt{s}} = 0.425.$$

This is also the boost of the B mesons (since they are produced almost at rest in the $\Upsilon(4S)$ frame), and hence the B mesons fly on average $0.2\,\mathrm{mm}$ before decaying.

The design current is $1.1\,\mathrm{A}$ in the HER and $2.6\,\mathrm{A}$ in the LER. A bunch contains

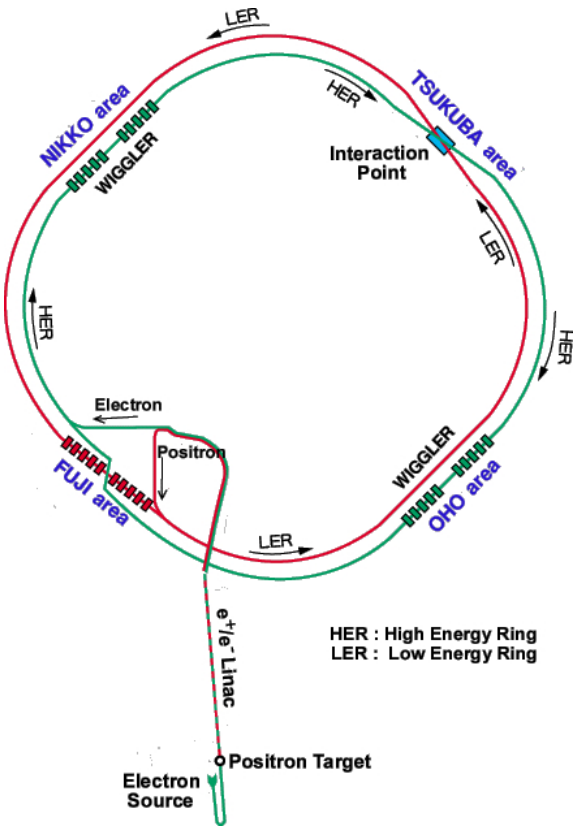


Figure 2.2: Schematic view of the KEKB and the Linac accelerator.

around 10^{10} particles and has a length $\sigma_z = 4.0\,\mathrm{mm}$. When the machine runs at its full capacity it hosts 5000 bunches per ring, corresponding to a bunch crossing rate of $2.1\,\mathrm{ns}$.

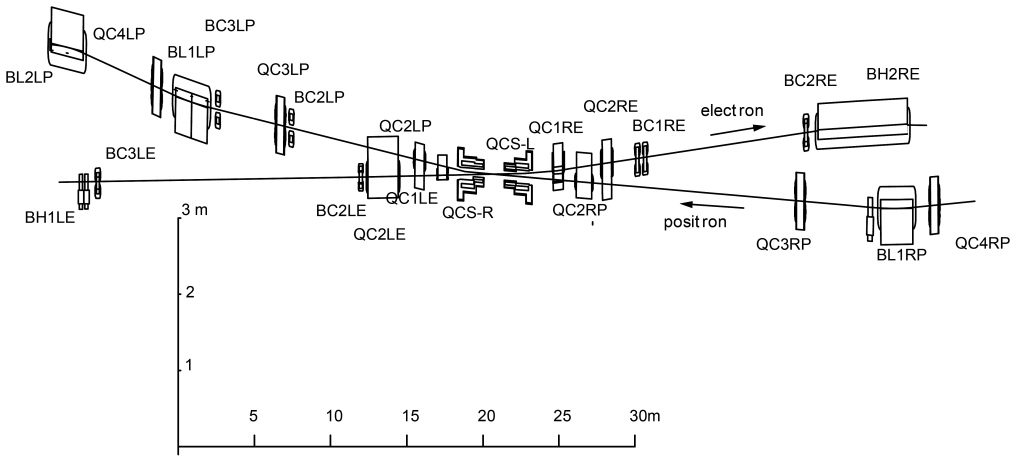


Figure 2.3: Schematic view [30] of the beam line and magnets around the interaction point.

The performance of the accelerator is measured by its delivered luminosity, L, which is defined as:

$$R = L \times \sigma,$$

where R is the rate of e^+e^- collisions and σ the total e^+e^- cross section. The luminosity measurement is described in Section 2.7.2. The peak luminosity is the maximum luminosity at the start of a fill. The specific luminosity refers to the luminosity per bunch divided by the product of bunch currents. Due to the exponential decrease of luminosity of the beams in time, they are only stored for a duration of eight hours, while the Belle detector takes data. After that the beams are dumped and HER and LER are refilled.

The design luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ was reached in May 2003 and since 2004 the machine has been operating in a "continuous" injection mode. Since then the luminosity has continued rising and in November 2006 a peak luminosity of $1.7 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ was reached.

At the interaction point (IP), the two beams collide with an angle of $\pm 11\,\mathrm{mrad}$ in order to reduce beam-interference background and to simplify the magnet construction. A schematic view of the magnets and beam line near the interaction point is shown in Figure 2.3. In January 2007 crab-cavities were installed to tilt the head-to-tail of the bunch just before the IP to increase the luminosity. This has increased the specific luminosity but the peak luminosity hasn't broken any previous records yet (see Figure 2.4).

The total integrated luminosity delivered by the KEKB accelerator reached $836\,\mathrm{fb}^{-1}$ on May 27th, 2008. The top plot in Figure 2.5 shows how the daily luminosity increased over the years, while the bottom plot shows the integrated luminosity as well as the different energy regions which were scanned. Finally Table 2.1 illustrates yet another example of the excellent performance of the accelerator, which has largely surpassed its design goals.

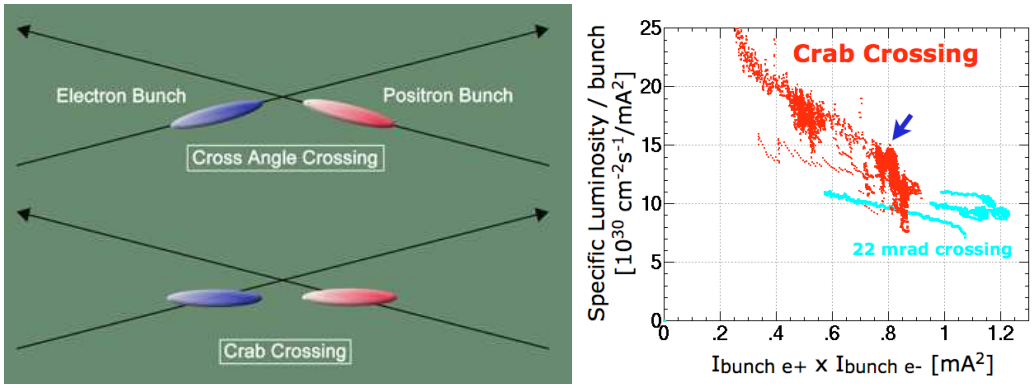


Figure 2.4: Left: bunch-crossing scheme, the top configuration shows the classical approach with the 22 mrad crossing angle, while the bottom plot of the figure shows how the crab cavities have tilted the bunches. Right: The specific luminosity with respect to the beam current. The light blue dots show the performance in the classical approach, the red when the crab cavities are installed.

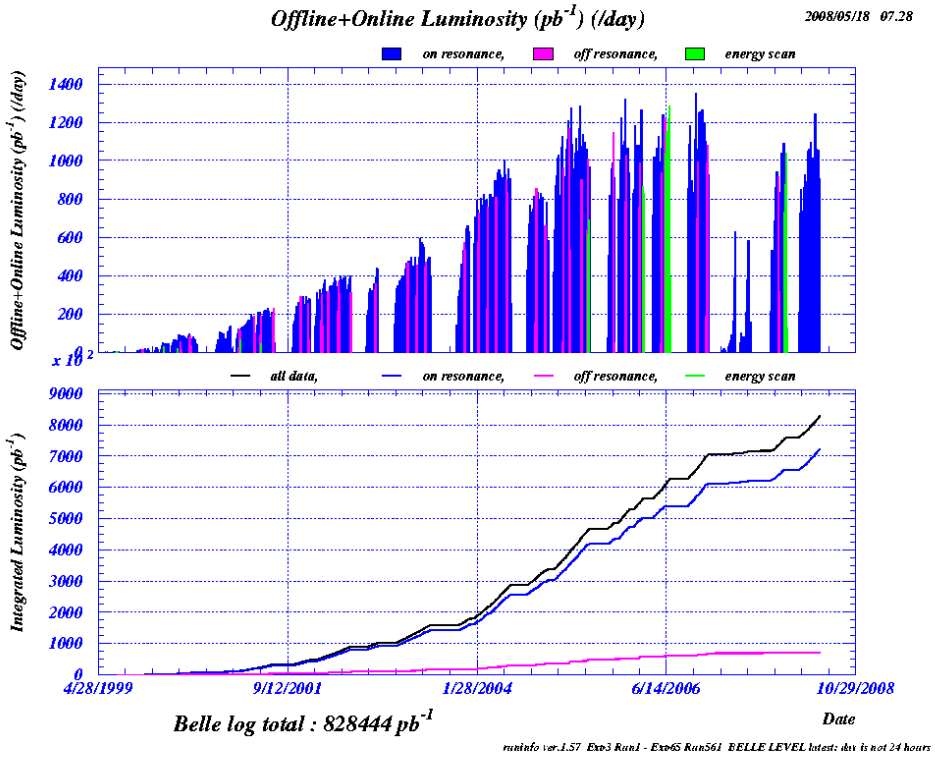


Figure 2.5: Daily (top) and total integrated luminosity (bottom) recorded by Belle since the beginning of data-taking.

Circumference	3016				m
	Design		**Record**		
Peak luminosity	1.0		1.712		$10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$
$\int$ Lum / day	~ 0.6		1.23		fb^{-1}
	Design		**19/05/2008**		
Bunch spacing	0.59		2.1		m
Number of bunches/ring	5000		1584		
RF frequency	~ 509		~ 509		MHz
	LER	HER	LER	HER	
Bunch current			1.0	0.6	mA
Beam current	2.6	1.1	1.6	0.9	A
Particles/bunch	3.3	1.4			10^{10}

Table 2.1: Design and actual machine parameters of KEKB [31].

2.3 The Belle detector

The next step after the creation of the initial $B^0\overline{B}^0$ state is the accurate measurement of the decay time and decay particles of both B mesons. In the study presented in this dissertation we reconstruct the decay of a neutral B meson to $D^{*+}D^{*-}$. This B meson is referred to as the B^{CP}, while the other B meson in the event is called the B^{tag}. The reconstruction of the decay particles allows us to identify the B_{CP} meson. Its flavor can be determined by the tagging method, explained in Section 2.7.4. The apparatus used to measure the properties of the decay particles is the Belle detector. We will first give an overview of the design of the detector, followed by a description of the techniques used to extract the relevant information of the measured data.

The detector is constructed around the interaction point of the accelerator to measure the particles which are created from the collision of the electron and positron beams. Most particles have a very short lifetime and decay before they can be detected. Only the lighter, final decay particles which live long enough such as protons, kaons, pions, electrons, neutrons, gammas and muons can leave a signal in the detector.

The detector's acceptance covers as much as 93% of the total solid angle. It is built in a typical "multi-layer" full-featured detector. Each subdetector has its own specialty but only by combining the different inputs can one obtain a meaningful physics event.

Let us define the detector coordinate-system as illustrated in Figure 2.6.

- The x axis is the radial horizontal axis from the center of the accelerator outwards.
- The y axis is the vertical axis.
- The z axis points anti-symmetric to the positron beam.

The three axes form a right-handed coordinate system. The origin is defined in the crossing point of the beams. Often cylindrical coordinates will be used to incorporate the detector's

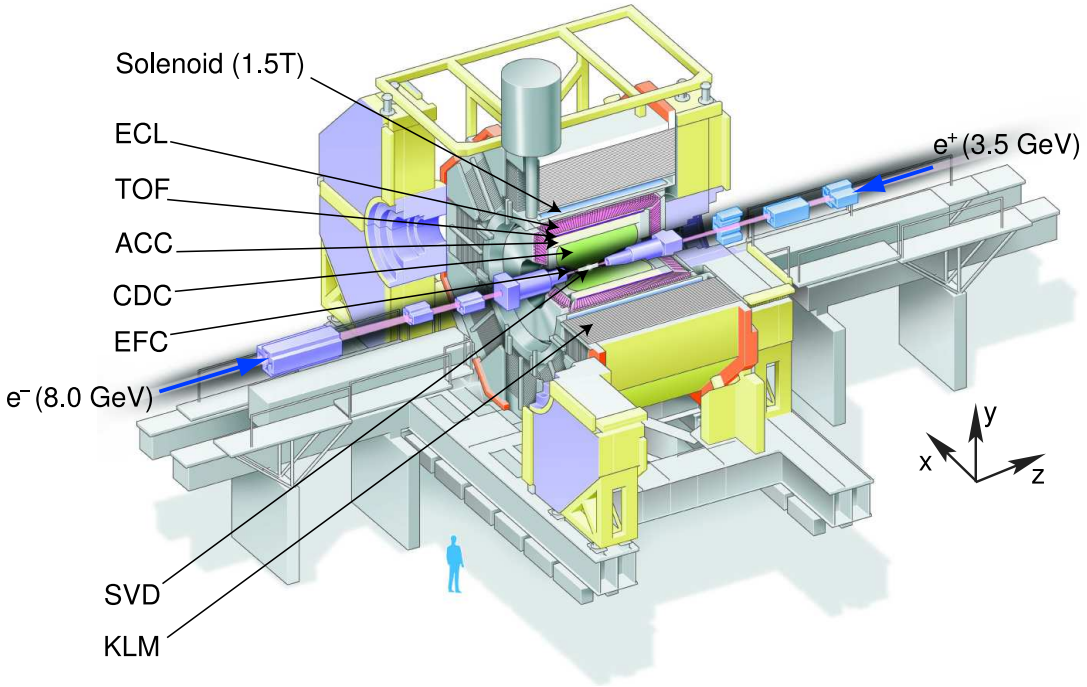


Figure 2.6: Design of the Belle detector.

symmetries. The radial distance is defined as $r = \sqrt{x^2 + y^2}$. It is useful to define two angles: the polar angle, θ, which is the angle measured with respect to the z axis and the azimuthal angle ϕ, measured with respect to the x axis. The detector clock is synchronized with that of the KEKB accelerator which is determined by the RF radio frequency. This also corresponds to the design bunch crossing which take place every $2\,\mathrm{ns}$.

The layout of the Belle detector is schematically shown in Figure 2.7. The trajectories of the charged particles are bent due to the $1.5\,\mathrm{T}$ magnetic field provided by the **superconducting solenoid** which is placed outside most of the subdetectors. The curvature is proportional to the particle's momentum. The track position along the particle's flight is measured by the **Silicon Vertex Detector (SVD)** and the **Central Drift Chamber (CDC)**, the two innermost subdetectors. The sensitive elements of the SVD start as close as $2.0\,\mathrm{cm}$ from the interaction point and are therefore also used to determine the decay vertex positions. Due to the large particle flux and the high vertex resolution required ($\sim 50\,\mu$m in z), the SVD is constructed with silicon sensors. The CDC measures, in addition to the charged track's position, the particle's energy loss which is used to identify pions, kaons and electrons. Two more detectors have as their main purpose particle identification (PID), the **Aerogel Cherenkov Counter (ACC)** and **Time-of-Flight detector (TOF)** subdetector. The ACC measures the Cherenkov light emitted by charged particles. Combined with the measured momentum it gives access to the mass of the particle and allows us to distinguish kaons from pions. The TOF measures the time between the interaction and the moment it detects the particle.

Around these light inner-detectors the **Electromagnetic calorimeter (ECL)** is installed, which has the purpose of measuring the energy of electrons and photons by making them interact with the dense detector medium where they deposit their energy in the form of a shower. The size and width of the shower is a measure of the energy of the particle.

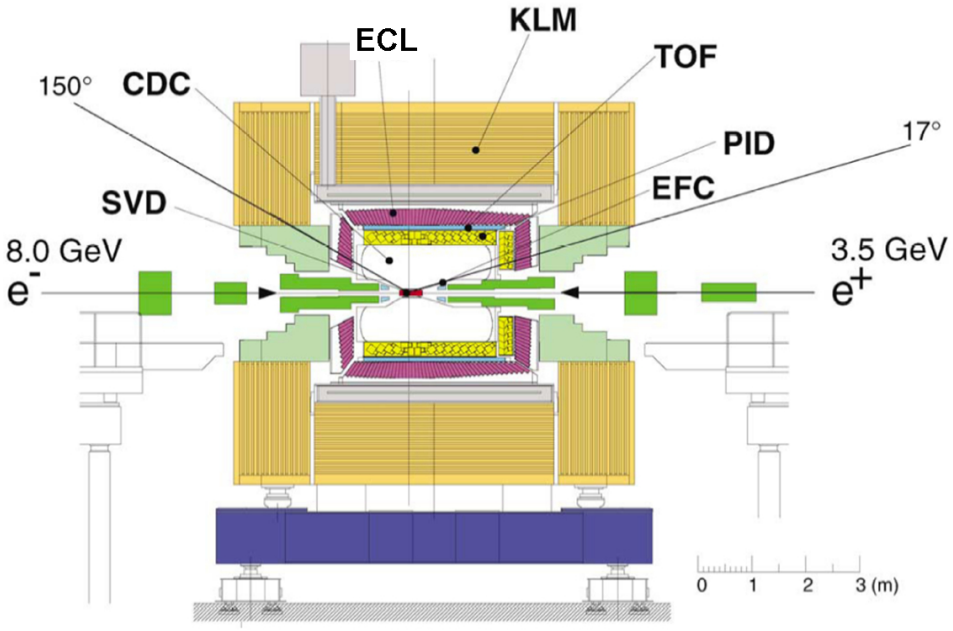


Figure 2.7: Side view of the Belle detector with its subdetectors.

Because the ECL is sensitive to electrons but hardly sensitive to hadrons, we can identify particles which are measured by the ECL and which left a signal in the tracking detectors as electrons. Photons can be separated from electrons by requiring signals in the ECL with no associated track measurement. Finally, outside the magnetic coil the **KLM** detector is installed, which detects K_L particles. The K_L particles decay into pions in this outer detector due to its long lifetime. Also muon tracks are measured after having traversed all the other subdetectors.

In the next section a more technical overview of each subdetector will be given but for a full and detailed description we refer to [32, 33] and [34]. The pictures on the following pages are taken from [32], unless stated otherwise.

2.3.1 The beam pipe

The vacuum around the beams is assured by the beam pipe. The two separate vacuum pipes around the LER and HER beams merge together at the level of the detector. The z vertex-position resolution of tracks, measured by the SVD detector, is limited by the multiple Coulomb scattering in the beam pipe wall. Therefore the beam pipe's material had to be reduced to a minimum. Furthermore the vertex resolution is inversely proportional to the distance between the first detection layers of the SVD and the interaction point. The need for a thin and small beam pipe radius is complicated by the beam-induced heating which can rise to a few hundred Watts. Therefore the beam pipe is made of a double-wall beryllium cylinder with an inner radius of just $2.0\,\mathrm{cm}$. Helium gas is flushed in the gap $(2.5\,\mathrm{mm})$ between the two layers for cooling. A cross section view of the beam pipe is shown in Figure 2.8. During the detector upgrade in 2003 the inner radius of the beam pipe was reduced to $1.5\,\mathrm{cm}$.

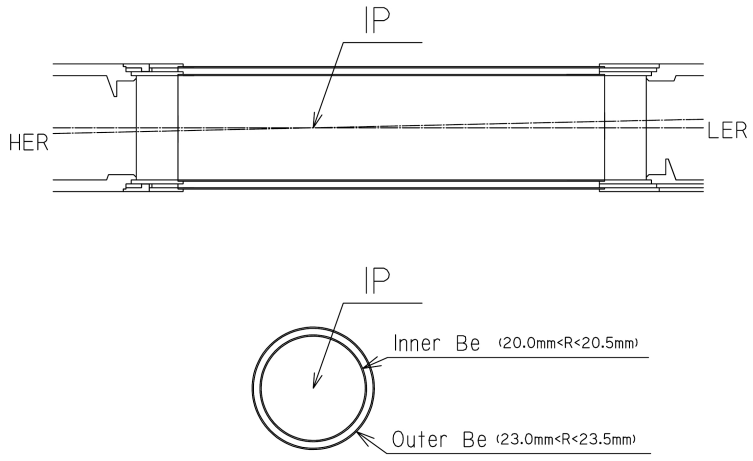


Figure 2.8: Cross-section view of the beam pipe before the detector upgrade.

2.3.2 The silicon vertex detector (SVD)

The Silicon Vertex Detector is an essential subdetector for time-dependent CP studies. Its key purpose is to provide a good measurement of the separation between the B^0 and the $\overline{B}^0$ decay vertices. The required z-position resolution is $\approx 100\,\mu$m. The SVD contributes also to the overall tracking performance. Because most of the particles of interest have a momentum lower than $1\,\mathrm{GeV}$, the resolution is dominated mainly by Coulomb scattering in the material on the particle's trajectory. The detector is therefore built of light materials and its readout electronics are kept outside the tracking volume. The detector's material also has to be resistant to the high radiation dose coming from the beam background. The structure of the vertex detector contains several layers of silicon sensors placed in a cylindrical configuration around the beam pipe. The first layer is as close as possible to the interaction point to improve the vertex resolution.

Double-Sided Silicon Detectors (DSSD) of n-type silicon are found to correspond best to the above mentioned criteria. A charged particle which traverses the depleted pn junction of the detector creates holes (e^+) and electrons along its trajectory (see Figure 2.9). These holes originate from the liberated valence electrons which go to the conduction band and leave behind a vacuum position. The electrons (holes) will migrate to the nearby p^+ (n^-) strip on the surface of the silicon layer. The induced current is then read out by the front-end electronics.

The silicon detectors, produced by Hamamatsu Photonics K.K.(Japan), were originally designed for the DELPHI micro-vertex detector at CERN [35]. The sensor's dimensions are $57.5 \times 33.5\,\mathrm{mm}^2$. The silicon modules contain 640 p^+ strips, parallel to the beam axis for the measurement of the azimuthal ϕ angle and 640 n^- strips, perpendicular to the beam axis for the tracks z-position measurement. The p^+ and n^- pitch-strip distance is $24\,\mu$m and $42\,\mu$m respectively.

The SVD detector was upgraded in 2003 and has since been referred to as the SVD2 detector [36] (and the previous version is now called SVD1 [37]). The main difference is the installation of an additional layer in the SVD2 detector which reduces the innermost radius from $30\,\mathrm{mm}$ to $20\,\mathrm{mm}$. To make this change possible the beam pipe radius

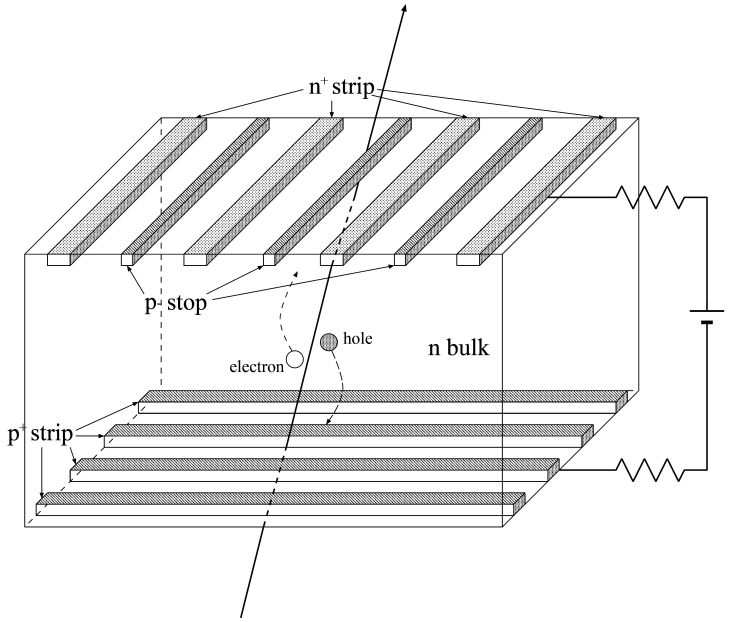


Figure 2.9: Illustration of a particle detection in a semiconductor.

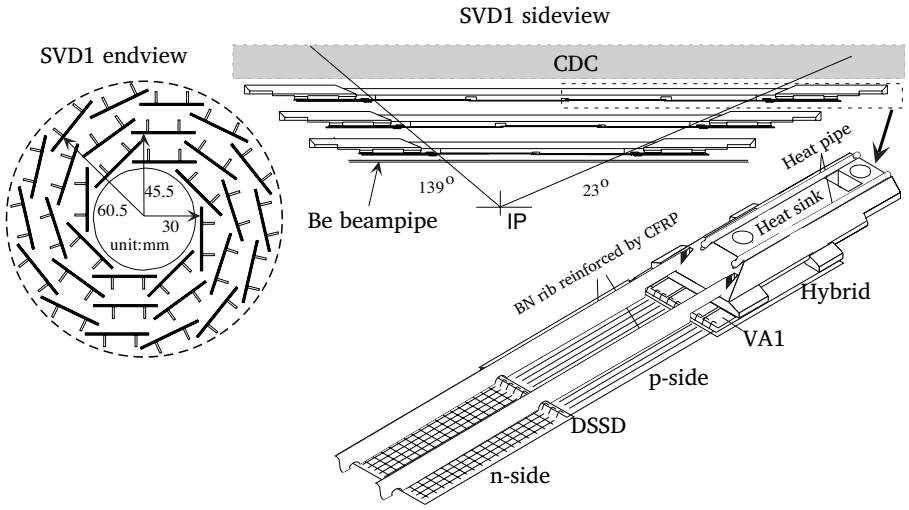


Figure 2.10: Design of the SVD1 subdetector.

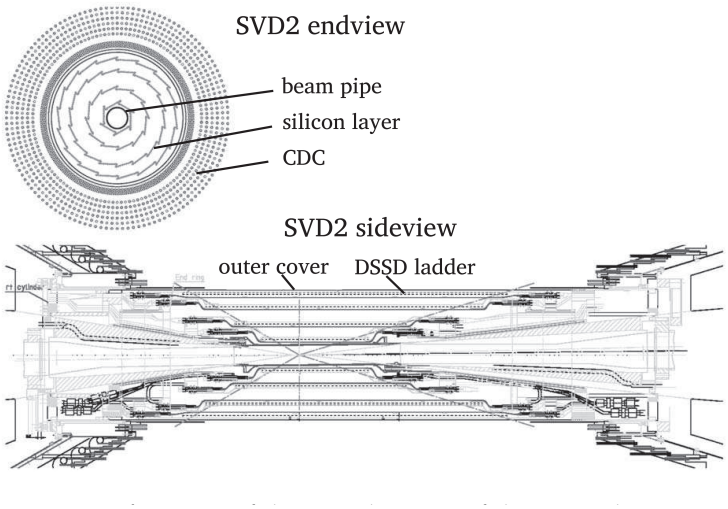


Figure 2.11: Design of the SVD2 subdetector.

had to be reduced from $20\,\mathrm{mm}$ to $15\,\mathrm{mm}$. The design of both SVD versions is shown in Figure 2.10 and Figure 2.11. Due to the extra detection layer in the SVD2 detector, the angular coverage was enlarged from $23^o < \theta < 139^o$ (86% of the total solid angle) for SVD1 to $17^o < \theta < 150^o$ (92% of the total solid angle) for SVD2. The SVD1 layers are built of 8, 10 or 14 ladders which are equipped with two or four DSSDs. The total number of readout channels is 81920 for SVD1 and 110592 for SVD2 corresponding to 102 DSSDs (SVD1) and 138 DSSDs (SVD2). The resolution on the impact parameter measured by the SVD1 detector is shown in Figure 2.12. The SVD upgrade increased the vertex resolution by 25% ($R\phi$ resolution becomes $12\,\mu\mathrm{m}$, z resolution becomes $19\,\mu\mathrm{m}$) and made it possible to reconstruct charged tracks using only SVD hits. The dead time is reduced due to the revised readout electronics and the radiation hardness is improved. Figure 2.13 shows the hits and tracks from a hadronic event in the SVD2 detector.

2.3.3 The central drift chamber (CDC)

The CDC detector measures the bended tracks of the charged particles which are used to determine precisely the corresponding momentum. The momentum resolution requirement for the physics analysis is: $\sigma_{p_t}/p_t \approx 0.5\%\sqrt{1+p_t^2}$ for charged particles with $p_t < 100\,\mathrm{MeV/c}$ in the polar region of $17^o \leq \theta \leq 150^o$ (p_t expressed in GeV units). The CDC is immersed in the uniform magnetic field which points in the z direction. When a charged particle moves through this magnetic field it follows a helix path along the field lines. The helix can be decoupled into a circular motion in one plane and a constant movement in the plane perpendicular to the circular plane. From this information the particle's momentum and the closest approach to the interaction point can be determined [38].

The design of the CDC detector is shown in the top plot of Figure 2.14. This small-cell drift chamber covers the region of $17^o \leq \theta \leq 150^o$ (corresponding to the coverage of the SVD2 detector) and consists of 50 anode sense-wire layers (for a total of 8400 readout channels) and three cathode strip-layers (corresponding to 1792 readout channels). The bottom plots of Figure 2.14 show the internal cell configuration. The detector is furthermore filled with a light mixture of helium and ethane gas. When a charged particle

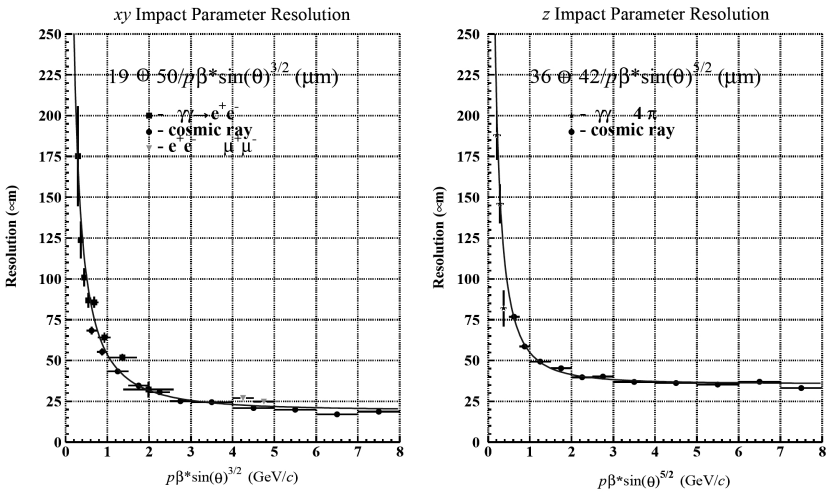


Figure 2.12: Resolution on the impact parameter in xy (left) and z (right) measured by the SVD1 detector.

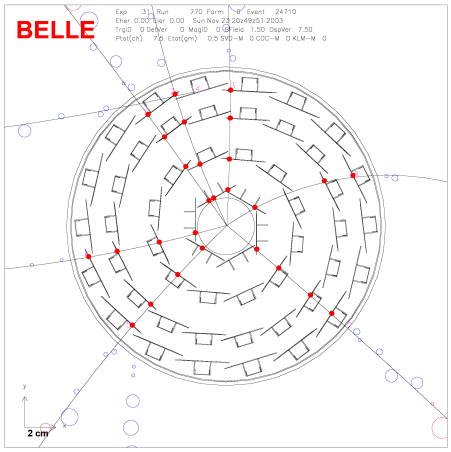


Figure 2.13: A hadronic event recorded in the SVD2 detector (front view).

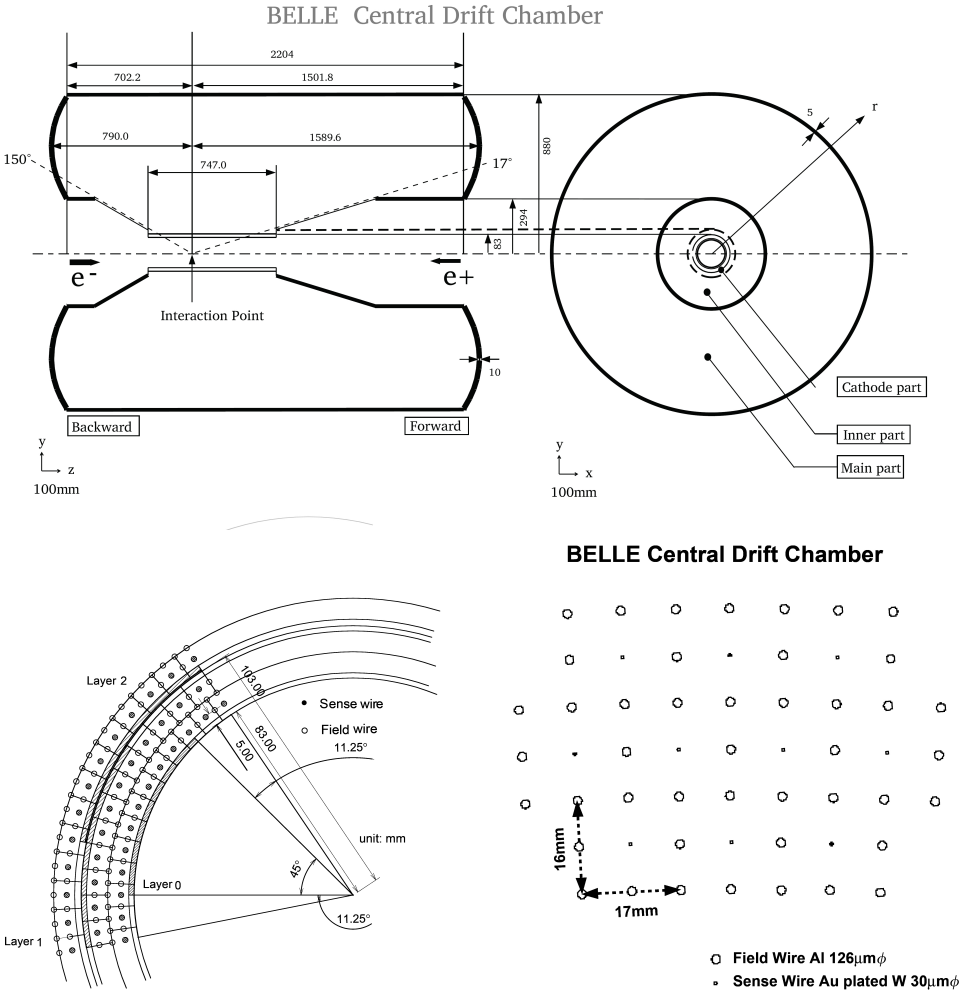


Figure 2.14: Top: general CDC detector layout. Bottom: wire arrangement in the CDC detector.

traverses the detector an electron on the outer layer of the gas atom can be kicked out due to the Coulomb interaction. These freed electrons are accelerated by the electrical field induced by the anode and cathode wire, in the direction of the closest anode wire. While being accelerated towards the wire, the electron in its turn will liberate more electrons. The induction of these accumulated charges provoke an electrical pulse.

The CDC detector also got upgraded during the summer of 2003. Due to the larger outer dimensions of the SVD2 detector the CDC's inner radius had to be increased. This is done by removing three inner layers of and replacing them by two layers of smaller so-called small-cell CDC (sCDC) modules, such that the inner radius became $140\,\mathrm{mm}$. The measurements provided by the CDC detector are also linked to the SVD hits in order to improve the overall resolution.

Both the pulse height and drift time are measured. The pulse height is related to the energy deposited through the ionization of the gas. The distance of the particle to the wire can be obtained by the drift time. From the measured energy deposition and distance to the wire we can calculate the energy loss dE/dx, which is a vital component for the particle identification. Figure 2.15 shows the distribution of dE/dx versus momentum. The separation between the different particles can clearly be seen. For kaons and pions with a momentum between 0.4 and $0.6\,\mathrm{GeV/c}$ a 3σ separation can be obtained.

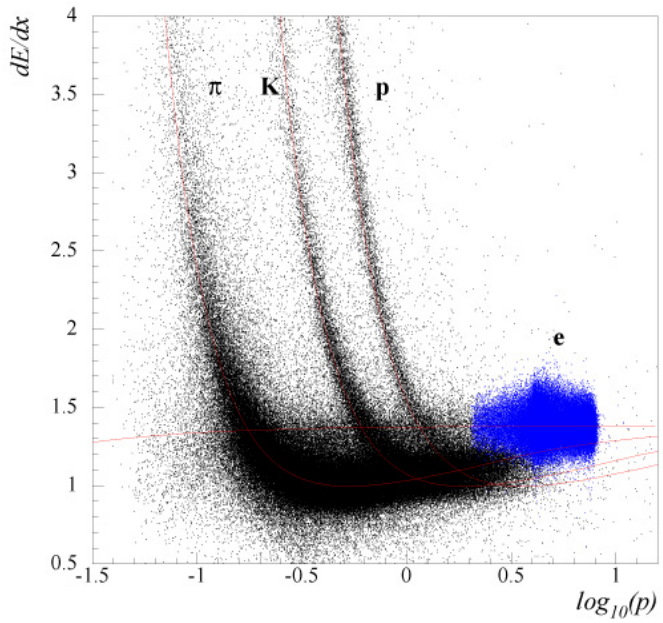


Figure 2.15: Particle identification from the CDC dE/dx measurement.

2.3.4 The aerogel Cherenkov counter (ACC)

The purpose of the aerogel Cherenkov counter is to enhance the particle identification by distinguishing pions from kaons, which is vital for many B physics measurements. The detector consists of an array of silica aerogel threshold Cherenkov-counters and extends the momentum coverage beyond the reach of the CDC dE/dx measurement (see Section 2.3.3) and the time-of-flight measurement (see Section 2.3.5).

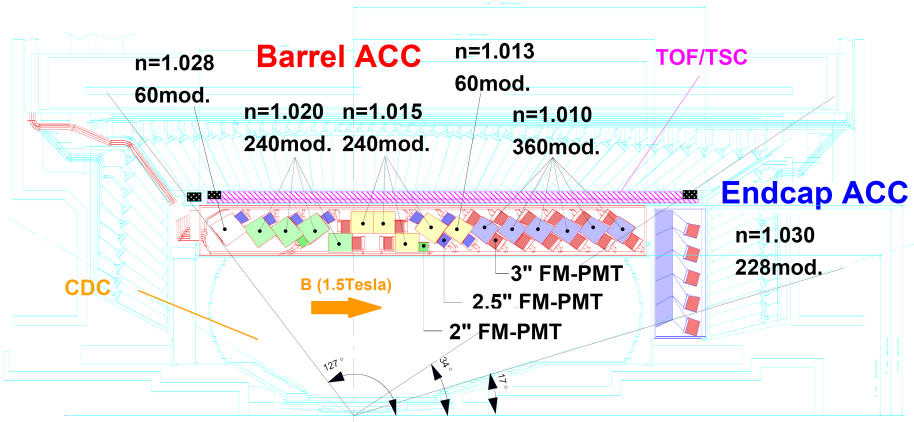


Figure 2.16: Design of the ACC detector (transverse cross section).

The principle of the particle identification method is based on the Cherenkov light emittance by particles traveling faster than the speed of light in that medium. Depending on the refractive index of the medium, the speed of light becomes

$$v_{\text{light in medium}} = c/n,$$

so when a particle's velocity v satisfies the following relationship:

$$n > c/v = \sqrt{1 + (mc/p)^2}$$

and a cone of light is emitted with an angle inversely proportional to its velocity. By choosing the appropriate refractive index, there will be a range of velocities for which pions emit Cherenkov radiation but kaons don't. The ACC allows particle identification for momenta between 1.2 to $3.5\,\text{GeV}/c$.

The detector is built of 960 aerogel counter modules in the barrel part, segmented in 60 cells in the ϕ direction and 228 modules in the forward end-cap which are arranged in five concentric layers, as shown in Figure 2.16. A typical counter module consists of five silica aerogel tiles inside a $0.2\,\text{mm}$ thick aluminum box. Each tile has a different refractive index, between 1.01 and 1.03, depending on its polar angle. The produced Cherenkov light is then read out by one or two fine mesh-type photomultiplier tubes attached directly to the box, as shown in Figure 2.17. The total number of readout channels in the barrel part is 1560. The end-cap adds another 228 channels.

For particles up to $4\,\text{GeV}$, the kaon identification efficiency is 80% or more while the pion fake rate remains below 10%. Below the pion threshold of $1\,\text{GeV}$ electron identification is possible as well.

2.3.5 The time-of-flight counter (TOF)

Particle identification with a time-of-flight detector is very efficient. It adds a third piece of information to the PID, namely in the energy regions where both the ACC and the CDD are less effective, i.e, below $1.2\,\text{GeV}$, which encloses 90% of the particles produced at the $\Upsilon(4S)$ resonance. In this range, the PID efficiency can be guaranteed as the flight length

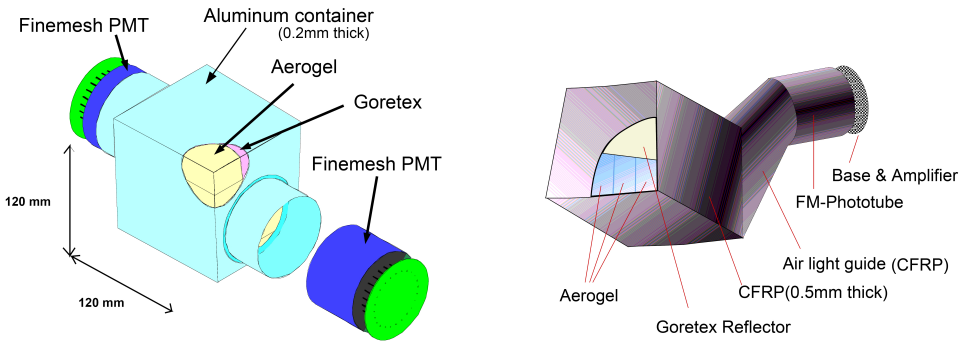


Figure 2.17: Design of an ACC barrel module (left) and end-cap module (right).

is $1.2\,\mathrm{m}$ and a time resolution is $100\,\mathrm{ps}$. The flight time, T, for a particle of mass m which travels a length L is given by

$$T = \frac{L}{c}\sqrt{1 + \left(\frac{mc}{p}\right)^2}.$$

The TOF measures the time the particle takes to travel from the interaction point to the TOF barrel. When combining the momentum measurement from the CDC with the measured velocity of the TOF, we can calculate the corresponding mass of the particle and identify it. The mass distribution calculated from TOF measurements is shown in Figure 2.19 and a clear separation between kaons, pions and protons can be seen.

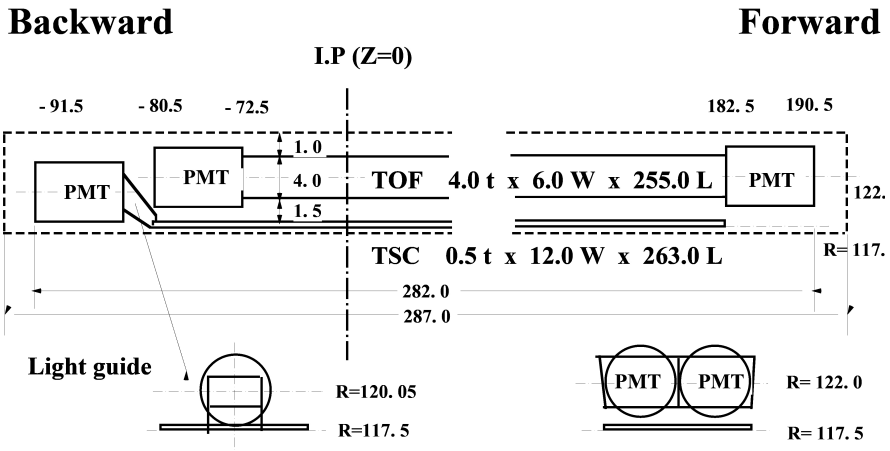


Figure 2.18: Configuration of the TOF module.

Due to its excellent time resolution the TOF detector also provides fast timing signals for the trigger. Simulation studies indicated however that to keep the fast trigger rate below $70\,\mathrm{kHz}$ in any beam background conditions, the TOF counters should be augmented by thin trigger scintillation counters (TSC).

The TOF counter and the TSC consists of fast scintillators and photomultiplier tubes, mounted directly on the scintillators. One module has two TOF counters and one TCS.

The $4.0\,\mathrm{cm}$ thick TOF counter is read out by two photomultipliers and the $0.5\,\mathrm{cm}$ thick TCS is read out by one photomultiplier. The configuration of a TOF module is shown in Figure 2.18. In the barrel region there are 64 TOF modules, which covers the polar region of $33^o < \theta < 121^o$. The TOF hit efficiency is 95% for single-end hits and 88% for both-end hits in $e^+e^- \to \mu^+\mu^-$ events. For more information on the TOF system, see Ref. [39].

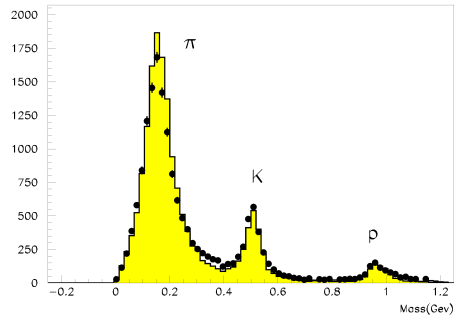


Figure 2.19: Particle identification using TOF measurements. The mass distribution shows a clear distinction between pions, kaons and protons.

When combining the information from the TOF measurement with the ACC and CDC information on dE/dx a 3σ separation between charged kaons and pions is obtained for a momentum range up to $3.5\,\mathrm{GeV/c}$

2.3.6 The electromagnetic calorimeter (ECL)

The ECL's first purpose is the detection of photons from π^0 decays and radiative B mesons, with a high efficiency and a good resolution in energy and position over a large energy range. Most photons are the end products of cascade decays and have thus a relatively low energy which asks for an excellent ECL performance below $500\,\mathrm{MeV}$. On the other hand, photons from decays such as $B \to K^*\gamma$ or $B \to \pi^0\pi^0$ produce merely photons with an energy above $4\,\mathrm{GeV}$. Figure 2.20 shows the relative energy and position resolution for photons measured by the ECL. A good resolution is also needed to separate those photons from the low energy background photons. For the detection of high momentum neutral pions, the energies of the two decay photons as well as their opening angle need to be measured accurately. This requires a fine-grained segmentation in the calorimeter.

The second purpose is providing information for the electron identification by testing if a signal in the ECL can be associated with a track with consistent energy and momentum.

The energy of the incident particle is measured by the shower it deposits in the CsI (Tl) crystals, through Bremsstrahlung and pair creation. The shape and total energy of electron showers differ greatly from hadronic showers. The latter only deposits a small amount of its total energy. The crystals are arranged in such a way that they all point roughly to the interaction point. The ECL design, shown in Figure 2.21, consists of 8736 tower-shaped crystals of which 6623 are in the barrel and the rest in the two end-caps. Each crystal has a depth of $30\,\mathrm{cm}$, corresponding to 16.2 radiation lengths and a surface of $5 \times 5\,\mathrm{cm}^2$. The crystal is then read out by a pair of silicon PIN [1] photodiodes mounted at the rear end of the crystal. A total solid angle of $12^o < \theta < 155^o$ is covered, which corresponds to 92% of the full solid angle.

[1] positive-intrinsic-negative

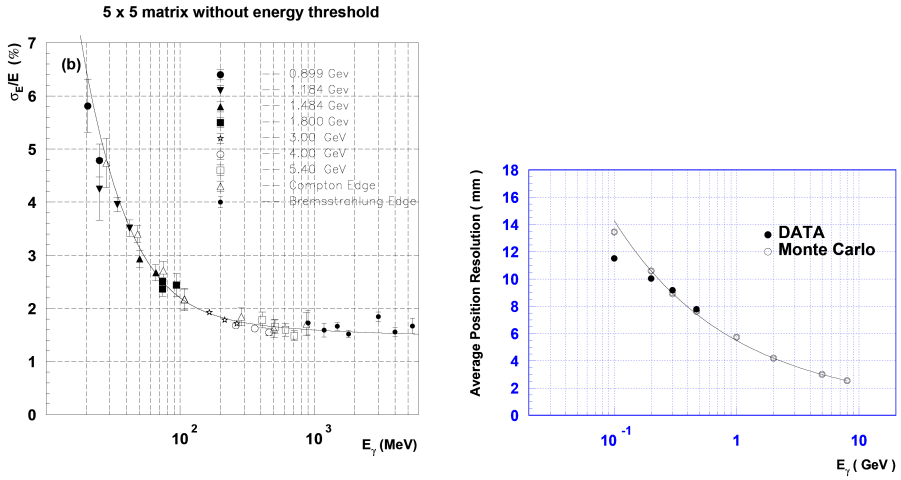


Figure 2.20: Relative energy resolution (left) and position resolution (right) for photons measured in the ECL detector.

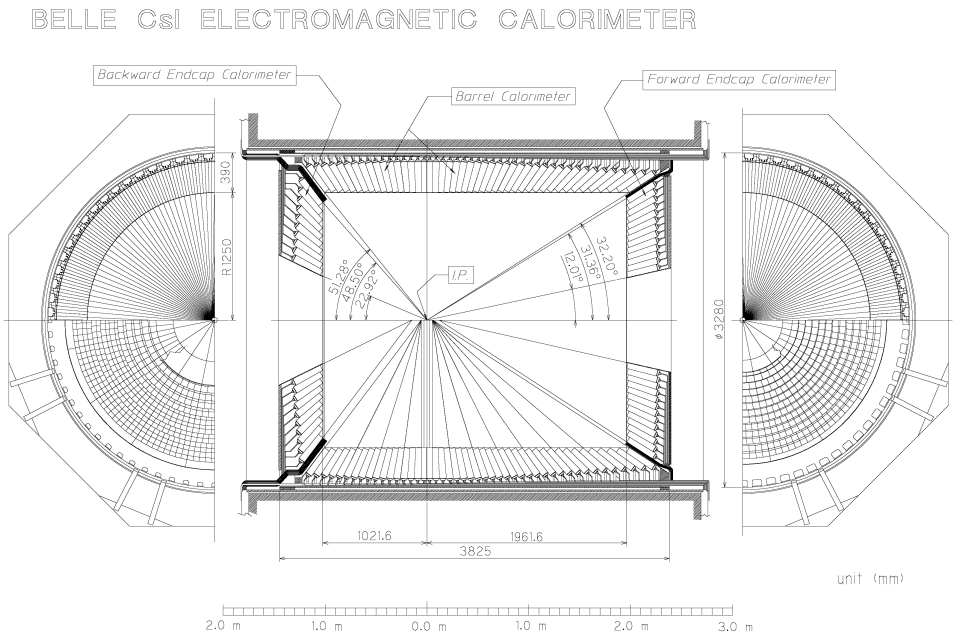


Figure 2.21: Configuration of the ECL detector.

2.3.7 The extreme forward calorimeter (EFC)

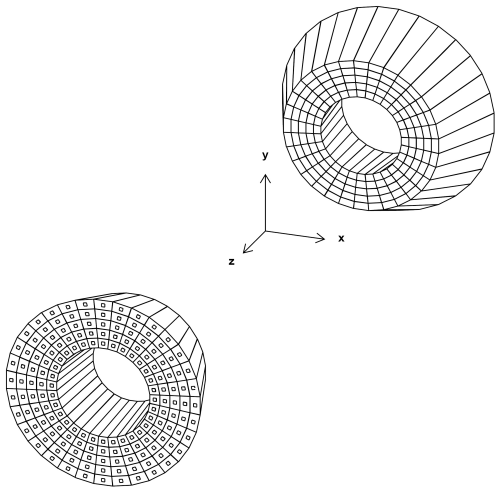


Figure 2.22: Configuration of the extreme forward calorimeter.

The EFC detector is used to extend the polar angle coverage of the ECL in the forward ($6.4^o < \theta < 11.5^o$) and backward ($163.3^o < \theta < 171.2^o$) direction. The detection of photons and electrons is based on the same principle as in the ECL. The EFC detector is installed around the beam pipe close to the IP and is subject to large radiation. Radiation hard BGO crystals (bismuth germanate $\mathrm{Bi_4Ge_3O_{12}}$) are chosen as the scintillating material, corresponding to 12 (11) radiation lengths in the forward (backward) direction. The produced scintillating light is captured by photodiodes. The 160 crystals are segmented in 32 (5) sections in ϕ (θ) as shown in Figure 2.22. The EFC has an energy resolution of 7.3% at $8\,\mathrm{GeV}$ and 5.8% at $3.5\,\mathrm{GeV}$. The EFC also has the function to mask the beam background for the CDC detector. Finally this detector plays an important role in the measurement of the delivered luminosity to Belle, as explained in Section 2.7.2.

2.3.8 The K_L and muon detector (KLM)

The KLM detector is the only Belle subdetector placed outside the magnet solenoid [40]. Its purpose is to detect K_L and muon particles. The K_L particle has a long lifetime and decays at the height of the KLM detector. The total material a K_L particle encounters while traversing the detector until it hits the KLM, corresponds to approximately one interaction length. The KLM detector therefore has to be denser in material. This is accounted for by iron plates which add 3.8 interaction lengths. The K_L particle traversing these iron plates will produce a shower of hadrons, producing a cluster of hits in the KLM. When a cluster cannot be associated with a charged track, we identify it as a K_L. Muons can be discriminated from K_L first, due to a different signal shape in the KLM detector, i.e, a line instead of a shower and second, because a muon particle leaves a track in the inner detectors. K_L and muons are detected with a high efficiency over a broad momentum range greater than $600\,\mathrm{MeV/c}$. For muons with a momentum larger than $1\,\mathrm{GeV/c}$ the detection efficiency is above 90% with a fake rate around 2%, due to non-interacting charged pions and kaons.

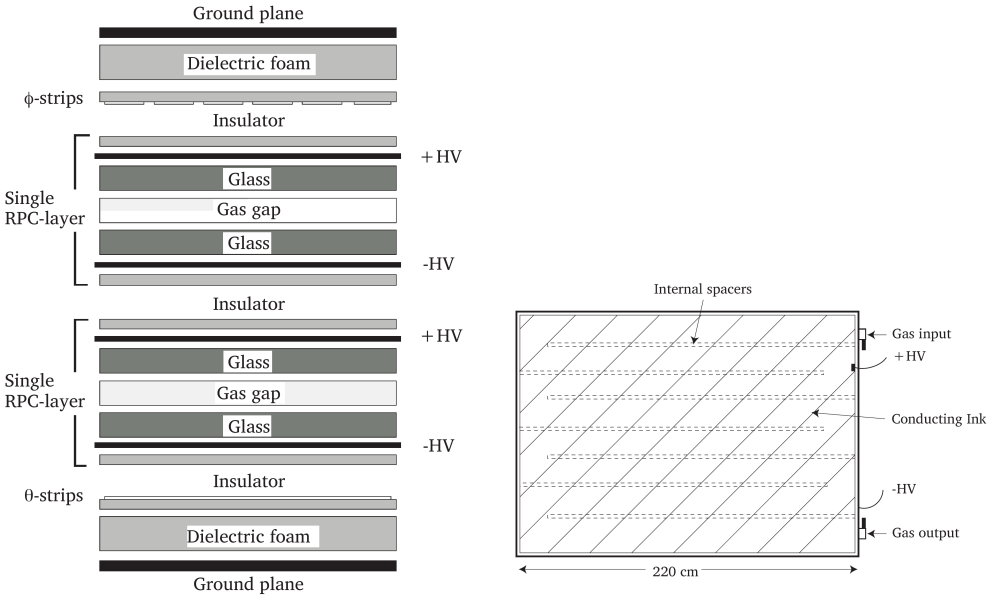


Figure 2.23: Configuration of the KLM detector (left) and its super-layers (right).

As shown in Figure 2.23 the detector consists of alternating layers of 15 Resistive Plate Counters (RPC) and 14 layers of $4.7\,\mathrm{cm}$ thick iron plates. In the end-cap region there are only 14 RPCs. The layers are then grouped in a so-called super layer, which consists of two RPC planes sandwiched between cathode strips that provide ϕ and θ information. The shower produced by the K_L allows us to measure the direction from the IP but the resolution on the energy deposition doesn't allow for an accurate energy determination. The position resolution for K_L mesons is $30\,\mathrm{mrad}$ in both angular directions. The time resolution is around a few nanoseconds. The iron plate is also used as return yoke for the magnetic field produced by the superconducting solenoid. The KLM detector covers a polar angle region of $20^o < \theta < 155^o$.

2.4 The trigger system

A trigger system decides if an event, recorded by the detector, should be saved by the data acquisition system. Most events recorded however are undesired, such as cosmic ray events, interactions in the beam pipe, synchrotron radiations or interactions between the beam and residual gas in the vacuum chamber. The production rate of these background events heavily depends on the beam quality. This calls for a trigger which is robust to unexpectedly high beam background rates. The trigger looks for typical signatures in the subdetectors which hint to an interesting event. During normal operation ($L = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$) the total rate of events is around $220\,\mathrm{Hz}$ which contains roughly $100\,\mathrm{Hz}$ of physically interesting events. Among the processes that are useful for physics analysis are the Bhabha and $\gamma\gamma$ processes but their rates are very large. The corresponding triggers need to be prescaled by 100, which is possible due to the distinct signal signature these events leave in the detector.

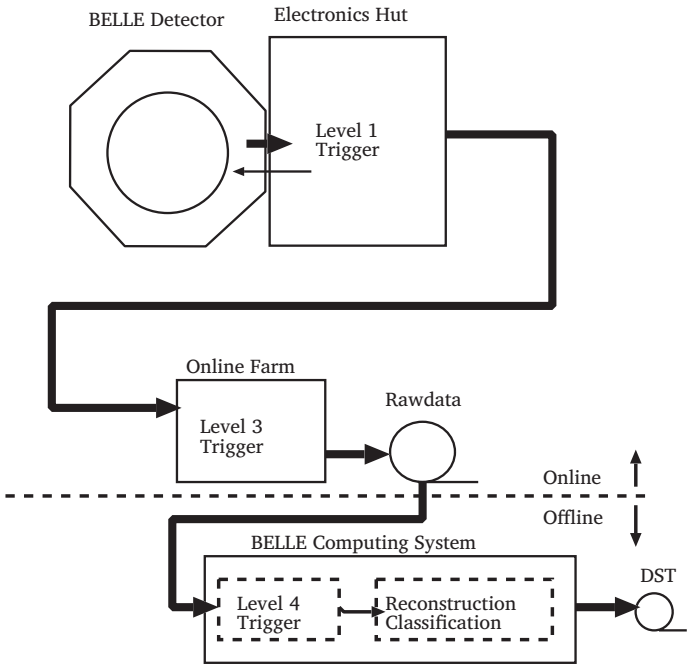


Figure 2.24: A schematic representation of the Belle trigger system.

The Belle trigger has an efficiency for hadronic events of more than 99.5%. It consists of a level-1 hardware-trigger, a level-3 software-trigger implemented in an online computer farm and a level-4 trigger which runs in the off-line Belle computing system as shown in Figure 2.24. The latter performs a more complex background reduction by fully reconstructing the event. In the next sections the different functions and designs of each subtrigger will be briefly explained.

2.4.1 The level-1 trigger

A central trigger-system, called Global Decision Logic (GDL), collects the trigger signals from each subdetector and issues the level-1 decision. An illustration is shown in Figure 2.25. The subtriggers arrive at the GDL $1.85\,\mu$s after the beam crossing, and the trigger decision is issued by the GLD $2.2\,\mu$s after the beam crossing. The CDC and TOF trigger on charged particles. The ECL trigger decision is based on the deposited energy and the number and timing of ECL cluster hits. The EFC subdetector triggers on Bhabha and two-photon events. Finally the KLM detector trigger provides a signal when muons are detected. In total 48 trigger signals are received by the GDL. For the hadronic trigger there are five main strategies:

- The two-track trigger requires two tracks with $r-\phi$ measurements and one measurement of a hit in the z-direction. The opening angle needs to be larger than 135^o. Finally there needs to be hits in the TOF and ECL clusters.

- The three-track trigger is similar to the two-track trigger but requires three or more $r-\phi$ CDC triggers. There exists more trigger types depending on the number of tracks.

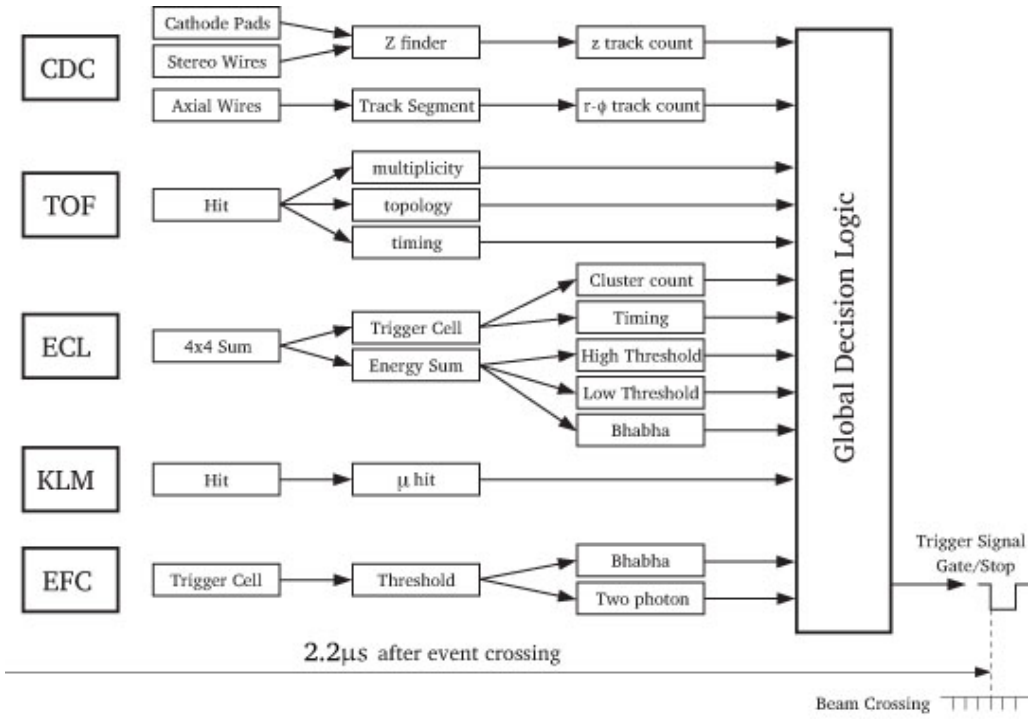


Figure 2.25: L1 trigger system layout.

- The isolated cluster counting trigger asks for four or more isolated ECL clusters to avoid Bhabha events.
- The total energy trigger requires the total sum of energy deposited in ECL to be larger than $1\,\mathrm{GeV}$. It is vetoed by the Bhabha events detected in the ECL and cosmic triggers.
- Combined trigger: this is a combination of the track, energy and cluster triggers.

When using the overlap of these triggers an efficiency of more than 99% is obtained for $B\overline{B}$ events. The individual trigger efficiency is between 90% and 97% for $B\overline{B}$. A detailed description of the Level-1 trigger can be found [41].

2.4.2 The level-3 and level-4 trigger

The purpose of the level-3 (L3) trigger is to reduce the number of events which passed the L1, first by verifying again the L1 conditions (except Bhabha and random trigger events) and then by performing a fast track reconstruction and discarding tracks for which the difference in the z direction is less than $5\,\mathrm{cm}$ from the IP. While the L3 keeps an efficiency of around 99% for τ and hadronic events, the data size is reduced by a factor of two.

The level-4 (L4) [42] trigger performs a full reconstruction using a fast tracker and rejecting tracks which do not point to the IP. The L4 trigger does not discard events as they remain in the raw data but reduces CPU time for DST production as explained in paragraph 2.6. The L4 trigger rejects around 78% of events for the DST production while keeping almost 100% of the $B\overline{B}$ events. Only the first experiments are processed using the level-4 trigger.

2.5 The data acquisition system (DAQ)

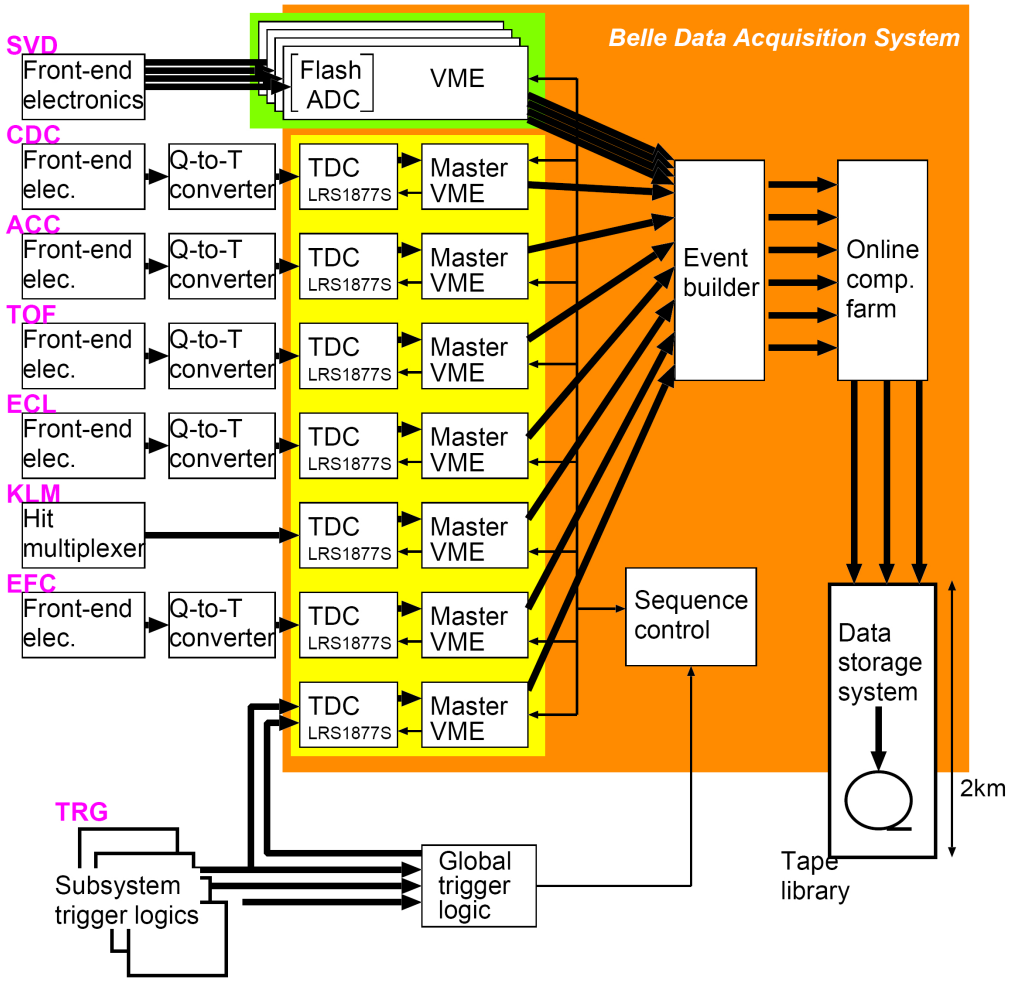


Figure 2.26: Data acquisition system.

The DAQ system needs to record events which passed the L1 to L3 requirements with a rate less than 500 Hz, while keeping the dead time below 10%. Therefore a distributed parallel-system has been set up, which is segmented into seven subsystems to handle the data from each of the subdetectors. A schematic overview of the DAQ system is shown in Figure 2.26. The data of most subsystems enter the DAQ system in the form of an electric signal. This then gets converted by a charge-to-time (Q-to-T) converter and then by a time-to-digital converter (TDC). The KLM signal's pulse shape does not provide useful information and thus no Q-to-T conversion is needed as shown in the scheme. For the SVD detector the signal is read out by flash analogue-to-digital converters (FADC) instead of TDC.

The readout timing is controlled by the sequence controller (SEQ) which gives a stop signal to the TDCs when the final trigger signal from the GDL is received. The signals from the subsystems are then transferred and combined by an event builder which records single events. The parallel data streams are converted from a "detector-by-detector" type

of information to an "event-by-event" data stream. The output is then sent to an online computer farm where the data is passed through the level-3 trigger filtering-system and combined to an offline event-format. The quality of the data is monitored by the DQM, an online data quality monitor, in the online farm. The data is then sent to the mass storage system of the KEK computing center via optical fibers and stored on tape. A typical hadronic $B\overline{B}$ or $q\bar{q}$ event has a data size of about $30\,\mathrm{kB}$ corresponding to a maximum data transfer rate of 15 MB/s.

2.6 The data processing

The data accepted by the L4 trigger are converted from raw data to fully reconstructed events and then stored on the data summary tapes (DST). At this stage the three momenta are calculated to create physics objects. The charged tracks are reconstructed from the signals of the CDC. They are then extrapolated outwards to the ACC, TOF, ECL and KLM detector or inwards to the SVD detector, to search for any matching signals. When signals are recorded in the KLM and ECL detector but there is no association possible to any CDC tracks, the particle will be considered as a neutral particle. Next the four-vectors are calculated as well as the PID likelihoods. Furthermore a series of flags and variables are determined which can be used for further analysis. All this information is stored in the DSTs using the PANTHER data format [43].

Events are further classified in so-called skims. These are subsamples of the total data set and are submitted to a handful of loose selection criteria in order to enhance the fraction of certain physics events. Most analysis, including this one, is performed on the HadronB(J) data skim which contains a larger fraction of standard hadron events [44]. The software for simulation and reconstruction of data as well as the analysis code is run in a C++ framework called the Belle Analysis Framework or BASF [45].

2.7 Tools for analyzes

2.7.1 The Monte Carlo generation

An important tool for analyzes is the Monte Carlo (MC) program. It generates physics events and simulates the detector response, which allows us to study the detector effects on our measurements. This is of vital importance for physics studies because, due to the complexity of the detector response and the large number of physics processes that comes into play when an event traverses the detector, it is impossible to disentangle the underlying physics event analytically.

Throughout this analysis we will use Monte Carlo samples many times: to define the selection cuts, to parametrize the detector response on certain variables or to compare data with MC distributions that contain considerably more events so that statistical fluctuations are reduced.

The production of MC data takes place in two stages: the generation of physics and the simulation of the detector response. In the first step particles are generated from the e^+e^- collision to the subsequent decays of very short-lived daughters. The EvtGen [46] package used for this purpose is an event generator written by the BABAR collaboration. The EvtGen module contains particle properties and event production rates as well as the relevant angular distributions for many decays. This information is collected from

several experiments in the form of world averages. Also relevant properties from the KEKB accelerator are incorporated, such as the electron and positron beam energies.

The second step is the simulation of the detector response by taking into account the detailed detector geometry, the response and inefficiencies. A BASF module called GSIM simulates the detector response using the CERN GEANT3 package [47]. The interactions between the final state particles and the detector are simulated and events are then reconstructed in the same way as the data events are. The produced Mini-DST (MDST) files [48] contain extra generation information, such as flags to identify the originally generated process. Background events are added by taking random trigger events with their corresponding noise hits. Finally evolutions in subdetectors, dead channels or the change in the size or position of the interaction region in time are also incorporated.

2.7.2 The determination of collected luminosity

The determination of the luminosity is done on two levels. The first one is referred to as online luminosity and is measured during data taking. It uses high statistics and provides a quick evaluation of the achieved luminosity. This essential feedback is necessary to optimize the beam parameters during operation. The second one is called the offline luminosity and is processed from the fully reconstructed DST files. It is dominated by the systematical error on the integrated luminosity. The offline luminosity is used for the normalization of experimental data for physics analysis and for collider performance optimization. The luminosity is a property independent of the physical processes but completely determined by the quality of the colliding beams. In e^+e^- colliders the online luminosity is mostly measured by determining the rate of $e^+e^- \to e^+e^-(\gamma)$ (Bhabha scattering) or $e^+e^- \to e^+e^-\gamma$ interactions (single Bremsstrahlung (SB)), as for these QED processes the cross sections can be calculated with a high precision (better than 1%). The photon in the Bhabha scattering gives rise to a radiation correction which limits the absolute accuracy of the luminosity measurement.

The outgoing Bhabha particles are produced under a very low angle, which the main detector does not cover [49]. This is the reason why the extra EFC detector was installed. The luminosity is determined by counting events where the scattered electron and positron simultaneously hit the forward and backward end-caps of the ECL. The Bhabha rate detected with the EFC is between 60 and $240\,\mathrm{Hz}$.

Single Bremsstrahlung interactions are characterized by a sharp angular distribution of the radiated photon, in the direction of the incoming $e^{\pm}$. This is measured by a zero degree luminosity monitor (ZDLM) [50], a device built just for this purpose.

For offline luminosity measurements a more complicated analysis is performed. Measurements of Bhabha events are taken over the whole phase-space, using not only the information from the ECL. MC techniques are used to estimate the detector coverage and experimental efficiency as well as the theoretical uncertainty on the physics model. The accuracy of the offline luminosity is around 1.4%. A more detailed explanation of these measurements is found in [51] and [52].

Finally we need to determine the number of $b\bar{b}$ events in the sample. The hadronic events produced at the $\Upsilon(4S)$ resonance are either a continuum $q\bar{q}$ event or a $b\bar{b}$ event, but the sum of both has to coincide with the total number of hadronic events collected. The number of continuum events can be estimated with the data samples collected *below* the $\Upsilon(4S)$ production energy, where no $b\bar{b}$ events are produced. However the cross-section of the continuum production is proportional to $1/\sqrt{s}$ on the CM energy. The correct relation-

ship is thus more complicated:

$$N_{\Upsilon}^{b\bar{b}} = \frac{1}{\epsilon_{\Upsilon}^{b\bar{b}}} \times \left(N_{\Upsilon}^{\mathrm{had,meas}} - N_{\mathrm{cont}}^{\mathrm{had,meas}} \times \frac{\mathcal{L}_{\Upsilon}}{\mathcal{L}_{\mathrm{cont}}} \times \frac{E_{\mathrm{cont}}^2}{E_{\Upsilon}^2} \times \frac{\epsilon_{\Upsilon}}{\epsilon_{\mathrm{cont}}} \right), \tag{2.1}$$

where $\mathcal{L}_{\Upsilon}$ is the luminosity at the $\Upsilon(4S)$ resonance and $\mathcal{L}_{\mathrm{cont}}$ that of the continuum. E is the CM energy at which events are produced, the off-resonance energy is denoted with E_{cont} and at the energy of the $\Upsilon(4S)$ resonance is denoted with $E_{\Upsilon(4S)}$. $N_{\Upsilon}^{b\bar{b}}$ is the produced number of $b\bar{b}$ events at the $\Upsilon(4S)$ resonance. $N_{\Upsilon}^{\mathrm{had,meas}}$ and $N_{\mathrm{cont}}^{\mathrm{had,meas}}$ are the number of hadrons in the on and off resonance respectively and ϵ is the reconstruction efficiency of the continuum background on- and off-resonance (ϵ_{Υ} and ϵ_{cont}) or the reconstruction efficiency of the $b\bar{b}$ events at the $\Upsilon(4S)$ resonance ($\epsilon_{\Upsilon}^{b\bar{b}}$).

2.7.3 Reconstruction of the decay length

As explained in Chapter 1, we can measure the CP asymmetry in this study by comparing the decay rate as a function of the decay length of the $B^0 \to D^{*+}D^{*-}$ with the decay rate of the $\overline{B}^0 \to D^{*-}D^{*+}$. It is therefore required that we can correctly identify the decay length as well as its flavor, which of course cannot be obtained from the final state as it is a CP eigenstate and accessible by both B mesons. In the decay chain $\Upsilon(4S) \to B^0\overline{B}^0 \to f_{CP}f_{\mathrm{tag}}$ one of the two B mesons decays at time t_{CP} to $D^{*+}D^{*-}$ and is called B^{CP}, while the other B meson, called B^{tag}, decays at time t_{tag} to a final state f_{tag}, which can be a flavor-specific decay. The wave function of the $B^0\overline{B}^0$ pair is [53],

$$|\psi\rangle = \frac{1}{\sqrt{2}}[|B^0\rangle_1 \otimes |\overline{B}^0\rangle_2 - |\overline{B}^0\rangle_1 \otimes |B^0\rangle_2], \tag{2.2}$$

such that the two B mesons keep opposite flavors at anytime until one of them decays. If the B^{tag} meson decays first to a flavor-specific decay, we can fix unambiguously the flavor of the B^{CP} meson at the time t_{tag}. This feature is specific for the B factories as the $B^0\overline{B}^0$ pair are fully flavor entangled, as they both arise from the $\Upsilon(4S)$ state. An illustration of the decay length reconstruction is shown in Figure 2.27

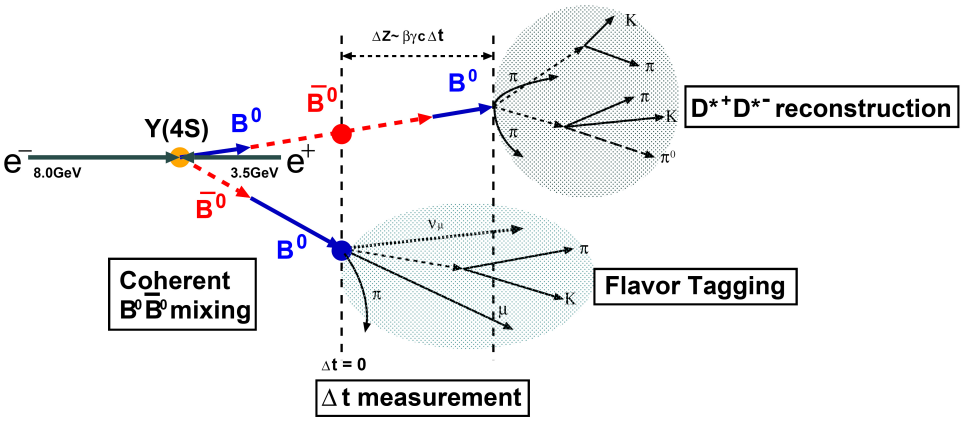


Figure 2.27: Illustration of the decay length and flavor reconstruction.

It is technically not possible to reconstruct the decay point of the $\Upsilon(4S)$. We can therefore not measure the absolute flight distance or flight time of the B^{CP}. However as the two B mesons are produced in a boosted system, we can measure the positions of the decay of both mesons and relate them to the proper-time *difference* between the B^{CP} and the $B^{\rm tag}$. And as said before we can only determine the flavor of B^{CP} at the moment $B^{\rm tag}$ decays. The CP asymmetries in Belle are therefore determined from the asymmetries in the proper-time difference between B^0 and $\overline{B}^0$ tagged events. Because the proper time has an exponential form, the proper-time difference has still the same form and we can extract the lifetime from the proper-time difference distribution as well. This principle also holds when the $B^{\rm tag}$ decays after the B^{CP}; we will then measure a negative proper-time difference.

2.7.4 The flavor tagging

The procedure of determining the flavor of the tagging B is called flavor tagging. Even for flavor specific decays i.e, decays where the flavor of the B decay can unambiguously be determined from the flavor of its decay products, there is only a fraction, ϵ which is called the tagging efficiency, which will correctly be tagged.

There are different tagging categories where the flavor of the accompanying B can be determined. We can look at the electric charge of:

- the high momentum lepton in $B^0 \to X l^+ \nu$ decays;
- kaons, as most of them originate from $B^0 \to K^+ X$ through the cascade $\overline{b} \to \overline{c} \to \overline{s}$;
- intermediate momentum leptons from $\overline{b} \to \overline{c} \to \overline{s} l^- \overline{\nu}$;
- slow pions from $B^0 \to D^{*-} X, D^{*-} \to \overline{D}^0 \pi^-$;
- high-momentum pions from $B^0 \to D^{(*)} \pi^+ X$;

or at the strangeness of Λ baryons from the cascade decays $\overline{b} \to \overline{c} \to \overline{s}$.

Sometimes no flavor can be obtained due to inefficiency in the particle detection and identification or when flavor-nonspecific decay processes took place (eg. $\overline{D}^0 \pi^0, \overline{D}^0 \to K^0 \pi^0$) or for processes which have very little information on the b flavor, for example, $b \to c\overline{u}d, \overline{K}^0 X$, for which the charged particles in the final state are all pions.

Incorrect assignments are caused by particle misidentification or from smaller physical processes that give a flavor estimate opposite to the dominate process, such as the charged kaon from $\overline{c}$ decay in $b \to c\overline{c}s$ processes or from $b \to c \to l$ leptons.

The fraction of B mesons which are incorrectly tagged is called the wrong tag fraction, w. The CP asymmetry measurement is diluted by a tagging dilution-factor $(1 - 2w)$. The effective flavor tagging efficiency on the direct CP-violating parameter is $\epsilon_{\rm eff} = \epsilon(1-2w)^2$. The average tagging efficiency at Belle is $\epsilon_{\rm eff} = 28.8 \pm 0.6\%$. Due to the dependency of the tagging quality on the decay of the $B^{\rm tag}$, events receive a quality flag, r ranging from zero, meaning no flavor information: ($w \cong 0.5$) to unity for unambiguous flavor assignment ($w \cong 0$). The tag flavor and the wrong tag fractions for each of the seven r-bins are computed using multi-dimensional likelihoods built from MC events [54, 55]. There exist different sets of parameters for data and MC and for SVD1 and SVD2 settings. Table 2.2 shows the parameters for the SVD2 data.

r-bin	range in r	w	Δw
1	$0-0.1$	0.5	0.0
2	$0.1-0.25$	0.429 ± 0.005	-0.039 ± 0.006
3	$0.25-0.5$	0.327 ± 0.006	-0.036 ± 0.006
4	$0.5-0.625$	$0.223^{+0.011}_{-0.006}$	0.018 ± 0.007
5	$0.625-0.75$	$0.161^{+0.010}_{-0.006}$	0.002 ± 0.006
6	$0.75-0.875$	0.105 ± 0.007	-0.027 ± 0.006
7	$0.875-1$	0.019 ± 0.005	0.001 ± 0.004

Table 2.2: Range of the different r bins with the corresponding wrong tag fraction parameters. The values are taken for the SVD2 data events.

The PDF of the proper-time difference for events where the B^{CP} is tagged as a B^0 or $\overline{B}^0$ now becomes diluted due to the wrong-tag fraction:

$$\begin{aligned}\mathcal{P}_{\rm sig}(B^{\rm tag}_{\rm meas}=B^0) &= (1-w_{B^0})\mathcal{P}_{\rm sig}(B^{\rm tag}_{\rm true}=B^0)+w_{\overline{B}^0}\mathcal{P}_{\rm sig}(B^{\rm tag}_{\rm true}=\overline{B}^0),\\ \mathcal{P}_{\rm sig}(B^{\rm tag}_{\rm meas}=\overline{B}^0) &= w_{B^0}\mathcal{P}_{\rm sig}(B^{\rm tag}_{\rm true}=B^0)+(1-w_{\overline{B}^0})\mathcal{P}_{\rm sig}(B^{\rm tag}_{\rm true}=\overline{B}^0). \qquad (2.3)\end{aligned}$$

where "true" means the actual flavor and "meas" means the flavor determined by the tagging algorithm. We then can define $w=\frac{w_{B^0}+w_{\overline{B}^0}}{2}$ and $\Delta w=(w_{B^0}-w_{\overline{B}^0})$. The proper time equation of (1.95) is now modified to incorporate the wrong-tag fraction and uses the proper-time difference between the B^0 and $\overline{B}^0$ tagged events instead of the absolute proper time:

$$\mathcal{P}_{\rm sig}(q,\Delta t)=\frac{e^{-|\Delta t|/\tau}}{4\tau}\Big(1-q\Delta w+q(1-2w)(\mathcal{S}\sin(\Delta m\Delta t)+\mathcal{A}\cos(\Delta m\Delta t))\Big) \qquad (2.4)$$

where $q=+1$ when the tagging B meson is identified as a B^0 and $q=-1$ when $B^{\rm tag}=\overline{B}^0$. This makes $\mathcal{P}_{\rm sig}(B^{\rm tag}_{\rm meas}=B^0)=\mathcal{P}_{\rm sig}(q=+1,\Delta t)$ and $\mathcal{P}_{\rm sig}(B^{\rm tag}_{\rm meas}=\overline{B}^0)=\mathcal{P}_{\rm sig}(q=-1,\Delta t)$

2.7.5 The detector resolution

A precise determination of Δt, the time between the decay of the two B mesons is yet another essential part in the CP analysis. As said before the difference in decay time of the two B mesons at Belle is measured by their displaced vertex in the z direction. To extract the true Δt from the measured Δz it is not enough to just apply the boost factor. The measurement of Δz is deteriorated by the detector resolution. A clear understanding of the detector response is indispensable as the detector resolution is of the same order as the average Δz at KEKB.

The detector resolution-function is determined with a large sample of decays such as, $\overline{B}^0\to D^+\pi^-, D^{*+}\pi^-, D^{*+}\rho^-, J/\psi K^0_S, J/\psi\overline{K}^{*0}, B^-\to D^0\pi^-$ and $J/\psi K^-$ where $J/\psi\to l^+l^-(l=e,\mu)$ and $D^{*+}\to D^0\pi^+$. More information on the selection and reconstruction requirements can be found in [56].

The resolution function $\mathcal{R}$ smears the true Δt distribution as follows:

$$\mathcal{P}(\Delta t)=\int_{-\infty}^{+\infty}\mathcal{P}(\Delta t')\mathcal{R}(\Delta t-\Delta t')d(\Delta t'). \qquad (2.5)$$

The resolution function $\mathcal{R}$ is a convolution of four different contributions:

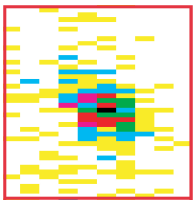
- the detector resolution $R_{\rm CP}$ of the CP-side B-vertex,
- the detector resolution $R_{\rm tag}$ of the tag-side B-vertex,
- a smearing due to the inclusion of tracks which do not originate from the B decay, $\mathcal{R}_{\rm NP}$ is mostly caused by charm and K_S decays,
- the kinematic approximation that the B mesons are at rest in the cms: $R_{\rm kin}$.

There is also a small number of events which have a very broad Δt distribution. These outlier events are described by a wide Gaussian function $\mathcal{P}_{\rm ol}(\Delta t)$.

A detailed description of the determination of each function will be further explained in Chapter 5.2. The general function used to describe the detector response might not always describe the resolution for the decay analyzed in this work. We will have to determine ourselves some parameters of the resolution function.

Chapter 3

Signal reconstruction and background study



In this Chapter we describe how the $B^0 \to D^{*+}D^{*-}$ events are reconstructed from the Belle data sample. We present the selection procedure, reconstruction performance and background study. Finally the signal yield on the full Belle data sample is extracted.

THE $B^0 \to D^{*+}D^{*-}$ decays are filtered out of the large DST files collected by the Belle detector using a set of selection criteria. Signal Monte Carlo (MC) events are used to study the selection cuts needed to enhance the purity of $B^0 \to D^{*+}D^{*-}$ events. While studying the selection criteria we only base ourselves on the MC samples. This so-called blind analysis method protects against biased results when obtaining selection cuts from a small data sample which is subject to large fluctuations. The goal of this chapter is to apply selection criteria to reject background and to obtain a good reconstruction efficiency such that a large number of signal events can be reconstructed on the data. Furthermore generic MC samples and off-resonance data sets are used to study the background behavior and to optimize the modeling of the signal events in the data. As mentioned in the introduction, obtaining a precise measurement of the branching fraction is not an objective of this analysis. Already in the previous analysis on $140\,\mathrm{fb}^{-1}$ the uncertainty of the result of the branching fraction was dominated not by statistics but by the systematic uncertainty which is mainly due to the tracking efficiency. Therefore even with more statistics, we will not enhance the total uncertainty on the branching fraction significantly. Therefore there is also no systematic study performed in this chapter.

3.1 Decay channels

In this analysis not all the $D^{*+}D^{*-}$ subdecays are reconstructed. Decays that have a very small contribution to the total yield are ignored. Also decays which would imply lot of combinatorial background are rejected. We have used the decay channels which were taken in the previous analysis. Only the $D^+ \to K_S K^+$ decay is discarded because of its

very low signal-to-background ratio.

We reconstruct $D^{*+}D^{*-}$ decays in the following combinations:

$$D^{*+}D^{*-} \to (D^0\pi^+)(\overline{D}^0\pi^-)$$
$$D^{*+}D^{*-} \to (D^0\pi^+)(D^-\pi^0)$$
$$D^{*+}D^{*-} \to (D^+\pi^0)(\overline{D}^0\pi^-).$$

The pions from the D^* decays are referred to as slow pions because of their low momentum (< 200 MeV in the CM of the $B^0\overline{B^0}$ system). Table 3.1 shows the D^0 and D^+ reconstructed decays as well as their branching fractions. The K_S decays are reconstructed in the $\pi^+\pi^-$ mode. Unless specified otherwise charged-conjugated decays are implied throughout.

	D^0 decay modes	$\mathcal{B}\,(\%)$		D^+ decay modes	$\mathcal{B}\,(\%)$
1	$D^0 \to K^-\pi^+$	3.8 ± 0.1	1	$D^+ \to K^-\pi^+\pi^+$	9.5 ± 0.3
2	$D^0 \to K^-\pi^+\pi^0$	14.1 ± 0.5	2	$D^+ \to K_S\pi^+$	1.5 ± 0.1
3	$D^0 \to K^-\pi^+\pi^+\pi^-$	7.7 ± 0.3	3	$D^+ \to K_S\pi^+\pi^0$	7.0 ± 0.5
4	$D^0 \to K_S\pi^+\pi^-$	2.9 ± 0.2	4	$D^+ \to K^+K^-\pi^+$	1.00 ± 0.04
5	$D^0 \to K_S\pi^+\pi^-\pi^0$	5.3 ± 0.6			
6	$D^0 \to K^+K^-$	0.38 ± 0.01			

Table 3.1: Reconstructed D^0 and D^+ decays with their branching fractions ($\mathcal{B}$).

Figure 3.1 shows the event-display of a $B^0 \to D^{*+}D^{*-}$ candidate as recorded in the detector. This particular example will be reconstructed through the $D^{*+} \to D^0\pi^+$ where $D^0 \to K^-\pi^+\pi^0$ and $D^{*-} \to \overline{D^0}\pi^-$ where $\overline{D^0} \to K^+\pi^-$. The figure is taken from [57].

3.2 Monte Carlo and data samples

The reconstructed data events which passed the selection criteria can be divided into four groups.

- Reconstructed signal events: these are $B^0 \to D^{*+}D^{*-}$ events which are reconstructed through the chosen subdecays.

- Combinatorial background events: these events come merely from $B\overline{B}$ events, other than $B^0 \to D^{*+}D^{*-}$ which passed through the selection criteria. There is also a small fraction of wrongly reconstructed $B^0 \to D^{*+}D^{*-}$ signal events, which can contribute to the combinatorial background. For example when a pion is identified as a kaon, the B^0 mass distribution will not peak around the B-mass. Depending on what we want to measure these events can still be useful (see Section 3.4).

- Peaking background events: these are background events which have a similar signal signature to the $B^0 \to D^{*+}D^{*-}$ when they are wrongly reconstructed and are thus a sub-category of the combinatorial background events. However these events show up with a signal-like distribution. This contribution will be studied in Section 3.6.

- Continuum background events: these are events coming from the $q\overline{q}$ ($q = u, c, d$ or s) continuum background under the $\Upsilon(4S)$ resonance peak (see Section 2.1). They contain no useful information for this analysis and the selection criteria should discard these events as much as possible.

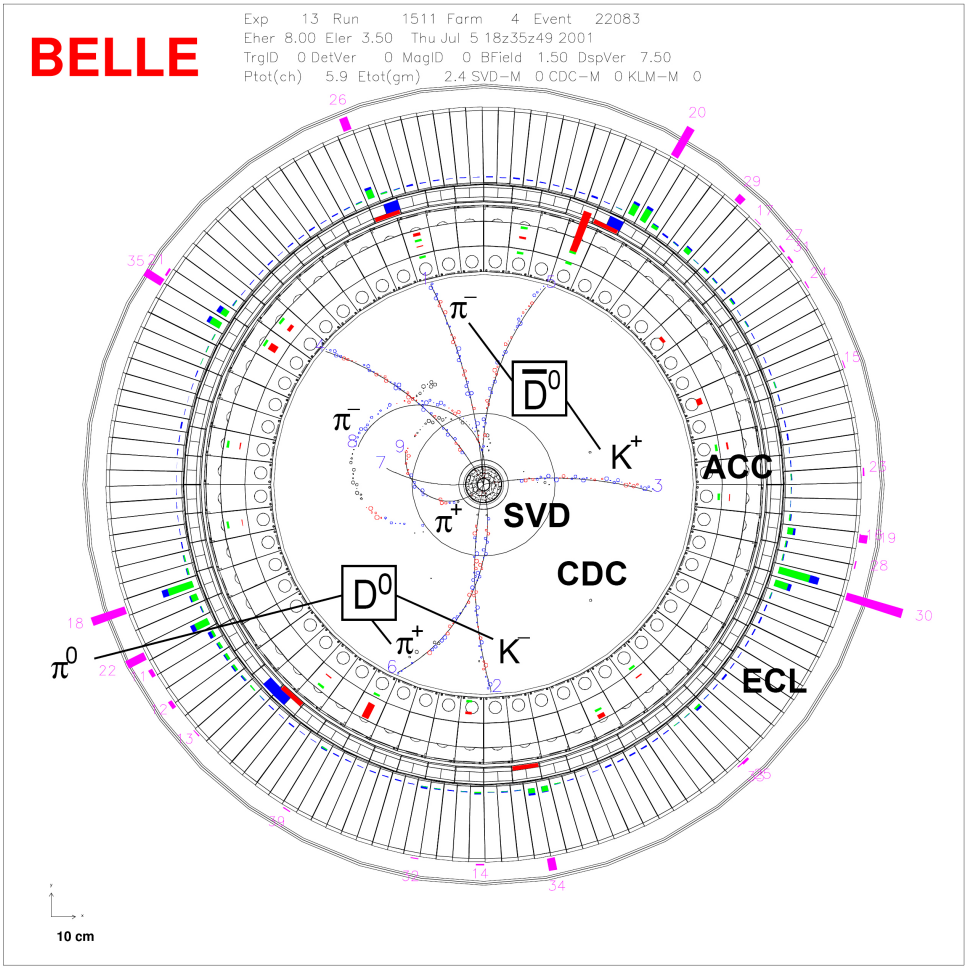


Figure 3.1: Event display of a $B^0 \to D^{*+}D^{*-}$ candidate as recorded in the detector. The curves with number symbolize the reconstructed tracks.

The selection criteria are chosen such that the purity $\mathcal{S}/(\mathcal{S}+\mathcal{B})$ is maximized. We use different Monte Carlo samples to study the effect of the selection criteria on the signal and background events.

In this study we use in total four types of data or MC samples.

- Signal MC: these are events generated by ourselves using EvtGen [46]. The samples contain only $B^0 \to D^{*+}D^{*-}$ signal events. The subdecays are also limited to the ones we reconstruct.

- Generic on-resonance MC: there are two generic on-resonance samples: one contains all the $\Upsilon(4S) \to B^0\overline{B^0}$ decay modes (also referred to as the mixed generic sample) and one which contains all the $\Upsilon(4S) \to B^+B^-$ modes (or charged generic sample). This sample is ideal to study background contributions.

- On-resonance data: this is the data set used for this analysis. It contains $605\,\mathrm{fb}^{-1}$ taken at the $\Upsilon(4S)$ resonance.

- Off-resonance data: as mentioned in Section 2.2 around 10% of data are taken below the $\Upsilon(4S)$ resonance and contain thus exclusively continuum background.

In the next sections we will describe each of these samples in more detail.

3.2.1 Signal MC

Three samples of 530k MC events are generated, each with a different polarization. The generated $\Upsilon(4S)$-meson decays exclusively in a $B^0\overline{B^0}$ pair. The neutral B-mesons are allowed to oscillate in time. The generated B^0 and $\overline{B^0}$ mesons are then each required to decay to a $D^{*+}D^{*-}$ state with 50% probability and with 50% probability they decay inclusively. Events where both B^0 and $\overline{B^0}$ or neither B-meson decays to a $D^{*+}D^{*-}$ state are rejected.

The generated events are then simulated through the detector as if they were collected over time with the corresponding detector and accelerator settings. The simulation was done using the GSIM package which calls the GEANT3 module.

3.2.2 Generic MC

There are two generic MC samples, one contains the generated $\Upsilon(4S) \to B^+B^-$ decays while the other contains $\Upsilon(4S) \to B^0\overline{B^0}$ decays where the $B^0\overline{B^0}$ pair undergoes coherent mixing. The generic $B^0\overline{B^0}$ sample contains a number of $B^0\overline{B^0}$ events equivalent to $1103\,\mathrm{fb}^{-1}$ and for the B^+B^- sample it is equivalent to $1185\,\mathrm{fb}^{-1}$. The generic MC files also contain a number of flags which make it possible to still retrieve information on the generation, such as which subdecay was generated.

3.2.3 On-resonance data

These are the data collected at the $\Upsilon(4S)$ resonance. This analysis is performed on $605\,\mathrm{fb}^{-1}$ which corresponds to $(657 \pm 9) \times 10^6$ $B\overline{B}$ events. The data is grouped in so-called "experiments". Table 3.2 summarizes the collected on- and off-resonance luminosity per experiment. The even experiment numbers are used for calibrations. The experiments before 7 only correspond to less than $1\,\mathrm{fb}^{-1}$ in total but is of bad quality, which is the reason they are not used for analysis. There is no experiment 29, to accentuate the transition from the

SVD1 to SVD2 data. Experiment 53 is also taken out of the table as it was performed on the $\Upsilon(5S)$ resonance.

SVD 1 data			SVD 2 data		
Exp	On-reson. (fb^{-1})	Off-reson. (fb^{-1})	Exp	On-reson. (fb^{-1})	Off-reson. (fb^{-1})
7	5.9	0.6	31	17.8	2.4
9	4.4	0	33	17.6	2.7
11	8.1	1.2	35	16.7	1.9
13	10.7	1.2	37	61.7	6.1
15	12.7	1.4	39	43.6	6.3
17	11.2	0.9	41	59.9	5.7
19	25.0	3.6	43	57.0	6.5
21	4.4	0.	45	13.0	2.3
23	6.3	1.4	47	37.6	3.4
25	25.8	1.7	49	27.3	2.6
27	25.4	3.7	51	38.9	4.8
			55	73.5	7.8
Total	139.9	15.7		464.7	52.6

Table 3.2: Collected on- and off-resonance luminosity per experiment.

3.2.4 Off-resonance data

The off-resonance samples contain $68\,\mathrm{fb}^{-1}$ of real data, collected $60\,\mathrm{MeV}$ below the $\Upsilon(4S)$ resonance. As $q\overline{q}$ ($q = u, c, d$ or s) events are produced at any CM-energy in e^+e^- collisions, they are also present in the on-resonance data. The off-resonance data contains only continuum events and offers therefore the opportunity to study exclusively the continuum background under the $\Upsilon(4S)$ resonance.

3.3 Selection criteria

The selection criteria are applied to the data and MC samples to enhance the $B^0 \to D^{*+}D^{*-}$ signal with respect to the other channels or continuum background. These criteria are first applied to the final-state daughter-particles which are in our case kaons and pions. The four-momentum of the candidate kaons and pions are then combined to form possible D mesons. More quality and selection requirements are applied to the D candidates before they are combined with the slow pion information to form D^* particles and finally B mesons. Furthermore filters are applied based on the event shape variables to separate $B\overline{B}$ events from continuum background events. Most selection criteria applied in this analysis are standard cuts, which are used in similar analyses [58, 59]. The following sections will discuss in more detail the chosen filters.

3.3.1 Track and event selection

- Tracks of charged particles are reconstructed requiring the transverse distance dr between the track and the interaction point (IP) to be less than $2.0\,\mathrm{cm}$ while the

absolute value of the longitudinal distance, (dz) has to be smaller than $4.0\,\mathrm{cm}$. The interaction point is the reconstructed three-dimensional Gaussian profile where the interaction between the electron and positron took place. A more detailed description of the interaction point is given in Section 5.1.

- In the center-of-mass system (CM) of the electron positron system, the $q\overline{q}$ events tend to consist of two back-to-back jets while the final particles emerging from a $B\overline{B}$ event are more isotropically distributed as both B mesons are almost produced at rest in the CM system. Fox-Wolfram moments [60] are used to characterize the event shape. We require the ratio of the second to the zeroth order Fox-Wolfram moments to be less than 0.4. The continuum background in this analysis is not the main source of background, as will be shown in Section 3.6.2.1. The cut is chosen such to keep most of the signal while rejecting some continuum background.

3.3.2 Particle identification for final state particles

- Charged kaons are separated from charged pions using a likelihood ratio constructed from ACC information, CDC dE/dx and TOF measurements. Charged 2-prong (3- or 4-prong) tracks are reconstructed as kaons when $\mathcal{L}_{K/\pi} = \mathcal{L}(K^{\pm})/(\mathcal{L}(K^{\pm}) + \mathcal{L}(\pi^{\pm})) > 0.1(0.6)$ and as pions when $\mathcal{L}_{K/\pi} < 0.9$.

- No particle identification is required for the slow pions.

- Neutral pions are reconstructed from two photons with energies above $30\,\mathrm{MeV}$ and with a total CM momentum $p_{\gamma\gamma}$ larger than $0.1\,\mathrm{GeV/c}$. The reconstructed invariant mass is required to be in the range $119\ \mathrm{MeV/c^2} < M_{\gamma\gamma} < 146\ \mathrm{MeV/c^2}$ as the mass of the neutral pion is $M_{\pi^0} = 135\,\mathrm{MeV/c^2}$.

- Neutral kaons are reconstructed via the decay $K^0_S \to \pi^+\pi^-$. The $\pi^+\pi^-$ invariant mass is required to be within $\pm 9\,\mathrm{MeV/c^2} (\approx \pm 3\sigma)$ of the K^0 mass. Cuts are applied to the displacement of the $\pi^+\pi^-$ vertex to the IP [61], this standard Belle algorithm for K^0_S reconstruction is referred to as the "GoodKs"-algorithm.

3.3.3 *D*-meson selection

Different D-meson subdecays have different invariant mass resolutions depending on the number and charges of the decay daughters. Neutral pions and kaons have a worse resolution as they leave no signal behind in the accurate tracking devices of the detector.

Figures 3.2 and 3.3 show the reconstructed D-mass distributions for each decay obtained on the signal MC sample. The events in the plots are required to be linked to the MC truth, meaning that the B meson and all its daughter particles are correctly reconstructed.

The signal events in the distributions are modeled with a double Gaussian. Table 3.3 shows the widths and fraction of both Gaussians for each decay mode.

The invariant mass of the D candidates is required to be within $\pm 6\sigma (\pm 3\sigma)$ of the nominal value for 2-prong (3- or 4-prong) decays, where σ is the width of the narrowest Gaussian of the channel dependent D mass resolution obtained from signal MC samples. This width ranges from $3.61\,\mathrm{MeV/c^2}$ to $11.0\,\mathrm{MeV/c^2}$.

In Figure 3.4 the mass difference between the D^{*+} and D^0 (top) or the D^{*+} and D^+ (bottom) is shown. These distributions contain only correct combinations in MC signal samples and are fitted with a double Gaussian function. The results are displayed

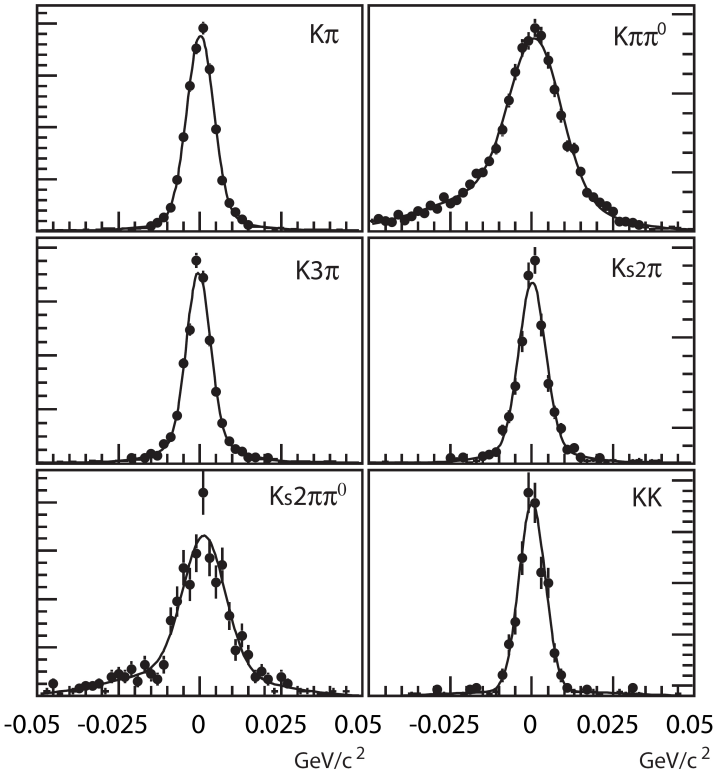


Figure 3.2: Fitted distributions of the reconstructed D^0 mass minus $1.8645\,\mathrm{GeV}/c^2$, for correct combinations in the MC signal samples.

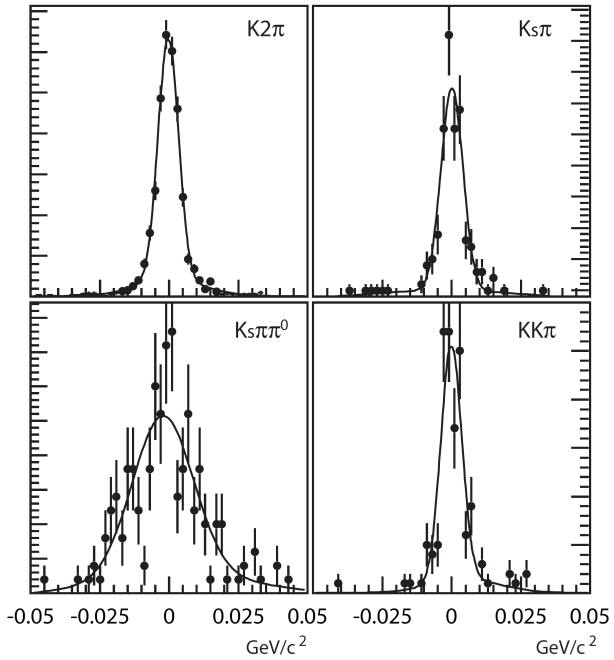


Figure 3.3: Fitted distributions of the reconstructed D^+ mass minus $1.8694\,\mathrm{GeV}/c^2$, for correct combinations in the MC signal samples.

decay mode	σ_1 (MeV/c^2)	f	σ_2 (MeV/c^2)
$D^0 \to K^-\pi^+$	4.1 ± 0.1	0.84 ± 0.02	12.4 ± 1.4
$D^0 \to K^-\pi^+\pi^0$	7.9 ± 0.3	0.57 ± 0.02	21.8 ± 0.8
$D^0 \to K\pi^+\pi^+\pi^-$	3.61 ± 0.1	0.74 ± 0.03	10.7 ± 0.7
$D^0 \to K_S\pi^+\pi^-$	4.1 ± 0.3	0.81 ± 0.06	12.0 ± 0.3
$D^0 \to K_S\pi^+\pi^-\pi^0$	6.4 ± 0.4	0.60 ± 0.05	21.6 ± 1.7
$D^0 \to K^+K^-$	4.0 ± 0.2	0.77 ± 0.03	31.9 ± 3.1
$D^+ \to K^+\pi^+\pi^-$	3.7 ± 0.2	0.82 ± 0.03	12.3 ± 1.8
$D^+ \to K_S\pi^+$	3.9 ± 0.3	0.64 ± 0.07	35.2 ± 7.2
$D^+ \to K_S\pi^+\pi^0$	11.0 ± 0.9	0.72 ± 0.06	32.1 ± 5.6
$D^+ \to K^+K^-\pi^+$	3.74 ± 0.4	0.84 ± 0.08	12.6 ± 2.8

Table 3.3: Width of the narrow (σ_1) and wide (σ_2) Gaussians and the fraction (f) of the narrow Gaussian of the fitted D^0 or D^+ mass distributions of signal MC truth.

in Table 3.4. The mass difference, $\Delta M = |M(D^*) - M(D)|$ is required to be within $\pm 3\,(2.25)$ MeV/c^2 of the nominal value for the D^0 (D^+) channel, which corresponds to a mass range of $\sim 3.5(2.0)\sigma$ where σ the width is of the larger Gaussian. The cuts are chosen such to remove wrongly reconstructed events as well as background events and is tighter for the $D^{*+} - D^+$ channel to discard the relatively larger background contributions from the slow neutral pions.

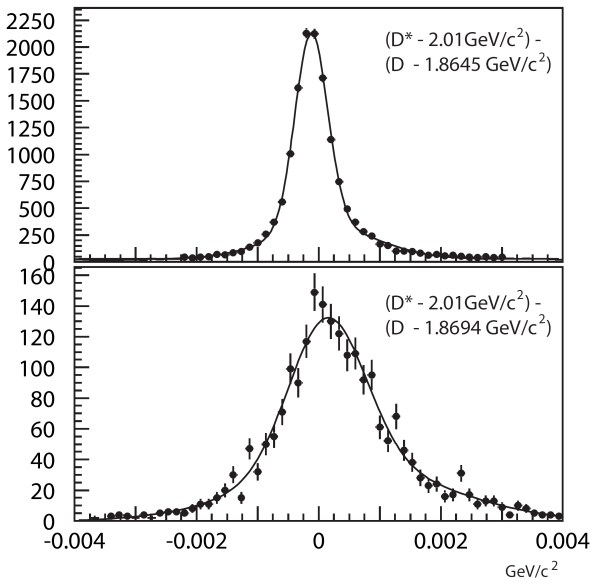


Figure 3.4: Fitted distributions of the reconstructed $D^* - D$ mass difference (minus its expected value) for correct combinations in MC signal samples. Top: $D^{*+} \to D^0\pi^+$. Bottom: $D^{*+} \to D^+\pi^0$.

D^{*+} decay mode	σ_1 (MeV/c^2)	f	σ_2 (MeV/c^2)
$D^{*+} \to D^0\pi^+$	0.266 ± 0.005	0.65 ± 0.01	0.849 ± 0.030
$D^{*+} \to D^+\pi^0$	0.625 ± 0.040	0.53 ± 0.05	1.595 ± 0.094

Table 3.4: Width of the narrow (σ_1) and wide (σ_2) Gaussians and the fraction (f) of the narrow Gaussian of the fitted D^* - D mass difference distributions of correct combinations in MC signal samples.

3.3.4 Kinematics fits

Vertex-constrained fits increase the quality of the tracks. Indeed requiring daughter particles to originate from the same point in space puts additional constraints on the tracks. The track's position and the corresponding momentum and mass is recalculated using this extra piece of information. Furthermore a mass-constrained fit can be applied which requires the daughter particles' momenta to be such that the invariant mass of the mother particle corresponds to its nominal value. The kinematic fits are only performed after having put constraints on the D mass region, as explained in the previous paragraph.

- The two photons which form a neutral pion candidate are required to emerge from the same point, assumed to be the IP. The photon directions are obtained from the ECL clusters and the IP.

- The charged pions forming a K^0_S candidate are subjected to a mass fit and a vertex fit.

- The daughter particles of the D meson are constrained with a vertex fit and by a mass-constrained fit.

- The slow pion track has a much worse resolution due to the multiple scattering effect. To correct this we use the fact that its track has to originate from the D^* decay vertex. The latter is obtained from the intersection of the IP and the D track's extrapolation to the IP. After that the slow pion is refitted to this point and the D^* mass is recalculated using a new slow pion momentum.

Events for which the (mass-)vertex fit didn't converge are discarded. The four-momenta of the D^* and B mesons are recalculated using the refitted daughter particles' four-momenta.

3.3.5 B-meson selection

Because of energy conservation the total energy of the two beams has to be equal to the total energy of the $B\overline{B}$ pair. We can therefore reconstruct the B-meson's mass by combining the beam energy in the CM system $E^{\rm CM}_{\rm beam}$, with the momentum of the B-meson: $M_{\rm bc} = \sqrt{(E^{\rm CM}_{\rm beam})^2 - (p^{\rm CM}_B)^2}$, which is called the beam-constrained mass. Another independent variable which is used to identify the B meson is the energy difference $\Delta E \equiv E^{\rm CM}_B - E^{\rm CM}_{\rm beam}$, where $E^{\rm CM}_B$ and $p^{\rm CM}_B$ are the CM energy and momentum of the B candidate. The momentum of the B meson is computed from the measured momenta of the daughter particles. The B^0 energy however is obtained from the momenta and the masses of the daughter particles. The masses are not directly measured, but assumed based on

the particle identification. The cut on the beam-constrained mass and energy-difference is:

$$\begin{aligned} 5.2\,\mathrm{GeV/c^2} &< M_{\mathrm{bc}} < 5.3\,\mathrm{GeV/c^2}, \\ -0.2\,\mathrm{GeV} &< \Delta E < 0.2\,\mathrm{GeV}. \end{aligned} \tag{3.1}$$

Furthermore we define two regions where we look for signal events:

1. The small signal region: $5.27\,\mathrm{GeV/c^2} < \mathrm{M_{bc}} < 5.3\,\mathrm{GeV/c^2}$ and $-0.04\,\mathrm{GeV} < \Delta\mathrm{E} < 0.04\,\mathrm{GeV}$
2. The large signal region: $5.23\,\mathrm{GeV/c^2} < \mathrm{M_{bc}} < 5.3\,\mathrm{GeV/c^2}$ and $-0.14\,\mathrm{GeV} < \Delta\mathrm{E} < 0.14\,\mathrm{GeV}$.

In Section 3.6 we will show that the definition of the large signal region is such that the tails of the contamination of mis-identification of hadrons are not contained in the large signal region.

3.3.6 Best candidate

We reconstruct 2290612 B^0 candidates in the MC signal sample, of which 46094 events have all the tracks and daughter tracks correctly identified and reconstructed (we call this MC truth). After applying these selection cuts, 55758 B^0 candidates are left of which 36183 are in the small signal region. 66.5% (or 43.4% in the small signal region) are true $B^0 \to D^{*+}D^{*-}$ decays. As we know that there is just one $B^0 \to D^{*+}D^{*-}$ decay per event, we accept only the B candidate with the smallest value of the quantity

$$\chi^2_{\mathrm{mass(D,D^*)}} = \sum_{i=1}^{2} \left(\frac{\Delta M_i - \Delta M^{\mathrm{PDG}}}{\sigma_{\Delta M_i}} \right)^2 + \sum_{i=1}^{2} \left(\frac{M_{D_i} - M_{D_i}^{\mathrm{PDG}}}{\sigma_{M_{D_i}}} \right)^2, \tag{3.2}$$

where $\sigma_{M_{D_i}}$ is the width of the narrow Gaussian of the channel-by-channel D mass distribution and $\sigma_{\Delta M_i}$ is the width of the wide Gaussian of the $D^* - D$ mass distribution. The PDG superscript refers to the nominal value of the particle according to the Particle Data Group world average [6].

After applying the best-candidate selection 27610 (35423) reconstructed B^0 events are left, of which 77.5% (60.8%) are true generated B^0 decays in the small (large) signal region. The other events are combinatorial background. In Section 3.4 it will be shown that events which cannot be associated with MC truth could still have to be considered as correct signal events, for example when they contain two correctly reconstructed D mesons. There are 24726 ± 157 such events in the signal MC sample, which make the final purity in the signal MC of 89.5%.

3.3.7 Reconstruction efficiency

The reconstruction efficiency (ϵ_{rec}) is a measure of the performance of the selection and reconstruction procedure. This information is crucial when determining the branching fraction from the yield measurement, which is not done in this analysis. However the relative differences in the reconstruction efficiency between data with different polarizations need to be known for the measurement of the CP-odd fraction, which is presented in the next chapter. The reconstruction efficiency for the generated decay channels can be

obtained from the signal MC as we know exactly how many events were generated. The reconstruction efficiency of a decay is defined as:

$$\epsilon_{\rm rec} = \frac{\text{number of reconstructed events}}{\text{number of generated events}}.$$

The reconstruction efficiency may depend on the subdecay, i.e. on the decay modes of the two D mesons. Therefore the efficiency defined above is merely an average over all subdecays, i, weighted by their branching fraction $\mathcal{B}_i$:

$$\epsilon_{\rm rec} = \frac{\sum_i \mathcal{B}_i \epsilon_{\rm rec,i}}{\sum_i \mathcal{B}_i} = \frac{\text{number of reconstructed events}}{\text{number of generated events}},$$

The number of reconstructed events is extracted from a two-dimensional maximum likelihood fit of the beam constrained mass and energy-difference distributions. Figure 3.5 shows the distribution of $M_{\rm bc}$ and ΔE for the signal MC sample. The red box shows the small signal region.

The $M_{\rm bc}$ distribution is described with a Gaussian function for the signal events and an ARGUS, threshold function [62] for the combinatorial background. We use a double Gaussian probability density function for the signal events in the ΔE distribution and a linear function to describe the background events. The two-dimensional probability density functions for signal events become:

$$\mathcal{P}(M_{\rm bc}, \Delta E) = \mathcal{G}(M_{\rm bc}, \mu, \sigma) \times \Big(f\mathcal{G}(\Delta E, \mu_1, \sigma_1) + (1-f)\mathcal{G}(\Delta E, \mu_2, \sigma_2) \Big), \qquad (3.3)$$

where $\mathcal{G}$ is a Gaussian function and f is the fraction of the narrow Gaussian function. The background PDF is:

$$\mathcal{P}(M_{\rm bc}, \Delta E) = ARGUS(M_{\rm bc}, s) \times \mathcal{P}_1(\Delta E, a), \qquad (3.4)$$

where the slope of the threshold function is s and $\mathcal{P}_1$ is a first order polynomial with a slope a.

The fit for signal MC events is performed on all the B^0 candidates, but for visualization purposes the result of the fit will always be shown projected on one of the two variables. The other variable will be required to be in the small signal region, as shown in Figure 3.6. The fitted number of signal events and background events (in the large signal region) are summarized in Table 3.5 and 3.6. The purity is defined as $\frac{\mathcal{S}}{\mathcal{S}+\mathcal{B}}$ in the small signal region.

events	quantity
signal ($\mathcal{S}$)	28038 ± 190
background ($\mathcal{B}$)	8019 ± 101
purity	94.5%

Table 3.5: Fitted signal and background yields in the signal MC sample.

We calculated the reconstruction efficiency for each polarization separately. The result is shown in Table 3.7. A_0 and $A_\parallel$ are the transversity amplitudes corresponding to the CP-even state. $A_\perp$ corresponds to the $l = 1$ angular momentum state, which is the CP-odd state. These polarization-dependent reconstruction efficiencies will be used in the angular analysis (see Section 4.1.2).

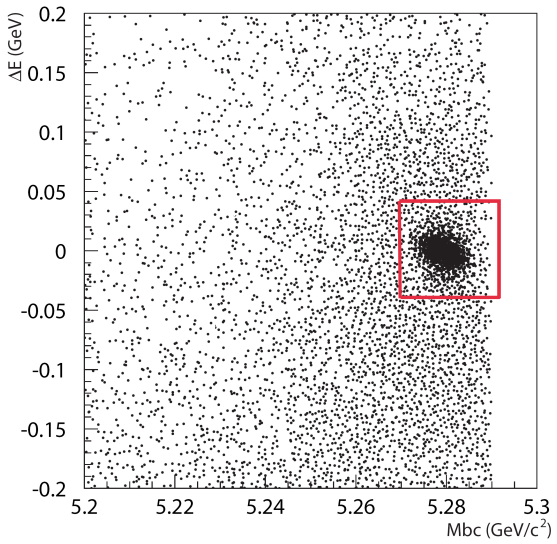


Figure 3.5: $M_{\rm bc} - \Delta E$ distribution of signal MC.

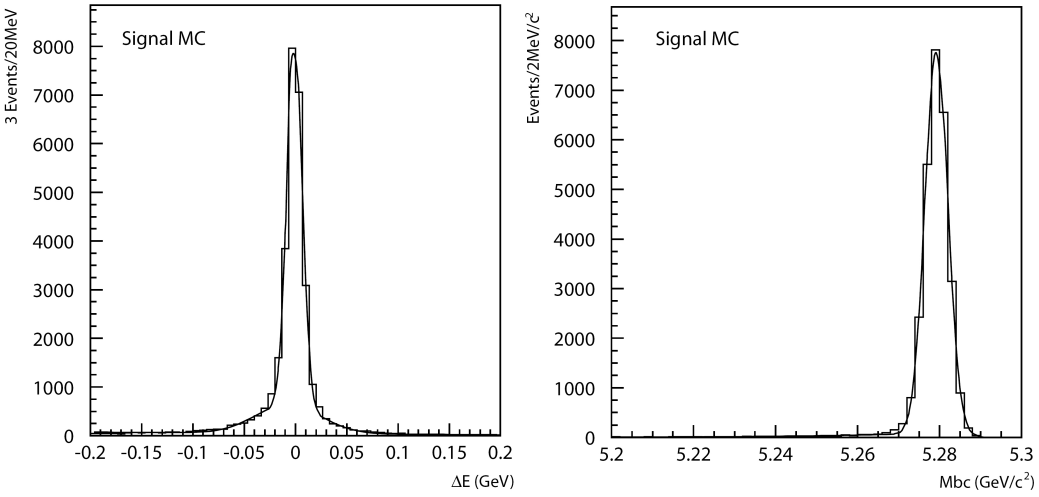


Figure 3.6: Projections of fitted ΔE and $M_{\rm bc}$ distribution of signal MC.

	ΔE			$M_{\rm bc}$	
μ_1	-0.78 ± 0.06	MeV	μ	5.2792	$\mathrm{GeV/c^2}$
σ_1	7.02 ± 0.07	MeV	σ	2.81 ± 0.01	$\mathrm{MeV/c^2}$
f	0.71 ± 0.01				
μ_2	-6.25 ± 0.49	MeV			
σ_2	28.18 ± 0.46	MeV			

Table 3.6: Fitted signal parameters of the 2D fit of ΔE and $M_{\rm bc}$ on signal MC.

	A_0	$A_\perp$	$A_\parallel$
generated events	526712	528306	513769
reconstructed events	9103 ± 99.7	9707 ± 94.8	9192 ± 99.1
ϵ (%)	1.73 ± 0.02	1.84 ± 0.02	1.79 ± 0.02

Table 3.7: Reconstruction efficiency for each polarization extracted from the signal MC.

3.4 Study of combinatorial background in signal MC

Some combinatorial background events in the signal MC can still have a signal-like shape in the CP fit. Combinatorial background events arise from wrongly identified particles or mistakes in the track reconstruction. In the CP analysis a fit is performed to the difference in vertex positions of the B meson which decays to the $D^{*+}D^{*-}$ mode and the other B meson in the event. The CP-side vertex is defined as the intersection of the beam profile and the tracks from the two D mesons. Therefore a signal event in which the two D-mesons are correctly reconstructed, but in which a mistake happens with the reconstruction of the slow pion, will still behave like a signal event in the distribution of the difference in vertex position. We want to check if these events show up as signal or as background events in the $M_{\rm bc}$ and ΔE distribution.

To investigate this we studied the ΔE, $M_{\rm bc}$ and the proper-time resolution shapes for events with badly reconstructed slow pions but still correctly reconstructed D mesons. Events where both D mesons are correctly reconstructed should be considered as signal events in the CP fit. Signal events with only one D correctly reconstructed will probably still have a signal behavior in Δt, although the vertex resolution might be of a worse quality. Finally a check is performed to see how many signal events have both D badly reconstructed.

Figure 3.7 shows the $M_{\rm bc}$ and ΔE distribution for signal-MC events which have both D mesons correctly reconstructed. The fit shown in these figures is a projection of a two-dimensional fit and the results are summarized in the left part of Table 3.8. We also fitted the proper-time resolution; this is the distribution of the difference between the generated and reconstructed vertex position. This distribution is fitted with a double Gaussian function and we added the widths and the fraction of the narrow Gaussian to the table. From the $M_{\rm bc}$ and ΔE distributions it can be seen that all events are modeled by the signal shape and there is no combinatorial background. This means that even though a slow pion might be poorly reconstructed, these events are considered as signal as they have a signal-like behavior in $M_{\rm bc}$ and ΔE. In the CP fit these events will be treated like signal events, which is correct as they will show indeed a signal-like behavior in Δt.

The $M_{\rm bc}$ and ΔE distribution for signal-MC events which have only one D meson correctly reconstructed is shown in Figure 3.8. The narrow Gaussian distribution that describes the signal shape in ΔE has a width which is comparable to the width of the wide Gaussian in the signal MC (see Section 3.3.7). This means that most of the events in the Gaussian shape will be correctly taken as signal events. The problem is the combinatorial background events in the small signal region. These events will be considered as background events in the $M_{\rm bc}$ and ΔE fit but will show up as signal in the Δt distribution, even though the resolution is slightly deteriorated. Luckily the fraction of events behaving like this (1.27%) is small and will represent few events in the final yield.

Finally the fraction of signal-MC events with both D-mesons wrongly reconstructed

	two true D mesons	one true D mesons	two wrong D mesons
signal events	24726 ± 157	3977 ± 81	155 ± 20
background events	163 ± 15	5316 ± 88	1084 ± 36
ΔE			
$\mu_{\rm main}$ (MeV)	-0.7 ± 0.1	-4.3 ± 0.8	-26.4 ± 1.0
$\sigma_{\rm main}$ (MeV)	6.9 ± 0.1	32.2 ± 0.8	49.7 ± 7.7
frac	0.77 ± 0.01	0.92 ± 0.01	0.78
$\mu_{\rm tail}$ (MeV)	-4.4 ± 0.3	-79.0 ± 0.3	-8.0
$\sigma_{\rm tail}$ (MeV)	22.8 ± 0.3	19.3 ± 2.1	15.9 ± 5.7
bkg slope	-2.16 ± 0.77	-3.00 ± 0.11	-2.77 ± 0.23
M_{bc}			
μ (GeV/c^2)	5.2792	5.2787	5.2801
σ (MeV/c^2)	2.8 ± 0.1	4.0 ± 0.1	4.3 ± 0.4
argus slope	78.2 ± 11.0	78.2 ± 11.0	96.2 ± 4.1
Δt			
$\sigma_{\rm main}$ (ps)	0.30 ± 0.01	0.34 ± 0.01	0.38 ± 0.03
frac	0.77 ± 0.01	0.68 ± 0.02	0.54 ± 0.07
$\sigma_{\rm tail}$ (ps)	0.86 ± 0.02	1.08 ± 0.04	1.11 ± 0.11

Table 3.8: Parameters of ΔE, $M_{\rm bc}$ and the proper-time resolution for correctly or wrongly reconstructed D mesons in signal MC.

can be neglected. The fit of these events in the $M_{\rm bc}$ and ΔE distribution shows that they represent $\sim 0.5\%$ of the total signal events in the signal MC, as demonstrated in Figure 3.9.

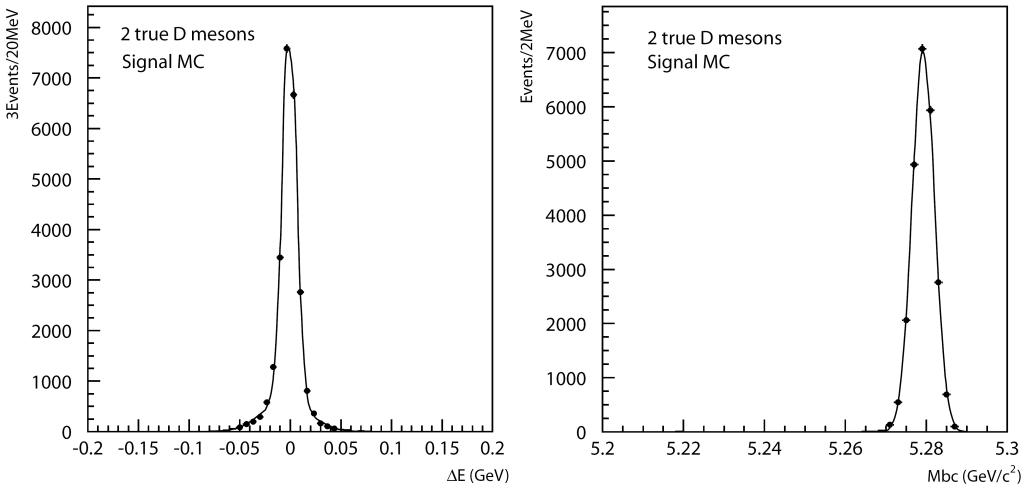


Figure 3.7: ΔE and $M_{\rm bc}$ distributions for signal MC events with both D mesons correctly reconstructed.

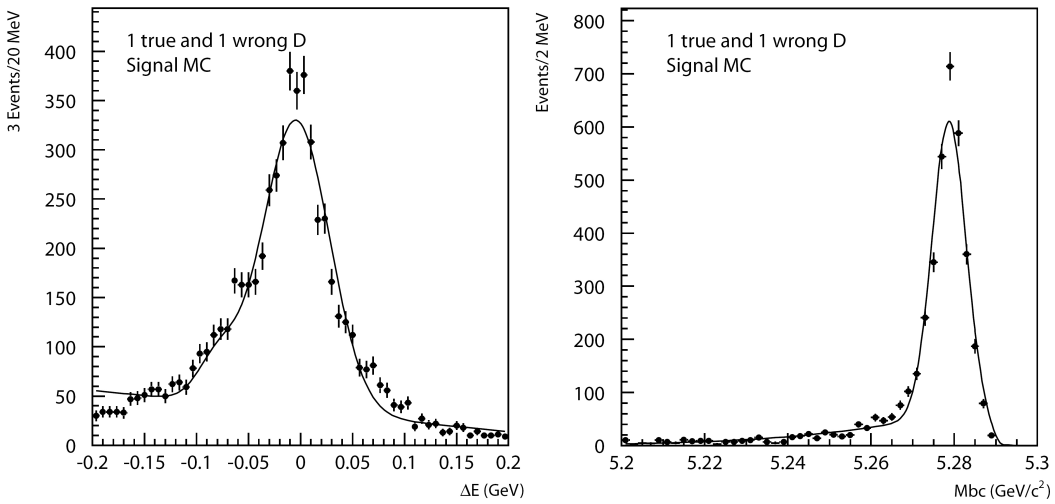


Figure 3.8: ΔE and $M_{\rm bc}$ distributions for signal MC events with only one D meson correctly reconstructed.

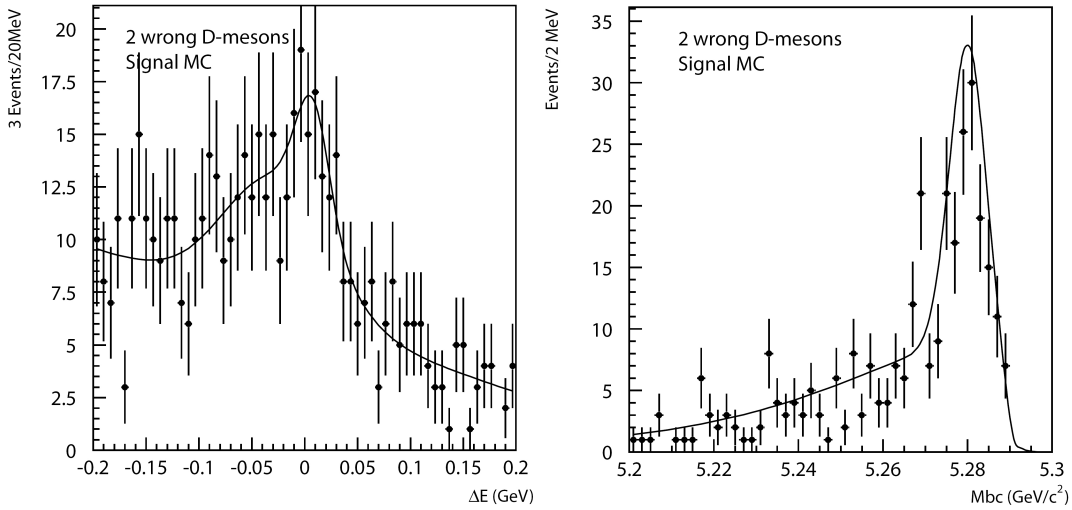


Figure 3.9: ΔE and $M_{\rm bc}$ distributions for signal MC events with two badly reconstructed D mesons.

3.5 Generic MC study

The generic MC sample is mainly used for background studies. It contains almost all physical processes that take place at the $B\overline{B}$ resonance, including the $B^0 \to D^{*+}D^{*-}$ decays. This sample can also be used to study the selection criteria or to determine the signal shape. This will give us complimentary information on the signal shape. The signal MC however has the advantage that only the signal decays of interest to us are generated, which allows us to create much larger statistics. The other difference is that the signal-MC sample contains only the D-meson subdecays which we wanted to reconstruct in the first place, while the generic-MC events are generated inclusively. Even with the smaller statistics it is still useful to study the signal events in the generic MC, as it compares better to the situation on the data.

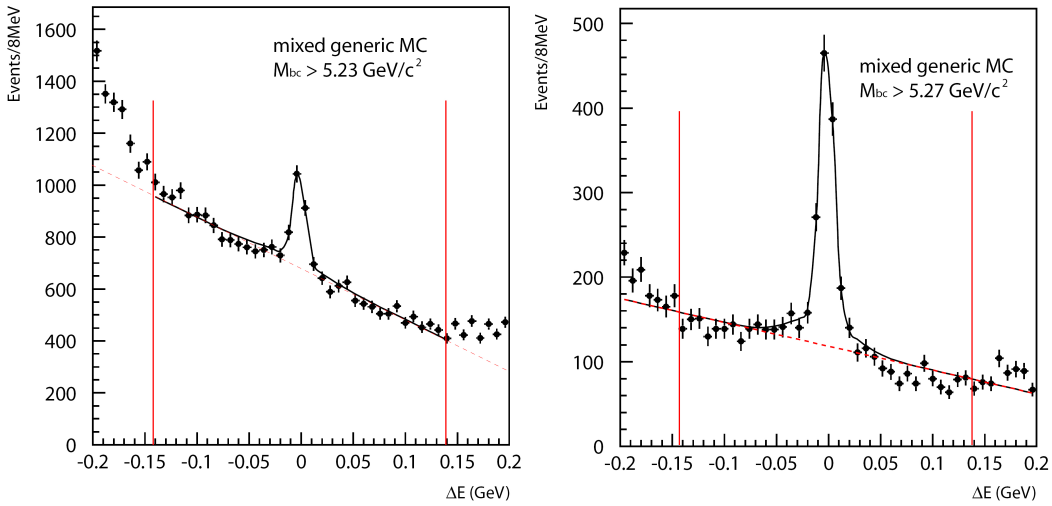


Figure 3.10: ΔE distribution of the $B^0\overline{B^0}$ generic MC events when $M_{\rm bc} > 5.23\,{\rm GeV/c^2}$ and $M_{\rm bc} > 5.27\,{\rm GeV/c^2}$.

The ΔE distribution of the generic B^0 $\overline{B^0}$ events for where $M_{\rm bc} > 5.23\,{\rm GeV/c^2}$ and $M_{\rm bc} > 5.27\,{\rm GeV/c^2}$ is shown in Figure 3.10. One notices that the background shape around $\pm 180\,{\rm MeV}$ does not correspond to the background model we used to describe the signal MC or data. These bumps come from hadron mis-identification. The energy of the particles is calculated from its measured momentum and its mass which is assigned according to the particle's identification. As the energy of the B-meson is the sum of the energies of the daughter particles, there can be an energy surplus of $\Delta E \simeq M_K^{\pm} - M_{\pi^{\pm}} \simeq 353\,{\rm MeV}$ when a pion is taken as a kaon in for example the $B^0 \to D_S^+ D^-$ decays. The behavior we see of the background near the border of the large signal region is merely due to the tails of the distribution. The bump on the left side of the ΔE spectrum is caused by the opposite effect, when a kaon is wrongly identified as a pion. Also on the data this effect can be seen outside the large signal region. The reason why this effect is barely noticeable on the signal MC is because only $B^0 \to D^{*+}D^{*-}$ decays were generated. The large signal region is chosen such that these bumps are not interfering when determining the data yield. This is also done in the analysis performed by the Babar collaboration [63]. Furthermore we can see from the distribution of ΔE that the background should be described by a higher

	generic MC	signal MC
ΔE		
$\mu_{\rm main}\,({\rm MeV})$	-1.98 ± 0.32	-0.78 ± 0.06
$\sigma_{\rm main}\,({\rm MeV})$	6.29 ± 0.35	7.02 ± 0.07
frac	0.60 ± 0.03	0.71 ± 0.01
$\mu_{\rm tail}\,({\rm MeV})$	-6.20 ± 1.89	-6.25 ± 0.49
$\sigma_{\rm tail}\,({\rm MeV})$	29.4 ± 2.02	28.18 ± 0.46
bkg slope	-3.47 ± 0.36	-2.85 ± 0.09
$M_{\rm bc}$		
$\mu\,({\rm GeV/c^2})$	5.2803	5.2792
$\sigma\,({\rm MeV/c^2})$	2.98 ± 0.08	2.81 ± 0.01
argus slope	88.3 ± 6.3	124.4 ± 1.6

Table 3.9: Parameters of the fitted ΔE and $M_{\rm bc}$ shape for the signal events in the generic MC and the signal MC.

order polynomial, instead of a first-order polynomial as is done in the signal MC. We will therefore model the generic MC and data background contribution with a second order polynomial.

The generic-MC samples also contain information on the originally generated event. We can therefore ask to see only these reconstructed events which are generated $B^0 \to D^{*+}D^{*-}$ decays, but still without any requirement on the $D^{(*)}$ decays. Figure 3.11 shows the fitted $M_{\rm bc}$ and ΔE distributions in the small signal region for such events. The two-dimensional fit was performed on the large signal region, i.e. $5.23\,{\rm GeV/c^2} < {\rm M_{bc}} < 5.3\,{\rm GeV/c^2}$ and $-0.14\,{\rm GeV} < \Delta{\rm E} < 0.14\,{\rm GeV}$. The fit parameters are summarized in Table 3.9.

When comparing the distribution and fit parameters of this generic MC sample, which contains only events which are generated $B^0 \to D^{*+}D^{*-}$ decays, with that of the signal MC sample, it is seen that there is more combinatorial background in the generic sample and that the width of the wider Gaussian is a bit smaller in the signal MC and contributes less to the total signal yield. The reason for this is the extra, dirtier decay channels of the D mesons which are contained in the generic-MC sample. The decay channels of the D meson, which are generated in the signal MC, are only those we decided to reconstruct, while in Figure 3.11 we only required the B^0 to decay to $D^{*+}D^{*-}$ without any requirements on the decay channels of the D mesons. The extra decays in this generic MC sample contain more often π^0 particles which contribute to the wider Gaussian in the ΔE distribution. As we know that the events in the generic distributions all come from $B^0 \to D^{*+}D^{*-}$ decays, this shape parametrization is more realistic to describe the distribution of the signal events in the data. We will therefore use this function to determine the fraction, mean and width of the large Gaussian in the ΔE distribution when determining the final signal yield on the data. The fitted number of signal events in the generic MC are summarized in Table 3.10. We will show in Section 3.7 that this number of signal events corresponds to the number of signal events in the data, when scaled to the corresponding luminosity.

signal events	1143.5 ± 38.3
background events	1001.1 ± 36.2

Table 3.10: Fitted signal and background events in the generic $B^0\ \overline{B^0}$ events .

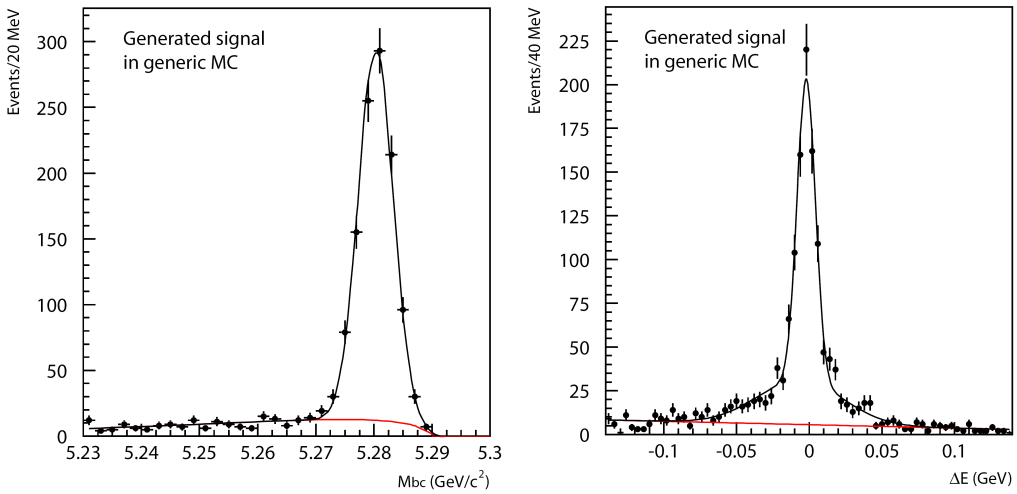


Figure 3.11: Projections of the fitted $M_{\rm bc}$ and ΔE distribution of the $B^0 \to D^{*+}D^{*-}$ events in the generic MC sample.

3.6 Background study

The background is studied to verify if there are any peaking background events. These are background events that are wrongly reconstructed but have a ΔE or $M_{\rm bc}$ distribution similar to that of signal. The peaking background from $B\overline{B}$ events is studied using the generic MC samples. Furthermore the background is studied to understand the nature of the events in the sideband region and under the signal peak in the data.

3.6.1 Peaking background

Peaking background events are caused by events that look similar in the detector to the $B^0 \to D^{*+}D^{*-}$ decays. For example when a $D_s \to KK\pi$ decay is reconstructed as a $D \to K\pi\pi$ its behavior in $M_{\rm bc}$ will be like a signal event. Its ΔE value however will be shifted, as the wrong mass is taken. Previous analysis from Belle and Babar have searched for peaking background events in the data samples. A summary of their conclusions is:

- Belle ($152 \times 10^6\, B\overline{B}$ events): No significant peaking background is observed [58].
- Babar ($383 \times 10^6\, B\overline{B}$ events): There is $(1.8 \pm 1.8)\%$ of peaking background in $M_{\rm bc}$ coming from the $B^+ \to \overline{D}^{*0}D^{*+}$ decay [63].

3.6.1.1 generic B^+ B^- events

The sample of reconstructed generic B^+ B^- events corresponds to $1185\,\mathrm{fb}^{-1}$. This means that the sample contains 1.959 times the number of B^+B^- events on the data. The ΔE and $M_{\rm bc}$ distributions are shown in Figure 3.12. They are separately fitted with the same PDFs as will be used on the data or signal MC to allow a different yield in both distributions which might be due to peaking background. As we don't expect many signal-like events, the shape of the signal PDF is fixed. The parameters that remain free are the background shape and the number of signal and background events. The fit was performed twice, once

while requiring the other variable to be within the small signal region and once requiring it to be in the large signal region. The fitted number of signal events is shown in Table 3.11. From the fit and the statistical errors on the obtained signal events we can conclude that we have no significant peaking background in the charged B decays.

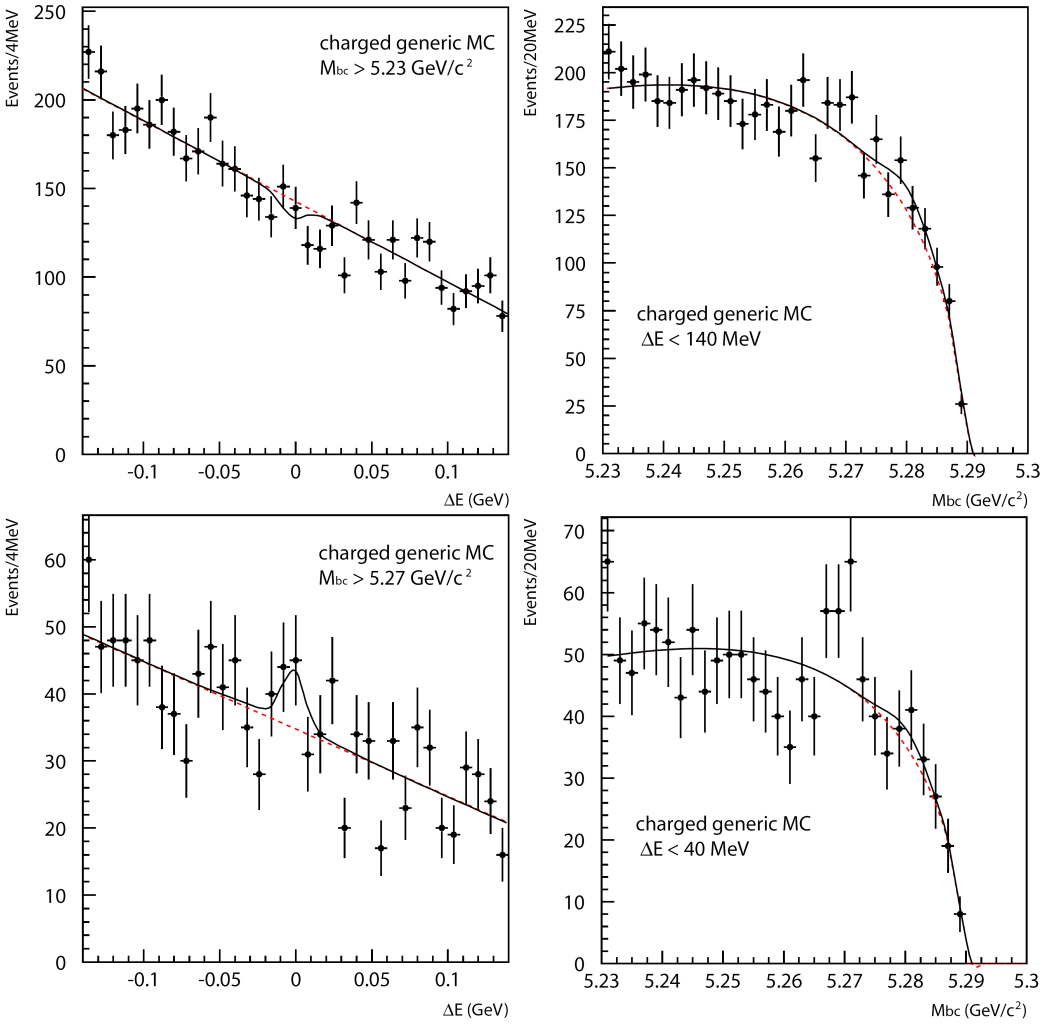


Figure 3.12: ΔE and $M_{\rm bc}$ distributions of the generic B^+ B^- events in the large (top) and small (bottom) signal region.

3.6.1.2 generic B^0 $\overline{B^0}$ events

The sample of generic B^0 $\overline{B^0}$ events contains $1103\,\mathrm{fb}^{-1}$ events or 1.825 times the number of $B^0\overline{B^0}$ interactions in the data sample. To study peaking background contributions in this sample we need to take out the actual $B^0 \to D^{*+}D^{*-}$ decays. This is possible by using the information on the generated decay. The resulting ΔE and $M_{\rm bc}$ distributions on the large and small signal region are shown in Figure 3.13. Like the charged generic sample, these distributions are fitted separately with the shape of the signal distribution fixed from the signal MC. The result is shown in Table 3.12.

	large signal region	small signal region
M_{bc}:	47.5 ± 36.8	11.5 ± 19.1
ΔE :	-28.8 ± 26.5	24.8 ± 14.5

Table 3.11: Fitted signal events in the generic B^+ B^- events.

	large signal region	small signal region
$M_{\rm bc}$	142.8 ± 61.7	70.1 ± 32.7
ΔE	55.7 ± 47.3	31.1 ± 24.1

Table 3.12: Fitted signal events in the generic B^0 $\overline{B^0}$ event sample.

From the figures and the fit result it seems that there is no significant peaking background in the ΔE distribution. However the result on the $M_{\rm bc}$ distribution cannot be neglected as it represents already $\frac{70.1 \pm 32.7}{1060} = 6.6 \pm 2.8\%$ of the signal events in the small signal region obtained from a one-dimensional fit on the $M_{\rm bc}$ distribution, where 1060 is the number of the true signal evens in the small signal region. This effect is reduced when obtaining the yield on the data sample as a two-dimensional fit is used.

Different checks are carried out to see if these 70 ± 33 peaking background events could be identified. The events are therefore separated between $D^{*+}D^{*-} \rightarrow (D^0\pi^+)(\overline{D}^0\pi^-)$ decays and $D^{*+}D^{*-} \rightarrow (D^0\pi^+)(D^-\pi^0)$ or $D^{*+}D^{*-} \rightarrow (D^+\pi^0)(\overline{D}^0\pi^-)$ channels. The result of the fit over these two samples in the small signal region is shown in Table 3.13.

It seems that the peaking background is merely caused by reconstruction of the $D^{*+}D^{*-} \rightarrow (D^0\pi^+)(D^-\pi^0)$ or $D^{*+}D^{*-} \rightarrow (D^+\pi^0)(\overline{D}^0\pi^-)$ channels. Each $D^{\pm}$ channel is therefore investigated one by one. Every subdecay contributed a little bit to the total peaking background and no decay contributed in a significantly larger amount.

More tests are done to check if the peaking background is caused by a D_s modes. The $D_s^{(*)}D^{(*)}$ control sample and the signal MC sample are used to investigate $D_s \rightarrow KK\pi$ and $D_s \rightarrow K_{\rm S}K$ which are reconstructed as $D_s \rightarrow K\pi\pi$ and $D_s \rightarrow K_{\rm S}\pi$ respectively. However the level of contamination of D_s particles in the sample cannot explain the $6.6 \pm 2.8\%$ peaking background. In the next section it will show that this peaking background is probably overestimated and will be taken into account in the systematic error.

3.6.2 Background level

From similar studies at Belle we expect the background composition on the data to be roughly distributed as follows:

	$D^0\overline{D^0}$	$D^+\overline{D^0}$ and D^0D^-
$M_{\rm bc}$	20.6 ± 21.3	54.8 ± 25.6
ΔE	41.9 ± 17.3	-3.2 ± 19.1

Table 3.13: Fitted signal events in the small signal region for separate decays, in the generic B^0 $\overline{B^0}$ event sample.

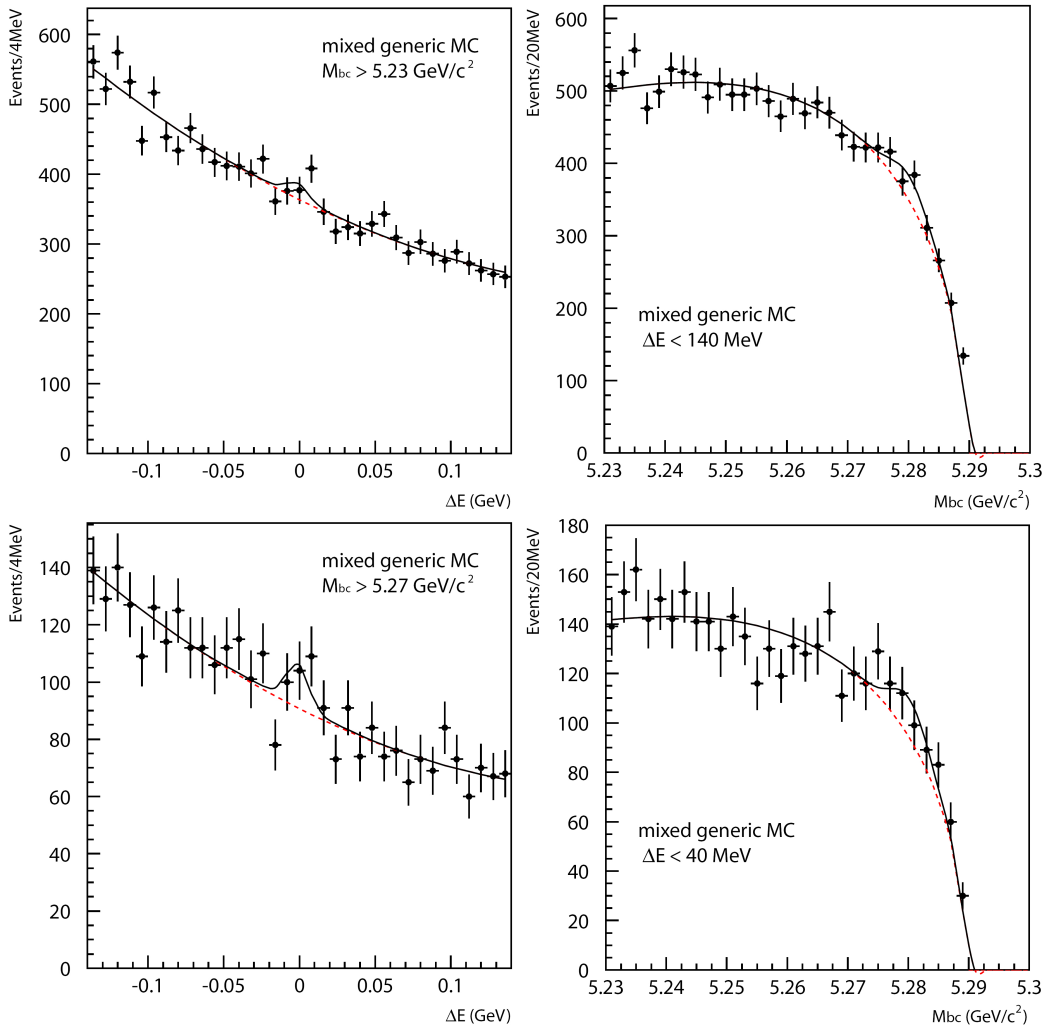


Figure 3.13: ΔE and $M_{\rm bc}$ distributions of the generic $B^0\ \overline{B^0}$ event sample in the large (top) and small (bottom) signal regions. The true $B^0 \to D^{*+}D^{*-}$ signal has been removed.

1. Continuum background: 50%;
2. Background from $B^0\overline{B^0}$: 30%;
3. Background from B^+B^-: 20%.

3.6.2.1 Continuum background

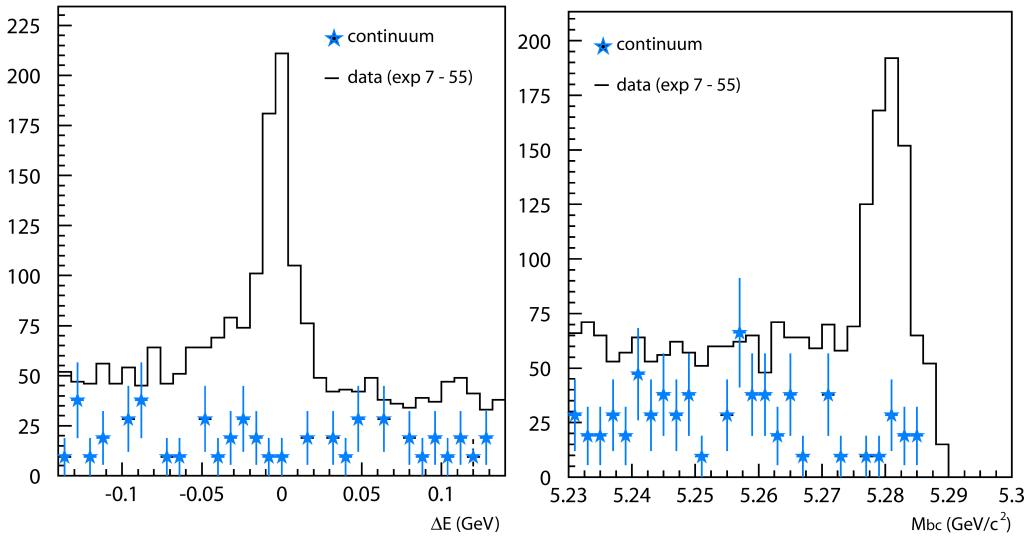


Figure 3.14: ΔE and $M_{\rm bc}$ distribution of the on-resonance data overlaid with the distribution of the off-resonance data normalized to the same luminosity.

The continuum background was studied on the off-resonance data sample of $68\,\mathrm{fb}^{-1}$ which corresponds approximately to 10% of the continuum background under the $\Upsilon(4S)$ resonance. We scaled the off-resonance data sample to $605\,\mathrm{fb}^{-1}$, to visualize the contribution to the data background. The statistical errors in each bin become large, but it is still possible to check if this background accounts roughly for half of the data background. Figure 3.14 shows the data distribution in the small signal region in ΔE and $M_{\rm bc}$ overlaid with the scaled continuum background. The scaled continuum contribution to the background has the expected level.

3.6.2.2 Generic $B^0\ \overline{B^0}$ and $B^+\ B^-$ events

The generic $B^0\ \overline{B^0}$ and $B^+\ B^-$ events are scaled and superimposed on the data distributions of ΔE and $M_{\rm bc}$ and the signal $B \to D^{*+}D^{*-}$ is taken out. Even though the number of signal events in the generic $B^0\overline{B^0}$ sample corresponds to the fitted number of signal events in the data when scaled to equal luminosity, the background level seems to be overestimated. The black dots on the histogram in Figure 3.15 represent the data distributions of $M_{\rm bc}$ and ΔE in the small signal region. We will describe in the next section how exactly the data events are obtained. The green histogram represents the scaled generic $B^+\ B^-$ events and the yellow one the scaled generic $B^0\ \overline{B^0}$ events where the signal has been removed. Each of these background components should roughly accommodate for

a third of the data background. However, the plots show that the generic background is overestimated, by a factor of two for the charged generic MC and around a factor of four for the $B^0\,\overline{B^0}$ background. To further investigate this we looked at the mass distributions of D^* and D^0 particles in the sideband region $M_{\rm bc} < 5.27\,{\rm GeV/c^2}$ (see Figure 3.16), to see if this extra background arises at the level of the D or D^* reconstruction. Both D and D^* mesons show a surplus in the generic MC with respect to the data background. This means that there might be a mistake in the branching fractions used to generate these events. However no other analysis noticed this.

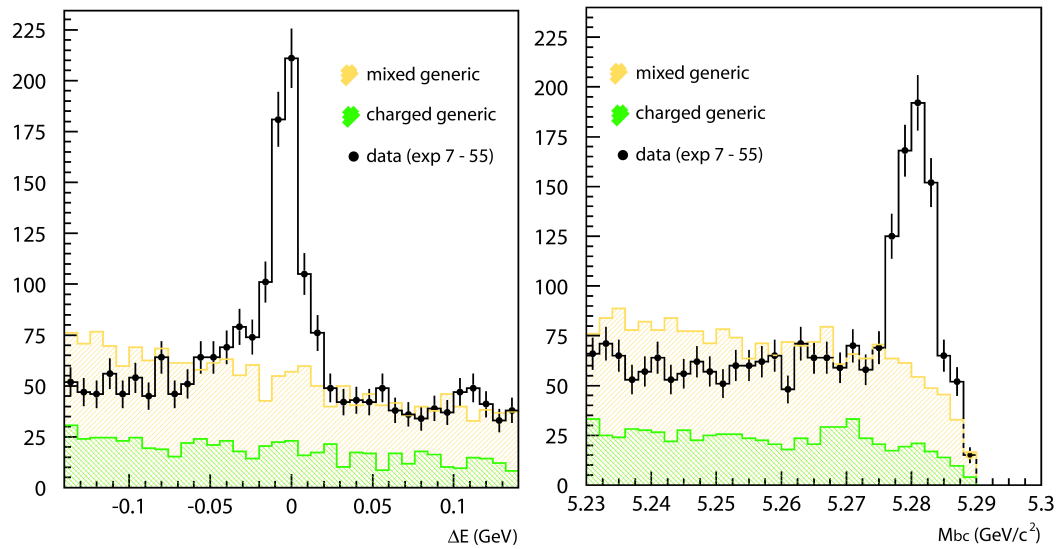


Figure 3.15: ΔE and $M_{\rm bc}$ data distributions of the on-resonance overlaid with the generic MC distributions normalized to the same luminosity.

Finally, Figure 3.17 summarizes the different contributions from the off-resonance, the B^+B^- and $B^0\,\overline{B^0}$ samples. The signal peak from the $B^0\,\overline{B^0}$ sample is now visible as well. The sum of the three background contributions is almost twice as large as the background level in the data. Different tests on this overestimation of the background in the generic MC samples, such as dividing the sample channel by channel, didn't lead to any trivial explanation. This means that the $6.6 \pm 2.8\%$ peaking background events of Section 3.6.1.2 is probably overestimated as well. We decided therefore to keep the peaking background level estimate as $6.6 \pm 2.8\%$ of the signal and to add this effect to the systematic error for the main results of this analysis instead of adding a peaking background contribution to the pdf of the $M_{\rm bc}$ distribution. With this conservative approach the peaking background contribution is probably overestimated, but as will be seen in Chapter 4 and 6 this does not give rise to an obvious overestimation of the systematic error.

3.7 Data

3.7.1 Skim

To study the on-resonance data sample of experiments 7-55, we use the events from the HadronB(J) skim. However, performing the reconstruction and applying selection criteria

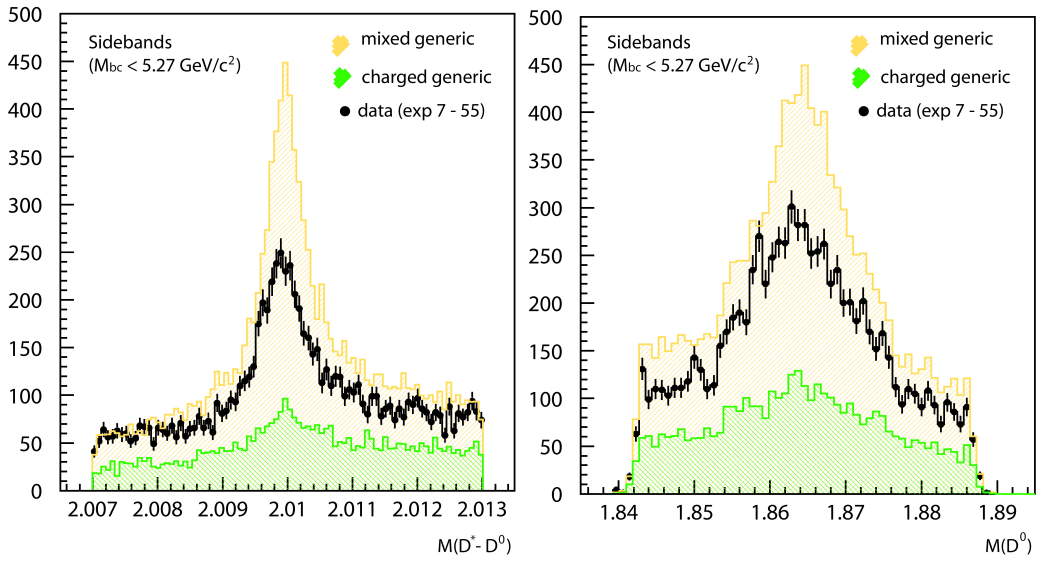


Figure 3.16: $D^{*+} - D^0 + 1.864\,\mathrm{GeV}$ and D^0 mass distributions in the data sideband region overlaid with the generic $B^0\ \overline{B^0}$ and $B^+\ B^-$ events normalized to the same luminosity.

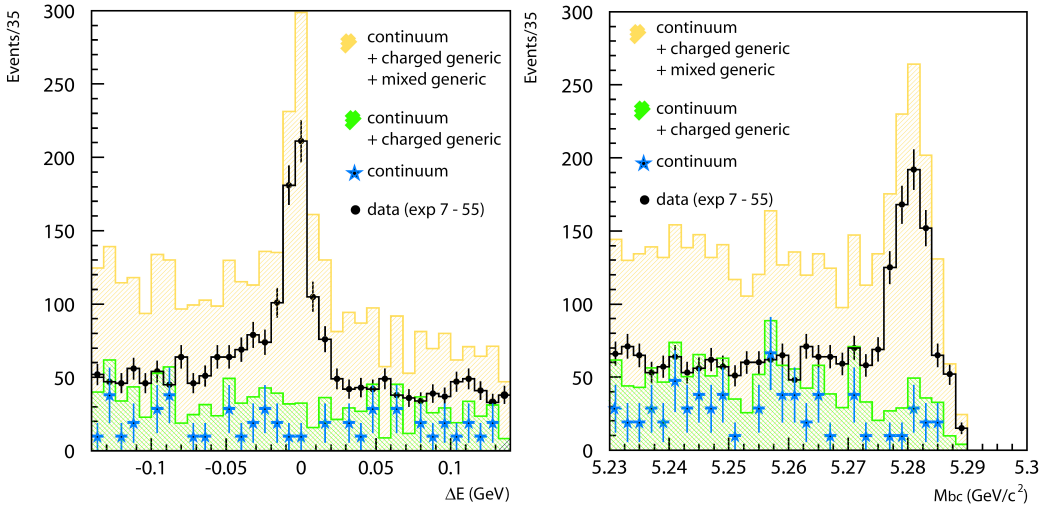


Figure 3.17: ΔE and M_{bc} data distributions superimposed with the sum of the different contributions estimated from the off-resonance data and generic MC.

on this sample still requires too much CPU time. Hence we need to pre-filter this sample further. This is done by applying a second skim, which enhances $B \to D^{*+}D^{*-}$ proportion in the sample by applying very loose cuts, which should barely discard any signal events.

The list of cuts used for this skim is shown in Table 3.14. This skim rejects approximately 99.8% of the data events from the HadronB(J) sample while keeping almost all the $B^0 \to D^{*+}D^{*-}$ events.

transverse distance (track w.r.t IP):	$dr < 2.0$ cm
longitudinal distance (track w.r.t IP):	$\|dz\| < 4.0$ cm
K identification	$\mathcal{L}(K^{\pm})/(\mathcal{L}(K^{\pm}) + \mathcal{L}(\pi^{\pm})) > 0.1$
π identification	$\mathcal{L}(K^{\pm})/(\mathcal{L}(K^{\pm}) + \mathcal{L}(\pi^{\pm})) < 0.9$
π^0 selection	$E_\gamma > 0.03$ GeV
$\mathrm{K_S}$ selection	goodKs [61]
D-mass window	$\pm 70\,\mathrm{MeV/c^2}$
D^*-D mass window	$\pm 25\,\mathrm{MeV/c^2}$
B window	$M_{bc} > 5.19\,\mathrm{GeV/c^2}$
	$\|\Delta E\| < 200\,\mathrm{MeV}$

Table 3.14: Selection criteria applied to the HadronB(J) skim.

3.7.2 Event reconstruction

The data which passed the skim criteria are reconstructed using the exact same reconstruction method as performed on the signal MC (see Section 3.3). Before the best candidate selection is applied there are on average 1.6 B^0 candidates per event. When the best candidate selection is applied we obtain 937 events in the small signal region for experiments 7-55.

3.8 Final signal yield in the data

The two-dimensional ΔE and M_{bc} distribution is shown in Figure 3.18. The signal yield on the total data set is obtained in the same way as explained for the signal MC, i.e. by fitting the reconstructed ΔE and M_{bc} distributions in a two-dimensional plane. The signal and background contributions are modeled by the functions given by Equations (3.3) and (3.4). The fraction and shape of the second Gaussian in the PDF of the ΔE distribution of the signal events are taken from the generic MC sample. The study in the previous section also showed that the ΔE background shape should be modeled with a second order polynomial:

$$\mathcal{P}(M_{\mathrm{bc}}, \Delta E) = ARGUS(M_{\mathrm{bc}}, s) \times \mathcal{P}_2(\Delta E, a, b), \tag{3.5}$$

with $\mathcal{P}_2 = N(a(\Delta E)^2 + b(\Delta E))$, where N is the normalization factor and the parameter s in the ARGUS function denotes its slope.

The ΔE and M_{bc} data distributions are fitted in the same way, with a two-dimensional unbinned maximum likelihood. The shape, number of signal events and number of background events are left floating during the fit, but due to the small statistics we fix the

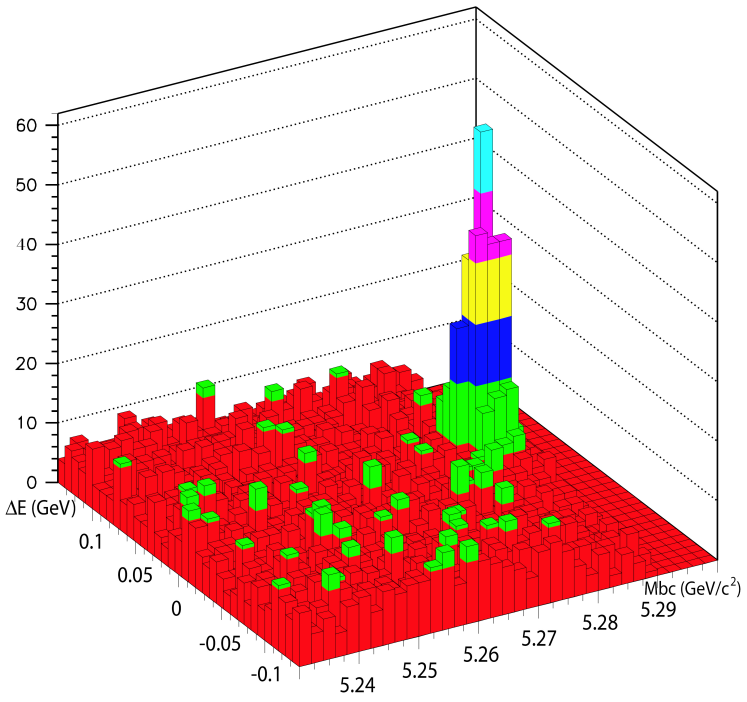


Figure 3.18: ΔE and $M_{\rm bc}$ data distribution in the two-dimensional plane.

fraction and shape of the second Gaussian in ΔE. The plots in Figure 3.19 show the projections on the small signal region of the fitted distributions as well as the background contributions in red. The parameters of the fit are summarized in Table 3.15. In a sample of 657 million $B\overline{B}$ events we reconstruct 554 ± 30 signal events. When requiring only events in the small signal region, we lose 7.3% signal events by cutting $\Delta E < 40\,\mathrm{MeV}$. The $M_{\rm bc} > 5.27\,\mathrm{GeV/c^2}$ cut removes only background events. From the 937 events in the small signal region, 513 ± 28 are signal events, which gives a signal purity in the small region of:

$$\frac{\mathcal{S}}{\mathcal{S}+\mathcal{B}} = \frac{513 \pm 28}{937} = 55 \pm 3.0\%.$$

The one-dimensional fits gave also compatible results.

	ΔE			M_{bc}	
μ_1	-3.37 ± 0.57	MeV	μ	5.2807	$\mathrm{GeV/c^2}$
σ_1	6.89 ± 0.47	MeV	σ	2.70 ± 0.14	$\mathrm{MeV/c^2}$
f	0.60*				
μ_2	-6.20*	MeV			
σ_2	29.4*	MeV			
a	-1.66 ± 0.16		s	41.0 ± 2.3	
b	2.26 ± 2.27				

Table 3.15: Fitted parameters of the 2D fit of the ΔE and $M_{\rm bc}$ data distribution. The parameters with a * are fixed to the signal MC values.

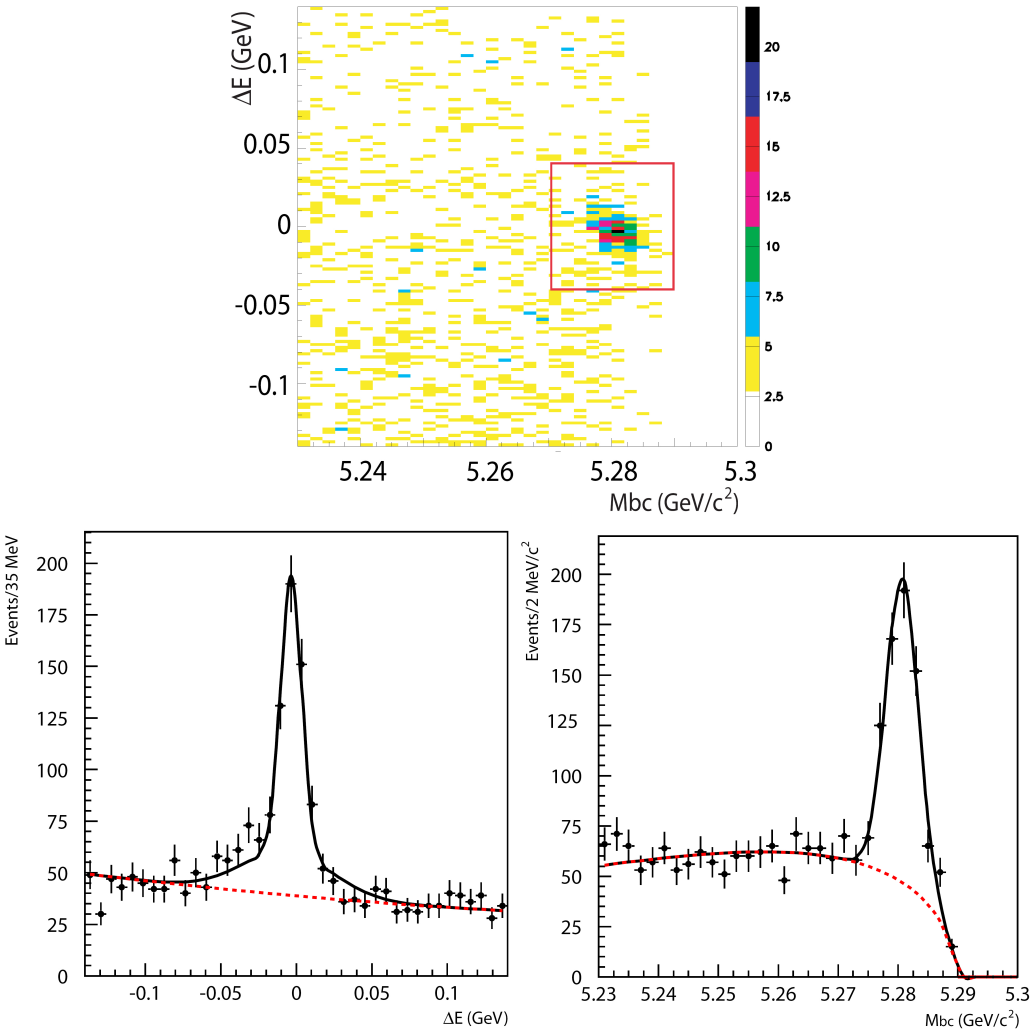


Figure 3.19: Top : Two-dimensional ΔE and $M_{\rm bc}$ data distributions. Bottom left: projection on ΔE. Bottom right: projection on $M_{\rm bc}$.

Finally we can estimate the branching fraction, using the following formula:

$$\begin{aligned}\mathcal{B}(B^0 \to D^{*+}D^{*-}) &= \frac{N_{\rm sig}^{\rm data}}{N_{\rm sig}^{\rm gen\,MC}} \times \frac{\mathcal{L}^{\rm gen\,MC}}{\mathcal{L}^{\rm data}} \mathcal{B}(\text{gen.MC}) \\ &= \frac{554 \pm 30}{1243 \pm 52} \times \frac{1103\,\text{fb}^{-1}}{656\,\text{fb}^{-1}} \times 1.10 \times 10^{-3} \\ &= (8.17 \pm 0.24) \times 10^{-4}, \end{aligned} \tag{3.6}$$

where $N_{\rm sig}^{\rm gen\,MC} = 1242.9 \pm 38.2$ is obtained from a two-dimensional fit on the total generic MC sample. This result is consistent with the world average [6] of $\mathcal{B}(B^0 \to D^{*+}D^{*-}) = (8.2 \pm 0.9) \times 10^{-4}$.

3.9 Discussion

An analysis of the $B^0 \to D^{*+}D^{*-}$ decay has also been performed by the BaBar collaboration [63, 64, 65]. The Babar analyses show a higher reconstruction efficiency than the Belle analysis. The latest BaBar analysis [65] is performed on $467 \times 10^6\,\text{B}\bar{\text{B}}$ events and yields 934 ± 10 signal events, which corresponds to a reconstruction efficiency which is more than a factor two higher than the one in this analysis. This effect is also seen in other analyses where slow pions enter the scene.

Table 3.16 shows the number of signal events reconstructed by BaBar and Belle in different double charm analyses. Figure 3.20 then represents the ratio of the BaBar and the Belle reconstruction efficiencies for these analyses. In the lastest $B^0 \to D^{*+}D^{*-}$ analysis [65] a more sophisticated track finding algorithm is used, which improved their tracking efficiency with another 10 to 20% for the reconstruction of $D^{*+}D^{*-}$ particles. From Figure 3.20 it can be seen that the BaBar performance relatively to the Belle performance increases when more slow particles enter the scene. The $D^{*+}D^{*-}K_{\rm S}^0$ reconstruction is added not because it has more slow charged pions but adding an the extra particle to the reconstruction is again more challenging for the Belle experiment.

		BaBar		Belle	
	Decay mode	Data sample	events	Data sample	events
1	$B^0 \to D^{*+}D^{*-}K_{\rm S}^0$	$230 \times 10^6\text{B}\bar{\text{B}}$	201 ± 17 [66]	$449 \times 10^6\text{B}\bar{\text{B}}$	131 ± 14 [67]
2	$B^0 \to D^{*+}D^{*-}$	$467 \times 10^6\text{B}\bar{\text{B}}$	934 ± 40 [65]	$657 \times 10^6\text{B}\bar{\text{B}}$	554 ± 30
3	$B^0 \to D^{*+}D^{*-}$	$383 \times 10^6\text{B}\bar{\text{B}}$	617 ± 33 [63]	$657 \times 10^6\text{B}\bar{\text{B}}$	554 ± 30
4	$B^0 \to D^{*+}D^{-}$	$467 \times 10^6\text{B}\bar{\text{B}}$	724 ± 37 [65]	$152 \times 10^6\text{B}\bar{\text{B}}$	155 ± 17 [68]
5	$B^0 \to D^{+}D^{-}$	$467 \times 10^6\text{B}\bar{\text{B}}$	152 ± 17 [65]	$537 \times 10^6\text{B}\bar{\text{B}}$	150 ± 15 [59]

Table 3.16: Luminosity and reconstructed signal events in different double charm analyses at BaBar and Belle.

The reason for the better reconstruction performance of low momentum particles at BaBar is the different configuration of the BaBar's innermost silicon detector (SVT). The BaBar SVT detector consists of five layers of double-sided silicon strip sensors [69], while the Belle innermost detector only has four layers of silicon. In the track finding algorithm of Belle, one searches first for correlated hits in the CDC detector to make a track candidate. In a second step SVD hits are associated to the CDC track; this improves the track

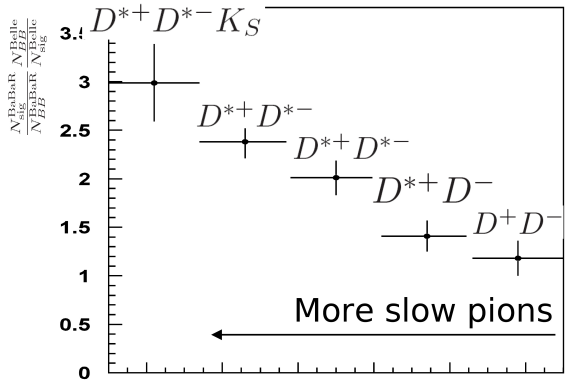


Figure 3.20: Ratio of the reconstruction efficiencies at BaBar and Belle for different double charm analyses, numbers extracted from Table 3.16.

resolution but not the track finding efficiency. Therefore low momentum charged particles which leave most of their tracks in the SVD detector and which are largely affected by multiple scattering are very difficult to find in the Belle detector.

Figure 3.21 shows the tracking efficiency for charged pions at Belle as a function of the momentum of the pion. The tracking algorithm of the BaBar experiment for low momentum tracks has an efficiency of 75% for particles with a momentum $p > 100\,\mathrm{MeV}/c$ and 90% for $p > 200\,\mathrm{MeV/c}$ [70], which is clearly better than the Belle performance in that region.

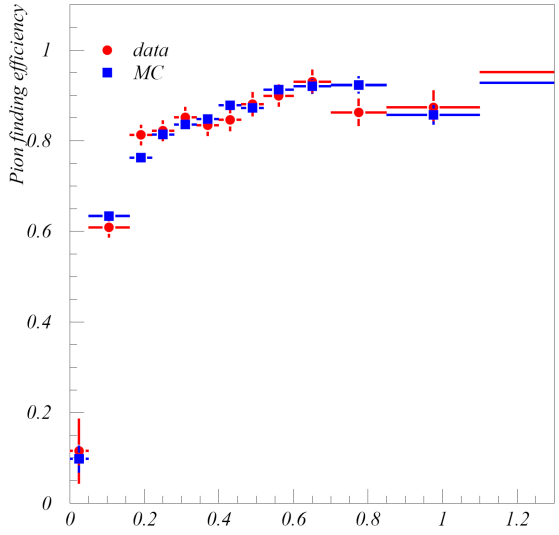
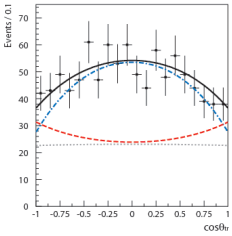


Figure 3.21: The tracking efficiency for charged pions at Belle as a function of the momentum (GeV/c) of the charged particle.

Chapter 4

Angular analysis



This chapter describes the model used to parametrize the angular distributions of the $B^0 \to D^{*+}D^{*-}$ decays in the transversity frame as well as the background events. A fast simulation is used to perform a linearity test and to estimate the error distribution. Finally, the extracted CP-odd fraction is presented followed by the results of a systematic study.

THE decay studied in this analysis consists of a spinless B-meson which decays in two D^* vector mesons. Because the D^* is a spin-1 particle ($J^P = 1^-$), the final states can exist with three different orbital angular momenta corresponding to $l = 0, 1$ or 2. As explained in Chapter 1 the CP eigenvalue depends on the orbital angular momentum in the following way: $CP = (-1)^l$. This means that two different CP eigenstates can exist depending on whether the orbital angular momentum is odd or even. The indirect CP-violating parameter of the CP-odd and CP-even states have an opposite sign, therefore these states need to be disentangled to prevent a dilution effect. In Section 1.8 we described how the directions of the D^* daughter particles can be used to separate the CP-odd from the CP-even contributions. This is done by working with the so-called transversity basis [71, 22].

The procedure of extracting the CP-odd fraction will be explained in the first section. Different fitting strategies are proposed and their performance is tested with fast MC samples. It should be pointed out that in the measurements of the CP-odd fraction, the errors are dominated by statistics, not systematics. Next a full MC sample is used to model the angular distributions of the three polarization states and to study the effect of the detector resolution. The shape of the background will be studied by looking at the sideband regions of the data. A fast MC is used to determine any possible bias in the fitting procedure and to study statistical fluctuations. Finally we measure the CP-odd fraction on the total data set. A study of the systematic error will be presented as well.

variable	$\mathcal{P}_0$	$\mathcal{P}_\perp$	$\mathcal{P}_\parallel$
$\theta_{\rm tr}$	$1-\cos^2\theta_{tr}$	$\cos^2\theta_{tr}$	$1-\cos^2\theta_{tr}$
θ_1	$\cos^2\theta_1$	$1-\cos^2\theta_1$	$1-\cos^2\theta_1$

Table 4.1: Theoretical shapes of the $\cos\theta_{\rm tr}$ and $\cos\theta_1$ distributions for the three polarization states.

4.1 Determining the angular model

4.1.1 Performance of the one-angle versus two-angle model

As a reminder we show the theoretical shapes of the $\theta_{\rm tr}$ and θ_1 distributions for the three polarization states in Table 4.1. From Eq. (1.137) it is seen that we can extract two independent polarization fractions, $R_\perp$ and R_0, by fitting the $\theta_{\rm tr}$ and θ_1 angular distribution. The one-angle transversity distribution of $\theta_{\rm tr}$ can be modeled with Eq. (1.140) $R_\perp$. A fast MC is used to test the performance of the two strategies: using two independent angles to extract $R_\perp$ and R_0, or only fitting one distribution to obtain just a value for $R_\perp$.

Different MC files are generated, each time using a different polarization fraction and number of events. The angular distributions are generated according to the theoretical expression (see Eq (1.138)) and no background contribution is added. Next the different fast MC samples are fitted with both the single-angle method and the double-angle method. The result is summarized in Table 4.2 and shows that when the angular shapes follow the theoretical distribution no method seems to offer a significantly better accuracy on the fitted CP-odd fraction.

Events	Generation		One-angle fit	Two-angle fit	
	$R_\perp$	R_0	$R_\perp$	$R_\perp$	R_0
100k	0.40	0.10	0.400 ± 0.002	0.400 ± 0.002	0.097 ± 0.002
100	0.43	0.17	0.416 ± 0.030	0.413 ± 0.031	0.202 ± 0.027
500	0.50	0.30	0.553 ± 0.079	0.570 ± 0.085	0.163 ± 0.060

Table 4.2: Generated and fitted parameters in the fast MC study using the one- or two-angle method.

When using only the $\cos\theta_{\rm tr}$ distribution to extract the CP-odd fraction there is no distinction made between the distributions corresponding to the A_0 and the $A_\parallel$ amplitudes. Theoretically this approach is correct as they have the same shape, but in practice a non-uniform reconstruction acceptance can distort the A_0 and the $A_\parallel$ distribution. This will be shown in Section 4.1.2. To parametrize the combined CP-even distribution the relative contribution of A_0 and $A_\parallel$ needs to be known.

The previous Belle analysis on the SVD1 data [58] used both the $\theta_{\rm tr}$ and the θ_1 distributions to extract $R_\perp$ and R_0. The parametrization of the efficiency in the $\theta_{\rm tr}$ and the θ_1 plane is not a trivial procedure. In the previous analysis the shape was determined from a binned two-dimensional histogram, which didn't allow an unbinned maximum likelihood procedure to extract $R_\perp$ and R_0. The latest Babar analysis [63] used the one-angle method where the CP-even distribution is constructed from the distributions corresponding to the A_0 and $A_\parallel$ amplitudes, summed in equal portions. They have taken the error on this assumption into account in the systematic uncertainty.

We preferred to perform a one-angle fit on the $\theta_{\rm tr}$ distribution to obtain only the CP-odd fraction, as done in the Babar analysis. Instead of adding the angular distributions corresponding to A_0 and $A_\parallel$ with equal weights we use the R_0 fraction determined by the previous Belle analysis. The error on this assumption will also be taken into account in the systematic error. This is justified as the systematic error is expected to be small.

For completeness we will study in this chapter the distributions of both the $\theta_{\rm tr}$ and θ_1 angles, even though the CP-odd fraction will only be extracted from the $\theta_{\rm tr}$ distribution.

4.1.2 Signal probability density function

In theory $\cos\theta_{\rm tr}$ and $\cos\theta_1$ are distributed like $1-\cos^2\theta$ or $\cos^2\theta$ depending on the polarization state of the two D^* particles, but as mentioned before in reality these distributions are affected by the reconstruction performance and angular resolution. The reconstruction efficiency depends on the kinematics of the particle, i.e, the reconstruction of low-momentum particles is more challenging and will often result in a worse performance. To illustrate this effect a signal MC sample is generated without any polarization. The shape in $\cos\theta_{\rm tr}$ and $\cos\theta_1$ should be completely flat if the reconstruction efficiency is independent of the angles. But Figure 4.1 shows clearly that this is not the case. It can be seen that the reconstruction efficiency is lower for events with $\cos\theta_1 \sim -1$ as they correspond to events with a slow pions decaying in the opposite direction of the momentum of the boosted D^* mother particle, in the CM system of the other D^*, which means that they have a relatively lower momentum than the pions decaying in the direction of their mother D^*. In the $\cos\theta_{\rm tr}$ distribution we observe the expected symmetrical shape around $\cos\theta_{\rm tr} \sim 0$, as this angle is defined between the decay direction of the slow pion and the z-axis of the transversity basis, which is perpendicular to boost of the D^*.
To take into account the reconstruction efficiency dependence when defining the probability density function of the $\cos\theta_{\rm tr}$ and $\cos\theta_1$ distributions, three signal MC samples are generated (corresponding to experiments 07 - 49), each with one of the three polarizations of D^{*+} and D^{*-}. The distribution of $\cos\theta_{\rm tr}$ and $\cos\theta_1$ are then fitted with a fourth-order polynomial function.

$$\mathcal{P}(\cos\theta) = N\Big[1 + a\cos\theta + b\cos^2\theta + c\cos^3\theta + d\cos^4\theta\Big], \tag{4.1}$$

where N is a normalization factor. Because of the symmetry in $\cos\theta_{\rm tr}$, this distribution is modeled with an even function i.e, a and c are set to zero. The fitted parameters are a, b, c and d, are summarized in Table 4.3 and the fitted distributions are shown in Figure 4.2

		a	b	c	d	χ^2/ndf
A_0	$\cos\theta_{\rm tr}$		-0.70 ± 0.03		-0.19 ± 0.04	1.27
	$\cos\theta_1$	0.95 ± 0.53	45.81 ± 5.00	6.73 ± 1.20	-7.80 ± 2.14	0.92
$A_\perp$	$\cos\theta_{\rm tr}$		8.39 ± 0.67		1.69 ± 0.49	1.34
	$\cos\theta_1$	0.07 ± 0.03	-1.00 ± 0.02	-0.07 ± 0.04	0.04 ± 0.02	0.80
$A_\parallel$	$\cos\theta_{\rm tr}$		-0.69 ± 0.04		-0.02 ± 0.04	1.27
	$\cos\theta_1$	0.12 ± 0.03	-1.02 ± 0.02	-0.13 ± 0.03	0.05 ± 0.02	0.85

Table 4.3: Parameters of the $\cos\theta_{\rm tr}$ and the $\cos\theta_1$ distributions for different polarizations obtained from the fit on the signal MC.

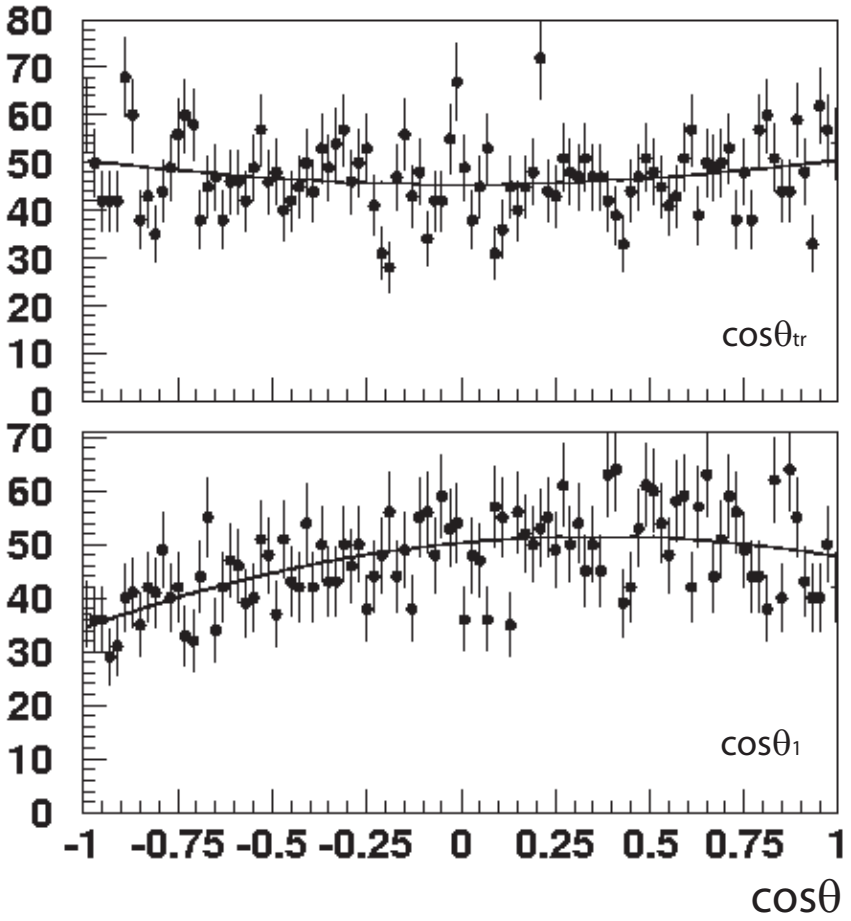


Figure 4.1: $\cos\theta_{\rm tr}$ (top) and $\cos\theta_1$ (bottom) distributions for a signal MC sample generated without any polarization.

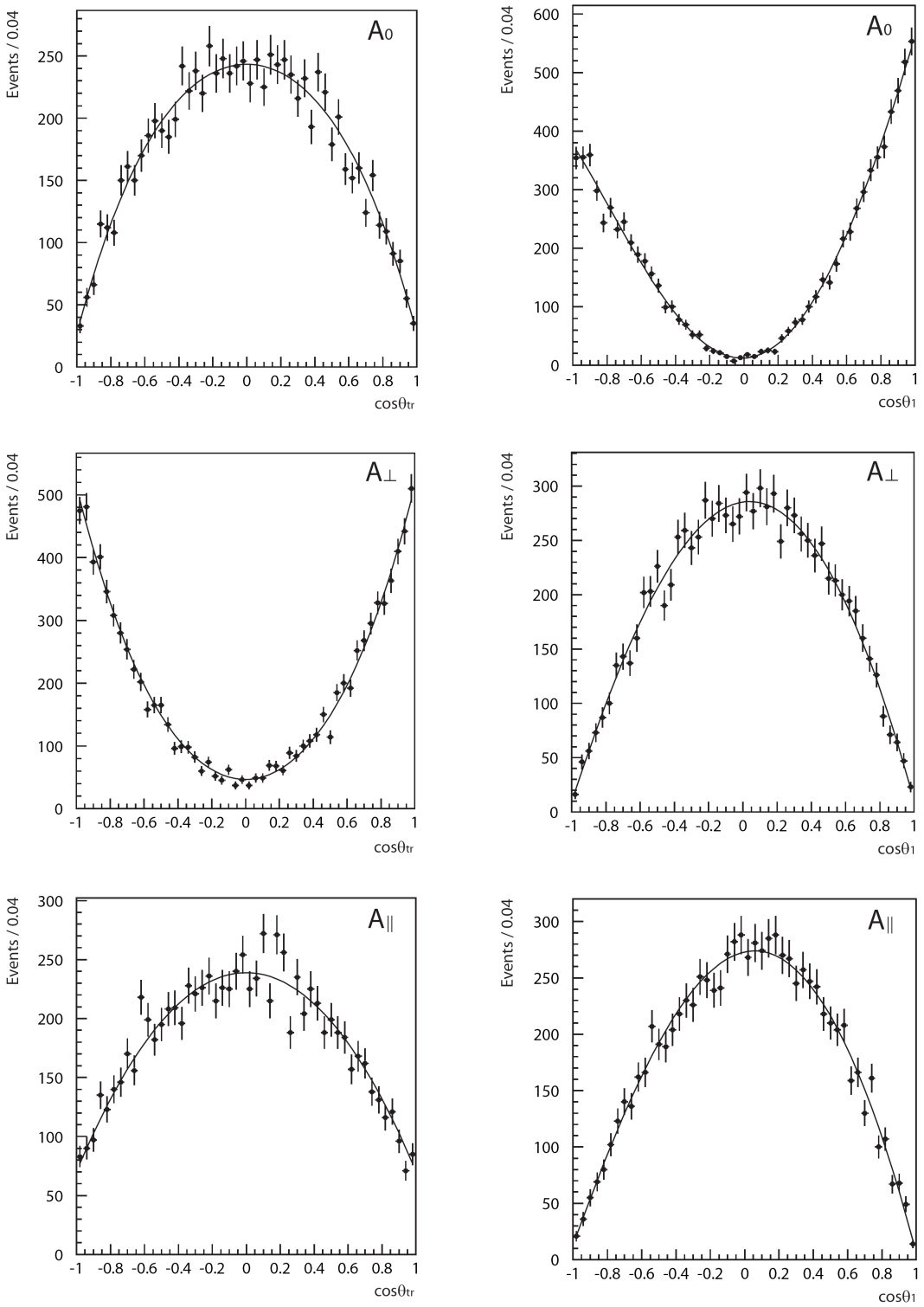


Figure 4.2: $\cos\theta_{\rm tr}$ and $\cos\theta_1$ distributions for signal MC files with different polarizations.

As only $R_\perp$ will be extracted from the $\cos\theta_{\rm tr}$ distribution we need to define the shape of the CP-even events using the A_0 and $A_\parallel$ distributions. The PDF of the CP-even events in $\cos\theta_{\rm tr}$ is modeled by adding the PDFs of the A_0 and $A_\parallel$ components as follows:

$$\mathcal{P}_{\rm even}(\cos\theta_{\rm tr}) = \frac{R_0}{R_0+R_\parallel}\mathcal{P}_0(\cos\theta_{\rm tr}) + \frac{R_\parallel}{R_0+R_\parallel}\mathcal{P}_\parallel(\cos\theta_{\rm tr}), \quad (4.2)$$

where $\mathcal{P}_0(\cos\theta_{\rm tr})$ is the shape of the A_0 component and $\mathcal{P}_\parallel$ is the shape of the $A_\parallel$ component, as obtained from the signal MC. The fraction of A_0 in the $\mathcal{P}_{\rm ev}(\cos\theta_{\rm tr})$ function is taken from the previous analysis [58], i.e,

$$\begin{aligned} R_0 &= 0.57 \pm 0.08, \\ R_\parallel &= 1 - R_0 - R_\perp, \\ &= 1 - (0.57 \pm 0.08) - (0.19 \pm 0.08), \\ &= 0.24 \pm 0.11, \\ \frac{R_0}{R_0+R_\parallel} &= 0.70 \pm 0.10, \end{aligned} \quad (4.3)$$

where the error on $\frac{R_0}{R_0+R_\parallel}$ is conservatively calculated assuming no correlation between R_0 and $R_\parallel$.

Not only the *s*hape of the $\cos\theta_{\rm tr}$ and $\cos\theta_1$ distributions are affected by the kinematics of the daughter particle, but also the absolute reconstruction efficiency ($= \frac{\text{number of reconstructed events}}{\text{number of generated events}}$) depends on the polarization. These dependencies are measured in Section 3.3.7. The relative efficiency reads:

$$\begin{aligned} \epsilon_0 &= 96.8 \pm 1.1\%, \\ \epsilon_\perp &= 102.9 \pm 1.0\%, \\ \epsilon_\parallel &= 100.2 \pm 1.1\%. \end{aligned} \quad (4.4)$$

Taking into account the relative reconstruction efficiency, the PDF of the CP-even events reads:

$$\mathcal{P}_{\rm ev}(\cos\theta_{\rm tr}) = \frac{R_0}{R_0+R_\parallel}\epsilon_0\mathcal{P}_0(\cos\theta_{\rm tr}) + \frac{R_\parallel}{R_0+R_\parallel}\epsilon_\parallel\mathcal{P}_\parallel(\cos\theta_{\rm tr}). \quad (4.5)$$

The relative reconstruction efficiency for the CP-even events, $\epsilon_{\rm ev}$ and the normalization factor, $N_{\rm eff}$, are defined as:

$$\epsilon_{\rm ev} = \frac{R_0}{R_0+R_\parallel}\epsilon_0 + \frac{R_\parallel}{R_0+R_\parallel}\epsilon_\parallel \quad (4.6)$$

$$N_{\rm eff} = R_\perp\epsilon_\perp + (1-R_\perp)\epsilon_{\rm ev}. \quad (4.7)$$

The total signal probability density function becomes:

$$\mathcal{P}_{\rm sig}(\cos\theta_{\rm tr}) = \frac{1}{N_{\rm eff}}\Big[R_\perp\epsilon_\perp\mathcal{P}_\perp(\cos\theta_{\rm tr}) + (1-R_\perp)\epsilon_{\rm ev}\mathcal{P}_{\rm ev}(\cos\theta_{\rm tr})\Big]. \quad (4.8)$$

With this function we will model the signal events on the data sample.

4.1.3 Background probability density function

The shape of the angular distribution of the background events can be determined beforehand from the sideband regions ($M_{\rm bc} < 5.27\,{\rm GeV/c^2}$ and $|\Delta E| > 40\,{\rm MeV}$) of the data if we assume that there is no peaking background i.e, that the behavior of background events in the sideband regions is the same as under the signal peak. The distributions are shown and fitted for information purposes only as the background shape will be extracted together with the CP-odd component in the final angular fit.

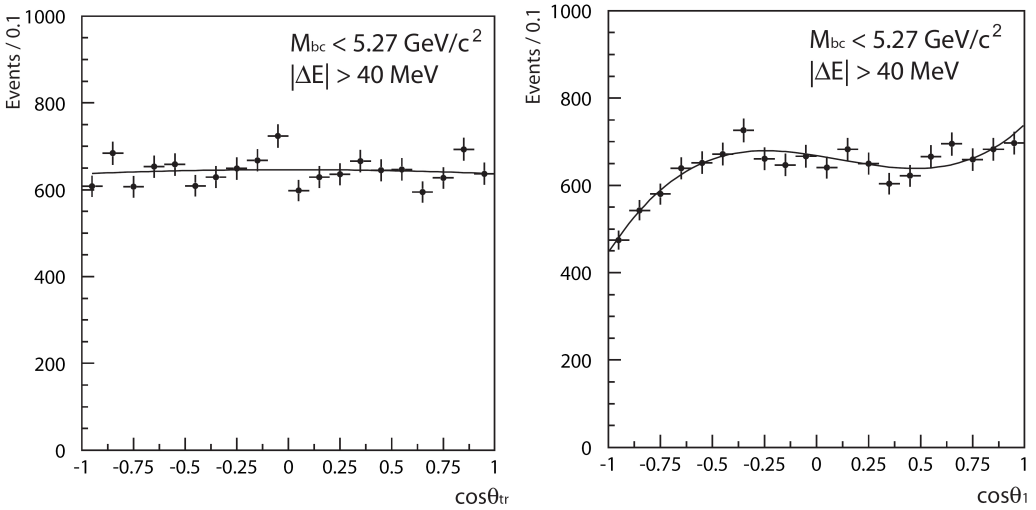


Figure 4.3: $\cos\theta_{\rm tr}$ and $\cos\theta_1$ distributions for the sideband regions of the data.

As shown in Figure 4.3, the $\cos\theta_{\rm tr}$ distribution of the sideband regions is fitted with a second-order polynomial while $\cos\theta_1$ is allowed to be an asymmetric third-order polynomial.

$$\mathcal{P}_{\rm bkg}(\cos\theta_{\rm tr}) \;=\; (1 + b_{\rm bkg}\cos^2\theta_{\rm tr}), \tag{4.9}$$

$$\mathcal{P}_{\rm bkg}(\cos\theta_1) \;=\; (1 + a_{\rm bkg}\cos\theta_1 + b_{\rm bkg}\cos^2\theta_1 + c_{\rm bkg}\cos^3\theta_1). \tag{4.10}$$

The fit on the sideband regions gives the following parameters:

	$\cos\theta_{\rm tr}$	$\cos\theta_1$
$a_{\rm bkg}$		-0.11 ± 0.04
$b_{\rm bkg}$	-0.02 ± 0.03	-0.11 ± 0.03
$c_{\rm bkg}$		0.33 ± 0.06
χ^2/ndf	1.76	0.97

Table 4.4: Fitted parameters of the $\cos\theta_{\rm tr}$ and $\cos\theta_1$ distributions for events in the sideband regions.

4.1.4 Background shape on the generic MC

Using the generic MC sample (as in Section 3.5) the shape of the background events under the signal peak can be studied. The black dots in Figure 4.4 represent the distribution of

$\cos\theta_{\rm tr}$ of the generic $B^0\overline{B^0}$ events in the small signal region. The distribution of the generic $B^0\overline{B^0}$ background events under the signal peak are shown in yellow. The generic B^+B^- event samples only contain background events and are represented by the green distribution. As can be seen from the figure, both the generated $B^0\overline{B^0}$ background as well as the B^+B^- have a flat behavior under the signal peak.

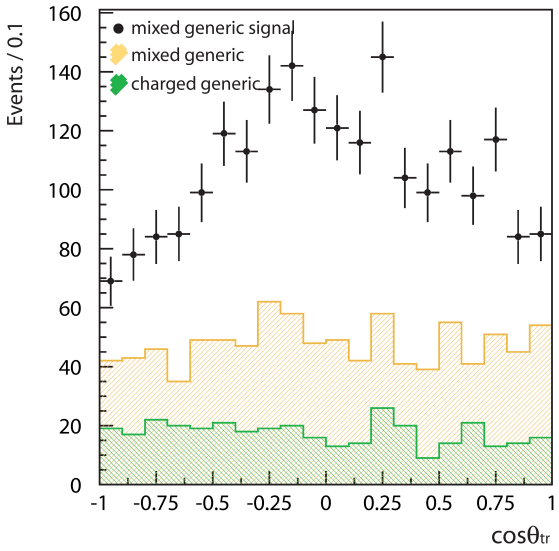


Figure 4.4: $\cos\theta_{\rm tr}$ distribution of the generic $B^0\overline{B^0}$ events superimposed on the generic $B^0\overline{B^0}$ and B^+B^- background events.

4.1.5 Total probability density function

The total probability density function used to extract the CP-odd fraction of the data can be chosen to model the distribution of $\cos\theta_{\rm tr}$ alone. This PDF will then consists of the signal probability density function, $\mathcal{P}_{\rm sig}$, obtained from the signal MC parametrization described by Eq. (4.8) and $\mathcal{P}_{\rm bkg}$, the background second-order polynomial function as shown in Eq. (4.9). The signal fraction, $f_{\rm sig}$ can be determined event-by-event from the $M_{\rm bc}$ and ΔE distribution. Another method is fitting the $M_{\rm bc}$ (and/or ΔE) distribution together with the $\cos\theta_{\rm tr}$ distribution, the signal fraction is then obtained during the fit. This method was used in the latest BaBar analysis. [63].

To study the performance of the fitting procedures a fast MC study is used, where to save computer processing power, we simulate only the distributions of the relevant parameters rather than simulating particle decays in the detector and the full detector response. We call such a simulation a fast MC. This technique will be used again to test the accuracy of other methods used in this analysis. The next paragraphs will explain the technical differences between the two fitting methods as well as the result from the fast MC.

4.1.5.1 Two dimensional binned fit on $\cos\theta_{\rm tr}$ and $M_{\rm bc}$

The two-dimensional PDF of $\cos\theta_{\rm tr}$ and $M_{\rm bc}$ has the following form:

$$\begin{array}{rcl} \mathcal{P}_{\rm tot}(M_{\rm bc},\cos\theta_{\rm tr}) & = & f_{\rm sig}\mathcal{P}_{\rm sig}(M_{\rm bc})\times\mathcal{P}_{\rm sig}(\cos\theta_{\rm tr}) \\ & + & (1-f_{\rm sig})\mathcal{P}_{\rm bkg}(M_{\rm bc})\times\mathcal{P}_{\rm bkg}(\cos\theta_{\rm tr}), \end{array} \tag{4.11}$$

where $\mathcal{P}_{\rm sig}(\cos\theta_{\rm tr})$ is defined in Eq. (4.8) and $\mathcal{P}_{\rm bkg}(\cos\theta_{\rm tr})$ in Eq. (4.9). As usual the functions described above are normalized over the area where they are fitted over. The PDF we use to describe the $M_{\rm bc}$ distribution is the same as the one described in the previous chapter.

$$\begin{array}{rcl} \mathcal{P}_{\rm sig}(M_{\rm bc}) & = & \mathcal{G}(M_{\rm bc},\mu,\sigma), \\ \mathcal{P}_{\rm bkg}(M_{\rm bc}) & = & {\rm ARGUS}(M_{\rm bc},s), \end{array} \tag{4.12}$$

The free parameters in the fit are the parameters describing the shape of the $M_{\rm bc}$ distribution (the mean and width of the Gaussian and the slope s, of the ARGUS function), the signal fraction, the background shape, $b_{\rm bkg}$ of $\cos\theta_{\rm tr}$ and of course the CP-odd fraction $R_\perp$. We chose to perform the two-dimensional fit on $\cos\theta_{\rm tr}$ and $M_{\rm bc}$ instead of ΔE to conform with the method described in [63]. Finally a three-dimensional fit could also have been performed, using the $\cos\theta_{\rm tr}$, $M_{\rm bc}$ and ΔE distribution. However performing a three-dimensional unbinned maximum likelihood is a procedure which becomes quite challenging, we will only investigate this solution if the performance of the other methods is unsatisfactory.

4.1.5.2 One dimensional unbinned ML fit on $\cos\theta_{\rm tr}$ using an event-by-event signal fraction from the two-dimensional $M_{\rm bc}$ and ΔE fit

While in the first method the signal fraction is determined during the fit, here we use the known shape of the $M_{\rm bc}$ and ΔE distribution which is obtained in the previous chapter. Even though the shape is not fitted again, the signal fraction is still calculated on an event-by-event basis depending on the value of $M_{\rm bc}$ and ΔE of each event. We use both $M_{\rm bc}$ and ΔE to determine the signal fraction of the event.

The signal fraction is thus for each event defined as:

$$f_{\rm sig}(M_{\rm bc},\Delta E) = \frac{\mathcal{P}_{\rm sig}(M_{\rm bc},\Delta E)}{\mathcal{P}_{\rm sig}(M_{\rm bc},\Delta E)+\mathcal{P}_{\rm bkg}(M_{\rm bc},\Delta E)}, \tag{4.13}$$

where $\mathcal{P}_{\rm sig}(M_{\rm bc},\Delta E)$ and $\mathcal{P}_{\rm bkg}(M_{\rm bc},\Delta E)$ are the fitted two-dimensional PDFs evaluated at the value of the event considered. The total probability density function reads:

$$\mathcal{P}_{\rm tot}(\cos\theta_{\rm tr}) = f_{sig}\mathcal{P}_{\rm sig}(\cos\theta_{\rm tr})+(1-f_{sig})\mathcal{P}_{\rm bkg}(\cos\theta_{\rm tr}), \tag{4.14}$$

where $\mathcal{P}_{\rm sig}(\cos\theta_{\rm tr})$ is defined in Eq. (4.8) and $\mathcal{P}_{\rm bkg}(\cos\theta_{\rm tr})$ in Eq. (4.9). Also here the background shape ($b_{\rm bkg}$) is left as a free parameter during the fit and is not fixed from the sideband regions. The other parameter that is left free during the fit is the CP-odd fraction, $R_\perp$.

Both fitting methods keep into account the relative differences of the reconstruction efficiency for each of the three polarizations by multiplying each function with its appropriate factor, given by Eq. (4.4).

4.1.5.3 Determining the preferred method using a fast MC

Different files are generated containing a large number of events using a fast MC technique. Each event is generated with a $M_{\rm bc}$, ΔE and $\cos\theta_{\rm tr}$ value. Each file contains a different CP-odd fraction. The parameters that are kept constant in the different files are:

- the fraction $\frac{R_0}{R_0+R_\parallel}$ which is fixed to the value obtained in the previous analysis shown in Eq. (4.3);
- the reconstruction efficiency for each polarization is generated according to Eq. (4.4);
- the background is generated according to the $\mathcal{P}_{\rm bkg}$ described in Eq. (4.9) with $b = -0.01$;
- the signal and background distributions of $M_{\rm bc}$ and ΔE are generated according to the distribution obtained on the data (see Section 3.7).

The generated files are fitted with both methods. To reduce the CPU time events are generated in the region $|\Delta E| < 40\,\mathrm{MeV}$. In the one-dimensional fit we require also $M_{\rm bc} > 5.27\,\mathrm{GeV/c^2}$ as the calculation of the event-by-event signal fraction using the ΔE and $M_{\rm bc}$ distribution still takes a lot of CPU time, as the files contain a large number of events. The two-dimensional fit is executed faster so that the $M_{\rm bc} > 5.2\,\mathrm{GeV/c^2}$ constraint is tight enough. The different cuts on $M_{\rm bc}$ should not affect the precision on the fitted $R_\perp$ value as it does not cut away signal events. The background shape however will be less precise in the one-dimensional fit than in the two-dimensional fit. The tight cuts on ΔE and $M_{\rm bc}$ are not applied on the data when performing the final fit, as the data sample is small enough.

generated			two dimensional fit		one dimensional fit	
sig-ev	bkg-ev	$R_\perp$ (%)gen	$R_\perp$ (%)	$\mathrm{a_{bkg}}$	$R_\perp$ (%)	$\mathrm{a_{bkg}}$
50k	125k	13.8	13.11 ± 0.40	-0.016 ± 0.022	13.26 ± 0.42	-0.038 ± 0.097
50k	125k	30.0	38.47 ± 7.20	-0.015 ± 0.007	$--$	$--$
50k	1250k	4.0	3.71 ± 0.47	-0.015 ± 0.007	3.51 ± 0.47	-0.024 ± 0.022
50k	1250k	13.8	13.34 ± 0.50	-0.016 ± 0.007	13.45 ± 0.51	-0.027 ± 0.022
50k	1250k	30.0	$--$	$--$	29.57 ± 0.56	-0.014 ± 0.022
50k	11250k	13.8	$--$	$--$	13.30 ± 0.84	-0.016 ± 0.006

Table 4.5: Fitted CP-odd and background shape parameter with the one-dimensional and two-dimensional methods.

Table 4.5 shows the fit results for both methods on the different MC files ($-$ denotes that this test is not performed). The result shows that the statistical errors of $R_\perp$ are not significantly smaller for any of the two methods. Therefore there is no statistical reason to decide on the best procedure. As in the one-dimensional unbinned maximum likelihood fit both the $M_{\rm bc}$ and ΔE information is used to determine the signal fraction, we have chosen this method to fit the data.

4.2 Fast MC and linearity test

To ascertain the accuracy of the fitting procedure a fast MC study is performed. This study is used for two different reasons. The first reason is the study of statistical fluctuations.

We generate many MC samples each containing the same number of signal events as in the data and with the same CP-odd fraction. When fitting these files one by one we can obtain the residual, pull and error distributions which give us information on how to interpret the result of the data. Secondly a linearity test can be performed. This is done by generating samples while varying the CP-odd fraction. Fitting these events will allow us to study a possible bias in the fitting procedure. The fast MC events are generated with the following characteristics:

- The ΔE and $\mathrm{M_{bc}}$ values of an event are generated according to the distribution obtained from the final fit of the data in the large signal region.

- The angular distribution of signal events is generated with the shape defined from the signal MC (see Section 3.8) such that detector smearing and reconstruction efficiency effects are taken into account.

- The angular distribution of background events is generated according to the shape of the data distribution in the sideband regions.

To study the residual, pull and error distributions of the fit, 1000 samples are generated with a CP-odd fraction of $R_\perp = 0.12$. Each file is fitted and the difference between the generated $R_\perp$ value and the fitted one is shown as a residual distribution in the left plot of Figure 4.5. This distribution should be centered around zero and the width is a measure of the statistical error to be expected on the fit parameter. When fitting this distribution with a Gaussian function we obtain the following result:

	residual
mean	0.001 ± 0.001
σ	0.043 ± 0.001.

The distribution is indeed centered at zero within errors. The expected statistical error is 0.043 ± 0.001.

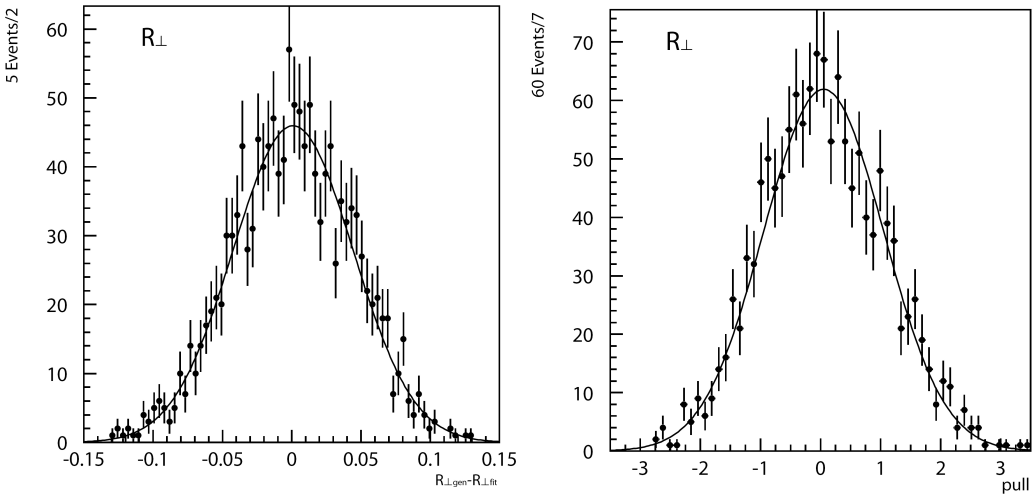


Figure 4.5: Left: the residual distribution of the toy MC. Right: the pull distribution.

The right plot of Figure 4.5 shows the pull distribution of the 1000 fitted samples. We define the pull as $\frac{R_\perp^{\rm gen}-R_\perp^{\rm fit}}{\rm error}$, where the error is that returned by the fit and $R_\perp^{\rm gen} = 0.12$ (note that the more conventional definition has an opposite sign as what we have used). The pull distribution is fitted with a Gaussian function. When there is no bias in the fitting procedure and the statistical errors are Gaussian distributed, the width of the distribution should be one and the mean value should be at zero. The obtained result is:

	pull
mean	0.052 ± 0.028
σ	1.000 ± 0.021

which is compatible with the expectations.

Finally we can study the error distribution, which is obtained from the statistical errors of the fitted results for each of the 1000 files. Figure 4.6 shows the error distribution fitted with a Gaussian function. The fit result reads:

	error
mean	0.0428 ± 0.0001
σ	0.0019 ± 0.0001

which means that the expected statistical error on the data is 0.0428 ± 0.0001. The error estimations also correspond to the width of the residual distributions, which means that the errors are correctly estimated.

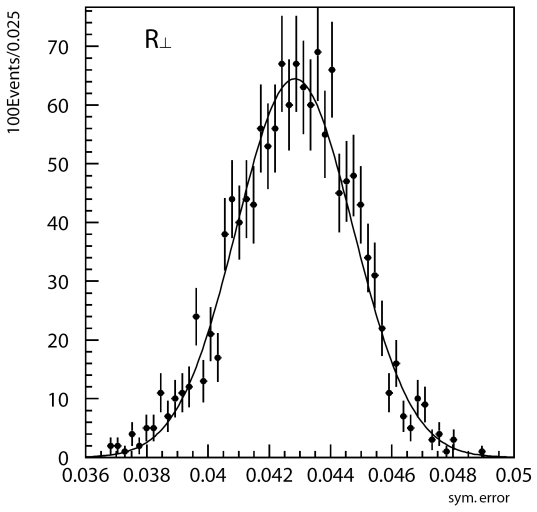


Figure 4.6: Error distribution of the fitted fast MC samples.

We can compare this result with the extrapolation of the statistical error of the CP-odd fraction of the previous analysis ($R_\perp^{\rm SVD1} = 0.19 \pm 0.08$) which was performed on 152×10^6 $B\overline{B}$ events. Without taking into account the dependence on the central value or any other possible changes in the reconstruction, we get:

$$\frac{\sqrt{152 \times 10^6} \times 0.08}{\sqrt{656 \times 10^6}} = 0.038. \tag{4.15}$$

The result of this rough test is close to the prediction of the more elaborate fast MC study.

A linearity test is performed to study the possible bias in the fitting procedure. For this study 1000 new files are generated, each containing the same number of signal events as on the data. The first 50 generated files have a CP-odd fraction of $R_\perp = 0$, but after every 50 generated files $R_\perp$ is increased with 0.05, such that the last 50 files are generated with $R_\perp = 0.95$. The files are then fitted one by one using the same procedure as on the data. The result is shown in Figure 4.7. The y value of each point in the figure represents the mean of the distribution of the fitted results of the 50 files corresponding to a same CP-odd fraction. The (vertical) error bars are too small to show up in the figure, but are obtained from the width of the distribution of these 50 files. The points in Figure 4.7 are then fitted with a first-order polynomial represented by the black line. For comparison a line with a slope equal to one and no offset is added in red to illustrate the situation without any bias. Visually one can already notice that the difference between the red and black line is very small. The result of the fit is shown in Table 4.6. The parameters are in close agreement with a situation with only a small bias. This effect is taken into account in the systematic uncertainty.

	linearity test
offset	-0.004 ± 0.002
slope	1.019 ± 0.005

Table 4.6: Fitted parameters of the linearity test.

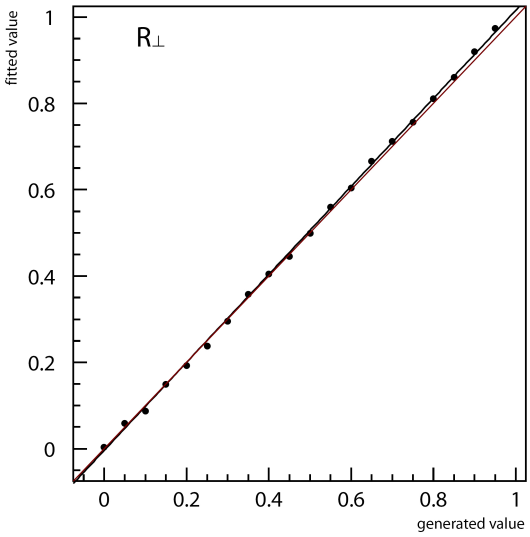


Figure 4.7: Linearity plot of $R_\perp$.

4.3 Extraction of the CP-odd fraction of the total data set

The extraction of the CP-odd fraction of the total data sample (experiments 07–55) is performed with an unbinned maximum likelihood fit on the $\cos\theta_{\rm tr}$ distribution, using an event-by-event signal fraction. The shape of the $M_{\rm bc}$ and ΔE distribution is obtained from

the final fit of the yield on the data (see Chapter 3.8). The background shape, $b_{\rm bkg}$ and $R_\perp$ are the two free parameters to be determined. The fit is performed over the large signal region and the result of the fit is shown in Figure 4.8. For visualization purposes only events in the small signal region ($5.27\,\mathrm{GeV/c^2} < \mathrm{M_{bc}} < 5.3\,\mathrm{GeV/c^2}$ and $|\Delta E| < 40$ MeV) are plotted. The solid line represents the total fitted function, the dotted gray line shows the background contribution. The CP-even and CP-odd contributions are indicated above the background with the dot-dashed blue and dotted red lines respectively.

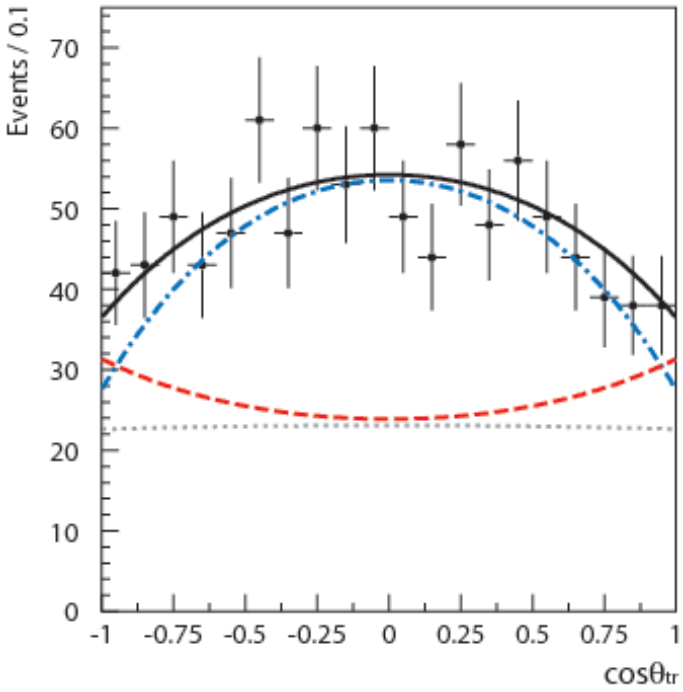


Figure 4.8: Fitted $\cos\theta_{\rm tr}$ distribution in the small signal region.

The fit gave the following result:

$$R_\perp = 0.125 \pm 0.043, \tag{4.16}$$
$$b_{\rm bkg} = -0.020 \pm 0.044. \tag{4.17}$$
$$\tag{4.18}$$

The extracted background shape is in agreement with the shape of the events in the side-band regions (see Section 4.1.3) and the generic MC shapes. The statistical errors are compatible with the expectation of the fast MC.

4.4 Systematic study

Systematic errors are biases in measurements which lead to measured values being systematically too high or too low. A systematic error is any biasing effect in the environment, methods of observation, or instruments used, which introduces an error in an experiment and is such that it always affects the results of an experiment in the same direction. The systematic error of the CP-odd fraction in this analysis is estimated by the deviation of the

$R_\perp$ measurement when varying the fixed parameters of the PDF within their one standard deviation error. The different deviations are then summed in quadrature to obtain the total systematic uncertainty on the CP-odd fraction of the data.

Table 4.7 shows the different sources of systematic errors on the polarization fraction. The systematic error due to the signal purity is obtained by changing the obtained signal yield by one sigma (see Chapter 3). Due to the different normalizations of the CP-odd and CP-even distribution, adding or removing one sigma to the signal yield should have the opposite effect on the fitted CP-odd fraction because the background shape is almost flat. This is what is observed, the averaged absolute value of the deviation reads 0.003. The reconstruction efficiency of the events with a perpendicular polarization is varied with one sigma and the deviation reads 0.003. The linearity test for the polarization fit is used to include a systematic error due to any bias in the fitting procedure. We refer to the work presented in Section 4.2. For $R_\perp = 0.125$ the linearity fit showed that a 0.002 deviation is obtained. Finally when adding or subtracting one standard deviation of the fraction of $R_0/(R_0 + R_\parallel) = 0.7037 \pm 0.1207$, a deviation of 0.009 is obtained. In the next chapter it will be shown that to perform a lifetime or CP fit on the data, tighter cuts on the reconstructed vertices are necessary. An overview of these vertex cuts is given in Section 5.1. When performing the angular analysis on the data sample which is subjected to the tighter vertex requirements a CP-odd fraction difference of 0.013 is obtained. Finally a peaking background contribution of 6.6% is added to the signal PDF to which conservatively a CP-odd behavior is given. The one-sided difference in central value is -0.016 but for elegance we will treat this as a symmetric, two-sided error.

The total systematic error for the polarization fraction yields: 0.023. As the shape of the background distribution is fitted during the polarization fit, no more contribution to the systematic error has to be included: the uncertainty in the background shape is included in the statistical error.

source	deviation
signal purity	0.0027
$\epsilon_\perp$	0.0031
fit bias	0.0016
$R_0/(R_0 + R_\parallel)$	0.0086
vertex cuts	0.0132
peaking background	0.0160
total	0.0228

Table 4.7: Different sources of the systematic uncertainties in the angular analysis.

4.5 Conclusion

We have measured the CP-odd fraction of the $B^0 \to D^{*+}D^{*-}$ decay in the Belle data sample of 657×10^6 $B\overline{B}$ events and found:

$$R_\perp \quad = \quad 0.125 \pm 0.043(\text{stat}) \pm 0.023(\text{syst}).$$

This result is consistent with the previous Belle result [58] as well as the BaBar result [63]. Figure 4.9 shows the BaBar result as of Spring 2008, the preliminary Belle result (as

shown at conferences during Spring 2008), and the average result obtained by the HFAG group [72]. The BaBar measurement is performed on 617 ± 33 signal events. It can be seen that both results are well compatible with each other.

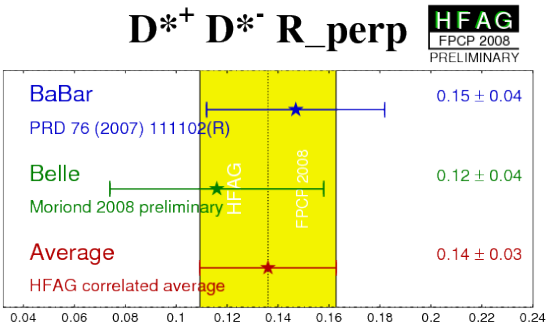
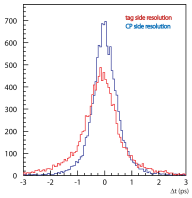


Figure 4.9: The BaBar and Belle preliminary results of the CP-odd fraction as of Spring 2008. The bottom line shows the average of both results.

Chapter 5

Lifetime measurement from the proper-time distribution



In this chapter we perform a fit to the proper-time difference of the neutral B meson decaying in the $D^{*+}D^{*-}$ mode and the other B meson in the event. By comparing the extracted B^0 lifetime to the precise world average we can test the accuracy of the resolution function and the background model.

We study in this chapter the proper-time distribution of the B^0 meson and extract its lifetime, τ_{B^0}. The proper time is obtained from the difference in z position of the vertex of the B meson which decays to $D^{*+}D^{*-}$ and the other B meson in the event. The difference in vertex positions is then converted to a proper-time difference according to:

$$\Delta t \simeq \frac{(z^{CP} - z^{\rm tag})}{\beta\gamma c}, \tag{5.1}$$

where $\beta\gamma$ the boost is of the $\Upsilon(4S)$ system, z^{CP} is the z coordinate of the vertex of the B meson reconstructed through $D^{*+}D^{*-}$, and $z^{\rm tag}$ refers to the other B meson called the $B^{\rm tag}$ meson. We will in the next chapters often refer to this proper-time *difference* as just the proper-time. Note that Δt can be positive or negative. The $|\Delta t|$ distribution can then be fitted with an exponential function to obtain the lifetime. The average Δz at KEKB is $c\tau_B(\beta\gamma)_\Upsilon \simeq 200\,\mu$m, where τ_B is the B-meson lifetime. However, the reconstructed z coordinate of the vertex is smeared by the detector resolution, the kinematic approximations and the non-primary vertices. To extract the true Δt distribution from the measured Δz we need to unfold the vertex detector resolution and (possible) biases in the measurement of Δz. The precise determination of the resolution function, $\mathcal{R}$, is essential as the vertex resolution is of the same order as the average Δz. The PDF function which describes the proper-time distribution of the signal events is obtained from a convolution of the theoretical Δt distribution with the resolution function:

$$\mathcal{P}(\Delta t) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{2\tau_{B^0}} \otimes \mathcal{R}. \tag{5.2}$$

We also show the determination of a background model for the Δt distribution using the events in the sideband region.

The lifetime of the B meson is already well established by previous measurements and the world average value, obtained by HFAG, reads $\tau_{\mathrm{B}^0} = 1.530 \pm 0.009\,\mathrm{ps}$ [72]. Due to the smaller statistics we cannot obtain an equally precise measurement, which is therefore also not the goal of this chapter. But by comparing the measured lifetime with the world average value we can test if the model for the resolution function and the background shape describes accurately our data. This is an important check as both functions will be used in the CP analysis as well.

The parameters in the resolution function are determined using a high-statistics sample of semi-leptonic and hadronic $b \to c$ decays [56]. These parameters are referred to as the "standard" Belle resolution parameters and are used in many analyses. However the topology of the $B^0 \to D^{*+}D^{*-}$ decay contains not only more charged tracks on average per event than the modes used to determine the standard resolution parameters but in the $B^0 \to D^{*+}D^{*-}$ vertex reconstruction there are no tracks used which originates directly from the B decay vertex, as the D^0 particles have a non-negligible lifetime. In Section 5.3.2 and Section 5.5, it will be shown that these standard parameters are not appropriate to describe the proper-time resolution of this analysis. We will therefore use a control sample to determine the main resolution parameters ourselves. Using these, we will then perform a fit to the proper-time distribution of the $B^0 \to D^{*+}D^{*-}$ data to obtain a measurement of the B^0 lifetime. As the lifetime measurement itself is not a main result in this analysis, there is no systematic study performed.

5.1 Vertex reconstruction

We will describe in this section how the vertex positions of the B^{CP} and B^{tag} are obtained. Each of the three (four) detection layers in the SVD1 (SVD2) detector can provide useful information about the tracks used to determine the vertex position. Reconstructed tracks which are associated to measurements in many layers are therefore more accurate than tracks obtained from only a few hits in the silicon layers. The tracks used to determine the vertex position are subject to more stringent requirements than the selection criteria of Chapter 3. Due to limited angular coverage, detector noise, and mistakes in the pattern recognition we cannot require all tracks to have a large minimal number of associated SVD hits as that would reduce the statistics too much. Even the standard criteria used in Belle analyses have been shown to be too strict for an analysis with as many charged tracks as this one. A precise study was performed in the $B^0 \to D^+D^-$ analysis [73] and showed that due to the many charged tracks in the final state, it is possible to loosen the requirements on the SVD hits. We therefore require the same criteria as have been proved to be optimal in the $B^0 \to D^+D^-$ analysis. At least one of the D mesons decays must have two tracks or more with more than one SVD hit in the $r\phi$ view and more than two SVD hits in the rz view.

The vertex position of the B^{CP} meson is determined from the IP profile and the reconstructed D meson tracks. No information from the slow pions or D^* meson is used. This is possible because the two D mesons originate at the B decay point. The slow pion tracks are generally of a poor quality due to worse performance of the tracking efficiency for low momentum particles.

The two reconstructed D meson tracks are fitted to a common vertex, which is obtained by finding the minimum of the χ^2 function, where $\chi^2 = \sum \Delta x^T V^{-1} \Delta x$. Δx denotes the difference between the updated and original parameters describing the particle track and V is the covariance matrix. The sum in the previous equation is taken over the two D

mesons.

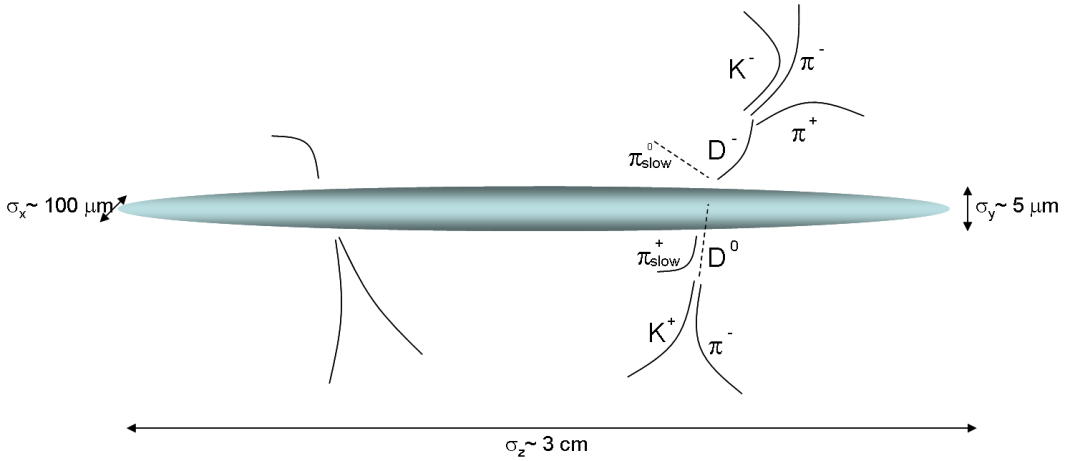


Figure 5.1: Schematic view of the IP-profile with the tracks of a B^0 (right) and an $\overline{B^0}$ decay (left).

The obtained vertex position is further constraint by fitting it to the interaction region, where the bunches from the LER and HER cross each other. Figure 5.1 shows a schematic view of the interaction point (IP) profile, which can be parametrized by a three-dimensional Gaussian. Fitting the obtained B vertex to the IP profile improves the z resolution of the vertex from $182\,\mu\mathrm{m}$ to $148\,\mu\mathrm{m}$. The IP profile is determined for every 10000 collected events, by extrapolating all the tracks in an hadronic event to a common region. In the x and y direction the obtained region is fitted with a double Gaussian while in the z direction a single Gaussian distribution is used. The beam-monitor detector from the accelerator gives additional information on the bunches. The resolution of the IP profile in y is fully determined by the accelerator information which has an accuracy of approximately $2\,\mu\mathrm{m}$ while the vertex fit only gives an accuracy of $70\,\mu\mathrm{m}$ in y. However in the z direction the vertex-fit gives a much better resolution than the beam-monitor. The obtained IP profile has typically the following dimensions: $\sigma_x = 100\,\mu$m, $\sigma_y = 5\,\mu$m and $\sigma_z = 3\,\mathrm{cm}$. This profile is furthermore smeared in the $r\phi$ plane to incorporate the transverse decay length of the B meson. The smearing is $21\,\mu\mathrm{m}$, corresponding to the width of the transverse proper-time distribution when it is approximated with a Gaussian distribution. For single-track vertices of the B^{tag}, it is found that when fitting the track to this IP profile a bias in the z position of the vertex is obtained. A study [74] showed that when replacing the IP profile by an "IP tube" this bias disappears. The IP tube is a virtual straight tube in the z direction which contains the IP profile. The IP tube has thus the same dimensions as the IP profile except in z where there is no boundary.

The fit of the vertices to the IP tube gives an additional term $\Delta r^T V_{\mathrm{IP}}^{-1} \Delta r$ to the χ^2 function, where Δr is the distance between the vertex and the IP tube and V_{IP} is the covariance matrix describing the IP tube. The quality of the fit can be measured by the obtained reduced χ^2 value (χ^2/ndf). Events with a poor fit quality are discarded from the lifetime and CP analysis. We therefore obtain the variable $\chi^2_{\mathrm{w/oIP}}/ndf$, which is the χ^2 calculated without the term related to the IP constraint, such that events with vertices that are close to the interaction region don't get enhanced. We require that events need to have a vertex fit with $\chi^2/ndf < 250$. Events for which the fit did not converge, where no

minimum is obtained, are discarded as well.

The decay vertex of the $B^{\rm tag}$ is determined from all the tracks which are not assigned to the B^{CP} decay. As for the CP side, tracks of bad quality are discarded and further requirements on the SVD hits per track are applied as well. The remaining tracks are fitted to a common vertex using the IP tube constraint. If the reduced χ^2 of the associated vertex fit exceeds 20, the track with the largest χ^2 contribution is removed and the vertex is refitted. If the track to be removed is a high-momentum lepton then it is nevertheless kept as it is likely to come from primary semi-leptonic B decays and the track with the second larges χ^2 is removed instead. This procedure is repeated until $\chi^2/ndf < 20$ or only one track is left.

The fitting range of the proper-time is $|\Delta t| < 70\,{\rm ps}$, which corresponds to approximately 45 times the B lifetime, such that it does not discard important events. The fitting functions described in this and the following chapter are then all normalized to one over the fitting range :

$$\int_{-T}^{T} \frac{d\mathcal{P}}{dt}(\Delta t) = 1,$$

where $T = 70\,{\rm ps}$.

The distributions of the vertex resolutions obtained in this analysis are shown for the signal MC in Figure 5.2 and for the data in the small signal region in Figure 5.3, separately for events recorded with the SVD1 and SVD2 detector configuration. A better vertex resolution is obtained for SVD2 as compared to SVD1, both on data and on MC. Furthermore the width of the distributions of the tag-side vertex resolution is larger than for the CP side. The shape of the MC distributions agrees well with the ones observed in data. The CP side vertex resolution is on average $\sim 75\,\mu{\rm m}$ and the average reconstruction efficiency is approximately $\sim 95\%$. For the tag side the average reconstruction efficiency is $\sim 93\%$ and the vertex resolution is $\sim 140\,\mu{\rm m}$.

The extra requirements on the events reduce the number of signal events in the data by 7.6% (from 554 ± 30 signal events to 511 ± 28 signal events) while the background decreases by approximately 20% in the large signal region.

5.2 Resolution function parameters

As explained in Section 2.7.5 the proper-time resolution function of the signal events, $\mathcal{R}(\Delta t)$, includes effects arising from the detector resolution (both for the CP side and tag side), the kinematic approximation and the parametrization of the smearing of the vertex due to non primary vertices (these are vertices obtained from tracks which are not coming directly from charged decay particles of the B meson):

$$\begin{aligned}\mathcal{R}_{\rm sig}(\Delta t) &= \int\int\int_{-\infty}^{\infty} \mathcal{R}_{\rm det}^{CP}(\Delta t - \Delta t')\mathcal{R}_{\rm det}^{\rm tag}(\Delta t' - \Delta t'') \\ &\quad \mathcal{R}_{\rm np}(\Delta t'' - \Delta t''')\mathcal{R}_{\rm kin}(\Delta t) d\Delta t' d\Delta t'' d\Delta t'''. \qquad (5.3)\end{aligned}$$

In the next paragraphs we will explain in more detail each of the contributions to the resolution as they play an important role in the accuracy of the lifetime and CP fit. In total 27 free parameters need to be defined to accurately describe this resolution function. The standard Belle resolution parameters are obtained by a study group and not by ourselves.

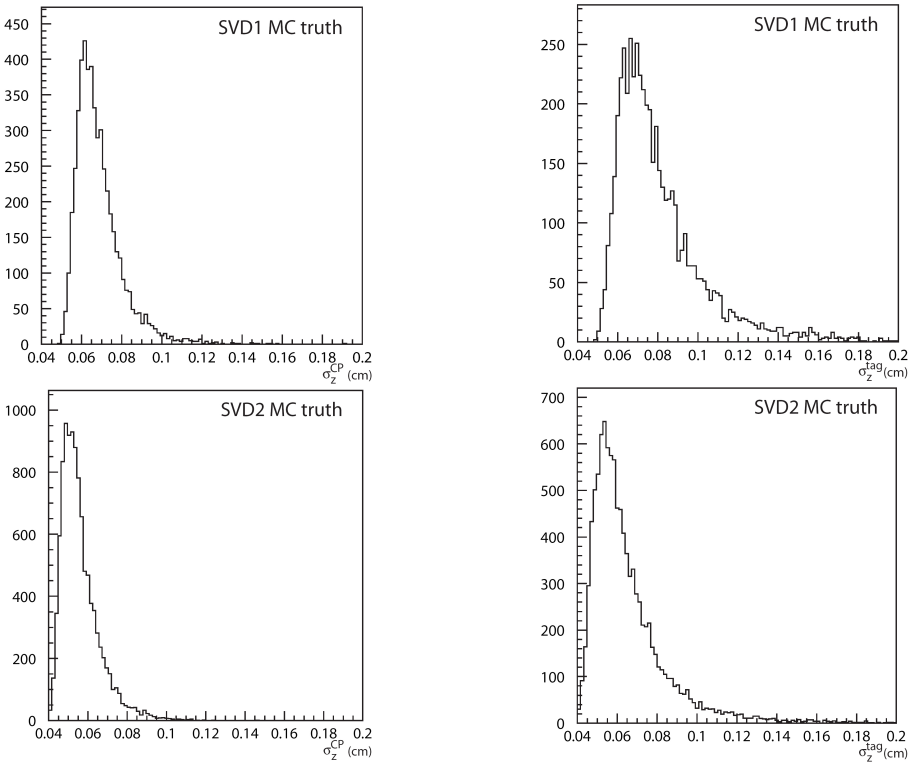


Figure 5.2: CP-side (left) and tag-side (right) vertex resolution, $\sigma_z^{CP,\mathrm{tag}}$, for true signal decays in MC samples, generated with SVD1 and SVD2 configurations.

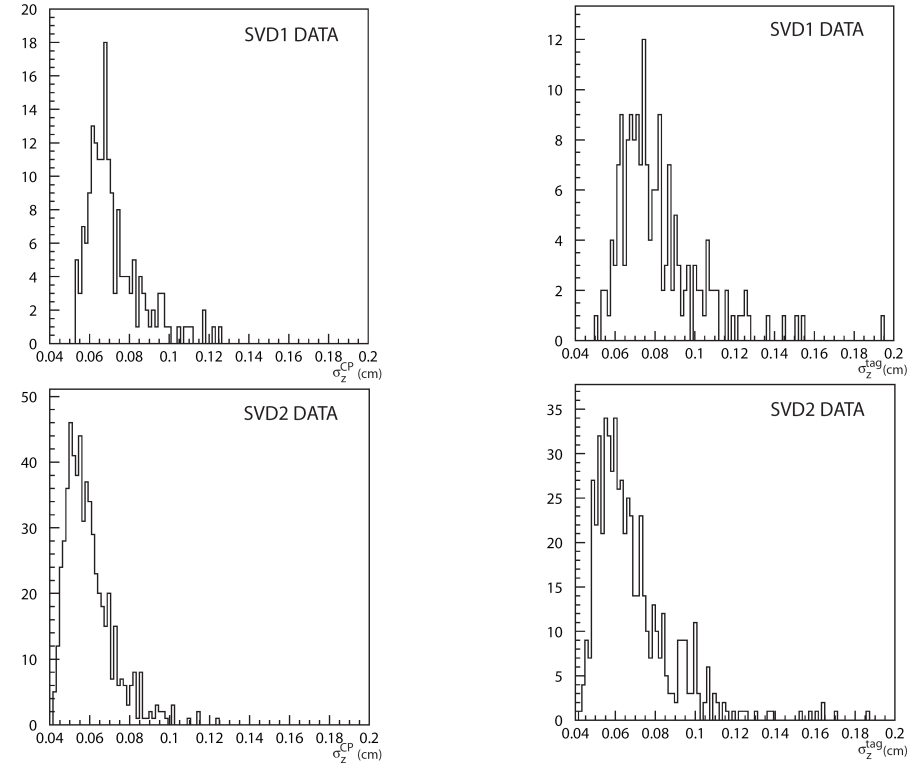


Figure 5.3: CP-side (left) and tag-side (right) vertex resolution, $\sigma_z^{CP,\mathrm{tag}}$, for SVD1 and SVD2 data in the small signal region.

We will briefly describe how these parameters are determined. For a detailed description of this study we refer to [56].

The following B decays are reconstructed: $\overline{B^0} \to J/\psi K^0_S$, $\overline{B^0} \to J/\psi \overline{K}^{*0}$, $\overline{B^0} \to D^{(*)+}\pi^-$, $\overline{B^0} \to D^{*+}\rho^-$, $B^- \to J/\psi K^-$, $B^- \to D^0\pi^-$ and $\overline{B^0} \to D^{*+}\ell^-\nu$, where $J/\psi \to \ell^+\ell^-$ ($\ell = e, \mu$) and $D^{*+} \to D^0\pi^+$. The functional forms of $\mathcal{R}_{\rm sig}$ are determined from a detailed MC simulation study. The obtained parameters have then been checked by performing lifetime fits on control samples.

A different set of parameters is also obtained for MC events. Both for the data and the MC parameters a distinction is made between SVD1 and SVD2 data. The overall resolution is estimated as the rms of the resolution function for data events and corresponds approximately to $1.4\,\mathrm{ps}$.

5.2.1 The kinematic correction

As said before we obtain the proper-time difference Δt from the difference in decay distances of the two B mesons and the boost of the $\Upsilon(4S)$ system. However by doing this we ignore the fact that the B mesons still has a small non-zero momentum in the center-of-mass system of the $\Upsilon(4S)$. The relationship between the boost of the B^{CP} meson in the lab frame, $(\beta\gamma)_{CP}$, and the boost of the $\Upsilon(4S)$ particle, $(\beta\gamma)_\Upsilon$, is given by:

$$\begin{aligned}(\beta\gamma)_{CP} &= \frac{(\beta\gamma)_\Upsilon E_B^{\rm cms} + \gamma_\Upsilon p_B^{\rm cms}\cos\theta_B^{\rm cms}}{m_B} \\ &= (\beta\gamma)_\Upsilon\Big(\frac{E_B^{\rm cms}}{m_B} + \frac{p_B^{\rm cms}\cos\theta_B^{\rm cms}}{\beta_\Upsilon m_B}\Big) \\ &\equiv (\beta\gamma)_\Upsilon(a_k + c_k). \end{aligned} \tag{5.4}$$

The same reasoning can be applied to the associated B meson. We obtain:

$$\begin{aligned}(\beta\gamma)_{\rm tag} &= (\beta\gamma)_\Upsilon\Big(\frac{E_B^{\rm cms}}{m_B} - \frac{p_B^{\rm cms}\cos\theta_B^{\rm cms}}{\beta_\Upsilon m_B}\Big) \\ &\equiv (\beta\gamma)_\Upsilon(a_k - c_k), \end{aligned} \tag{5.5}$$

where $E_B^{\rm cms} \approx 5.292\,\mathrm{GeV}$, $p_B^{\rm cms} \approx 0.340\,\mathrm{GeV/c}$ and $\cos\theta_B^{\rm cms}$ are the energy, the momentum and the angle with respect to the beam direction for the B^{CP} meson in the $\Upsilon(4S)$ rest frame and m_B is the mass of the neutral B meson. The velocity of the $\Upsilon(4S)$ is $\beta_\Upsilon = 0.391\,c$. The proper-time difference should be:

$$\Delta t^{\rm true} = z_{CP}/c(\beta\gamma)_{CP} - z_{\rm tag}/c(\beta\gamma)_{\rm tag},$$

where we call $\Delta t^{\rm true}$ the true the proper-time difference, which is actually only correct when the transverse momentum of the B mesons is neglected. However the proper-time difference used in this analysis (Δt) has a much larger approximation:

$$\begin{aligned}\Delta t &= \Delta z/c(\beta\gamma)_\Upsilon \\ &= (a_k + c_k)t_{CP}^{\rm true} - (a_k - c_k)t_{\rm tag}^{\rm true} \\ &= a_k\Delta t_{\rm true} + c_k(t_{CP} + t_{\rm tag})\end{aligned}$$

which gives a difference of:

$$\begin{aligned}\Delta t - \Delta t_{\rm true} &= (a_k - 1)\Delta t_{\rm true} + c_k(t_{CP} + t_{\rm tag}) \\ \overset{a_k\simeq 1, c_k\simeq 0.165\cos\theta}{\Longrightarrow} &\simeq 0.165\cos\theta_B^{\rm cms}(t_{CP} + t_{\rm tag}).\end{aligned} \tag{5.6}$$

This smearing effect can be analytically calculated once the angle θ of the B mesons is known. The kinematic correction therefore does not give rise to additional resolution parameters.

5.2.2 The detector resolution function

The detector resolution function, $\mathcal{R}_{\rm det}$ describes the effect of the detector response on the reconstructed vertex. The function consists of two parts:

$$\mathcal{R}_{\rm det} = \mathcal{R}^{CP}_{\rm det} \otimes \mathcal{R}^{\rm tag}_{\rm det}, \tag{5.7}$$

where the first (second) term is the detector resolution of the CP-side (tag-side) B vertex. Both sides are parametrized with the same functional form, but with different parameters for the CP side and the tag side. A study of the high statistics MC sample described in Section 5.2 is used to obtain the parameters of the detector resolution function. In the MC simulation the short-lived ($\tau < 10^{-9}\,\mathrm{s}$) secondary particles are forced to decay with zero lifetime at the B meson decay point, such that the smearing effect from non-primary particles is excluded. The residual vertex distribution i.e, the distribution of the difference between the reconstructed and the generated vertex positions is studied to define the shape of the detector resolution function. The residual distributions of the fully reconstructed B^{CP} and the associated $B^{\rm tag}$ meson are defined as:

$$\begin{aligned} \delta z_{CP} &\equiv z^{\rm rec}_{CP} - z^{\rm gen}_{CP}, \\ \delta z_{\rm tag} &\equiv z^{\rm rec}_{\rm tag} - z^{\rm gen}_{\rm tag}. \end{aligned}$$

A study [75, 76] showed that the residual distributions are best parametrized with a double Gaussian function centered at zero and with an event-by-event width proportional to the error estimate of the vertex z-coordinate, $\sigma_{\rm vtx}$. Furthermore a dependence on ξ^2 is observed, where ξ^2 is a χ^2-like measure of the quality of the vertex fit. A MC study had shown that the fit χ^2 is correlated with the B decay length due to the tight IP constraint in the transverse plane. To avoid this correlation, the variable ξ^2 is introduced, which uses the z information only. This goodness-of-fit is defined as the reduced χ^2 (χ^2 divided by the number of degrees of freedom) projected onto the beam direction:

$$\xi \equiv (1/2n) \sum_{1}^{n} [(z^i_{\rm after} - z^i_{\rm before})/\epsilon^i_{\rm before}]^2, \tag{5.8}$$

where n is the number of tracks used in the fit (and $2n$ the number of degrees of freedom), $z^i_{\rm before}$ and $z^i_{\rm after}$ are the z positions of each track at the closest approach to the origin, before and after the vertex fit respectively, $\epsilon^i_{\rm before}$ is the error of $z^i_{\rm before}$. The shape of the residual reads:

$$\begin{aligned} \mathcal{R}_{\rm det}(\delta z) &= (1 - f_{\rm tail})\mathcal{G}(\delta z, 0, (s_{\rm main} + \xi s_{\rm tail})\sigma_{\rm vtx}) \\ &+ f_{\rm tail}\,\mathcal{G}(\delta z, 0, s_{\rm tlml}(s_{\rm main} + \xi s_{\rm tail})\sigma_{\rm vtx}), \end{aligned} \tag{5.9}$$

where $\mathcal{G}(x, \mu, \sigma)$ is a Gaussian distribution of the variable x, with mean μ and width σ and $\sigma_{\rm vtx}$ is the vertex resolution of the B^{CP} or $B^{\rm tag}$ for $\mathcal{R}^{CP}$ or $\mathcal{R}^{\rm tag}$ respectively. The four parameters that determine this distribution are marked in red. Both the B^{CP} vertex and the tag-side vertex are fitted with this distribution. However for events where the $B^{\rm tag}$ is only reconstructed with one track no ξ dependence can be obtained as there is no

goodness of fit defined for a single-track vertex. The single-track vertex is defined by the interaction point between the IP tube and the track. The detector resolution function for such events is given by Eq. (5.9) where ξ is set to zero. The standard detector resolution parameters for the B^{CP} vertex of the data and MC events are summarized in Table 5.1.

SVD1

parameter	data	MC
$s_{\rm main}$	$0.91^{+0.46}_{-0.80}$	1.09
$s_{\rm tail}$	$2.37^{+0.43}_{-0.39} \times 10^{-2}$	1.36×10^{-2}
$f_{\rm tail}$	0.0	4.47×10^{-4}
$s_{\rm tlml}$	1.0	6.65

SVD2

parameter	data	MC
$s_{\rm main}$	$0.66^{+0.45}_{-0.09}$	0.69
$s_{\rm tail}$	$3.50^{-0.85}_{-0.60} \times 10^{-2}$	2.95×10^{-2}
$f_{\rm tail}$	$1.02^{+0.04}_{-0.05} \times 10^{-1}$	9.98×10^{-2}
$s_{\rm tlml}$	$4.75^{+0.49}_{-1.03}$	5.57

Table 5.1: Standard ICHEP06 detector resolution parameters for the B^{CP} vertex of SVD1 and SVD2 data and MC.

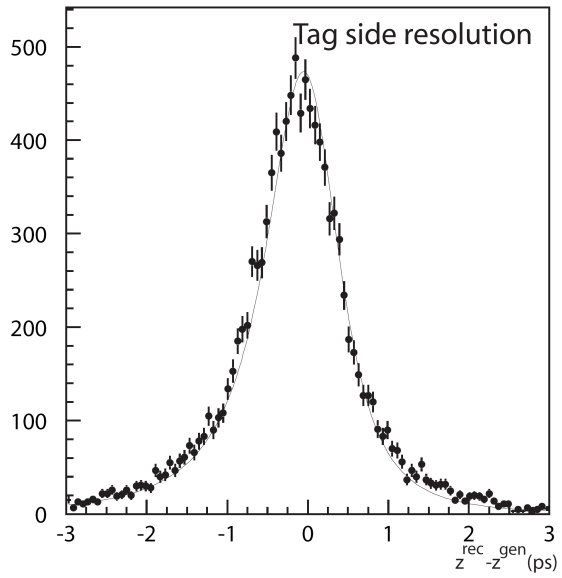


Figure 5.4: $B^{\rm tag}$ vertex residual distribution obtained in $B^0 \to D^{*+}D^{*-}$ MC events with the standard tag-side detector resolution function superimposed.

There are no analysis-specific constraints applied to the reconstruction of the tag-side B-vertex, therefore the standard parametrization of $\mathcal{R}_{\rm tag}$ should be adequate to describe the tag-side vertex resolution in this analysis as well. Figure 5.4 shows the standard resolution function for the $B^{\rm tag}$ vertex, drawn over its residual distribution and it looks satisfactory. As shown in Figure 5.5, the tag-side resolution is worse than the CP-side resolution,

which means that the tag-side is the limiting factor on the quality of the total detector resolution.

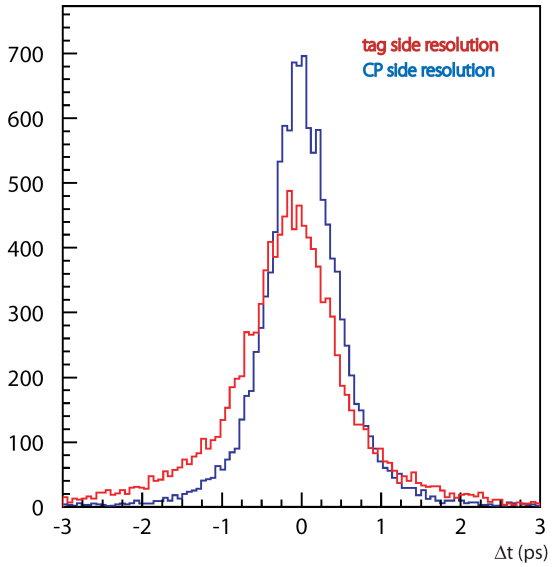


Figure 5.5: Tag-side (red) and CP-side (blue) vertex resolution distribution.

There still exist a very long tail in the residual and proper-time distributions that cannot be described by the detector resolution function. An additional outlier term is introduced to parametrize these events. Due to the very large width ($\sim 40\,\mathrm{ps}$) this function is not convoluted with the lifetime distribution or other resolution components, but it is kept as an independent term, defined as:

$$P_{\mathrm{ol}}(\Delta t) = \mathcal{G}(\Delta t, 0, \sigma_{\mathrm{ol}}). \tag{5.10}$$

The fraction of outlier events, f_{ol}, is obtained separately for one-track vertices and multi-track vertices.

The resolution function of the vertex position in this $B \to D^{*+}D^{*-}$ analysis can be different from the standard parametrization described in this section due to the large tracks multiplicity. Therefore it is not a priori obvious if the standard parametrization for the CP side can be used throughout this analysis. In principle, determining the resolution parameters ourselves from the residual distribution would be possible but the obtained parameters would be less precise than the standard ones. In the $B \to D^{+}D^{-}$ analysis [59] the detector resolution parameters for the CP side were obtained specifically for that analysis. Due to the large resemblance between the vertex construction of that analysis and ours, a study is performed on the SVD1 cross-check analysis, using their obtained parameters to describe the resolution (see Chapter 7). But again larger uncertainties remain on the $D^{+}D^{-}$ parameters than on the standard ones. Another option, which we chose for this analysis, is to use a control mode leaving a similar signature in the detector as the $B \to D^{*+}D^{*-}$ modes, but with a larger branching fraction such as to obtain the resolution parameters with a smaller uncertainty. This is described in Section 5.5.

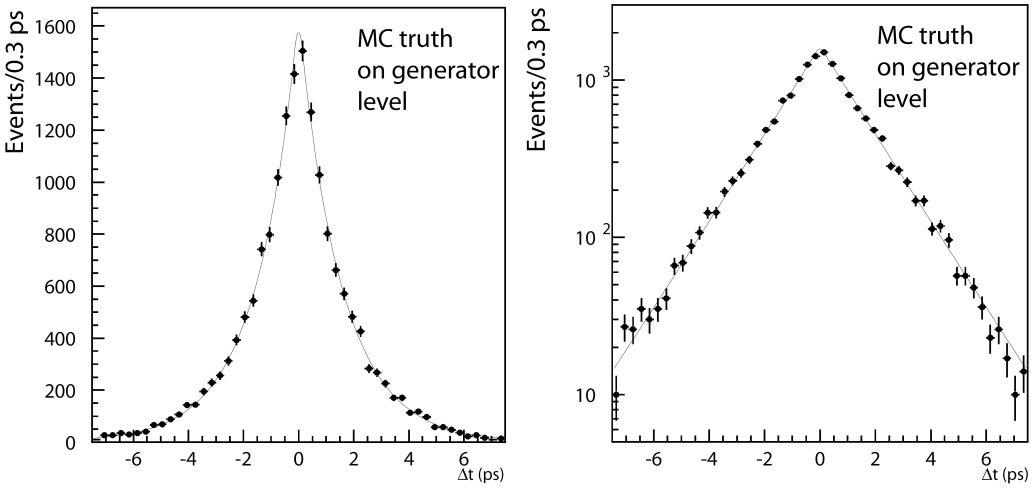


Figure 5.6: Fitted proper-time distribution of the generated signal MC.

5.2.3 The effect of non-primary decay vertices

When tracks that don't belong to the primary decay-vertex of the associated B meson are used to determine its decay vertex, an additional smearing is obtained. For the study of this effect two MC samples were used, a special sample where all short-lived secondary particles are forced to decay with zero lifetime at the B decay points, and a normal MC sample. The resolution function with which the proper-time distribution needs to be convoluted is found to be best described by a prompt component and a component that accounts for smearing due to charm and strange decays. The latter is studied using the difference in $z_{\rm tag}$ between the two MC samples. The resolution function reads

$$\mathcal{R}_{\rm np}(\Delta z_{\rm tag}) = f_\delta \delta^{\rm Dirac}(\Delta z_{\rm tag}) + (1-f_\delta)[f_{\rm p} E_{\rm np}(\Delta z_{\rm tag};\tau_{\rm p}^{\rm np}) + (1-f_{\rm p})E_{\rm np}(-\Delta z_{\rm tag};\tau_{\rm n}^{\rm np})],$$

where the parameters are indicated in red and $E_{\rm np}$ is:

$$E_{\rm np}(x;\tau) = \begin{cases} \frac{1}{\tau}e^{-x/\tau} & \text{if } x \geq 0, \\ 0 & \text{if } x < 0 \end{cases}.$$

The variable f_p is the fraction of the positive (negative) exponential part. The slope of the exponential is parametrized as:

$$\tau_{\rm p(n)}^{\rm np)} = \tau_{\rm p(n)}^0 + \tau_{\rm p(n)}^1 (s^0 + s^1 \xi_{\rm tag}^z)\sigma_{\rm tag}^z/(\beta\gamma c)\Upsilon. \tag{5.11}$$

The six parameters $f_\delta, f_{\rm p}, \tau_{\rm p}^0, \tau_{\rm p}^1, \tau_{\rm n}^0$ and $\tau_{\rm n}^1$ and the two scale factors s^0 and s^1 are obtained in the study presented in [75]. These standard parameters will be used throughout this analysis, which is justified as $\mathcal{R}_{\rm np}$ is specific to the tag-side vertex only.

5.3 Proper-time fit on signal MC

5.3.1 Signal MC at generation level

The signal MC sample described in Section 3.2.1 is generated with a B^0 lifetime of $1.534\,{\rm ps}^{-1}$. In Figure 5.6 the true proper-time distribution, $\Delta t = t_{\rm gen}^{CP} - t_{\rm gen}^{\rm tag}$ is shown using the generator information of the reconstructed events. This is therefore not affected by the detector

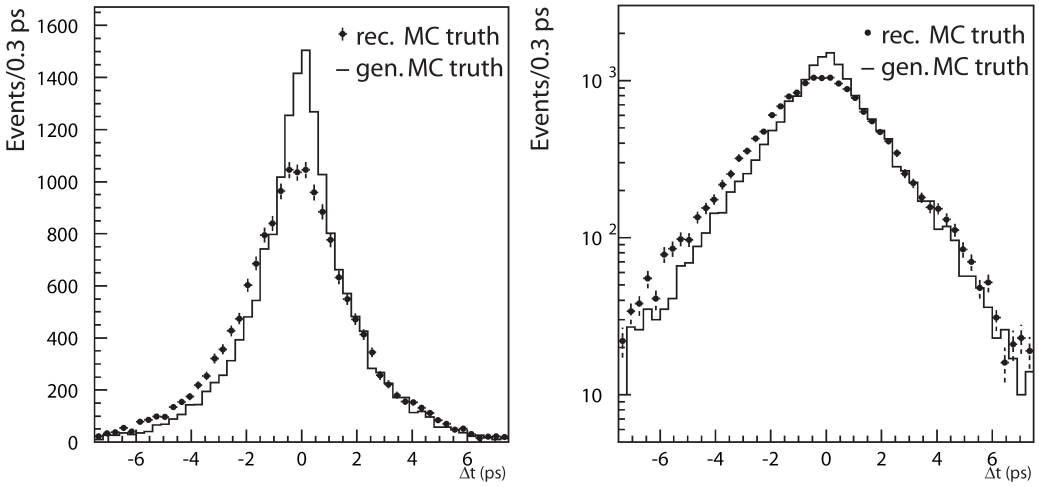


Figure 5.7: Fitted proper-time distribution of the generated and reconstructed signal MC.

response. However the kinematic approximation is still used. The distribution is therefore modeled with:

$$\mathcal{P}(\Delta t) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{2\tau_{B^0}} \otimes \mathcal{R}_k. \tag{5.12}$$

The fitted lifetime τ_{B^0} obtained from the true generated proper-time distribution of the reconstructed signal MC events is:

$$\tau_{B^0}^{\mathrm{MC,true}} = 1.552 \pm 0.012\,\mathrm{ps}.$$

5.3.2 Signal MC at reconstruction level

The reconstructed proper-time distribution of selected signal MC events is affected by the smearing which is illustrated in Figure 5.7. As described in the previous section, the theoretical function now needs to be convoluted with the resolution function $\mathcal{R}$ in order to model correctly $\Delta t = t^{CP} - t^{\mathrm{tag}}$. The PDF of the signal events is:

$$\mathcal{P}(\Delta t)_{\mathrm{sig}} = (1 - f_{\mathrm{ol}})\Big[\frac{e^{-|\Delta t|/\tau_{B^0}}}{2\tau_{B^0}} \otimes \mathcal{R}\Big] + f_{\mathrm{ol}}\mathcal{P}_{\mathrm{ol}}, \tag{5.13}$$

where $\mathcal{P}_{\mathrm{ol}}$ describes the outlier events. As the distribution only contains true signal events, there is no need to include a background component in the fit. The proper-time distribution of the reconstructed signal MC truth events is fitted using the standard resolution parameters for MC events and shown in Figure 5.8. Throughout this chapter we will present the fitted Δt distributions in a linear and log scale, where the latter accentuates the distributions of the tails where only a few events are present. The fitted lifetime of the signal MC reads:

$$\tau_{B^0}^{\mathrm{MC,rec}} \quad = \quad 1.69 \pm 0.02\,\mathrm{ps}.$$

The standard parameters for the signal MC do not seem to adequately describe the resolution of the B^{CP} vertex distribution obtained from $D^{*+}D^{*-}$ decays. We therefore fitted

simultaneously the main parameter, $s_{\rm main}$, of the detector resolution of the CP side with the B^0 lifetime. We chose to leave this parameter floating as it is the most significant parameter in $\mathcal{R}_{CP}$, which is the function responsible for the largest error on the vertex position. The result is summarized in Table 5.2. The fitted resolution parameter is larger than the resolution parameter in the standard settings. The overestimated lifetime now shifts more towards the generated value, but it still seems too high. This procedure is repeated separately for SVD1 signal MC and SVD2 signal MC. But no satisfactory result is obtained, as shown in Table 5.2.

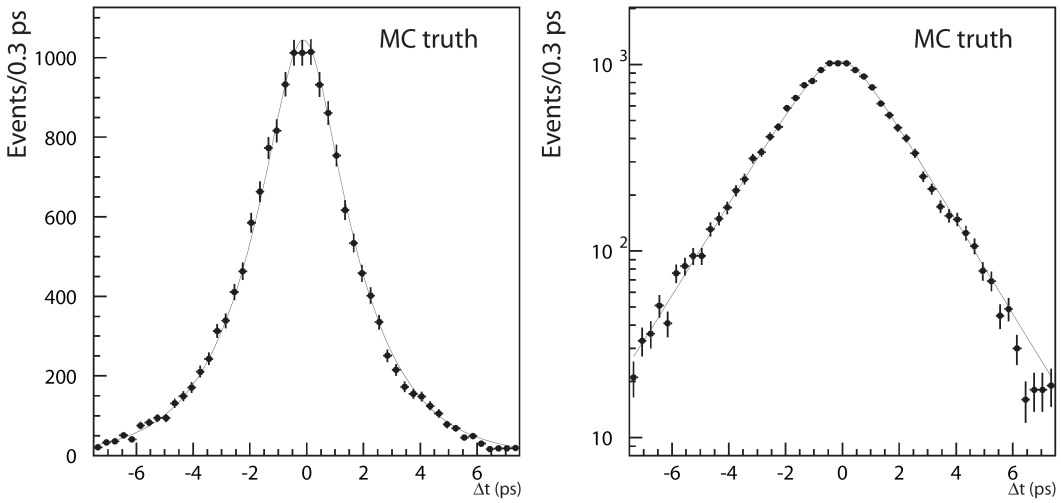


Figure 5.8: Fitted proper-time distribution of the reconstructed signal MC using the standard resolution function.

Parameter	fitted value
$s_{\rm main}$(SVD1)	1.70 ± 0.21
$s_{\rm main}$(SVD2)	1.51 ± 0.17
τ_{B^0}(SVD1 + SVD2)	$1.61 \pm 0.02\,\mathrm{ps}$
τ_{B^0}(SVD1)	$1.652 \pm 0.036\,\mathrm{ps}$
τ_{B^0}(SVD2)	$1.589 \pm 0.028\,\mathrm{ps}$

Table 5.2: Measured lifetime of reconstructed signal MC events when leaving the main resolution parameter free.

The standard resolution parameters on MC level do not seem to model the detector effect for $B \to D^{*+}D^{*-}$ decays very well. However there exist different sets of resolution parameters for the MC as for the data. Therefor we cannot draw any conclusions yet on the quality of the detector parametrization for data events.

5.4 Background probability density function

The proper-time distribution of the data events will be modeled with a signal PDF and a background PDF:

$$\mathcal{P}(\Delta t) = (1 - f_{\rm ol})(f_{\rm sig}\mathcal{P}_{\rm sig}(\Delta t) + (1 - f_{\rm sig})\mathcal{P}_{\rm bkg}(\Delta t)) + f_{\rm ol}\mathcal{P}_{\rm ol}, \tag{5.14}$$

where $f_{\rm sig}$ is the signal purity. $\mathcal{P}_{\rm bkg}(\Delta t)$ is the background probability density function which consists of a physics-like part convoluted with a resolution-like function:

$$\mathcal{P}_{\rm bkg}(\Delta t) \;=\; \int_{-\infty}^{+\infty} \mathcal{P}^{\rm true}_{\rm bkg}(\Delta t')\mathcal{R}_{\rm bkg}(\Delta t - \Delta t')d\Delta t'. \tag{5.15}$$

The physics part of the background function consists of a prompt component and a lifetime component with a lifetime $\tau_{\rm bkg}$. The prompt component models the Δt shape for $q\overline{q}$ events where q is an u, d, c or s quark. In these events all tracks originate from the same point leading to a Dirac delta function. The second part of the physics component describes the $B^0\overline{B^0}$ combinatorial background events whose proper-time distribution can be modeled by an exponential behavior with a $\tau_{\rm bkg}$ lifetime.

$$\mathcal{P}^{\rm true}_{\rm bkg}(\Delta t) \;=\; (1 - f_\delta)\frac{e^{-\Delta t/\tau_{\rm bkg}}}{2\tau_{\rm bkg}} + f_\delta\delta(\Delta t - \mu_\delta). \tag{5.16}$$

The resolution function is again parametrized with a Gaussian function depending on the vertex resolution:

$$\begin{aligned} R_{\rm bkg} \;=\; & \Big[(1.0 - f^{\rm tail}_{\rm bkg})\mathcal{G}(\Delta t, \mu_{\rm bkg}, s^{\rm main}_{\rm bkg}\sigma_{\rm vtx}) \\ & +\; f^{\rm tail}_{\rm bkg}\mathcal{G}(\Delta t, \mu_{\rm bkg}, s^{\rm tail}_{\rm bkg}s^{\rm main}_{\rm bkg}\sigma_{\rm vtx})\Big]. \end{aligned} \tag{5.17}$$

The parameters that need to be determined are, f_δ, μ_δ, $\tau_{\rm bkg}$, $f^{\rm tail}_{\rm bkg}$, $\mu_{\rm bkg}$, $s^{\rm tail}_{\rm bkg}$ and $s^{\rm main}_{\rm bkg}$. We assume that the background events in the sideband region have a similar shape as the background events under the signal peak. This assumption will be taken into account in the study of the systematic error (see Chapter 6). The background parameters are thus obtained by fitting the sideband region of the data sample, which contains data events with $\mathrm{M_{bc}} < 5.27\,\mathrm{GeV/c^2}$. A Gaussian function with a large width is again added to describe outliers events.

The proper-time distribution of data events in the sideband region is shown in Figure 5.9. The solid line shows the fitted background PDF, while the blue line is the contribution from the δ-function and the green line is the background lifetime function. The obtained values for the parameters are shown in Table 5.3. With these parameters the background PDF is constructed to fit the background events in our data sample.

Other models of the background shape are examined as well. A simpler model contains less parameters: only one Gaussian function is used to describe the resolution of the background shape, furthermore μ_δ is fixed to zero. Visually this function is found to describe the distribution less accurately than the function described above. The next model tested contains a separation between events where the tag-side vertex is composed of a single track or several tracks. Due to the correlation between the parameters, the relative statistical errors become large, therefore this last model is discarded as well.

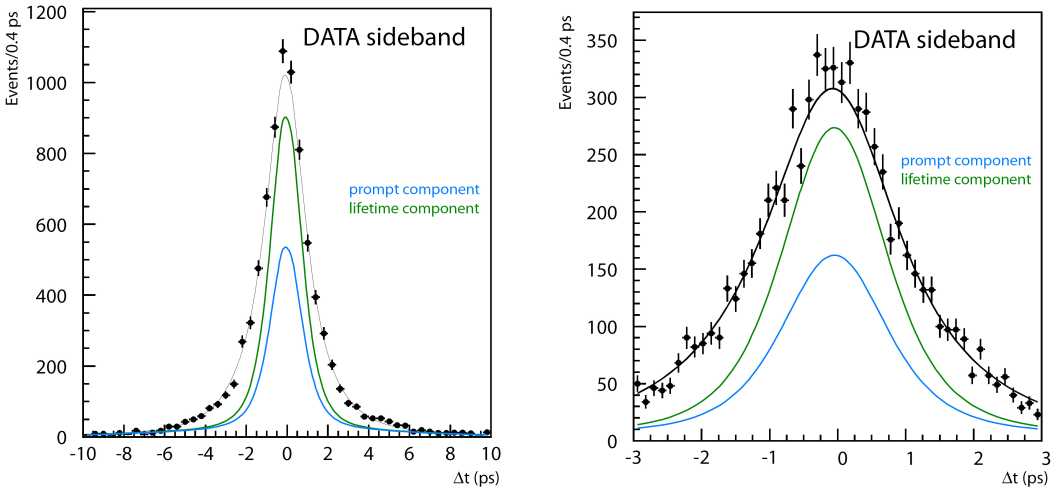


Figure 5.9: Fitted proper-time distribution of the events in the sideband region.

parameter	fitted value
$\tau_{\rm bkg}$	$1.10 \pm 0.07\,{\rm ps}$
μ_δ	$0.11 \pm 0.07\,{\rm ps}$
f_δ	0.37 ± 0.04
$\mu_{\rm bkg}$	$-0.16 \pm 0.04\,{\rm ps}$
$f_{\rm bkg}^{\rm tail}$	0.11 ± 0.01
$s_{\rm bkg}^{\rm main}$	1.72 ± 0.06
$sm_{\rm bkg}^{\rm tail}$	6.71 ± 0.37

Table 5.3: Fitted background parameters from events in the sideband region.

5.5 Control sample

A control sample is often used to study fitting procedures and resolutions on data events with decays that are similar to the decay under study but which have a larger branching fraction. Fit biases or problems in the resolution functions are often not visible from a fit on the data distribution due to low statistics.

The reconstructed decay modes chosen for the control sample in this analysis are:

- $B^0 \to D_s^+ D^-$
- $B^0 \to D_s^+ D^{*-}$
- $B^0 \to D_s^{*+} D^-$
- $B^0 \to D_s^{*+} D^{*-}$

and we further reconstruct $D_s^* \to D_s\gamma$ and $D_s^+ \to K_S^0 K^+$, $D_s^+ \to K^+K^-\pi^+$ and $D_s^+ \to K^+\pi^-\pi^+$.

We will perform two studies using the control sample. First the lifetime of the B^0 is measured in order to check if the standard resolution parameters properly describe the data. If not, the main resolution parameter can be obtained from the control sample. The second check is performing a CP fit on the control sample, which will be shown in the next chapter.

The selection cuts applied to define the control sample are similar to those used in the $B^0 \to D^{*+}D^{*-}$ analysis. We obtain 9667.8 ± 113.2 signal events, which is a factor of 20 more than the signal events in $B^0 \to D^{*+}D^{*-}$. The purity in the small signal region is 82%. The fitted $M_{\rm bc}$ distribution is shown in Figure 5.10. In this plot the ΔE distribution is required to be in the small signal region.

The proper-time distribution of the control sample is obtained with the same procedure as described in Section 5.1: the vertex position of the B^{CP} meson is obtained from the intersection point of the IP-tube and the tracks of both D-mesons and the vertex position of the $B^{\rm tag}$ meson is obtained from the intersection point of the remaining tracks in the event.

The background shape is determined with events from the sideband region of ΔE and $M_{\rm bc}$, using the model described in Section 5.4. The obtained parameters are summarized in Table 5.4. For comparison the background parameters of the proper-time distribution of $B^0 \to D^{*+}D^{*-}$ are shown in the table as well, to illustrate the large resemblance between the two samples. The obtained Δt shape of the background events is shown in Figure 5.11, superimposed on the proper-time distribution of the events in the sideband region.

	parameter	$B \to D_s^{(*)+} D^{(*)-}$	$B \to D^{*+}D^{*-}$
1	$\tau_{\rm bkg}$ (ps)	1.42 ± 0.03	1.10 ± 0.07
2	μ_δ (ps)	0.07 ± 0.03	0.11 ± 0.07
3	f_δ	0.31 ± 0.02	0.37 ± 0.04
4	$\mu_{\rm bkg}$ (ps)	-0.12 ± 0.02	-0.16 ± 0.04
5	$f_{\rm bkg}^{\rm tail}$	0.06 ± 0.01	0.11 ± 0.01
6	$s_{\rm bkg}^{\rm main}$	1.68 ± 0.03	1.72 ± 0.06
7	$s_{\rm bkg}^{\rm tail}$	8.02 ± 0.32	6.71 ± 0.37

Table 5.4: Fitted background parameters of events in the sideband region of the control sample (left) and, for comparison, of the $B \to D^{*+}D^{*-}$ sample (right).

The lifetime is extracted from the proper-time fit of the total control sample in an analogue way as will be done for the $B \to D^{*+}D^{*-}$ events. The fit is performed on the

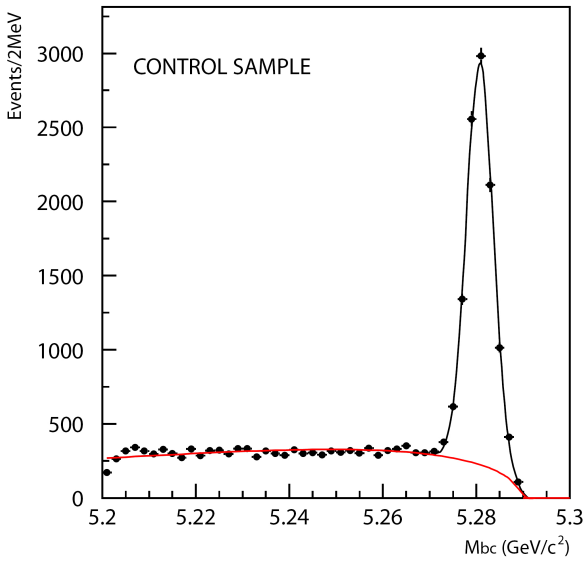


Figure 5.10: The $M_{\rm bc}$ projection of the fitted distribution of the control sample.

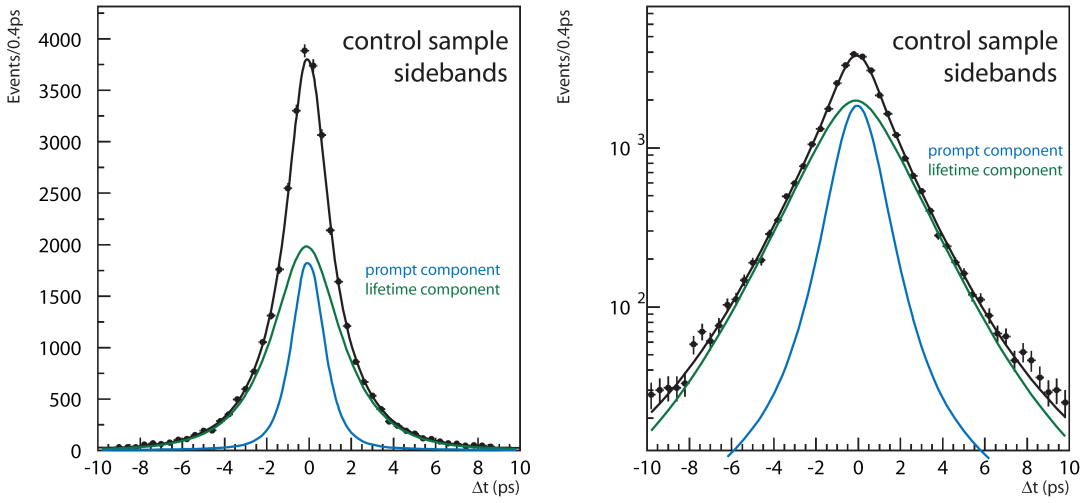


Figure 5.11: Fitted proper-time distribution of the events in the sideband region of the control sample. The prompt- and lifetime-component of the background PDF are added in blue and green respectively.

events in the large signal region and the purity is calculated on an event-by-event basis. The resolution parameters for the signal events are the standard parameters. The fitted lifetime of the control sample reads:

$$\tau_{B^0}^{\rm data}\ (B^0 \to D_s^{(*)+} D^{(*)-}) \quad = \quad 1.71 \pm 0.02\,{\rm ps}.$$

The world average value for the B^0 lifetime is $1.53 \pm 0.01\,{\rm ps}$, which means that when using the standard resolution parameters the fitted lifetime is more than eight standard deviations away from the expectation. A similar trend is observed on the signal MC for the $B \to D^{*+} D^{*-}$ events. g

Due to the high statistics of the control sample we can determine the dominating resolution parameter ourselves. Therefore we fit the proper-time distribution again while leaving free the parameter, $s_{\rm main}$, which is the main scaling factor of the width of the Gaussian function of the detector resolution (see Eq. (5.9)). The fitted distribution is shown in Figure 5.12 and the fitted parameters are

$$\begin{aligned} \tau_{B^0}^{\rm data}\ (B^0 \to D_s^{(*)+} D^{(*)-}) \quad &= \quad 1.58 \pm 0.03\,{\rm ps}, \\ {\rm SVD1:}\ \ {\rm s_{main}}\ ({\rm data}, {\rm B^0} \to {\rm D_s^{(*)+} D^{(*)-}}) \quad &= \quad 1.81 \pm 0.26, \\ {\rm SVD2:}\ \ {\rm s_{main}}\ ({\rm data}, {\rm B^0} \to {\rm D_s^{(*)+} D^{(*)-}}) \quad &= \quad 1.92 \pm 0.20. \end{aligned}$$

The value of the lifetime is now satisfactory as it is less than two standard deviations away

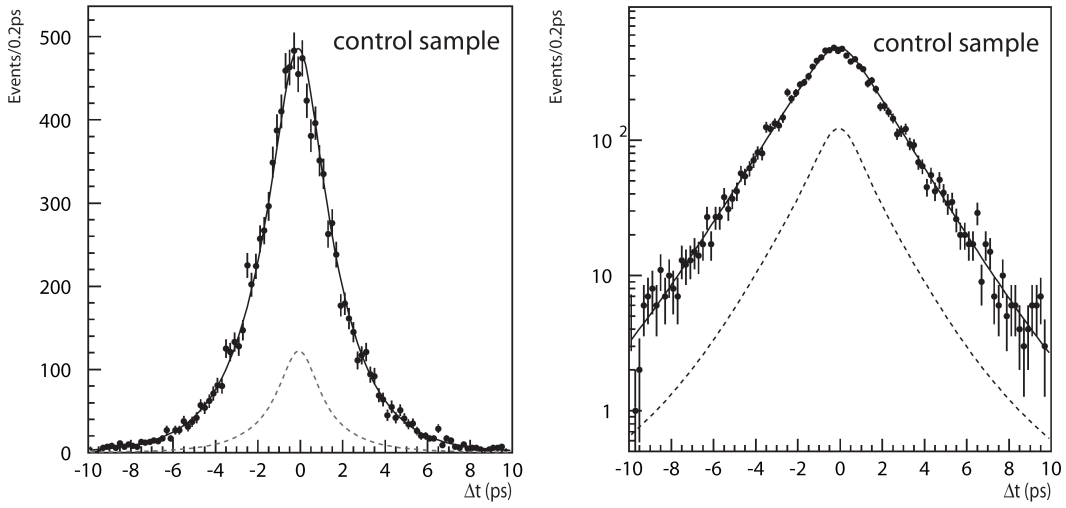


Figure 5.12: Fitted proper-time distribution of the control sample leaving the main resolution parameter free. The dotted line represents the background contribution.

from the world average. Like in the signal MC, the main resolution parameters are shifted to a higher value compared to the standard parameters.

This study on the control sample showed that the standard resolution parameters don't describe the vertex resolution when long-lived tracks are used, very precisely. The main detector resolution parameter will therefore be replaced by the obtained resolution parameter of the control sample, i.e, $s_{\rm main}\ ({\rm SVD1}) = 1.81$ and $s_{\rm main}\ ({\rm SVD2}) = 1.92$.

If the standard parametrization does not describe the resolution on the vertex of the data events, the MC events in the $B \to D^{*+} D^{*-}$ analysis will probably not be appropriately

described as well, which might explain the problematic results of the lifetime on the signal MC. We can scale the main detector resolution parameter to the same factor as obtained on the data. This gives us the following result for the MC:

$$s_{\rm main}\ ({\rm data}, {\rm B}^0 \to {\rm D}^{*+}{\rm D}^{*-}) = s_{\rm main}\ ({\rm data}, {\rm B}^0 \to {\rm D}_{\rm s}^{(*)+}{\rm D}^{(*)-}), \quad (5.18)$$

$$s_{\rm main}\ ({\rm MC}, {\rm B}^0 \to {\rm D}^{*+}{\rm D}^{*-}) = \frac{s_{\rm main}\ ({\rm MC}, {\rm standard})}{s_{\rm main}\ ({\rm data}, {\rm standard})} s_{\rm main}\ ({\rm data}, {\rm B}^0 \to {\rm D}_{\rm s}^{(*)+}{\rm D}^{(*)-}). \quad (5.19)$$

An overview of the values of $s_{\rm main}$ is given in Table 5.5. We can refit the proper-time dis-

		$s_{\rm main}$	
settings		SVD1	SVD2
data	standard	$0.919^{+0.46}_{-0.80}$	$0.66^{+0.45}_{-0.09}$
data	$B^0 \to D_s^{(*)+} D^{(*)-}$	1.81 ± 0.26	1.92 ± 0.20
data	$B^0 \to D^{*+} D^{*-}$	1.81 ± 0.26	1.92 ± 0.20
MC	standard	1.087	0.690
MC	$B^0 \to D^{*+} D^{*-}$	2.165	2.022

Table 5.5: Main detector resolution parameter, $s_{\rm main}$, for data and MC in the different samples.

tribution of the signal MC with the standard resolution parameters but replacing the main parameter, $s_{\rm main}\ ({\rm MC})$ by the scaled main resolution parameter. The obtained lifetime on the signal MC now gives the satisfactory result of:

$$\tau_{B^0}^{\rm MC} = 1.55 \pm 0.02\,{\rm ps}. \quad (5.20)$$

The fitted distribution is shown in Figure 5.13, however the improvement of the fit is not visible by eye.

5.6 Lifetime fit on the total data sample

The probability density function used to fit the data sample consists of a signal part and a background part, described already in the previous sections. The signal ratio is determined from the two dimensional fit of ΔE and $M_{\rm bc}$, obtained in Section 3.8 and is calculated event-by-event.

The proper-time distribution of the data events in the $B \to D^{*+}D^{*-}$ analysis should be fitted with the main resolution parameter, $s_{\rm main}$, set to the obtained result from the control sample while the other parameters are still set to the standard ones for data events.

In Figure 5.14 the proper-time distribution is shown together with the fit using $s_{\rm main}$ from the control sample (red line). The fit is performed on the large signal region, but for visualization purposes the plot shows the result in the small signal region. The extracted lifetime in the $B \to D^{*+}D^{*-}$ sample reads:

$$\tau_{B^0}^{\rm data} = 1.49 \pm 0.10\,{\rm ps}. \quad (5.21)$$

which is within the statistical uncertainty, consistent with the world average.

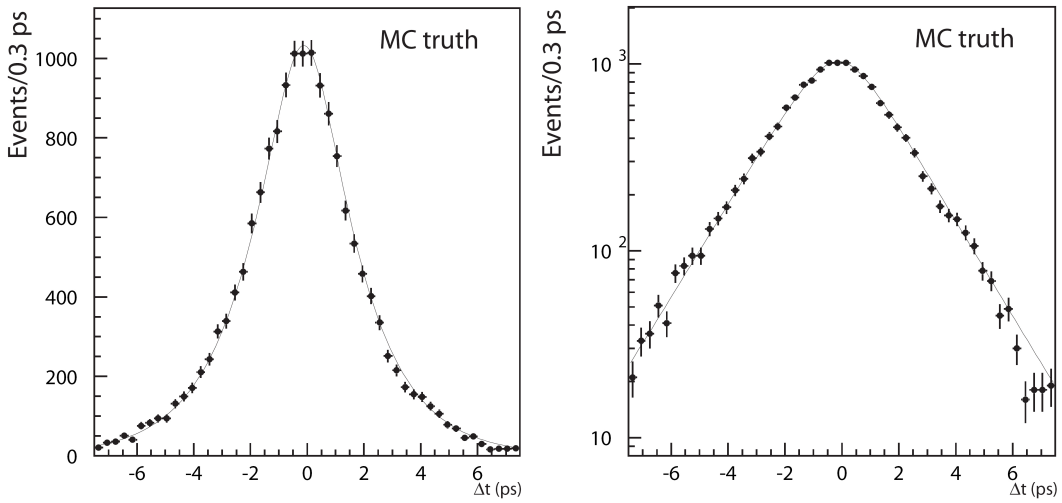


Figure 5.13: Fitted proper-time distribution of the signal MC sample using the standard resolution parameters but for the main detector resolution the scaled parameters are used.

For information we also present the extracted lifetime if we use only the standard parameters. The green curve in Figure 5.14 shows again the proper-time distribution together with the fit using only standard resolution parameters. Visually we cannot observe a large difference, however the measured lifetime using the standard parameters becomes:

$$\tau_{B^0}^{\rm data} \quad = \quad 1.62 \pm 0.10\,{\rm ps}. \tag{5.22}$$

Due to the large statistical error it is not obvious how well the data are described using the standard parameters, however the result using standard resolution parameters has a central value slightly further away from the world average than when using the parameters of the control sample.

We can therefore conclude that the background shape of the proper-time distribution and the resolution parameters are well defined with the use of the control sample. The fitted lifetime of the B^0 in the $B^0 \to D^{*+}D^{*-}$ mode is

$$\tau_{B^0} = 1.49 \pm 0.10\,{\rm ps}, \tag{5.23}$$

which is within errors compatible with the world average. We can therefore use the proper-time model in the CP analysis.

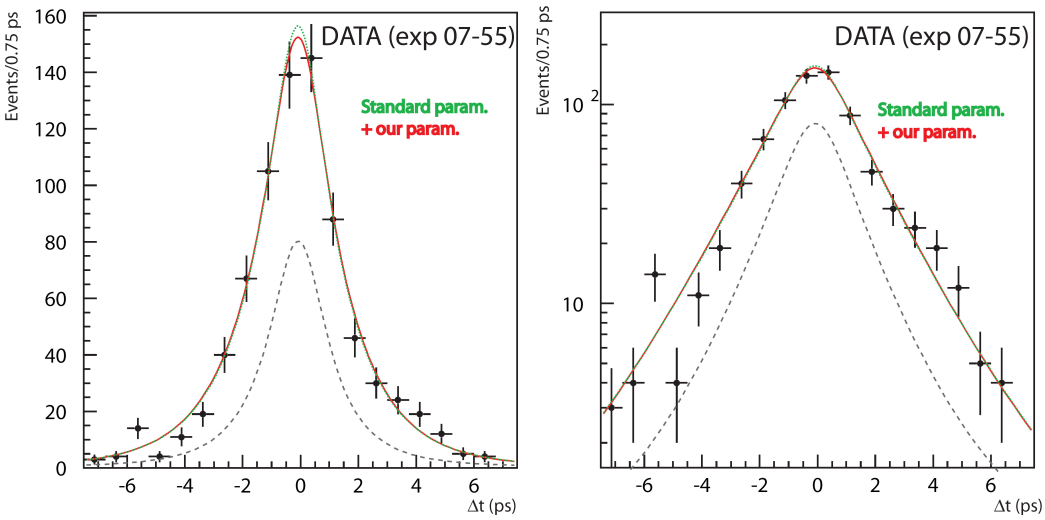
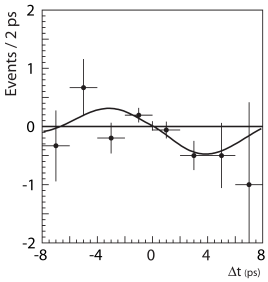


Figure 5.14: Fitted proper-time distribution of the data events of $B \to D^{*+}D^{*-}$, using all the standard resolution parameters (green) or using the main detector resolution parameter from the control sample (red). The dotted line represents the background contribution.

Chapter 6

Measurement of the CP asymmetry



In this chapter the main measurement of the analysis will be done by performing a fit on the proper-time distribution of the B^0 and the $\overline{B^0}$ tagged events to extract the CP violating parameters. Different studies are performed on the fitting procedure's accuracy, using an extensive fast MC sample and the signal MC sample. Finally a consistency check is performed on the control sample and the systematic study is presented as well.

We will remind the reader briefly in this chapter how we can theoretically extract the CP-violating parameters from the proper-time distribution, as discussed in Chapter 1. However, in practice we will also have to incorporate the effect of incorrect flavor assignment as well as the CP-odd dilution. The next sections will show different tests performed on fast-toy MC samples and the signal MC sample to study the possible effects of a bias in the fitting procedure and to estimate the statistical uncertainty on the CP-violating parameters. Also the CP asymmetry of the control sample is extracted as a final check of the fitting procedure. We then perform the CP fit on the total data set for the $D^{*+}D^{*-}$ reconstructed B modes, which is the main goal of this analysis. A study of the likelihood function will tell us the significance of the obtained CP asymmetry. Finally a systematic study will be performed to estimate the uncertainty of our measurement.

6.1 The probability density function

The term in the theoretical PDF of Δt for signal events, that contains the CP-violation parameters has opposite signs for B^0-tagged and $\overline{B^0}$-tagged events, so a distinction between the two CP conjugated processes can be made. The decay-time distribution for B^0-tagged and $\overline{B^0}$-tagged events is described in Chapter 1 and Chapter 2 as:

$$\mathcal{P}_{\rm sig}(\Delta t) \;=\; \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}}(1+q(\mathcal{S}\sin(\Delta m\Delta t)+\mathcal{A}\cos(\Delta m\Delta t))), \tag{6.1}$$

where $q=+1(-1)$ when the other B meson in the event decays as a B^0 ($\overline{B^0}$) and $\Delta t = t_{CP}-t_{tag}$ is the proper-time difference between the two decays. τ_{B^0} is the neutral B

lifetime, Δm the mass difference between the two B^0 mass eigenstates. $\mathcal{S}$ and $\mathcal{A}$ are the CP-violating parameters

$$\mathcal{S} = \frac{2\Im(\lambda)}{|\lambda|^2+1}, \quad \mathcal{A} = \frac{|\lambda|^2-1}{|\lambda|^2+1},$$

where λ is a complex parameter depending on the $B^0 - \overline{B^0}$ mixing as well as on the decay amplitudes for both B^0 and $\overline{B^0}$ to the CP eigenstate, as explained in Chapter 1.

This theoretical PDF is shown for different values of $\mathcal{S}$ and $\mathcal{A}$ in Figure 6.1. In case of no direct CP violation ($\mathcal{A} = 0$), the two distributions corresponding to B^0-tagged events or $\overline{B^0}$-tagged events are symmetric one with respect to the other, as shown in the top left plot of Figure 6.1. When there is no mixing-induced CP violation ($\mathcal{S} = 0$) the two curves are symmetric around $\Delta t = 0$, as shown in the top right plot. Notice that no time-dependent CP analysis is needed to extract the direct CP violating parameter $\mathcal{A}$. Merely counting the number of tagged B^0 and tagged $\overline{B^0}$ events is sufficient.

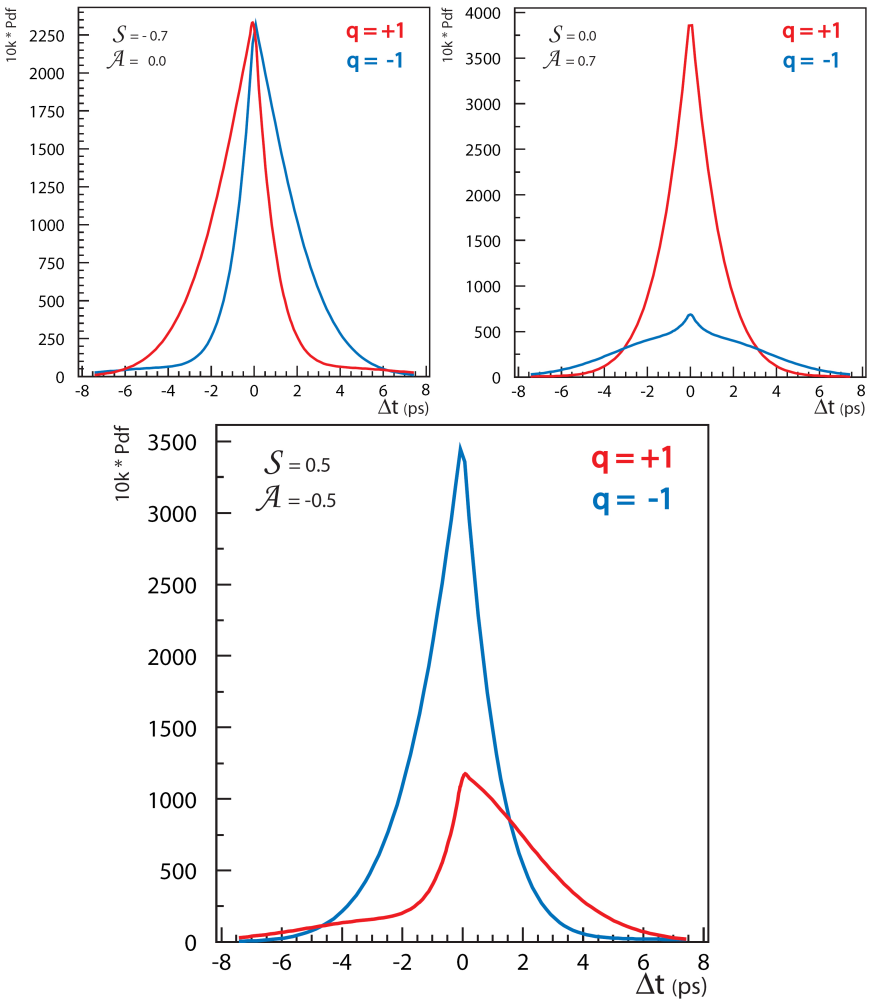


Figure 6.1: Proper-time PDF for B^0- and $\overline{B^0}$-tagged events without any detector smearing, tagging or polarization dilution for different values of the CP-violating parameters.

We define the raw CP asymmetry in each Δt bin as $a = (N_+ - N_-)/(N_+ + N_-)$, where $N_+(N_-)$ is the number of observed candidates with $q = +1(-1)$, such that:

$$a = \mathcal{S}\sin(\Delta m \Delta t) + \mathcal{A}\cos(\Delta m \Delta t), \tag{6.2}$$

when no smearing, tagging or polarization dilution is taking into account. The CP-violating parameters can be read directly from the asymmetry plot. The value of the PDF at $\Delta t = 0$ shows $\mathcal{A}$ and at $\Delta t \sim \pi$ shows $\mathcal{S}$ as illustrated in Figure 6.2.

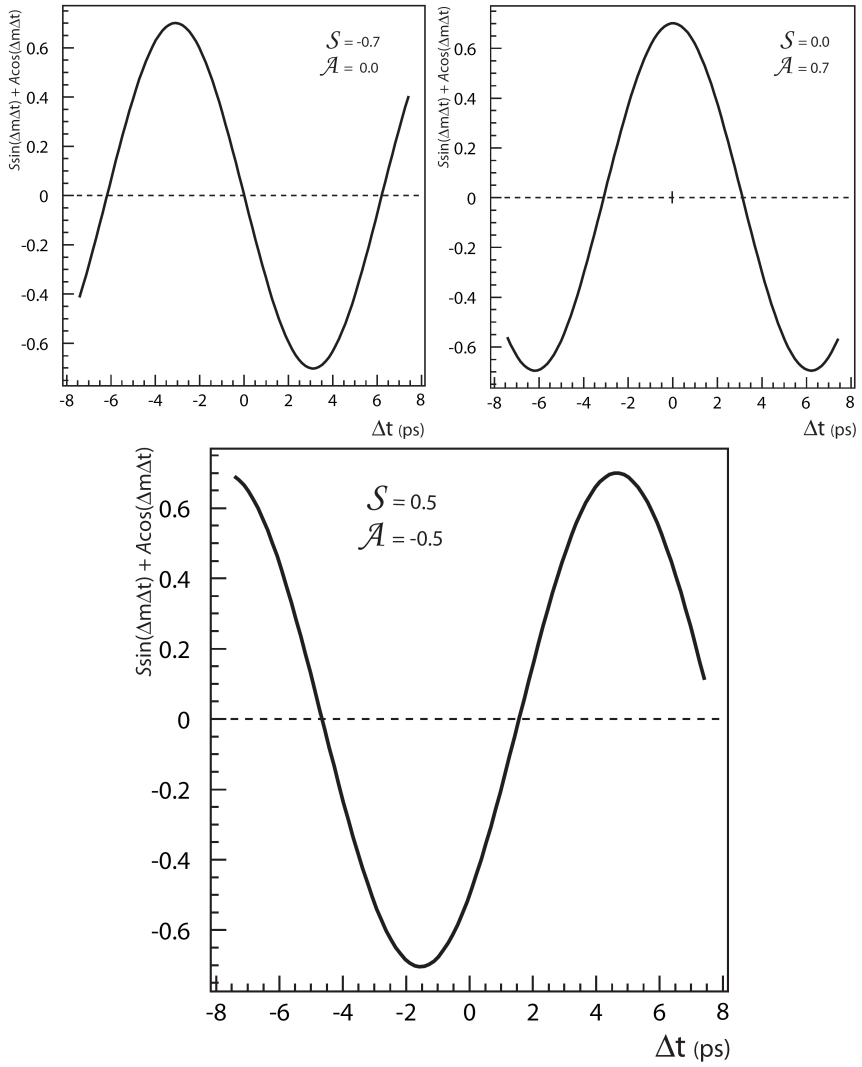


Figure 6.2: CP asymmetry PDF without any detector smearing, tagging or polarization dilution for different values of the CP-violating parameters.

Due to imperfections in the flavor tagging, an average wrong tag fraction, w, is introduced and a quantity Δw is used to describe the difference in wrong tag fraction for a B^0 and a $\overline{B^0}$ event, as explained in Section 2.7.4. The values of $w(r)$ and $\Delta w(r)$, which depend on the quality of the flavor tagging (grouped in r-bins), are obtained by the flavor tagging work group [77] and using a procedure described in Chapter 2. When incorporat-

ing the incorrect flavor assignment the PDF becomes

$$\mathcal{P}_{\rm sig}(\Delta t) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}}\Big[1 - q\Delta w + q(1-2w)\Big(\mathcal{S}\sin(\Delta m\Delta t) + \mathcal{A}\cos(\Delta m\Delta t)\Big)\Big].$$

The standard model prediction in absence of penguins is $\mathcal{A} = 0$ and $\mathcal{S} = -\eta \sin 2\beta$, where η is the CP eigenstate of $D^{*+}D^{*-}$. Notice that when no flavor tag is available ($q = 0$), the above equation becomes equal to Eq. (5.1), which describes the lifetime distribution.

The parameter η is $+1$ for CP-even events and -1 for CP-odd events. The probability of an $D^{*+}D^{*-}$ state being CP-even or CP-odd is determined with an angular analysis (see Chapter 4) and needs to be taken into account in the CP fit to avoid a dilution of the measurement of $\sin 2\beta$. This CP-odd probability is calculated on an event-by-event basis, using its $\cos\theta_{\rm tr}$ value and the known overall $\cos\theta_{\rm tr}$ shape, obtained from the angular analysis. For CP-odd events we switch the sign in the $\sin(\Delta m\Delta t)$, such that a measurement of $\sin 2\beta$ can be performed. We can therefore not use the factor $\mathcal{S}$ anymore but use

$$\mathcal{S}' = \mathcal{S}/\eta \stackrel{\rm SM}{=} -\sin 2\beta. \tag{6.3}$$

Including this property in the previous equation, the signal probability density function becomes:

$$\begin{aligned}\mathcal{P}_{\rm sig}(\Delta t) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}}\Big[1 - q\Delta w + q(1-2w) \\ \Big((1 - 2f_{\rm odd}(\cos\theta_{\rm tr}))\mathcal{S}'\sin(\Delta m\Delta t) + \mathcal{A}\cos(\Delta m\Delta t)\Big)\Big],\end{aligned} \tag{6.4}$$

where $f_{\rm odd}(\cos\theta_{\rm tr})$ is

$$f_{\rm odd}(\cos\theta_{\rm tr}) = \frac{\mathcal{R}_\perp\epsilon_\perp\mathcal{P}_\perp(\cos\theta_{\rm tr})}{\mathcal{R}_\perp\epsilon_\perp\mathcal{P}_\perp(\cos\theta_{\rm tr}) + (1.0 - \mathcal{R}_\perp)\epsilon_{\rm ev}\mathcal{P}_{\rm ev}(\cos\theta_{\rm tr})}. \tag{6.5}$$

The parameters $\epsilon_{\rm ev}$ and $\epsilon_\perp$ are the relative reconstruction efficiencies defined in Section 3.3.7 and $\mathcal{P}_\perp(\cos\theta_{\rm tr})$ ($\mathcal{P}_{\rm ev}(\cos\theta_{\rm tr})$ is the shape of the CP-odd (CP-even) distribution defined in Section 4.1.2. In Eq. (6.5) the normalization factor is not written for visualization purposes, but the functions are normalized over the fitted area.

Notice the difference between $f_{\rm odd}(\cos\theta_{\rm tr})$, the event-by-event probability of being a CP-odd event and $\mathcal{R}_\perp$, which is an overall parameter extracted from the angular analysis and is the average CP-odd fraction of the data sample, i.e. $\mathcal{R}_\perp = 0.125 \pm 0.043\,({\rm stat}) \pm 0.023\,({\rm syst})$.

The signal probability density function is convoluted with the detector resolution function described in Section 5.2. As in the lifetime fit the main resolution parameter is extracted from a control sample and a wide Gaussian is added to model outlier events.

The total probability density function consists of a signal PDF and a background PDF, with an event-by-event signal fraction which is determined from a three-dimensional distribution of ΔE, $M_{\rm bc}$ and $\cos\theta{\rm tr}$. The fit will be performed on the large signal box while we will show the results in the small signal box for a better visualization. The function describing the background events is the same as the one used in the lifetime fit and consists of a physics part convoluted with a resolution function (see Section 5.4). The total PDF of the proper-time distribution of the CP fit now reads:

$$\mathcal{P}(\Delta t;\mathcal{A},\mathcal{S}') = (1 - f_{\rm ol})\Big(f_{\rm sig}\mathcal{P}_{\rm sig}(\Delta t;\mathcal{A},\mathcal{S}')\otimes\mathcal{R} + (1 - f_{\rm sig})\mathcal{P}_{\rm bkg}(\Delta t)\Big) + f_{\rm ol}\mathcal{P}_{\rm ol}, \tag{6.6}$$

where $\mathcal{P}_{\rm ol}$ is the PDF of the outlier events. The free parameters of the fit are $\mathcal{A}$ and $\mathcal{S}'$.

6.2 Fast MC studies

A fast MC study is used to survey the statistical properties of the CP-violating parameters when extracted from the proper-time difference, as is done for the study of the CP-odd fraction (see Section 4.2). We inspect if the fitting procedure shows any bias and estimate the size of the statistical errors of the extracted parameters.

6.2.1 Generation

The generated samples are aimed to be as realistic as possible. The different features that are included in the fast MC are summarized here.

We generate three types of events: the signal events, the background events and the outlier events.

The ΔE and $M_{\rm bc}$ values

- The ΔE and $M_{\rm bc}$ values are generated in a signal region of $|\Delta E| < 60\,{\rm MeV}$ and $M_{\rm bc} > 5.25\,{\rm GeV/c^2}$.

- The values of the signal events and the background events of ΔE and $M_{\rm bc}$ are generated according to the observed distribution in the data (see Chapter 3).

- The signal purity is the same as on the data.

- A fraction of the events, $f_{\rm ol}$, is generated as outlier events. The ΔE and $M_{\rm bc}$ values of these outlier events are generated as if they are background events.

The proper-time distribution smeared with the detector resolution

- The function of the proper-time distribution of the signal events is an exponential decay distribution with a lifetime of $\tau = 1.53\,{\rm ps}$. This distribution is convoluted with the resolution function. The kinematic approximation depends on the $\cos\theta_B$. The detector resolution function and the smearing due to non-primary vertices depend on the vertex resolution, the number of tracks used to determine the vertex and the χ^2 distribution of the vertex fit and on the detector configuration.

- Because of the dependency of the proper-time resolutions on the detector settings, we generate for each event an experiment number. The probability of each experiment number is proportional to the luminosity of data taken per experiment.

- The distribution of the vertex resolution, the χ^2 distribution of the vertex fit, the number of degrees of freedom of the vertex fit and the number of tracks per vertex for the tag side are generated according to their distributions in the data. We generate just one or two tracks for the tag-side vertex, for the single- or multi-track vertex resolution parameters. The fraction of one-track vertices is taken from the data. The vertex resolution distribution of the data is obtained by fitting this with a pseudo-threshold function. The reduced χ-distribution is fitted with three exponential functions and the signal events in this fast MC are generated according to this function.

- The exponential function describing the proper-time distributions for signal events is convoluted with the resolution function, taking into account dependencies stated above. The smearing is done using the "standard" detector resolution parameters for data events; and depending on the generated experiment number, the parameters for SVD1 or SVD2 data are used.

- The values of the proper-time of the signal events are generated according to this function convoluted with the detector resolution.

- The proper-time value of the background events is generated according to the PDF described in Section 5.4.

- The proper-time values of the outlier events are generated according to a wide Gaussian centered at zero. The width , $\sigma_{\rm ol}$ and $f_{\rm ol}$, are taken from the detector resolution parameters.

The polarization

- The $\cos\theta_{\rm tr}$ value of the signal events are generated according to the $\cos\theta_{\rm tr}$ distribution of the signal events in the data, meaning with a probability $R_\perp$ to be CP-odd.

- The $\cos\theta_{\rm tr}$ value of the background events and outlier events is generated according to the distribution of the background events obtained from the data 4.1.3.

The flavor of the $B^{\rm tag}$

To determine the function according to which the flavor is generated we need to generate an r bin, $w(r)$ and $\Delta w(r)$ value.

- The r-bin value for signal events is generated according to the distribution obtained from the signal MC. For background events the distribution is obtained from the data events in the sideband region.

- A value for the wrong tag fraction $w(r)$ and $\Delta w(r)$ is generated according to the generated r-bin. The wrong tag fraction and $\Delta w(r)$ per r bin are set to the standard values.

- The value of the flavor of the $B^{\rm tag}$ is set to $+1$ with a probability ($\mathcal{P}$) depending on: $\cos\theta_{\rm tr}$, Δt, $w(r)$ and $\Delta w(r)$ and to -1 with $1-\mathcal{P}$. This probability reads:

$$\mathcal{P} = \Delta w + (1-2w)\Big((1-2f_{\rm odd}(\cos\theta_{\rm tr}))\mathcal{S}'\sin(\Delta m\Delta t) + \mathcal{A}\cos(\Delta m\Delta t)\Big)\Big], \quad (6.7)$$

  where $\Delta m = 0.507 ps^{-1}$ and $S' = -\sin 2\beta$ (instead of $S = -\eta\sin 2\beta$) .

- The flavor value of the $B^{\rm tag}$ of the background and outlier events is generated with equal probability of being $+1$ or -1.

In this section of the fast MC study we will refer to $\mathcal{S}'$ as just $\mathcal{S}$ for simplicity. We generated 6400 different toy-MC samples each containing 500 signal events, which corresponds roughly to the number of $B^0 \to D^{*+}D^{*-}$ signal events we have in the data sample. The generated CP variables $\mathcal{S}$ and $\mathcal{A}$ are changed in each sample in steps of 0.025 between $[-1,1]$, such as to fill a two-dimensional grid, as illustrated in Figure 6.3. These samples are then fitted one by one according to the procedure described in Section 6.1.

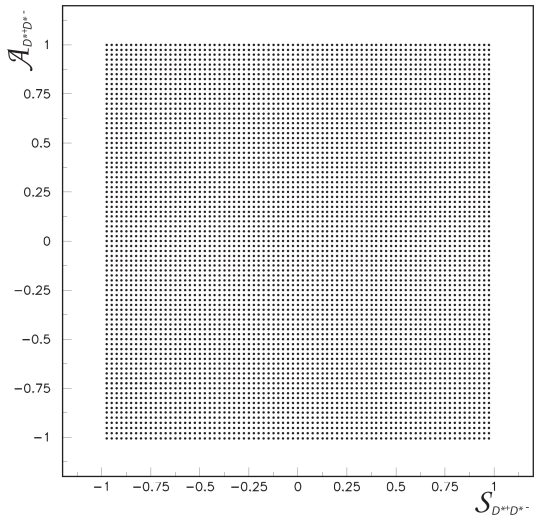


Figure 6.3: Generated $\mathcal{S}$ and $\mathcal{A}$ values in the fast MC.

6.2.2 Residual, error and pull distributions

6.2.2.1 The residual distribution

Figure 6.4 shows the residual distributions for both fitted CP parameters. We define the residual distribution as the difference between the generated and the fitted value (which is the opposite of what is used in most analysis). When the fitting procedure does not contain a bias, the residual distribution should be centered around zero. The width of the distribution is the measure of the statistical accuracy of the fitted parameters. The residual distributions are fitted with a Gaussian function and the results are summarized in Table 6.1. Besides the visual appearance of the residual distribution also the results of the fit show no bias for any of the two parameters.

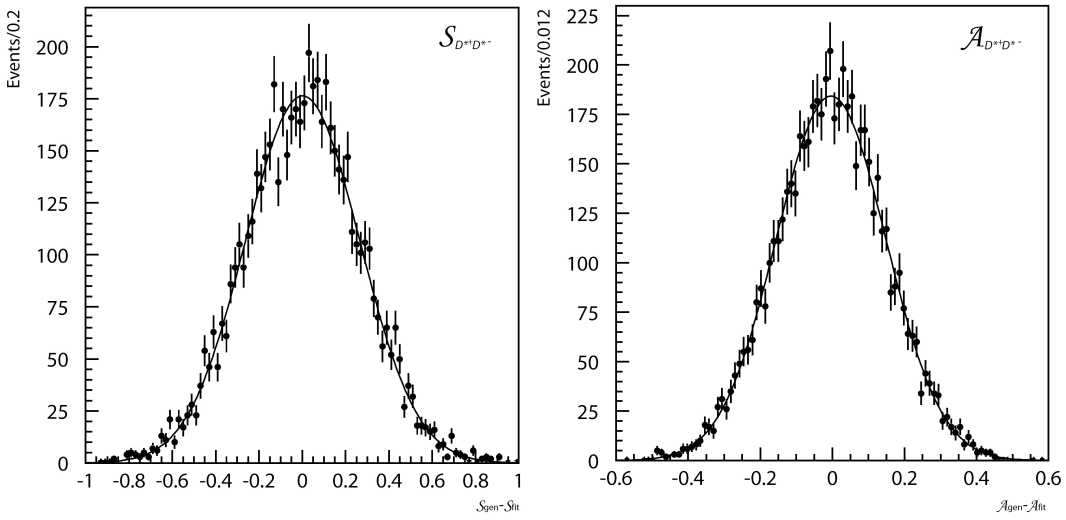


Figure 6.4: Distributions of $\mathcal{S}_{\rm gen} - \mathcal{S}_{\rm fit}$ (left) and $\mathcal{A}_{\rm gen} - \mathcal{A}_{\rm fit}$ (right).

	$\mathcal{S}_{\rm gen} - \mathcal{S}_{\rm fit}$	$\mathcal{A}_{\rm gen} - \mathcal{A}_{\rm fit}$
mean	-0.001 ± 0.003	-0.004 ± 0.002
σ	0.263 ± 0.003	0.153 ± 0.002

Table 6.1: Parameters of the Gaussian fit of $\mathcal{S}_{\rm gen} - \mathcal{S}_{\rm fit}$ and $\mathcal{A}_{\rm gen} - \mathcal{A}_{\rm fit}$.

6.2.2.2 The error distributions

One of the main reasons to perform a fast MC is to study the error distribution. In Figure 6.5 the symmetric errors of the fitted $\mathcal{S}$ and $\mathcal{A}$ values are shown, while in Figure 6.6 the asymmetric errors are shown. The symmetric errors are obtained by the "Migrad" algorithm of the Minuit program [78], while the asymmetric errors are calculated with the "Minos" algorithm of Minuit.

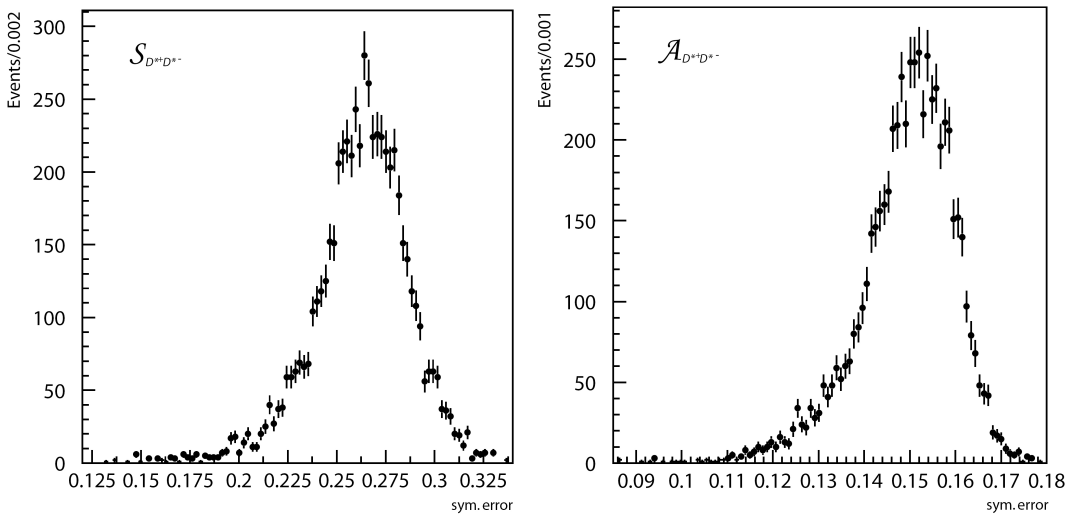


Figure 6.5: Symmetric errors on $\mathcal{S}$(left) and $\mathcal{A}$(right) as returned by the fit.

For both CP-violating parameters the distributions show tails at lower values which is due to the dependence of the error on the central value. The closer the absolute value of the CP-violating parameters approaches 1, the smaller the error on the fit is, as illustrated in Figure 6.7. This is because the fitted parameter reaches the physical limit, the area where the PDF is no longer defined. The figures also show that the error on the $\mathcal{S}$ parameter is on average larger than the error on the $\mathcal{A}$ parameter. This is a known phenomena which is partially caused by the uncertainty on the dilution due to polarization, which only weakens the precision of $\mathcal{S}$ but has no influence on $\mathcal{A}$.

In order to estimate the statistical errors on the CP parameters of the data we create 2500 extra samples, each containing 500 signal events, but now both the generated CP-violating parameters are fixed to their Standard Model values, i.e. $\mathcal{S} = -0.7$ and $\mathcal{A} = 0$. The symmetric errors of the fits are shown in Figure 6.8. Both distributions are then fitted with a Gaussian function and the parameters of the fit are summarized in Table 6.2. The error distribution is not necessarily a Gaussian distribution. The fit was just used to estimate in an easy way the average error, however visually it is seen that the error

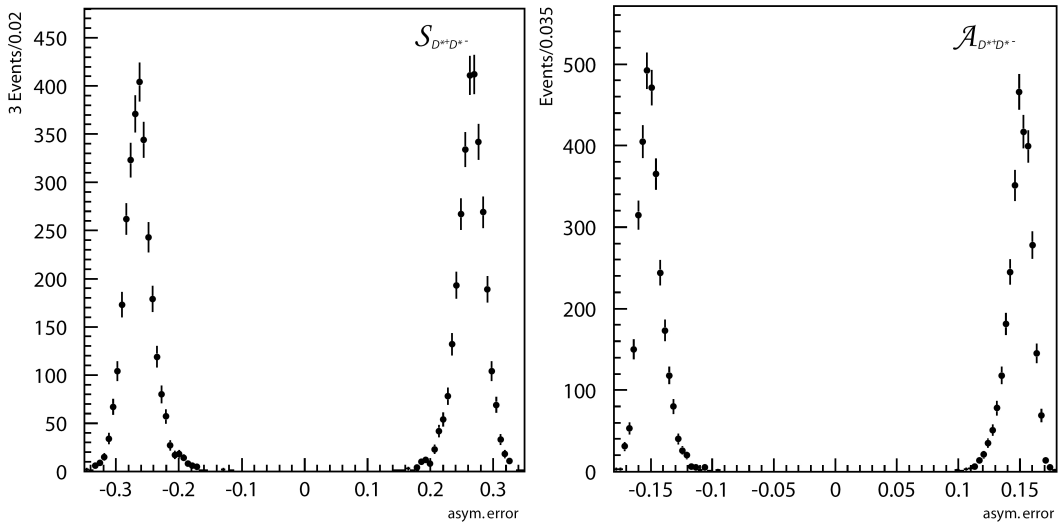


Figure 6.6: Asymmetric errors on $\mathcal{S}$(left) and $\mathcal{A}$(right) as returned by the fit.

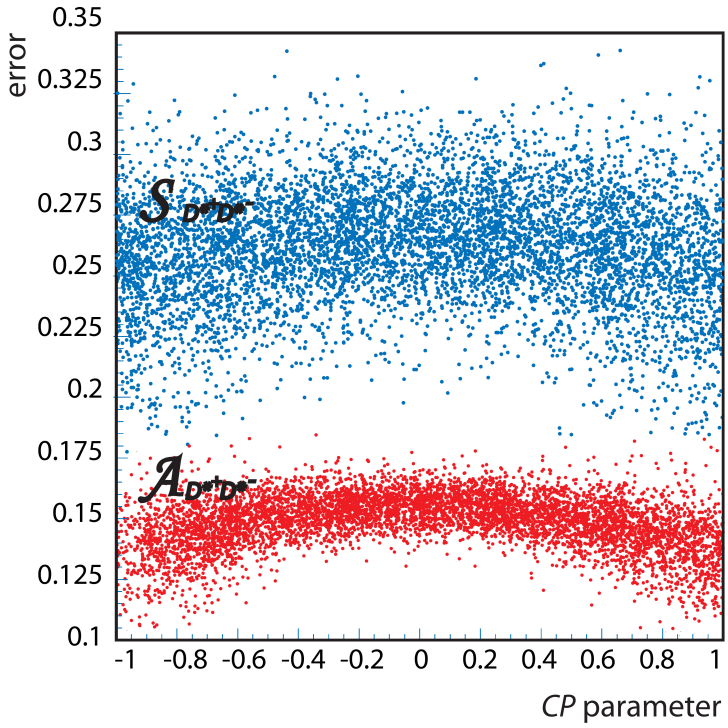


Figure 6.7: Error on $\mathcal{S}$ (blue) and $\mathcal{A}$ (red) as returned by the fit as a function of the fitted central value.

distribution is very similar to a Gaussian distribution. We also verified that these estimated errors of $\mathcal{S}$ and $\mathcal{A}$ correspond to the width of the residual distributions, which means that the errors are correctly estimated.

	error on $\mathcal{S}$	error on $\mathcal{A}$
mean	0.2632 ± 0.0004	0.1552 ± 0.0001
σ	0.022 ± 0.004	0.006 ± 0.001

Table 6.2: Parameters of the Gaussian fit of the error distribution of the generated Standard Model CP-violating parameters.

The expected errors of the Standard Model CP-violating parameters are $\sigma_S = 0.263$ and $\sigma_A = 0.155$ where the statistical uncertainty is dependent only on the generated signal events in the samples. These results also correspond approximately to the naive extrapolation based on the results of the previous Belle analysis on SVD1 data [58], which is:

$$\begin{aligned} \mathcal{A}_{\mathrm{SVD1}} &= -0.26 \pm 0.26\,(\mathrm{stat}), \\ \mathcal{S}_{\mathrm{SVD1}} &= -0.75 \pm 0.56\,(\mathrm{stat}). \end{aligned}$$

Without taking into account the dependence of the error on the central value, or any possible changes in the reconstruction software, the extrapolation of the statistical error reads:

$$\begin{aligned} \frac{\sqrt{152 \times 10^6} \times 0.26}{\sqrt{656 \times 10^6}} &= 0.13, \\ \frac{\sqrt{152 \times 10^6} \times 0.56}{\sqrt{656 \times 10^6}} &= 0.27, \end{aligned}$$

as the statistical error is inversely proportional to the square root of the number of $B\overline{B}$ events.

6.2.2.3 The pull distribution

We defined the pull of a measurement of $\mathcal{S}$ as:

$$\mathrm{Pull}(\mathcal{S}) = \frac{\mathcal{S}_{\mathrm{gen}} - \mathcal{S}_{\mathrm{fit}}}{\mathrm{error\ on}\,\mathcal{S}_{\mathrm{fit}}},$$

and of course the same definition holds for the $\mathcal{A}$ parameter. In this definition the pull has the opposite sign as what is usually used. Figures 6.9 and 6.10 show the pull distributions for $\mathcal{S}$ and $\mathcal{A}$ computed with the symmetric and asymmetric errors, respectively.

When the errors and the residuals are Gaussianly distributed the pull as well should be a Gaussian function, centered at 0 when no bias occurs in the fitting process and with a width of $\sigma = 1$ when the errors are properly estimated. The distributions are therefore fitted with a Gaussian function and the parameters from the fit are summarized in Table 6.3. The mean of the distribution is each time compatible with zero, indicating no or a very small bias. The widths of the distributions are compatible with one, indicating that the error and residual distributions are properly estimated.

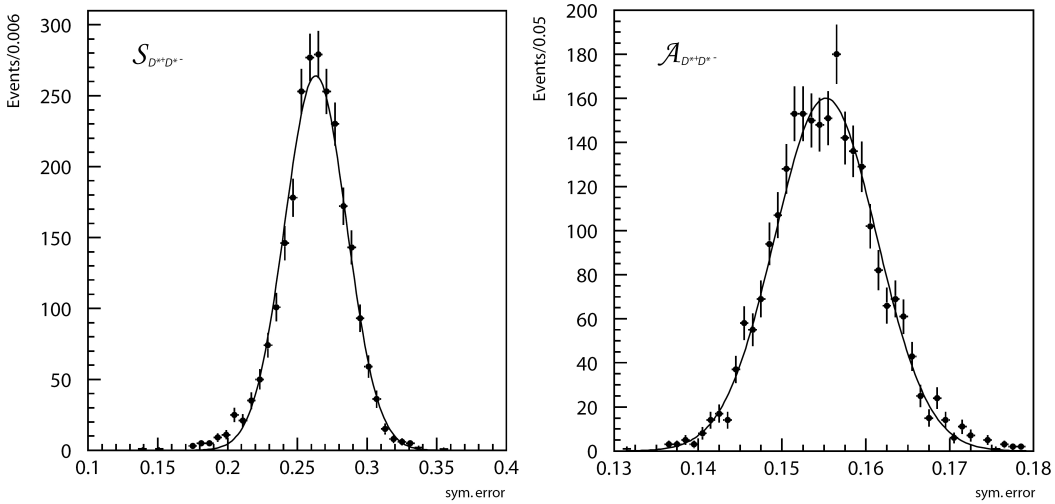


Figure 6.8: Symmetric errors of $\mathcal{S}$(left) and $\mathcal{A}$(right) as returned by the fit of the generated Standard Model CP-violating parameters.

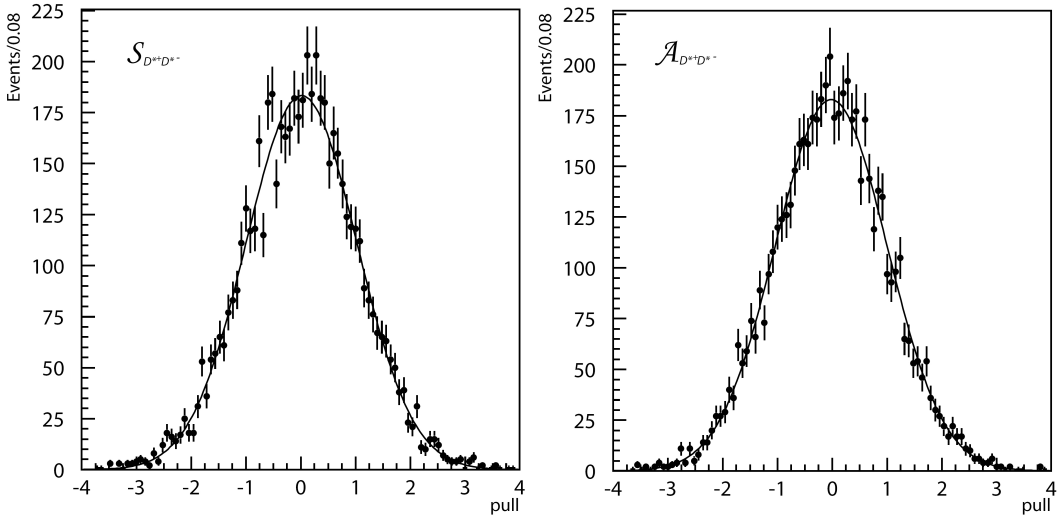


Figure 6.9: Pull distributions of $\mathcal{S}$(left) and $\mathcal{A}$(right) with symmetric errors.

	pull $\mathcal{S}$	pull $\mathcal{A}$
	symmetric error	
mean	-0.01 ± 0.01	-0.02 ± 0.01
σ	1.02 ± 0.01	1.02 ± 0.01
χ^2/ndf	1.19	0.78
	asymmetric error	
mean	-0.01 ± 0.01	-0.03 ± 0.01
σ	1.00 ± 0.01	1.00 ± 0.01
χ^2/ndf	0.94	1.05

Table 6.3: Parameters of the Gaussian fit of the pull distributions of the generated Standard Model CP-violating parameters.

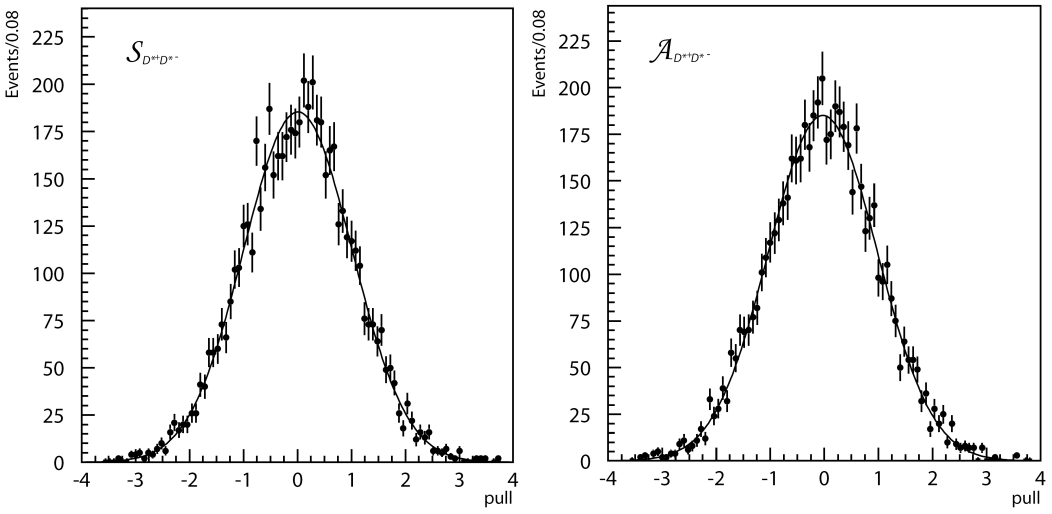


Figure 6.10: Pull distributions of $\mathcal{S}$(left) and $\mathcal{A}$(right) with asymmetric errors.

6.2.3 Correlation

The fast MC sample, which is generated with the Standard Model values for the CP-violating parameters, is also used to check if there is any correlation between the fitted $\mathcal{S}$ value and the fitted $\mathcal{A}$ value. In theory those are two independent variables and no correlation should be observed. Figure 6.11 shows a scatter plot of the fitted CP-parameters and indeed no correlation is observed.

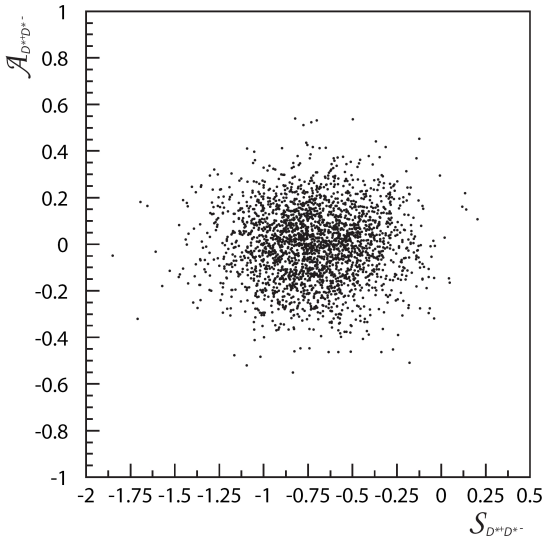


Figure 6.11: The fitted $\mathcal{S}$ and $\mathcal{A}$ values of the generated Standard Model CP-violating parameters.

	linearity $\mathcal{S}_{\text{toy}}$	linearity $\mathcal{A}_{\text{toy}}$
constant	0.0007 ± 0.0043	0.0057 ± 0.0022
slope	1.034 ± 0.007	0.995 ± 0.003

Table 6.4: Parameters of the fit of the linearity plot.

6.2.4 Linearity test

We perform a test to check if there exist any linear effect between the residual distribution and the CP parameter. For the linearity test on $\mathcal{S}$ we generate 4410 samples each containing 500 signal events. The generated CP parameter is changed after each 210 generated files, the first 210 files are generated with $\mathcal{S} = -1$, the last 210 with $\mathcal{S} = +1$, while $\mathcal{A} = 0$ is kept as a constant. A similar procedure is applied to generate 4410 files where $\mathcal{A}$ varies between -1 and $+1$ while $\mathcal{S}$ is kept constant to zero. The CP parameters of the 4410 files are then fitted one by one, with the same function as will be used for the data fit. The distribution of the fitted parameters, grouped per generated CP parameter are fitted with a Gaussian function. The mean of the distribution of these 210 events, corresponding to the same generated CP parameter, is then plotted in the linearity plot in Figure 6.12. The error bars are the errors as returned by the fit on the mean of the distribution.

The obtained data points are then fitted with a first-order polynomial to see if there is any bias. The result of the fits are summarized in Table 6.4 and drawn in black on the plot. For visualization purposes a red reference line with slope 1.0 and constant 0.0 is shown as well, which corresponds to the situation with no bias. The fit results show barely any bias for $\mathcal{A}$ and a small deviation for the $\mathcal{S}$ parameter as the slopes of both fits are close to 1.0. This deviation will be taken into account in the systematic study but is negligible compared to, for example, systematic errors induced by the resolution on the vertex position.

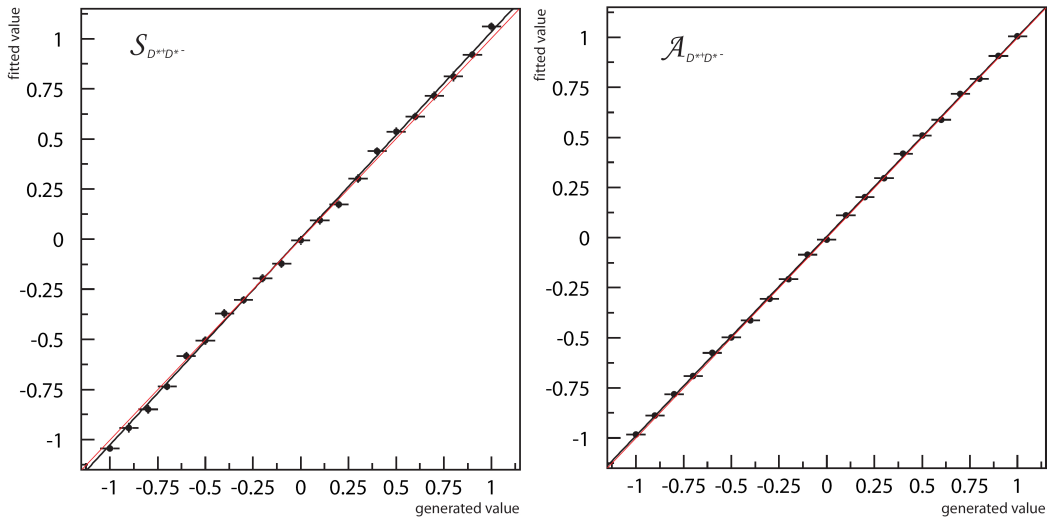


Figure 6.12: Constructed linearity plots for $\mathcal{S}$ (left) and $\mathcal{A}$ (right), the error bars represent the error on the fitted mean of the distribution.

6.3 CP measurement on signal MC sample

We also perform the CP fit on the signal MC sample for which the detector effects on the measurements are taken into account in the most detailed way. We used 13 statistically independent samples, which each contain 500 signal-MC events. As signal MC only contains the signal events we have added background events from the toy MC from the previous section, with a similar purity as on the data. The signal events are generated with the Standard Model CP-violating parameters, i.e. $\mathcal{A}_{\rm MC} = 0.0$ and $\mathcal{S}'_{\rm MC} = -0.7$ and with the same CP-odd fraction as on the data. These samples are reconstructed after passing through GEANT, and fitted with the same CP-fit code as will be used on data. We employ the "standard" resolution parameters for MC events except that the main resolution parameter is set to the scaled MC resolution parameter obtained from the control sample (see Section 5.5). In the next paragraphs we will study the same distributions as examined in the previous section: the residual distribution, the error and pull distribution. The signal MC files are fitted with Eq. (6.6), where free parameters of the fit are $\mathcal{A}$ and $\mathcal{S}'$. In this section we will use the notation $\mathcal{S}$ when we actually refer to $\mathcal{S}'$. This means that this fitted parameter corresponds to a situation with only CP-even events.

Figure 6.13 shows the 13 fitted values of the CP-violating parameters. This distribution is again fitted with a Gaussian function and the result is summarized in Table 6.5. We obtain a mean of the distribution which is very close to the generated result.

The (symmetric) error distributions are shown in the plots in Figure 6.14. The distribution is fitted with a Gaussian function and the results are summarized in Table 6.6. The mean of the distribution corresponds well to the mean of the error distribution of the large sample of fast MC events.

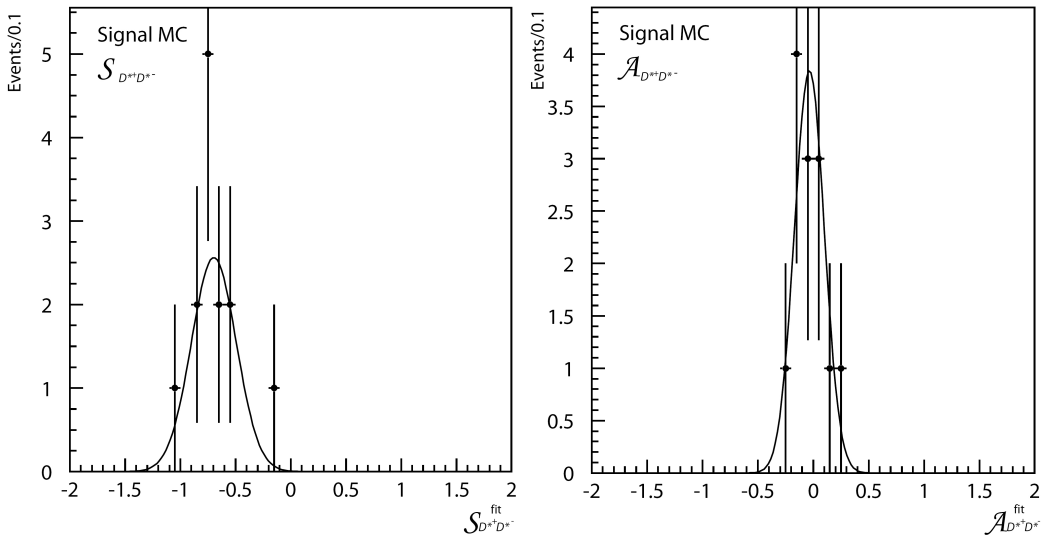


Figure 6.13: Distributions of $\mathcal{S}$(left) and $\mathcal{A}$(right) as returned by the fit of the signal MC samples.

	fitted $\mathcal{S}_{\rm MC}$	fitted $\mathcal{A}_{\rm MC}$
mean	-0.696 ± 0.056	-0.035 ± 0.037
σ	0.202 ± 0.039	0.135 ± 0.026

Table 6.5: Parameters of the Gaussian fit of $\mathcal{S}$(left) and $\mathcal{A}$(right) as returned by the fit of the signal MC samples.

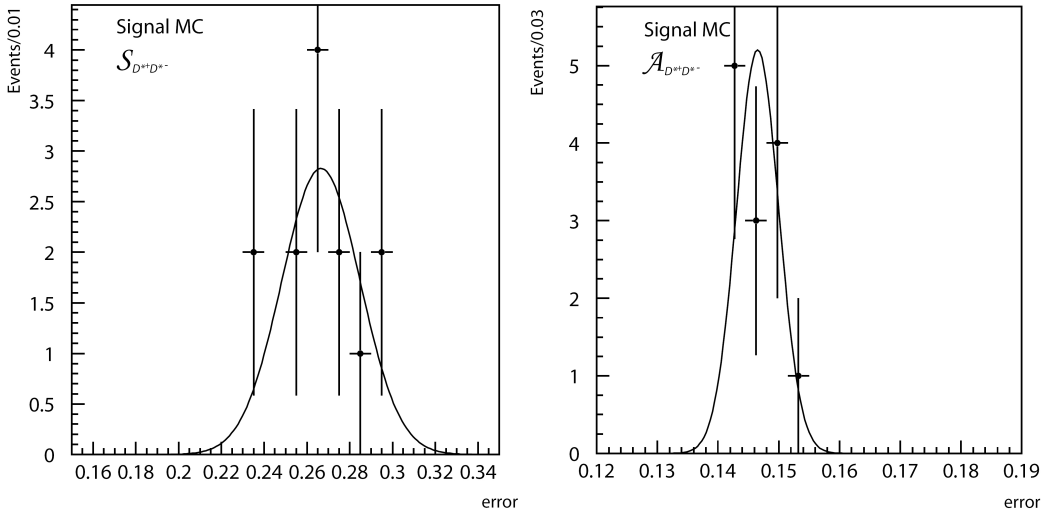


Figure 6.14: Symmetric errors on $\mathcal{S}$(left) and $\mathcal{A}$(right) as returned by the fit on the signal MC samples.

	errors $\mathcal{S}_{\rm MC}$	errors $\mathcal{A}_{\rm MC}$
mean	0.267 ± 0.005	0.146 ± 0.001
σ	0.018 ± 0.002	0.003 ± 0.001

Table 6.6: Parameters of the Gaussian fit of the error distribution of the CP-violating parameters of the signal MC samples.

	pull $\mathcal{S}_{\rm MC}$	pull $\mathcal{A}_{\rm MC}$
mean	-0.012 ± 0.231	-0.267 ± 0.236
σ	0.830 ± 0.165	0.848 ± 0.170

Table 6.7: Parameters of the Gaussian fit of the pull distribution of the CP-violating parameters of the signal MC samples.

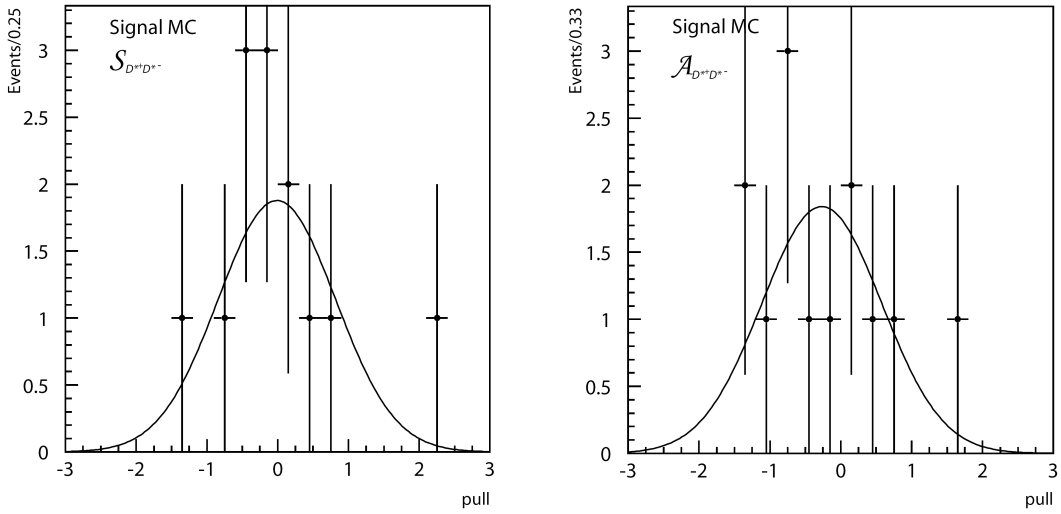


Figure 6.15: Fitted $\mathcal{S}$(left) and $\mathcal{A}$(right) pull distributions.

Finally, the pull distributions are shown in Figure 6.15. They are again fitted with a Gaussian function of which the mean should be close to zero and the width close to one. The mean and width of the fitted Gaussian is summarized in Table 6.7.

We can conclude that, even with the small number of signal-MC samples simulated with the detailed detector response taken into account, the results of the error and the pull distribution are very satisfactory.

6.4 CP measurement on the control sample

As explained in the previous chapter the control sample of $B^0 \to D_s^{(*)+} D^{(*)-}$ decays leaves a similar signal in the detector as the $B^0 \to D^{*+}D^{*-}$ decay but it has the advantage that due to the larger branching fraction we reconstruct around 20 times more signal events.

The Feynman diagrams of the control channels are identical to those of $B^0 \to D^{*+}D^{*-}$, except that the down anti-quark is replaced with a strange anti-quark. The CKM matrix elements corresponding to this transition are $V_{\rm cb}^* V_{\rm cs}$ for the tree diagram and $V_{\rm tb}^* V_{\rm ts}$, $V_{\rm cb}^* V_{\rm cs}$ and $V_{\rm ub}^* V_{\rm us}$ for the penguin diagrams. As the final state is not accessible from B^0 and $\overline{B^0}$ decays there is no mixing-induced CP violation possible in this decay. The decay amplitudes A_f and $\overline{A}_{\overline{f}}$, where A_f ($\overline{A}_{\overline{f}}$) represents the decay amplitude of a B^0 ($\overline{B^0}$) particle and f and $\overline{f}$ are the final states, have the same absolute value. The amplitudes $\overline{A}_f$ and $A_{\overline{f}}$ are negligible as they occur only through higher-order diagrams. Therefore no direct CP violation is expected in the control sample either.

We will use this high statistics data sample to perform a last consistency check. The same PDF is used as the one that will be used to describe the proper-time difference in the $B^0 \to D^{*+}D^{*-}$ fit, although no angular analysis or separation into CP-even or CP-odd states is required. We use the standard resolution parameters except for the main resolution parameter which is obtained from the lifetime measurement of the control sample (see Section 5.5). The signal fraction is calculated on an event-by-event basis from the ΔE and $M_{\rm bc}$ distribution. The fitted distribution is shown in Figure 6.16, where the top plot shows the events which are tagged as $\overline{B^0}$ and the bottom plot the events tagged

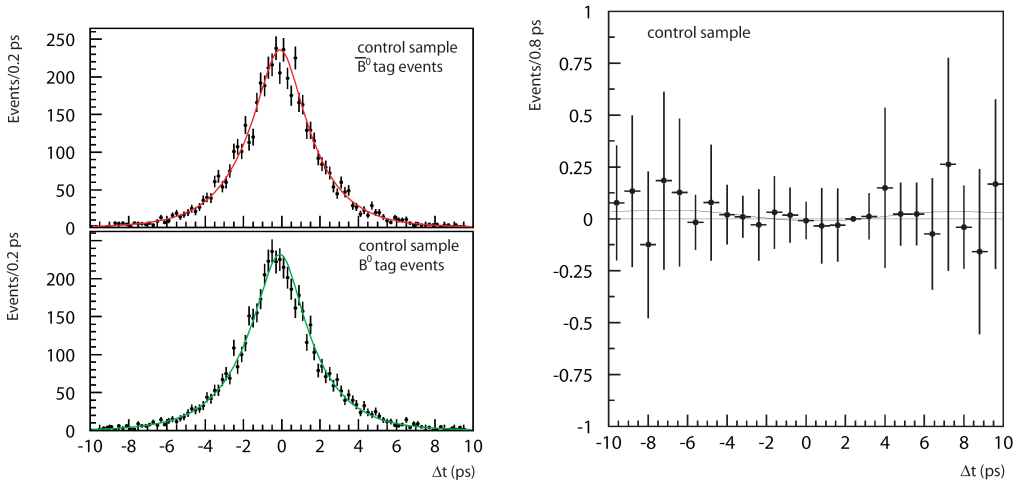


Figure 6.16: Left: The fitted proper-time distribution of well-tagged $\overline{B^0}$ events (top) and B^0 events (bottom) of the control sample. Right: The raw proper-time asymmetry distribution.

as B^0. The result of the CP measurement on the control sample is:

$$\begin{aligned} \mathcal{S}(B \to D_s^{(*)} D^{(*)}) &= -0.07 \pm 0.04, \\ \mathcal{A}(B \to D_s^{(*)} D^{(*)}) &= -0.02 \pm 0.03, \end{aligned}$$

which is within one standard deviation the expected result for $\mathcal{A}$ and within two standard deviations the expected result for $\mathcal{S}$.

6.5 CP measurement on the data sample

The data sample is the same one from which the lifetime measurement was performed i.e, subject to the more stringent requirements on the vertices. The PDF used to describe the data is explained in Section 6.1. The CP-violation parameters for $B^0 \to D^{*+}D^{*-}$ are extracted from an unbinned maximum likelihood fit of the proper-time difference using Eq. (6.6) and reads

$$\begin{aligned} \mathcal{S}'(B \to D^{*+}D^{*-}) &= -0.96 \pm 0.25, \\ \mathcal{A}(B \to D^{*+}D^{*-}) &= 0.15 \pm 0.13, \end{aligned}$$

where the errors are statistical only. The parameter $\mathcal{S}'$ again is defined as if all events are CP even, such that there is no dilution of the eigenvalue and $\mathcal{S}'$ corresponds, in the Standard Model and in absence of penguin diagrams to $\mathcal{S}' = -\sin 2\beta$. The statistical correlation of the two measurements is 10.7%. The result is still in the physical region as $\sqrt{\mathcal{S}'^2 + \mathcal{A}^2} = 0.97 \pm 0.25 < 1$. The statistical errors are close to the prediction from the fast MC study.

Figure 6.17 shows the fitted proper-time distribution of well-tagged $B^0 \to D^{*+}D^{*-}$ candidates ($r \geq 0.5$) for $q = +1$ (right) and $q = -1$ (left). The dotted area is the background contribution while the thick curves are the sum of signal and background PDF.

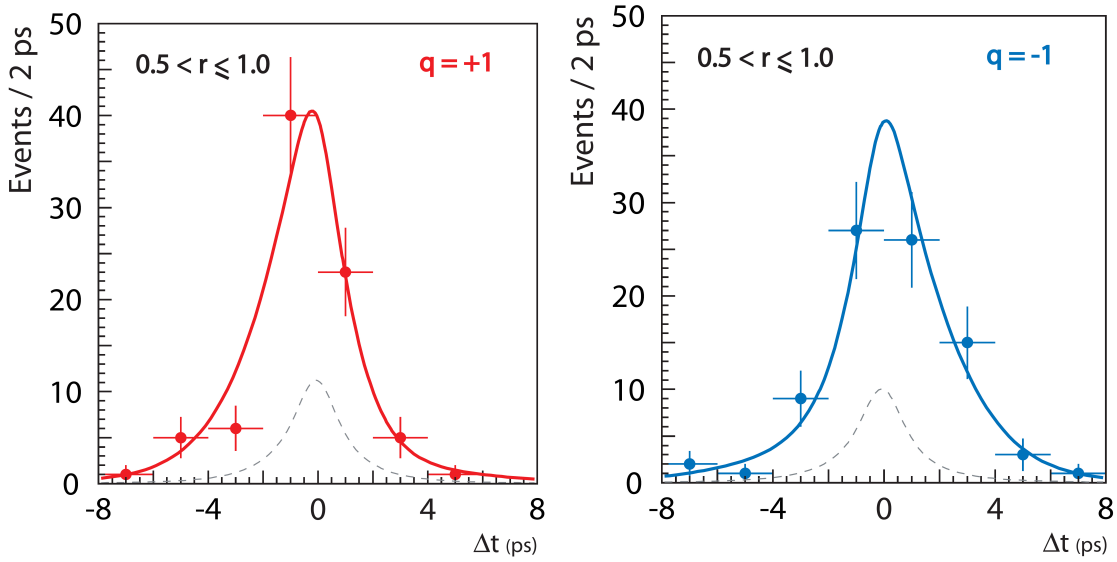


Figure 6.17: The fitted proper-time distribution of well-tagged $\overline{B^0}$ events (left) and B^0 events (right) of the data.

For visualization we also show the result superimposed on the same plot (see left plot in Figure 6.18) with an average background contribution.

We define the raw asymmetry in each Δt bin as $a = (N_+ - N_-)/(N_+ + N_-)$, where $N_+(N_-)$ is the number of observed candidates with $q = +1(-1)$, such that:

$$a = -\Delta w + (1 - 2w)\Big((1 - 2R_{\rm odd})\mathcal{S}' \sin(\Delta m \Delta t) + \mathcal{A}\cos(\Delta m \Delta t)\Big). \tag{6.8}$$

The diluted CP-violating parameters can be read directly from the plot. The value of the PDF at $\Delta t = 0$ shows $-\Delta w + (1 - 2w)\mathcal{A}$ and at $\Delta t \Delta m = \pi/2$, or roughly at $\Delta t = \pi$ shows $-q\Delta w + q(1 - 2w)\Big((1 - 2R_{\rm odd})\mathcal{S}'\Big)$. The raw fitted asymmetry is shown in the right plot of Figure 6.18.

We check the consistency of this result by fitting separately the SVD1 and SVD2 data, which provides a comparison with the results from the previous analysis. We obtain the following results for the fit on SVD1 data only:

$$\begin{aligned} \mathcal{S}'(B \to D^{*+}D^{*-}) &= -0.76 \pm 0.54, \\ \mathcal{A}(B \to D^{*+}D^{*-}) &= -0.20 \pm 0.33, \end{aligned}$$

which are compatible with the previous Belle analysis:

$$\begin{aligned} \mathcal{S}'(B \to D^{*+}D^{*-}) &= -0.75 \pm 0.56, \\ \mathcal{A}(B \to D^{*+}D^{*-}) &= -0.26 \pm 0.26. \end{aligned}$$

The fit on SVD2 data only gives:

$$\begin{aligned} \mathcal{S}'(B \to D^{*+}D^{*-}) &= -1.05 \pm 0.29, \\ \mathcal{A}(B \to D^{*+}D^{*-}) &= +0.24 \pm 0.15. \end{aligned}$$

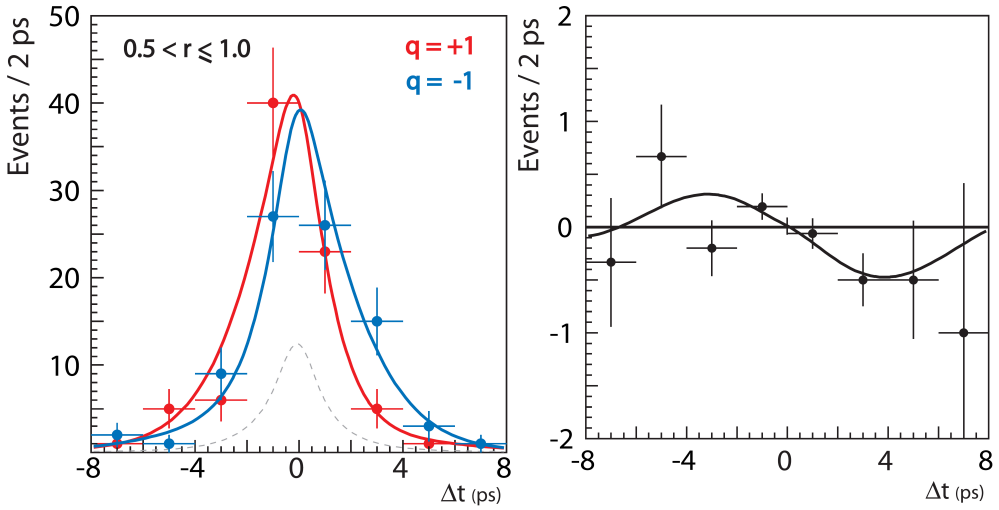


Figure 6.18: Left: the fitted proper-time distribution of well-tagged $B^0 \to D^{*+}D^{*-}$ candidates ($r \geq 0.5$) for $q = +1$ and $q = -1$. The gray curve is the background contribution while the thick curves are the sum of signal and background. Right: fitted raw CP asymmetry of the same candidates.

6.6 CP-violation significance

We calculate the significance of CP violation of our measurement in this section. The fitted CP-violation parameters are obtained from a maximum likelihood method where the likelihood function is defined as:

$$\mathcal{L}(\mathcal{A}, \mathcal{S}') = \prod_{i=1}^{N} \mathcal{P}(\Delta t_i; \mathcal{A}, \mathcal{S}'),$$

where the product is taken over all the measurements of Δt. The obtained fit results correspond to the values of $\mathcal{A}$ and $\mathcal{S}'$ which maximize $\ln \mathcal{L}$. For large samples, or for samples where the error distribution is Gaussian (which is approximately the case in this analysis, see Section 6.2.2.2), $\mathcal{L}$ has a Gaussian form and $\ln \mathcal{L}$ is parabolic. The numerically equivalent of a s-standard deviation error can be obtained from the contour of $\mathcal{L}(\mathcal{A}, \mathcal{S}')$ for a certain $\mathcal{A}, \mathcal{S}'$ such that:

$$\ln \mathcal{L}(\mathcal{A}, \mathcal{S}') = \ln \mathcal{L}_{\max} - s^2/2,$$

where $\mathcal{L}_{\max}$ is the value of $\ln \mathcal{L}$ at the solution point.

The plots in Figure 6.19 show $2(\ln \mathcal{L}(\mathcal{A}, \mathcal{S}') - \ln \mathcal{L}_{\max})$ in one dimension around its minimum, while fixing the other CP-violating parameter to its solution point, such that the vertical axis gives directly a measure of the number of standard deviations away from the minimum solution.

Figure 6.20 shows contours of the likelihood function. This two-dimensional function takes into account the correlations between the two parameters being plotted. The contours in the figure are plotted around the solution point (green) and then for the four first standard deviations away. The situation of no CP violation is marked with the red dot. To calculate the significance (S) of the measured CP violation we have to assume

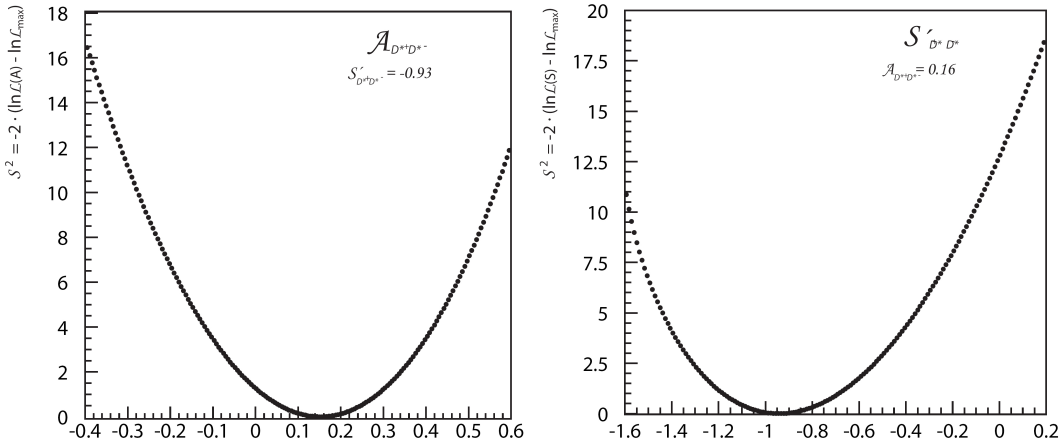


Figure 6.19: Likelihood scan as a function of $\mathcal{A}$ (left) and $\mathcal{S}'$ (right).

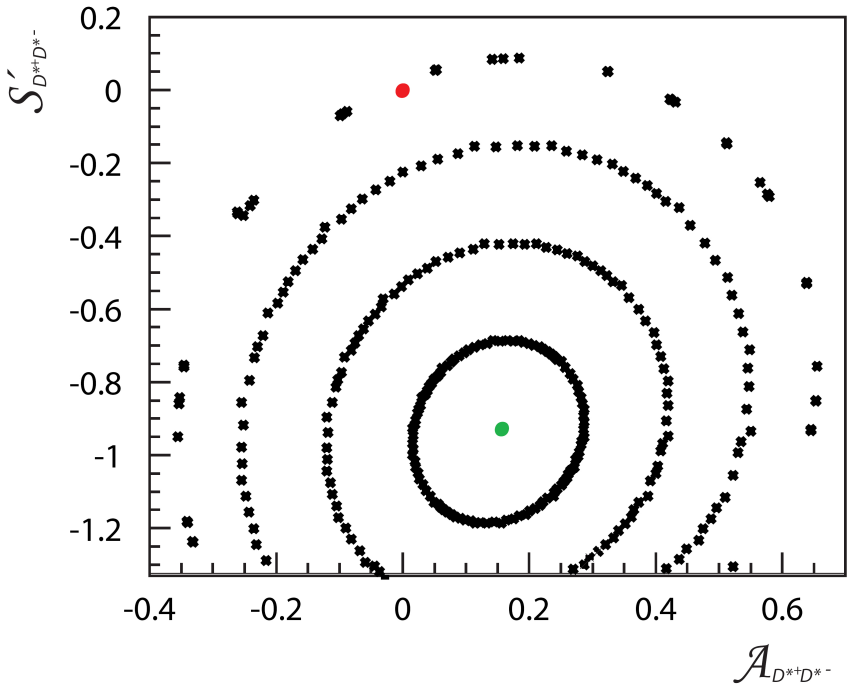


Figure 6.20: 1σ, 2σ, 3σ and 4σ contours of the likelihood function. The green dot represents the point of maximum likelihood, the red dot represents the situation of no CP violation.

that the likelihood profile follows a χ^2 distribution with 2 degrees of freedom ($\mathcal{A}$ and $\mathcal{S}$) constraint. We obtain an evidence for CP violation with a significance of 3.5.

6.7 Systematic study

The same method is used to estimate the systematic uncertainty of the CP-violating parameters, as described in the angular analysis (see Section 4.4): the fixed parameters are varied within their standard deviation and the different resulting deviations on the measured parameters are added in quadrature to obtain the total systematic error. Due to the complexity of the time-dependent PDF the effect of changing a parameter on the CP-violation parameters can be asymmetric, or different for $\mathcal{A}$ and $\mathcal{S}'$. The results are summarized in Table 6.7.

source	$\mathcal{A}$	$\mathcal{S}'$
signal purity	± 0.002	± 0.004
standard resolution function	± 0.004	$+0.000 \quad -0.102$
resolution from control sample	± 0.002	± 0.030
background shape	± 0.000	± 0.006
CP-odd fraction $R_\perp$	± 0.004	± 0.109
fit bias	± 0.010	± 0.031
Δm, τ_{B^0}	± 0.002	± 0.004
flavour tagging	± 0.011	± 0.020
vertex cuts	± 0.003	± 0.028
Δt fit range	± 0.010	± 0.004
peaking background	± 0.010	$+0.000 \quad -0.027$
tag side interference	$\pm$ 0.034	$\pm$0.007
total	± 0.040	$+0.123 \quad -0.162$

Table 6.8: Systematic errors on the CP-violating parameters for $B^0 \to D^{*+}D^{*-}$ decays.

The first source examined is the signal yield. Varying this value also changes the composition of the $\cos\theta_{\rm tr}$ distribution. Therefore both the yield and $R_\perp$ are varied in a correlated way. The number of signal events is varied by ± 1 standard deviation and $R_\perp$ by ± 0.0031. The contribution of the resolution function is estimated by simultaneously varying all the standard parameters in a correlated manner by $\pm 1\sigma$ while the main resolution parameter, obtained from the control sample, is kept as a constant. Next the main resolution parameter is varied while the standard parameters are kept constant.

The parameters describing the background shape are varied by one standard deviation. The largest contribution comes from the $R_\perp$ fraction which only affects ${\mathcal{S}'}_{D^{*+}D^{*-}}$. Varying the resolution parameters moves $\mathcal{S}_{D^{*+}D^{*-}}$ further away from zero. The results from the linearity study with the fast MC are also taken into account, the bias of the CP violating parameters at the measured values being included in the table. The two constants in Eq. (6.4), Δm and τ_{B0} are varied around their world averages [72].

Systematic errors due to wrong tag fractions are estimated by varying the parameters $w(r)$ and $\Delta w(r)$ in each r region by their $\pm 1\sigma$ errors. The vertex quality cut is changed to $\xi^2 < 125$ and the effect is added in the table. The Δt fit range is been changed from $|\Delta t| < 70\,\text{ps}$ to $|\Delta t| < 10\,\text{ps}$. A peaking background contribution is added with no CP

violation. Finally a tag side interference uncertainty is added. The different sources are added in quadrature to yield ± 0.04 for $\mathcal{A}$ and $^{+0.12}_{-0.16}$ for $\mathcal{S}'$.

6.8 Discussion and conclusion

We can conclude that the measured CP-violation parameters in the $B^0 \to D^{*+}D^{*-}$ decay are:

$$\begin{aligned} \mathcal{A}_{D^{*+}D^{*-}} &= 0.15 \pm 0.13(\text{stat}) \pm 0.04(\text{syst}), \\ \mathcal{S}'_{D^{*+}D^{*-}} &= -0.96 \pm 0.25(\text{stat})^{+0.12}_{-0.16}(\text{syst}), \end{aligned}$$

using 657×10^6 $B\overline{B}$ events. Evidence of CP violation is obtained with a $3.5\,\sigma$ significance. These measurements are consistent with and supersede the previous Belle result [58]. Figure 6.21 show the results of the CP violating parameters measured by BaBar (as of Spring 2008) and Belle (the preliminary result as shown at conferences during Spring 2008). For the direct CP violation the notation $\mathcal{C} = -\mathcal{A}$ is used. The BaBar result is obtained using 617 ± 33 signal events. As can be seen from the figures both measurements are compatible with each other. The smaller systematic error of BaBar with respect to our final result is partially due to their analysis method. BaBar performed a three dimensional fit of the proper-time difference, M_{bc} and $\cos\theta_{\text{tr}}$ distribution, such that the systematic errors due to the signal purity and the CP-odd fraction are accounted for in the statistical error. The bottom line shows the average result provided by the HFAG group [72].

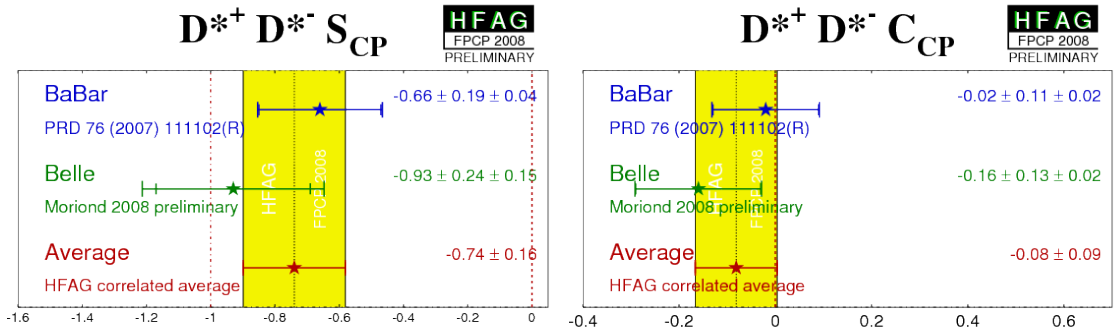
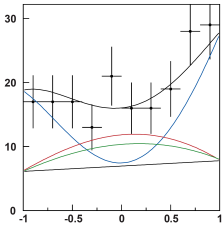


Figure 6.21: The BaBar and Belle preliminary results of the CP-violating parameters, $\mathcal{S}'$ (left) and $\mathcal{C} = -\mathcal{A}$ (right) as of Spring 2008. The bottom line shows the average of both results.

Chapter 7

Repetition of analysis on SVD1 data



An analysis of the first $140\,\mathrm{fb}^{-1}$ of data is repeated in this chapter; from the extraction of the yield to the CP-odd fraction and the CP fit of the proper-time distribution. The aim is to employ as much as possible the same procedure as in the previous Belle analysis so that a comparison of the results can be made.

THE first step of this analysis was to reproduce the previous Belle result of $B^0 \to D^{*+}D^{*-}$ on the first $140\,\mathrm{fb}^{-1}$ data sample [58], which corresponds to the data taken with the SVD1 detector configuration. The cross-check presented in this chapter does not intend to show a more precise measurement on the SVD1 data, but is performed to reassure that our procedure is correct and consistent.

As this was the first work performed for this analysis, we have taken the opportunity to study certain aspects of the analysis in more detail, while on other occasions we have simplified the procedure not to get overwhelmed with technical issues. The previous analysis was documented in [57]. No matter how complete that note is, there are always some details that are not mentioned. It can also be that older software is not available anymore so that reproducing exactly the previous analysis today becomes unrealistic. A result that is fairly close to the previous result will therefore be accepted. We will also not perform any systematics study on these results, the errors quoted in this chapter are statistical only.

This chapter is organized in the same way as the main analysis. The first section will describe the decay channels which are reconstructed and the selection cuts applied to the data. A comparison with the previous analysis will be made for the reconstruction efficiency and the yield. In the next section we repeat the angular analysis on SVD1 data with the procedure employed in the previous analysis. The proper-time distribution and the resolution function is studied in more detail. Finally in the last section a repetition of the CP fit on SVD1 data is performed.

7.1 Signal reconstruction and yield

7.1.1 Selection and reconstruction efficiency

The analysis is repeated using the selection cuts and fitting procedures which are as close as possible to those of the previous analysis. The cuts are analogue but not completely identical to the criteria used for the analysis in the full data set.

Because this check was performed at an early stage of our analysis, only a subset of the subdecays used in the previous analysis are reconstructed, namely those that do not suffer from too much background (see Table 7.2). Therefore, if we are able to reproduce the efficiency of the previous analysis, we expect to extract only a fraction of the total signal previously observed in the SVD1 data set, estimated to be 85%.

Table 7.1 summarizes the selection criteria applied in the previous analysis and the ones used now for the cross-check analysis. The variables that are used to select the signal events have been described in Section 3.3. The wink symbol ($\surd$) means that the same cut is used as in the previous analysis. For comparison we also included the selection criteria of the final analysis.

The generated signal MC samples are obtained using the procedure described in Section 3.2.1 with the difference that the B^0 meson decays only to $D^{*+}D^{*-}$ (and nothing else) while the $\overline{B^0}$ decays inclusively. Other technicalities that were used in the previous analysis and are therefore also repeated here include:

- the generated events have no polarization;
- the generated D^* mass has a width of $0.01\,\mathrm{MeV}$;
- each decay mode is generated and reconstructed separately, avoiding possible cross feed;
- the number of reconstructed events are counted in the region $|\Delta E| < 0.1\,\mathrm{GeV}$ and $M_{\mathrm{bc}} < 5.27\,\mathrm{GeV/c^2}$.

These four features are not implemented in the final analysis, in order to obtain a more realistic reconstruction efficiency.

The ratio of the obtained reconstruction efficiency with respect to the reconstruction efficiency of the previous analysis is shown in Table 7.2, for the different sub modes accompanying the $\overline{D^0} \to K^+\pi^-$ decay. We can conclude that when applying selection criteria which barely differ from the previous analysis, we obtain the same reconstruction efficiency on the signal MC.

7.1.2 Data yield

For the cross-check analysis we directly used existing skims, which were made for the previous analysis. The selection criteria for these skims are similar to those described in Section 3.7.1.

In the previous analysis $129.5^{+12.9}_{-12.0}$ signal events were found in experiments 07$-$27. We know now that the reconstruction efficiency in this cross-check analysis is similar to that of the previous analysis, but we only expect to find 85% of the events, due to the smaller number of selected decay channels. We should therefore find $129.5 \times 0.85 = 110$ events.

previous analysis	cross-check analysis	final analysis
$dr < 0.4$ cm if no SVD hits	$dr < 2.0\,\mathrm{cm}$ [1]	$dr < 2.0\,\mathrm{cm}$ [1]
$\|dz\| < 4.0$ cm if no SVD hits	$\|dz\| < 4.0\,\mathrm{cm}$ [1]	$\|dz\| < 4.0\,\mathrm{cm}$ [1]
$dr < 6.0$ cm if SVD hits	$dr < 2.0\,\mathrm{cm}$ [1]	$dr < 2.0\,\mathrm{cm}$ [1]
$\|dz\| < 5.0$ cm if SVD hits	$\|dz\| < 4.0\,\mathrm{cm}$ [1]	$\|dz\| < 4.0\,\mathrm{cm}$ [1]
$R - 2 < 0.4$	$\surd$	$\surd$
Kaon id. for 2 prong $\mathcal{L}_{K/\pi} > 0.1$	$\surd$	$\surd$
Kaon id. for 3 or 4 prong $\mathcal{L}_{K/\pi} > 0.2$	> 0.1 [1]	> 0.6
Pion id. $\mathcal{L}_{K/\pi} < 0.9$	$\surd$	$\surd$
no π_{slow}-id.	$\surd$	$\surd$
π^0: $119\ \mathrm{MeV/c^2} < \mathrm{M}_{\gamma\gamma} < 146\ \mathrm{MeV/c^2}$	$\surd$	$\surd$
π^0: $E_\gamma > 0.03$ GeV	$\surd$	$\surd$
$p_{\pi^0} > 0.1$ GeV/c	$\surd$	$\surd$
$K_\mathrm{S} \to \pi^+\pi^-$: goodKs [61]	$\surd$	$\surd$
MV-fit of all D candidates	V fit for D's, MV fit for ΔE [2]	MV-fit of all
π_{slow} recalculated with IP constraint	$\surd$	$\surd$
π_{slow} recalculated with associated B vertex	- [3]	- [3]
2-prong: D mass within 6σ of nom. value	$\surd$	$\surd$
3- or 4-prong: D mass within 3σ	$\surd$	$\surd$
$M(D^*) - M(D^0) < 3\ \mathrm{MeV/c^2}$	$\surd$	$\surd$
$M(D^*) - M(D^+) < 2.25\ \mathrm{MeV/c^2}$	$\surd$	$\surd$
$D^{0,+}$: $\chi^2_{\mathrm{vertex}} < 40$	$\surd$	$\surd$
$\chi^2_{\mathrm{mass(D,D^*)}} < 5$	Not [4]	Not [4]
Best candidate based on $\chi^2_{\mathrm{mass(D,D^*)}}$	$\surd$	$\surd$
small signal region: $\mid \Delta E \mid < 40$ MeV	< 30 MeV	< 40 MeV
large signal region: $\mid \Delta E \mid < 200$ MeV	< 100 MeV	< 200 MeV
small signal region: $M_{\mathrm{bc}} < 5.27\,\mathrm{GeV/c^2}$	$\surd$	$\surd$
large signal region: $M_{\mathrm{bc}} < 5.2\,\mathrm{GeV/c^2}$	$\surd$	$\surd$

[1] : To unify the cuts.
[2] : To avoid tails that appear when applying a MV fit to the D mesons. As it is only a small effect the MV-fit is anyhow applied in the full analysis.
[3] : For simplicity (this cut is not used in the full analysis either as it is not a standard cut)
[4] : To avoid a cut on a variable which is used to choose the best candidate.

Table 7.1: Selection criteria used in the previous Belle analysis, in the cross-check analysis and in the final analysis. The differences are explained or justified in the footnotes.

The statistical error on this expectation is the standard deviation of a binomial distribution with probability of $p = 85\%$ and $N = 130$ events:

$$\sigma = \sqrt{Np(1-p)} = \sqrt{130 \times 0.85 \times (1 - 0.85)} = 4. \tag{7.1}$$

We thus expect to find 110 ± 4 signal events in the SVD1 data sample.

Figure 7.1 shows the scatter plot of ΔE versus M_{bc}. A two-dimensional maximum likelihood fit is performed on this data with the same PDF as described in Eq. (3.3) and (3.4). The mean and width of ΔE and M_{bc} are fixed to the corresponding value from the signal MC, as well as the fraction between the two Gaussian functions used to describe the ΔE

		efficiency ratio (%)		
1	$D^0 \to K^-\pi^+$	99.4	$\pm$	3.2
2	$D^0 \to K^-\pi^+\pi^0$	108.8	$\pm$	3.3
3	$D^0 \to K^-\pi^+\pi^+\pi^-$	96.7	$\pm$	4.0
4	$D^0 \to K_S\pi^+\pi^-$	83.2	$\pm$	4.9
5	$D^0 \to K^+K^-$	89.4	$\pm$	10.5
1	$D^+ \to K^-\pi^+\pi^+$	103.8	$\pm$	5.4

Table 7.2: Ratio of reconstruction efficiencies of this cross-check analysis with respect to the previous analysis.

distribution. The parameters extracted from the signal MC, which are used to fit the signal shape of the data are summarized in Table 7.3. Note that more parameters are fixed than in the full analysis as we have smaller statistics in SVD1 data. A different fitting region of $|\Delta E|$ is used in the cross-check analysis compared to the previous analysis or the analysis on the full data set, i.e. $-100\,\mathrm{MeV} < \Delta\mathrm{E} < 100\,\mathrm{MeV}$. From this fit we obtain 106.6 ± 12.0 signal events, in good agreement with expectations. The sliced projections in the small signal region on each of the variables is shown in Figure 7.2.

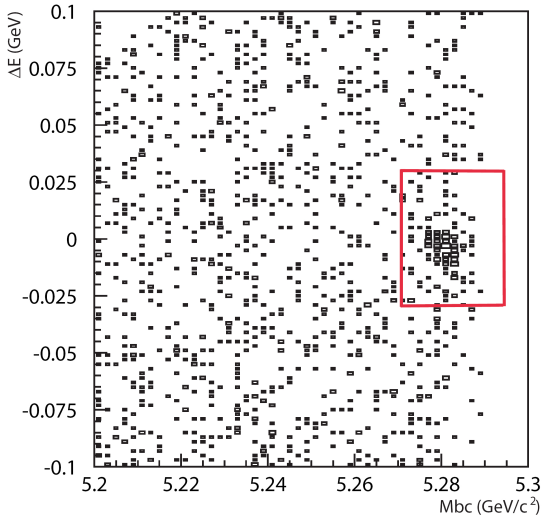


Figure 7.1: $\Delta E - M_{\mathrm{bc}}$ distribution of SVD1 data.

7.2 Angular analysis

In the previous analysis of SVD1 data, a two-dimensional binned fit of the distribution of $\cos\theta_{\mathrm{tr}}$ and $\cos\theta_1$ was performed . In order to do this the correlation between the two angles, the reconstruction efficiency and the background shape need to be intensively studied. However the purpose of repeating this analysis is to briefly check that the central

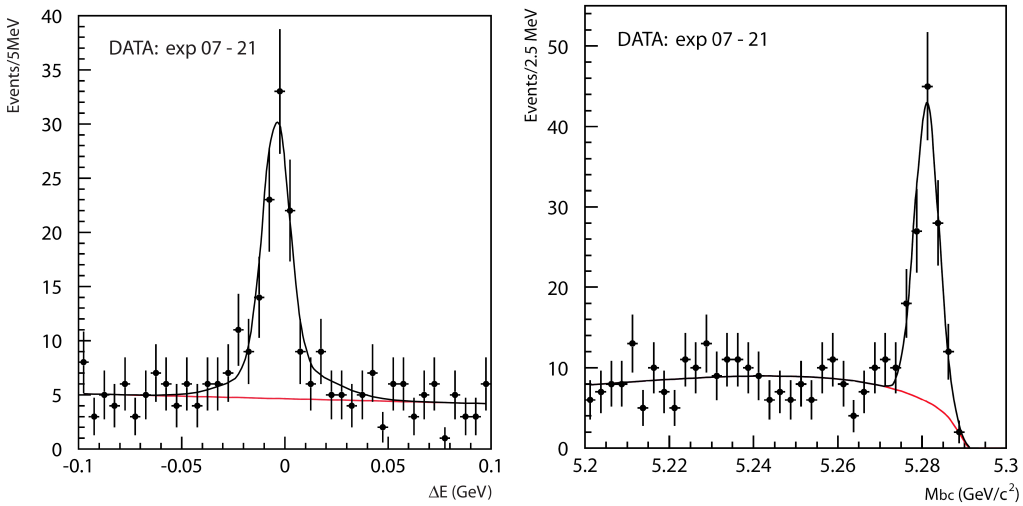


Figure 7.2: Projections of the fitted ΔE and $M_{\rm bc}$ distribution of SVD1 data.

	ΔE			$M_{\rm bc}$	
$\mu_{\rm main}$	-0.3 ± 0.1	MeV	μ	5.2795	$\mathrm{GeV/c^2}$
$\sigma_{\rm main}$	6.4 ± 0.1	MeV	σ	2.81 ± 0.03	$\mathrm{MeV/c^2}$
f	0.68 ± 0.01				
$\mu_{\rm tail}$	-4.2 ± 0.6	MeV			
$\sigma_{\rm tail}$	20.9 ± 0.9	MeV			

Table 7.3: The fitted signal parameters of the 2D fit of ΔE and $M_{\rm bc}$ on signal MC.

value of $R_\perp$ and R_0 can be reproduced, rather than optimizing the statistical error. Therefore a similar but simpler method is chosen: fitting both angles simultaneously but not in two dimensions.

Because the angular distributions of $\cos\theta_{\rm tr}$ and $\cos\theta_1$ are used we can extract $R_\perp$ and R_0. The three polarizations are modeled separately using the shapes determined from the signal MC as has been done for the full analysis (see Section 4.1.2). However now we do not need to define a CP-even shape. The angular distributions are then fitted simultaneously but each one-dimensionally. During the simultaneous fit $R_\perp$ and R_0 are common parameters. The signal PDF of each of the two angles reads:

$$\begin{aligned}\mathcal{P}^{\rm tr}_{\rm sig}(\cos\theta_{\rm tr}) &= R_0\mathcal{P}^{\rm tr}_0(\cos\theta_{\rm tr}) + R_\perp\mathcal{P}^{\rm tr}_\perp(\cos\theta_{\rm tr}) \\ &\quad +(1-R_0-R_\perp)\mathcal{P}^{\rm tr}_\parallel(\cos\theta_{\rm tr}), \qquad (7.2)\\ \mathcal{P}^{1}_{\rm sig}(\cos\theta_1) &= R_0\mathcal{P}^{1}_0(\cos\theta_1) + R_\perp\mathcal{P}^{1}_\perp(\cos\theta_1) \\ &\quad +(1-R_0-R_\perp)\mathcal{P}^{1}_\parallel(\cos\theta_1), \qquad (7.3)\end{aligned}$$

where $\mathcal{P}_0$, $\mathcal{P}_\perp$ and $\mathcal{P}_\parallel$ are the probability density function of $\cos\theta_{\rm tr}$ or $\cos\theta_1$. As in the previous analysis, the shape of the background events is not left as a free parameter in the fit but is determined beforehand. This is done by fitting the data in the sideband region ($M_{\rm bc} < 5.27\,{\rm GeV/c^2}$) with a constant for $\cos\theta_{\rm tr}$ and a linear function for $\cos\theta_1$. The fit parameters are summarized in Table 7.4 and the fitted $\cos\theta_{\rm tr}$ and $\cos\theta_1$ distributions are shown in Figure 7.3

	$\cos\theta_{\rm tr}$	$\cos\theta_1$
slope		0.12 ± 0.06
χ^2	0.51	0.79

Table 7.4: Fitted parameters of the $\cos\theta_{\rm tr}$ and $\cos\theta_1$ distributions of the SVD1 data sideband region.

The total probability density function is constructed from the signal and background PDFs:

$$\mathcal{P}_{\rm tot} = f_{\rm sig}\mathcal{P}_{\rm sig} + (1-f_{\rm sig})\mathcal{P}_{\rm bkg}. \qquad (7.4)$$

In this cross-check analysis $f_{\rm sig}$ is kept as a constant corresponding to the average signal fraction in the small signal region of the SVD1 data. This fraction is obtained from the yield fit on the SVD1 data, shown in Section 7.1. The previous analysis however, used an event-by-event signal fraction. The fit in this analysis is performed on the small signal region, which is justified because the background shape is determined beforehand from the data of the sideband regions.

The fitted angular distributions of the SVD1 data are shown in Figure 7.4 and to be compared with those published in the previous analysis shown in Figure 7.5. The fitted R_0 and $R_\perp$ fractions are:

$$\begin{aligned}R_\perp &= 0.27 \pm 0.10,\\ R_0 &= 0.54 \pm 0.10,\end{aligned}$$

while the previous analysis obtained:

$$\begin{aligned}R_\perp &= 0.19 \pm 0.08,\\ R_0 &= 0.57 \pm 0.08,\end{aligned}$$

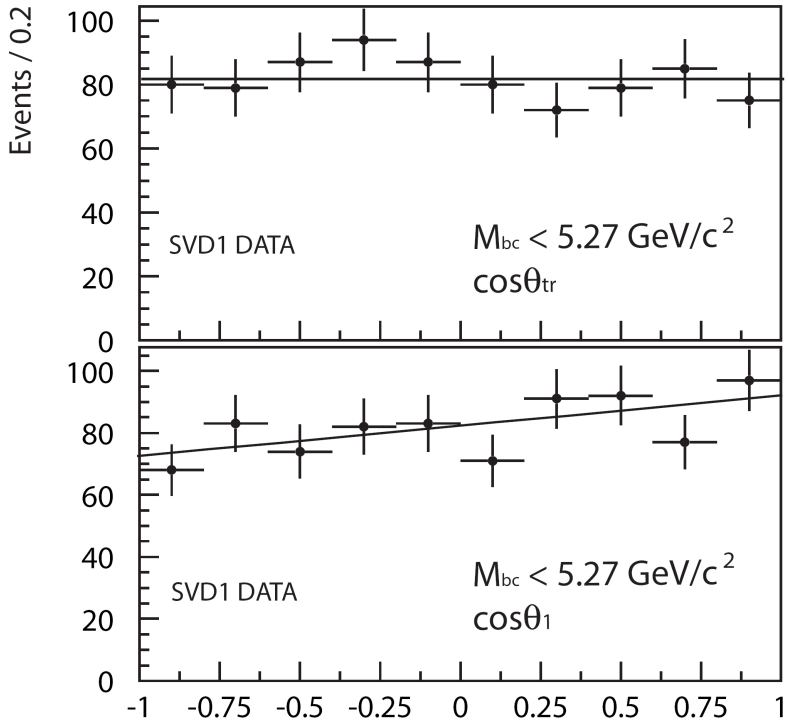


Figure 7.3: Fitted $\cos\theta_{\rm tr}$ (top) and $\cos\theta_1$ (bottom) distributions of the sideband region of SVD1 data.

However, we cannot interpret this one sigma deviation as usual as the results are obtained from the same data set, which means that the results are correlated. Still considering the fact that this analysis is done with a simplified fitting method and that different software versions are used, we conclude that the result is satisfying enough.

7.3 Lifetime measurement

Neither the resolution parameters nor the relevant software versions used in the previous analysis are still available. Therefore the cross-check of the lifetime is performed with the software version available at the time of this test (ICHEP04). Note that for the total analysis even more recent software and resolution parameters are used (ICHEP06). In this cross-check analysis, the B^{CP} vertex is reconstructed from the two D mesons, as in the full analysis. However the IP tube constraint is only implemented in the very last software version. To be compatible with the previous analysis we use the IP profile as a constraint instead of the IP tube. The vertex of the tag-side is obtained in the same way as explained in the full analysis. The data are again subject to more stringent conditions on the vertex quality. We require $\Delta t < 70\,\mathrm{ps}$ and $\chi^2/ndf < 100$, which leaves us with 108 ± 14 signal events.

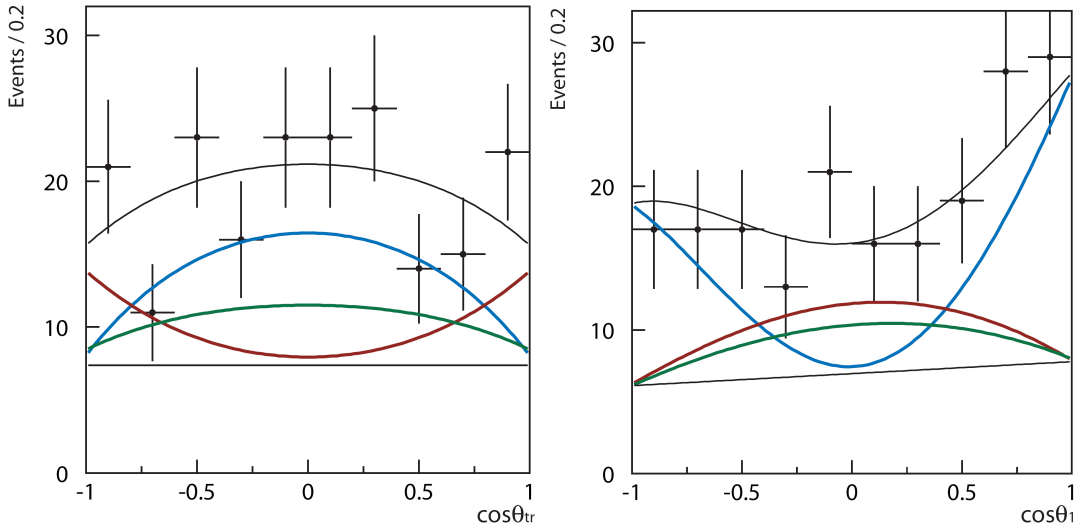


Figure 7.4: Fitted $\cos\theta_{\rm tr}$ and $\cos\theta_1$ distributions of the SVD1 data. The red, blue and green curves represent the contributions of the A_0, $A_\perp$ and $A_\parallel$ signal amplitudes, shown above the background level.

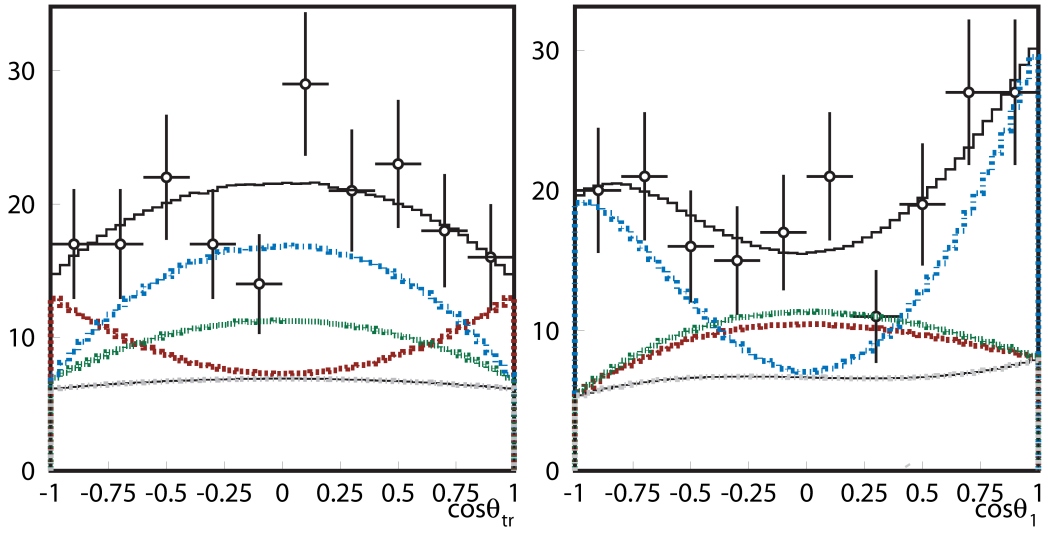


Figure 7.5: Fitted $\cos\theta_{\rm tr}$ and $\cos\theta_1$ distribution of the SVD1 data in the previous analysis. The red, blue and green curves represent the contributions of the A_0, $A_\perp$ and $A_\parallel$ signal amplitudes, shown above the background level.

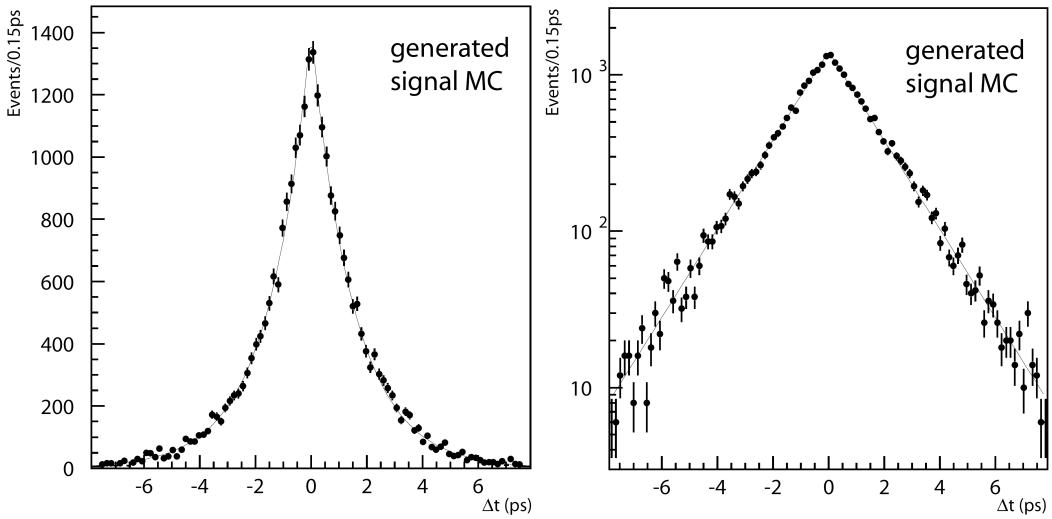


Figure 7.6: Fitted true proper-time distribution of the signal MC.

7.3.1 Signal MC

Figure 7.6 shows the true proper-time distribution of the signal MC. When fitting this distribution with just an exponential function, we obtain the following lifetime:

$$\tau_{\rm gen} = 1.532 \pm 0.009\,{\rm ps}, \tag{7.5}$$

which corresponds to the generated lifetime of $1.53\,{\rm ps}$.

The reconstructed proper-time distribution of the signal MC events is shown in Figure 7.7. This distribution is fitted with an exponential function convoluted with the resolution function, and a wide Gaussian is added to model outlier events. The form of the PDF for the signal MC events is identical to the one employed in the full analysis. The resolution parameters are set to the "standard" ones and the extracted lifetime is

$$\tau_{\rm MC,stand} = 1.60 \pm 0.02\,{\rm ps}, \tag{7.6}$$

Because this result is more than three sigma away from the true value some effort is put into looking at these resolution parameters.

Only the D mesons and the beam profile are used to obtain the B^{CP} vertex, therefore there is a large resemblance with the vertex properties of the $B^0 \to D^+D^-$ study. In the latest $B \to D^+D^-$ analysis from Belle [59], the resolution parameters are determined specifically for that decay mode by fitting the residual distribution. Due to the similarity of the two analyses, we also studied the proper-time fit of the signal MC using the resolution parameters of the $B \to D^+D^-$ analysis, in order to understand the size of the impact of these parameters on the extracted lifetime value. The result, using the parameters from the $B \to D^+D^-$ analysis is

$$\tau_{\rm MC,D^+D^-} = 1.51 \pm 0.02\,{\rm ps} \tag{7.7}$$

and the fitted distribution is shown in Figure 7.8.

Because of a rather large dependence on the resolution parameters, we studied the resolution distribution in more detail by looking, for example, at the vertex residual distribution of the B^{CP} meson (see left plot of Figure 7.9). This distribution is constructed from

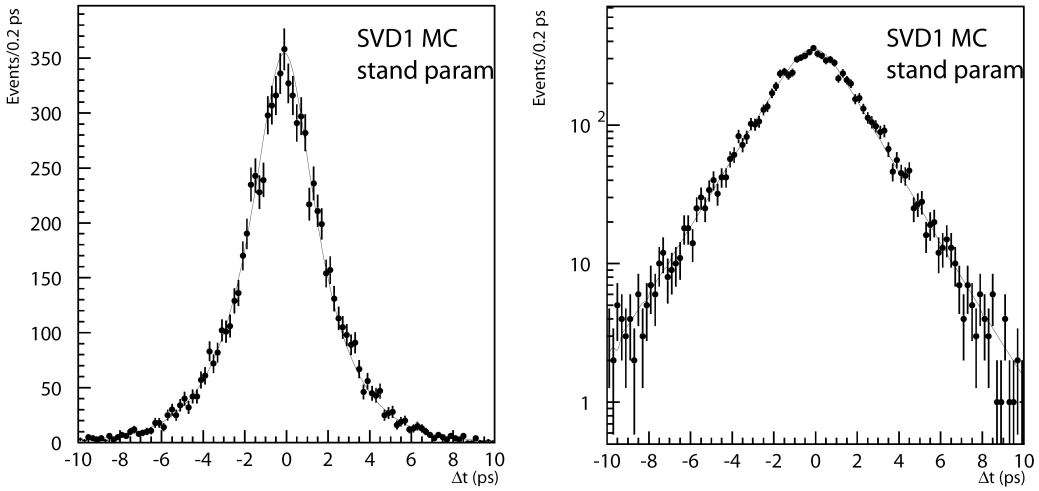


Figure 7.7: Fitted proper-time distribution of the signal MC using the standard resolution function.

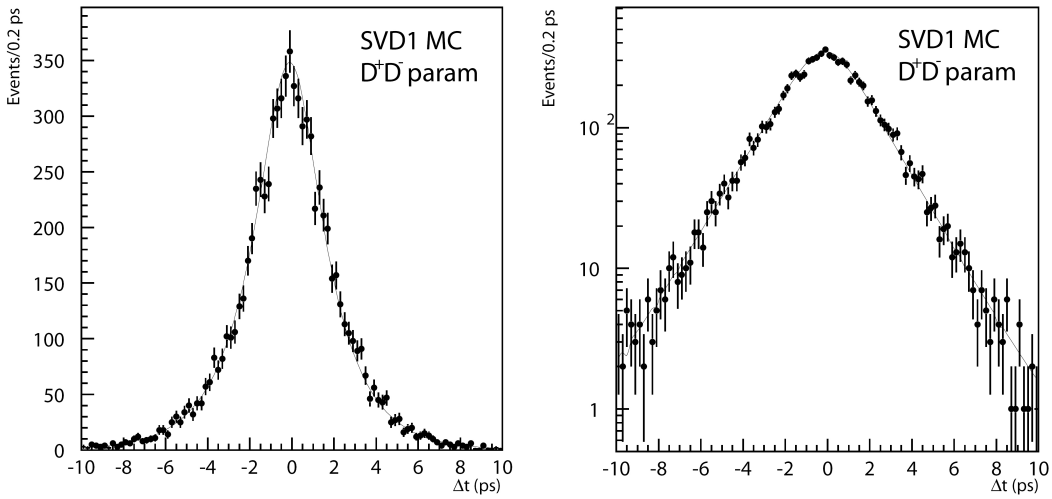


Figure 7.8: Fitted proper-time distribution of the signal MC using the resolution parameters of the $B \to D^+D^-$ analysis.

the difference between the true and reconstructed proper times, i.e. $\Delta t = t_{\rm gen}^{CP} - t_{\rm rec}^{CP}$. In this plot we have merely superimposed the resolution function on this distribution using the standard parameters. In the right plot of Figure 7.9 the resolution distribution is fitted with the detector resolution function which is explained in Section 5.2:

$$\begin{aligned} \mathcal{R}_{\rm det}(\Delta t) &= (1 - f_{\rm tail})\mathcal{G}(\Delta t, (s_{\rm main} + \xi s_{\rm tail})\sigma_{\rm vtx}) \\ &+ f_{\rm tail}\mathcal{G}(\Delta t, s_{\rm tlml}(s_{\rm main} + \xi s_{\rm tail})\sigma_{\rm vtx}). \end{aligned} \tag{7.8}$$

The parameters in red are obtained from the fit. The results of the fit are summarized in the first column of Table 7.5. For comparison the standard parameters for SVD1 and the parameters from the $B \to D^+D^-$ analysis are also displayed in the same table.

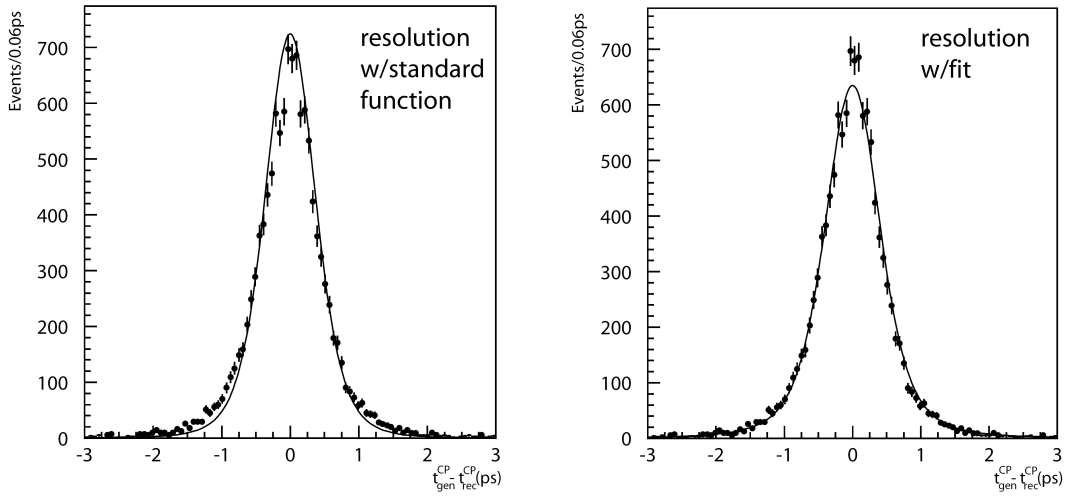


Figure 7.9: Vertex residual distribution with standard resolution function (left) and fitted resolution function (right).

	$B^0 \to D^{*+}D^{*-}$	Standard	$B^0 \to D^+D^-$
$s_{\rm main}$	1.07 ± 0.01	1.036	1.16 ± 0.02
$s_{\rm tail}$	0.13 ± 0.01	0.07	0.19 ± 0.02
$f_{\rm tail}$	0.05 ± 0.01	0	0.07 ± 0.01
$s_{\rm tlml}$	2.92 ± 0.13	0	3.54 ± 0.17

Table 7.5: Detector resolution parameters obtained from the residual distribution (left), the standard parameters for SVD1 MC (center) and the parameters obtained from the $B \to D^+D^-$ analysis (right).

When now fitting the signal MC proper-time distribution with the resolution parameters obtained from the residual distribution we obtain a lifetime of

$$\tau_{\rm MC,stand} = 1.56 \pm 0.02\,{\rm ps}, \tag{7.9}$$

The reason why this does not give exactly the generated lifetime is (besides fluctuations from the simulation) also due to the fact that only the detector resolution parameters have been tuned by ourselves. The other parameters of the resolution function are still the

standard ones. In the previous analysis the standard resolution parameters were used but scale factors were obtained from similar analyses (eg. $B^0 \to D^{*+}D^-$) and from a control sample, however the exact value of these parameters is not stated in the Belle note.

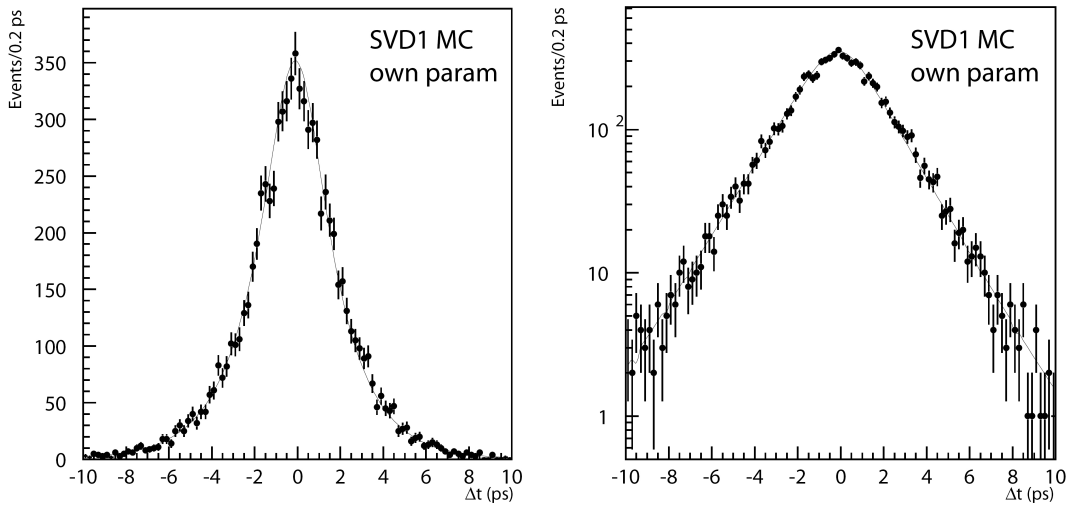


Figure 7.10: Fitted proper-time distribution of the signal MC using the resolution parameters obtained from the residual distribution.

7.3.2 Proper-time distribution for background events

The probability function used to describe background events in SVD1 data is obtained by fitting events in the sideband region, i.e, $M_{\rm bc} < 5.27\,{\rm GeV/c^2}$. The function used to parametrize these background events is a simplified version of what is described in the previous analysis, where a distinction was made between one-track and multi-track vertices. However here we use

$$\mathcal{P}(\Delta t) \;=\; \left((1-f_\delta)\frac{1}{2\tau_{\rm bkg}}e^{-\Delta t/\tau_{\rm bkg}} + f_\delta\delta(\Delta t)\right) \otimes \mathcal{G}(\mu, \sigma^{\rm bkg}). \tag{7.10}$$

The red parameters are obtained by fitting the sideband events. The results are shown in Table 7.6 and the fitted distributions are shown in Figure 7.11.

f_δ	0.30 ± 0.18
$\tau_{\rm bkg}$	$1.38 \pm 0.18\,{\rm ps}$
$\mu_{\rm bkg}$	$-0.05 \pm 0.05\,{\rm ps}$
$\sigma_{\rm bkg}$	$0.72 \pm 0.19\,{\rm ps}$

Table 7.6: Fitted parameters of the sideband events of SVD1 data.

7.3.3 Lifetime fit

The lifetime of the B^0 meson is now fitted using only SVD1 data. This is not a new measurement so we can use it to study the effect of the resolution parameters on the

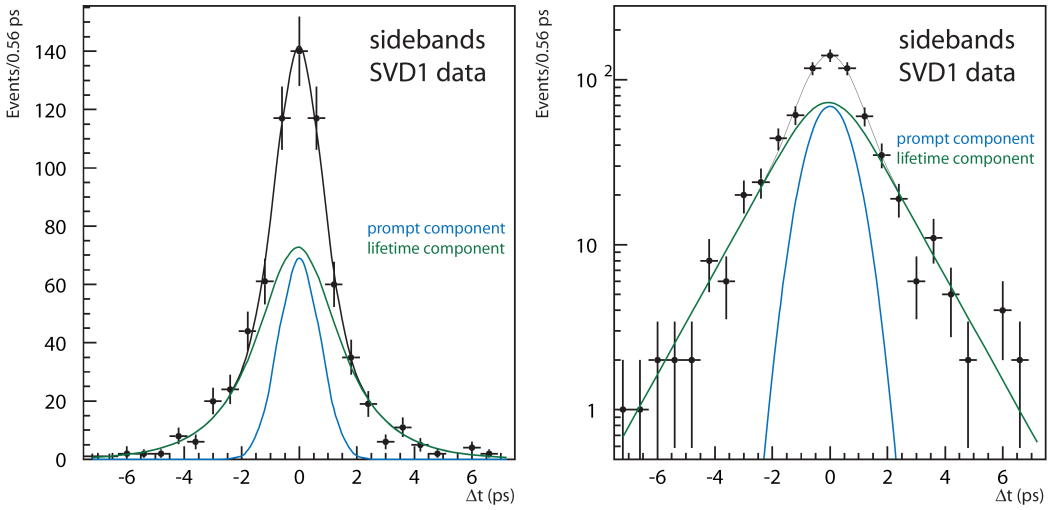


Figure 7.11: Fitted proper-time distribution of the sideband events of SVD1 data.

data. Figures 7.12, 7.13 and 7.14 show the fitted lifetime distribution using the standard resolution parameters, our own obtained resolution parameters and the parameters from the $B \to D^+D^-$ analysis, respectively.

In the Belle note of the previous $B \to D^{*+}D^{*-}$ analysis, a lifetime of $\tau_{\mathrm{SVD1-old}} = 1.7 \pm 0.2\,\mathrm{ps}$ is measured.

The measured B lifetimes for the three different settings are shown in Table 7.7. The variations between the results are well below one standard deviation. The result is compatible with the value $\tau_{\mathrm{SVD1-old}} = 1.7 \pm 0.2\,\mathrm{ps}$ obtained in the previous analysis. Remarkably we obtain a smaller statistical error than in the previous analysis. This is probably due to improvements in the reconstruction algorithms of the general Belle software since the time of the previous analysis.

	Standard	$B \to D^+D^-$	$B \to D^{*+}D^{*-}$
τ_{SVD1}	1.68 ± 0.14 ps	1.60 ± 0.16 ps	1.64 ± 0.15 ps

Table 7.7: Fitted lifetime in SVD1 data using different resolution parameters.

7.4 CP measurement

The repetition of the CP fit is performed in a much simpler way than what is done in the full analysis or than what was done in the previous analysis. We rely on many software functions of the Belle software library such as the tagging algorithm and the algorithm which determines the tag vertex. The software versions used in the previous analysis are not available anymore which makes a detailed comparison study even more challenging for this CP measurement.

More effort is spent on testing the software packages, the normalization procedure and plotting of the fitted PDF on the data. A fast MC is used to check that the basic fitting

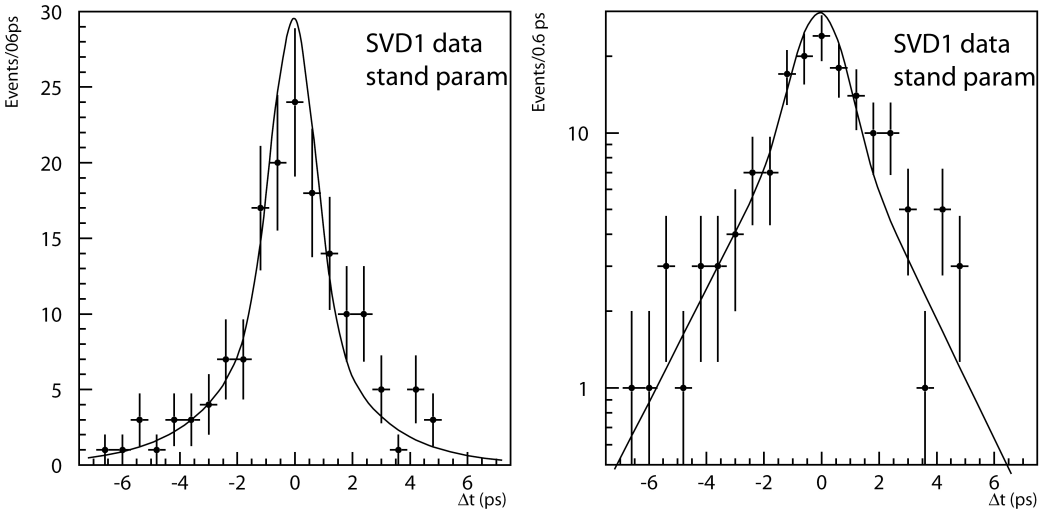


Figure 7.12: Fitted proper-time distribution using standard resolution parameters, on SVD1 data.

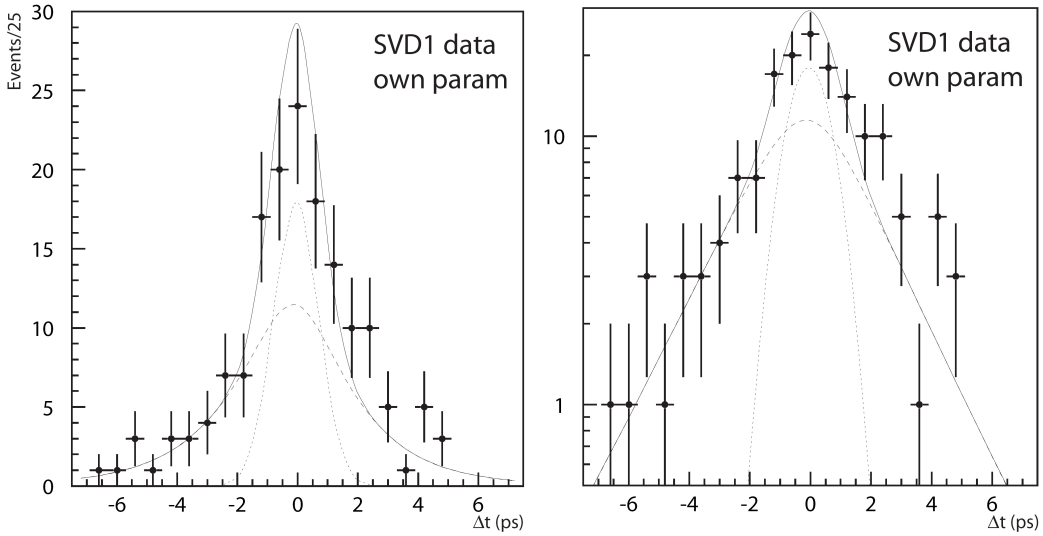


Figure 7.13: Fitted proper-time distribution using our own obtained resolution parameters, on SVD1 data.

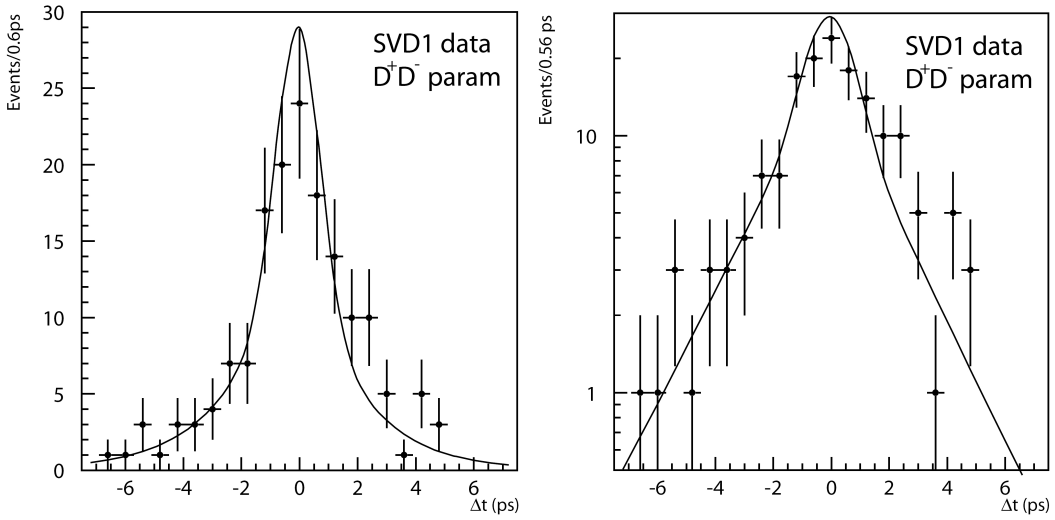


Figure 7.14: Fitted proper-time distribution using the resolution parameters from the $B \to D^+D^-$ analysis, on SVD1 data.

procedure is correct. However this is not intended to be used for a linearity test or for statistical examinations of the error distribution because the model used is too simplified.

Fast MC samples are generated with a proper-time distribution according to:

$$\mathcal{P}(\Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau}\left[1 \pm \left(\mathcal{S}'\sin(\Delta m \Delta t) + \mathcal{A}\cos(\Delta m \Delta t)\right)\right] \otimes \mathcal{R},$$

where $\mathcal{R}$ is the standard resolution function and $\mathcal{S}' = \mathcal{S}/\eta \stackrel{SM}{=} -\sin 2\beta$. A wrong tag fraction of $w = 0.23$ is taken into account but the dependence on r is ignored. Background events are generated according to the function described in Eq. (7.10). The signal fraction (f) is not calculated on an event-by-event basis but a constant ratio is used. The MC samples are generated with a high statistics of 20000 events per sample in the small signal region. The generated data are then fitted with the following function:

$$\begin{aligned}\mathcal{P}(\Delta t) &= f\frac{e^{-|\Delta t|/\tau}}{4\tau}\Big[1 + q(1-2w)\Big((1-2R_\perp)\mathcal{S}'\sin(\Delta m\Delta t) + \mathcal{A}\cos(\Delta m\Delta t)\Big)\Big] \\ &+ (1-f)\mathcal{P}_{\rm bkg}(\Delta t),\end{aligned} \tag{7.11}$$

where $R_\perp$, w and f are now average values. The CP-odd and the signal fraction are set to be the observed result of the SVD1 analysis. Figure 7.15 shows an example of a fitted fast MC sample generated with $\mathcal{A} = 0.3$ and $\mathcal{S}' = 0.6$. The fit gave the following result:

$$\begin{aligned}\mathcal{A} &= 0.29 \pm 0.02, \\ \mathcal{S}' &= 0.58 \pm 0.03.\end{aligned}$$

Many different settings for the CP-violating parameters are generated and fitted. Each time the results are found to be compatible with the generated values.

A study is also performed on 780 signal MC events in the small signal region, using the fully reconstructed MC sample for the signal events, which contains the detector response. The generated CP-violation parameters are set to the Standard Model predictions, i.e.

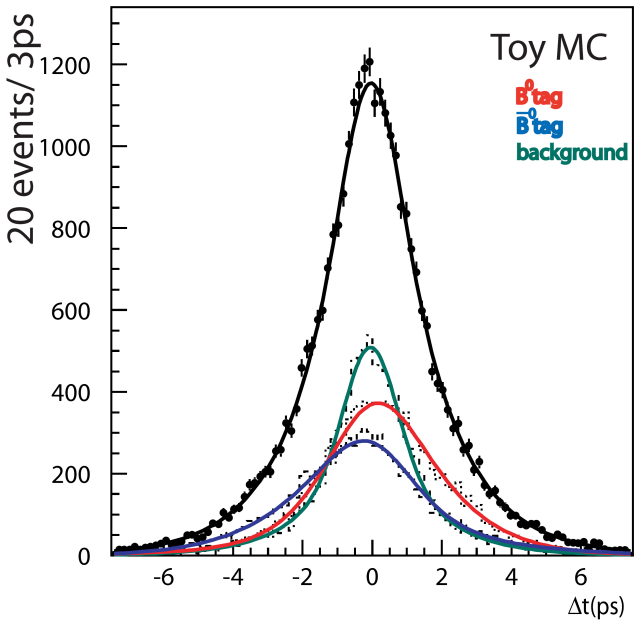


Figure 7.15: CP fit of the toy MC, the green line shows the background contribution over the background events, the red, blue and black line represent the B^0, $\overline{B^0}$ and total PDF, plotted over the corresponding distributions, respectively.

$\mathcal{A} = 0.00$ and $\mathcal{S} = -0.72$. Background events are generated with the fast MC according to the shape of the proper-time distribution in the sideband region. The proper-time is again fitted with the PDF described in Eq. (7.11) and the results are

$$\begin{aligned} \mathcal{A}\ \text{signal MC} &= -0.03 \pm 0.10, \\ \mathcal{S}'\ \text{signal MC} &= -0.80 \pm 0.13. \end{aligned}$$

The fitted proper-time distribution is shown in Figure 7.16. The results are very close to the generated values which shows that the fitting procedure in this simplified model is satisfactory. Finally the CP fit is performed on the data. In the previous analysis the CP fit was performed on 129 signal events and the model included event-by-event CP-odd fractions and signal fractions. All these differences will result in a larger statistical error on the CP parameters, which is fine as it is not our aim to improve the accuracy of the result. Figure 7.17 shows the fitted SVD1 data of the cross-check analysis. The result is

$$\begin{aligned} \mathcal{A} &= -0.46 \pm 0.33, \\ \mathcal{S}' &= -0.66 \pm 0.74, \end{aligned}$$

which is compatible with the Standard Model prediction. The result of the previous analysis is

$$\begin{aligned} \mathcal{A} &= -0.26 \pm 0.26, \\ \mathcal{S}' &= -0.75 \pm 0.56. \end{aligned}$$

Given the different fitting functions and software used, and the different values for the CP-odd fraction, we still obtain a satisfactory result for the cross check of the CP parameters:

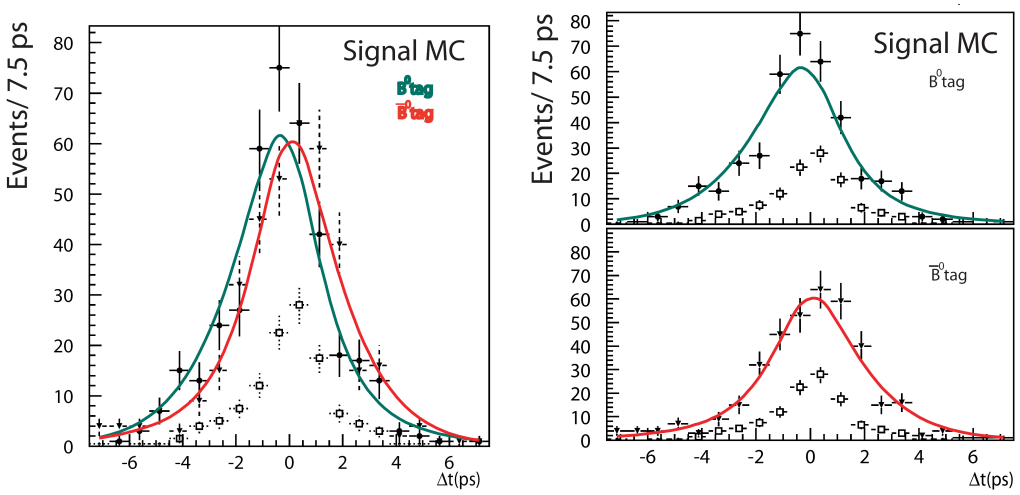


Figure 7.16: Fitted proper-time distribution of tagged $\overline{B^0}$ events (red) and B^0 events (green) of the signal sample for SVD1 configuration. The square symbols show the background events.

the direct CP-violation parameter within one standard deviation from each other and the indirect CP-violation parameter very close to each other. Note however that the result of the CP fit on SVD1 data only in the main analysis is closer to the previous Belle result than the result presented in this Chapter. This can be due to larger overlap of decay channels reconstructed, or due to the larger resemblance of the PDFs of the proper-time.

We can conclude that we have obtained the same reconstruction efficiency and yield in this cross-check analysis with respect to the previous analysis. The comparison of the CP-odd fraction, lifetime measurement and CP-violation parameters are consistent as well.

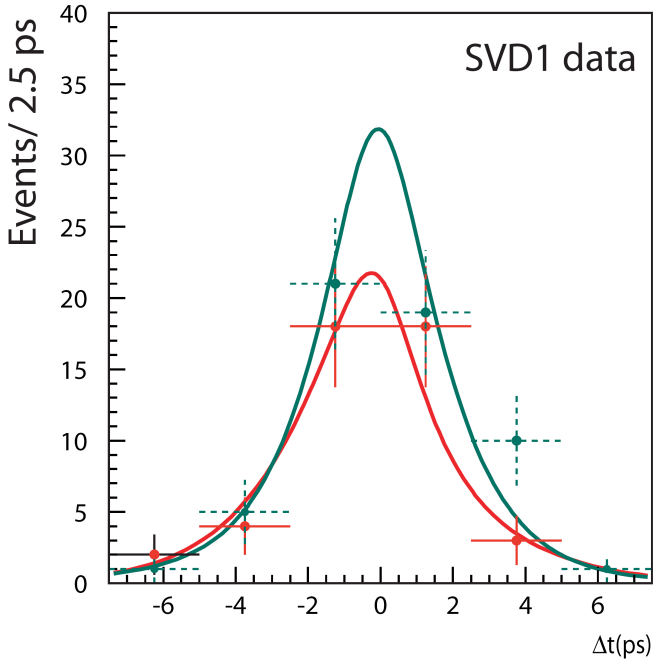


Figure 7.17: Fitted proper-time distribution of tagged $\overline{B^0}$ events (red) and B^0 events (green) of the SVD1 data.

Conclusion

In this dissertation we have presented a study of the CP asymmetry in the $B^0 \to D^{*+}D^{*-}$ decay, based on a data sample of $657 \times 10^6 B\overline{B}$ events recorded with the Belle detector. The decays are produced by the $\Upsilon(4S)$ resonance at the electron-positron asymmetric-energy accelerator KEKB at Tsukuba, Japan.

We enhanced the fraction of $B^0 \to D^{*+}D^{*-}$ events in the sample by applying selection criteria and extract 554 ± 30 signal events, with a purity of 55%, from an unbinned two-dimensional maximum likelihood fit of the beam-constrained mass and energy of the reconstructed B^0.

As the final state consists of two vector mesons the CP eigenvalue can be $+1$ when the decay proceeds through the CP-even S and D waves or -1 for the CP-odd P wave. This superposition dilutes the measurement of the CP asymmetry. Therefore we have performed a time-integrated angular analysis in the transversity basis and obtained a CP-odd fraction of:

$$R_\perp \;=\; 0.125 \pm 0.043(\mathrm{stat}) \pm 0.023(\mathrm{syst}),$$

which is consistent with the prediction from theory as well as the previous measurements.

The time-dependent decay rates of a neutral B meson to a CP eigenstate is given by

$$\mathcal{P}(\Delta t) \;=\; \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}}\Big\{1 + q\Big[\mathcal{S}\sin(\Delta m \Delta t) + \mathcal{A}\cos(\Delta m \Delta t)\Big]\Big\},$$

where $q = +1(-1)$ when the other B meson in the event decays, as a B^0 ($\overline{B^0}$) and $\Delta t = t_{CP} - t_{tag}$ is the proper time difference between the two decays. τ_{B^0} is the neutral B lifetime, Δm the mass difference between the two B^0 mass eigenstates and $\mathcal{S}$ and $\mathcal{A}$ are the CP-violating parameters:

$$\mathcal{S} = \frac{2Im(\lambda)}{|\lambda|^2 + 1}, \quad \mathcal{A} = \frac{|\lambda|^2 - 1}{|\lambda|^2 + 1},$$

where λ is a complex parameter depending on the $B^0 - \overline{B^0}$ mixing as well as on the decay amplitudes for both B^0 and $\overline{B^0}$ to the CP eigenstate.

At the quark level the $B^0 \to D^{*+}D^{*-}$ decay is a $b \to c\bar{c}d$ transition, where the tree amplitude is CKM-suppressed. The contribution of penguin diagrams in this decay is estimated to be at the percent level [1]. When ignoring penguin corrections, the Standard Model prediction for the CP parameters is $\mathcal{A}_{D^{*+}D^{*-}} = 0$ and $\mathcal{S}_{D^{*+}D^{*-}} = -\eta_{D^{*+}D^{*-}} \sin 2\beta$, where $\beta = \arg[-V_{cd}V_{cb}^*]/[V_{td}V_{tb}^*]$ an angle of the unitarity triangle and η is the CP eigenvalue of the $D^{*+}D^{*-}$ eigenstate. Any large measured deviation with respect to the prediction can be a sign of New Physics.

The B^0-meson decay-vertices are obtained using the reconstructed D-meson trajectories and the beam spot profile. The model of the time-dependent decay rate is modified to incorporate the effect of incorrect flavor assignment, the CP-odd dilution and the background. We performed various cross-checks such as a fit to the CP asymmetry of the control sample, $B^0 \to D^{(*)+}D_s^{(*)-}$ which gives $\mathcal{A} = -0.02 \pm 0.03(\mathrm{stat})$ and $S = -0.07 \pm 0.04(\mathrm{stat})$ and is consistent with no CP asymmetry. The lifetime fit to the $B^0 \to D^{*+}D^{*-}$ is consistent with the world average value. Finally, a repetition of the analysis on the first 152×10^6 $B\overline{B}$ events was performed and the results are within the uncertainty consistent with the previous Belle analysis.

The result of the maximum unbinned likelihood fit of the CP-violating parameters in the $B^0 \to D^{*+}D^{*-}$ channel on the full data set available by the summer of 2008 is:

$$\begin{aligned} \mathcal{S}'_{D^{*+}D^{*-}} &= -0.96 \pm 0.25(\mathrm{stat})^{+0.12}_{-0.16}(\mathrm{syst}), \\ \mathcal{A}_{D^{*+}D^{*-}} &= +0.15 \pm 0.13(\mathrm{stat}) \pm 0.04(\mathrm{syst}), \end{aligned}$$

where $\mathcal{S}'_{D^{*+}D^{*-}} = \frac{1}{\eta}\mathcal{S}_{D^{*+}D^{*-}}$. This result corresponds to an evidence of CP violation with a $3.5\,\sigma$ significance. These measurements are consistent with and more precise than the previous results. They are also in agreement with the Standard Model prediction for tree decays. To extract a conclusion on the existence of New Physics, the statistical error on the CP violating parameters needs to be reduced. Therefore an experiment which provides even more luminosity is needed. This could be achieved with the new generation experiment at KEK, which will collect larger statistics: the super-Belle experiment. Furthermore the theoretical understanding of the large-distance QCD effects of the final state hadrons needs to be improved in order to distinguish New Physics effects from final state effects. This problem however can be addressed by measuring proper-time differences in the $B_s \to D_s^{*+}D_s^{*-}$ decay simultaneously with the $B^0 \to D^{*+}D^{*-}$ decay as they are related through the U-spin symmetry such that the CP violating effect should be similar in both decays. We wait therefore for the LHCb experiment which, contrary to Belle, can probe both the B_s as well as the B^0 decays.

Bibliography

[1] Z. Z. XING. CP violation in $B_d \to D^+D^-, D^{*+}D^-, D^+D^{*-}$ and $D^{*+}D^{*-}$. *Phys. Rev. D*, 61:014010, 2000.

[2] E. NOETHER. Invariante Variationsprobleme. *Nachr. Ges. Wiss. Goettingen*, 2:235, 1918.

[3] C. S. WU. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, 1957.

[4] A. D. SAKHAROV. Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe. *Soviet Physics - JPTP Letters*, 5:24–27, 1967.

[5] F. HALZEN AND A.D. MARTIN. Quarks and leptons: An introductory course in modern particle physics. 5:24–27, 1967.

[6] W.-M. YAO ET AL. Review of Particle Physics. *Particle Data Group, Journal of Physics*, G 33:1, 2006. `http://pdg.lbl.gov`.

[7] D0 Collaboration, S. ABACHI ET AL. Observation of top quark production in $p\bar{p}$ collisions with the collider detector at Fermilab. *Phys. Rev. Lett.*, 74:2632, 1995.

[8] CDF Collaboration, F. ABE ET AL. [CDF COLLABORATION]. Observation of top quark production in $p\bar{p}$ collisions with the collider detector at Fermilab. *Phys. Rev. Lett.*, 74:2626, 1995.

[9] VICTOR F. WEISSKOPF AND EUGENE P. WIGNER. Calculation of the natural brightness of spectral lines on the basis of Dirac's theory. *Z. Phys.*, 63:54–73, 1930.

[10] T. NAKADA. CP violation and B physics. `http://nakada.web.cern.ch/nakada/epfl/`, 2006.

[11] R. FLEICHER. Flavor physics and CP violation. `http://physicschool.web.cern.ch/physicschool/LatAmSchool/2007/Info/Info_Participants.asp/`, 2006, Vina del Mar, Chile.

[12] NICOLA CABIBBO. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10:531–533, 1963.

[13] M. KOBAYASHI AND T. MASKAWA. CP violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.*, 49:652, 1973.

[14] LINCOLN WOLFENSTEIN. Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 51:1945–1947, 1983.

[15] J. CHARLES ET AL. CKM FITTER GROUP. *Eur. Phys. J. C.*, 41:1–131, 2005.

[16] Y. SAKAI A. CECCUCCI, Z. LIGETI. The CKM quark-mixing matrix. Review in [6].

[17] OLIVIER SCHNEIDER. B^0–$\bar{B}^0$ mixing. Review in [6], hep-ex/0405012.

[18] GUSTAVO C. BRANCO, LUÍS LAVOURA, AND JOÃO P. SILVA. *CP Violation*. Oxford University Press, New York, 1999.

[19] A. LENZ AND U. NIERSTE. Theoretical update of $B_s - \overline{B}_s$ mixing. *JHEP*, 072:0706, 2007.

[20] L. FERNANDEZ. *Exclusive trigger selections and sensitivity to the $B_s - \overline{B}_s$ mixing phase at LHCb*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne, Switzerland, 2006. "http://doc.cern.ch/archive/electronic/cern/preprints/thesis/thesis-2006-042.pdf".

[21] X. Y. PHAM AND Z.Z XING. CP asymmetries in $B_d \to D^{*+}D^{*-}$ and $B_s \to D_s^{*+}D_s^{*-}$ decays: P-wave dilution, penguin and rescattering effects. *Phys. Lett. B*, 458:375, 1999.

[22] BABAR COLLABORATION. *The BABAR physics book: physics at an asymmetric B-factory*. SLAC-R-0504, 1997.

[23] Y. GROSSMAN AND M. P. WORAH. CP asymmetries in B decays with New Physis in decay amplitudes. *Phys. Lett. B*, 407:307–313, 1997.

[24] R. FLEICHER. Exploring CP Violation and Penguin Effects through $B_d^0 \to D^+D^-$ and $B_s^0 \to D_s^+D_s^-$. *arXiv:0705.4421v1*, 2007.

[25] H. MIYAKE, M. HAZUMI. Measurement of branching fraction, polarizations and time dependent CP asymmetries in $B \to D^{*+}D^{*-}$ decays. Belle Note 553, 2005.

[26] CFS COLLABORATION. Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collision. *Phys. Rev. Lett.*, 39(5):252–255, 1977.

[27] T. BROWDER AND K. HONSCHEID. B Mesons. *Prog. Part. Nucl. Phys.*, 35:81–220, 1995.

[28] CLEO Collaboration, D. BESSON ET AL. Observation of New Structure in the e^+e^- Cross Section above the $\Upsilon(4S)$. *Phys. Rev. Lett.*, 54:381, 1982.

[29] S. KUROKAWA AND E. KIKUTANI. Overview of the KEKB accelerators. *Nucl. Instr. Meth. A*, 499:1, 2003.

[30] The KEKB design report. *"http://www-acc.kek.jp/KEKB/publication/KEKB_design_report/KEKB%20Design%20Report.html"*.

[31] The KEKB parameters. *http://www-acc.kek.jp/KEKB/Commissioning/parameter_history.html*.

[32] Belle Collaboration, ABASHIAN ET AL. A. The Belle detector. *Nucl. Instr. Meth. A*, 479:117, 2002.

[33] J. WICHT. *Study of the radiative penguin $B_s \to \phi\gamma$ and $B_s \to \gamma\gamma$ decays and the resonant $B^+ \to K^+\gamma\gamma$ decays at Belle*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne, Switzerland, 2008. "http://belle.kek.jp/belle/theses/doctor/2008/Wicht.pdf".

[34] C. JACOBY. *Measurements of the rare radiative B meson decays $B \to \phi K\gamma$ and $B \to \omega K\gamma$ at the Belle experiment and contribution to software development at the LHCb experiment*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne, Switzerland, 2007. "http://belle.kek.jp/belle/theses/doctor/2007/Jacoby.ps.gz".

[35] Dephi Collaboration, P. ABREU ET AL. Performance of the Delphi detector. *Nucl. Instr. Meth. A*, 379:57–100, 1996.

[36] Belle SVD2 Group, Z. NATKANIEC ET AL. Status of the Belle silicon vertex detector. *Nucl. Instr. Meth. A*, 560:1, 2006.

[37] G. TAYLOR. The Belle silicon vertex detector: present performance and upgrade plan. *Nucl. Instr. Meth. A*, 501:22, 2003.

[38] Y. OHNISHI. Track Parametrization. Belle Note 148, 1996.

[39] H. KICHIMI ET AL. The Belle TOF system. *Nucl. Instr. and Meth. A*, 453:315, 2000.

[40] A. ABASHIAN ET AL. The K_L^0/μ detector subsystem for the Belle experiment at the KEK B-factory. *Nucl. Instr. Meth. A*, 449:112, 2000.

[41] Y. USHIRODA ET AL. Devellopment of the central trigger system for the Belle detector at the KEK B-factory. *Nucl. Instr. Meth. A*, 438:460, 1999.

[42] Belle Collaboration, H.KAKUNO K.HANAGAKI, M.HAZUMI. Level 4 software trigger at Belle. Belle Note 299, 2000.

[43] S. NAGAYAMA. PANTHER - User's manual, reference manual and primer. Belle Note 130 , Belle Note 131 and Belle Note 132, 1996.

[44] B. CASEY. HadronB. Belle Note 390, 2001.

[45] R. ITOH. BASF User's Manual. Belle Note 161, 1997.

[46] D.J. LANGE. The EVTGEN particle decay simulation package. *Nucl. Instr. Meth. A*, 462:152–155, 2001. http://www.slac.stanford.edu/~lange/EvtGen/.

[47] R. BRUN ET AL. GEANT 3.21. *CERN Report DD/EE/84-1*, 1984.

[48] H. OZAKI. Mini-DST Tables V0.0. Belle Note 146, 1996.

[49] C.M. CARLONI CALAME, C. LUNARDINI, G. MONTAGNA, O. NICROSINI, AND F. PICCININI. Large-angle Bhabha scattering and luminosity at flavour factories. *Nucl. Phys. B*, 584:459, 2000.

[50] S. UEHARA. Devellopment study of zero-degree luminosity monitor using a GSO scintillator for the Belle experiment. Belle Note 174, 1997.

[51] V. ZHILICH ET AL. Offline luminosity measurements from the Belle detector. Belle Note 453, 2001.

[52] D. BEYLIN ET AL. Online luminosity measurements from the Belle detector. Belle Note 465, 2001.

[53] Belle Collaboration, A. GO, A. BAY ET AL. Measurement of EPR-type flavor entanglement in $\Upsilon(4S) \to B^0\overline{B}^0$ decays. *Phys. Rev. Lett.*, 99:131802, 2007.

[54] H. KAKUNO. Flavor tagging by Multidimensional likelihood method. Belle Note 384, 2001.

[55] H. KAKUNO ET AL. Neutral B flavor tagging for the measurement of mixing-induced CP violation at Belle. *Nucl. Instr. Meth. A*, 533:516, 2004.

[56] H. TAJIMA ET AL. Proper-time resolution function of measurements of time evolution of B mesons at the KEK B-factory. *Nucl. Instr. Meth. A*, 533:370, 2004.

[57] H. MIYAKE. *Measurement of polarization and time-dependent CP asymmetry parameters in $B^0 \to D^{*+}D^{*-}$ decays*. PhD thesis, Osaka University, Japan, 2005. "`http://osksn2.hep.sci.osaka-u.ac.jp/theses/doctor/2004/2004_Eng.html`".

[58] Belle Collaboration, H. MIYAKE, M. HAZUMI ET AL. Branching Fraction, Polarization and CP-Violating Asymmetries in $B^0 \to D^{*+}D^{*-}$ Decays. *Phys. Lett. B*, 618:34, 2005.

[59] Belle Collaboration, S. FRATINA ET AL. Evidence for CP Violation in $B^0 \to D^+D^-$ Decays. *Phys. Rev. Lett.*, 98:221802, 2007.

[60] G. C. FOX AND S. WOLFRAM. Observables for the analysis of event shapes in e^+e^- annihilation and other processes. *Nucl. Rev. Lett*, 41:1581, 1978.

[61] F. FANG. Study of $K_{\rm S} \to \pi^+\pi^-$ selection. *Belle Note 323*, 2000.

[62] H. ALBRECHT ET AL. (ARGUS COLLABORATION). Search for hadronic $b \to u$ decays. *Phys. Rev. Lett. B*, 241:278, 1990.

[63] BABAR Collaboration, B. AUBERT ET AL. Improved Measurement of Time-Dependent CP Asymmetries and the CP-odd fraction in the decay $B^0 \to D^{*+}D^{*-}$. *Phys. Rev. D*, 76:111102, 2007.

[64] BABAR Collaboration, 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 γ. *Phys. Rev. D*, 73:112004, 2006.

[65] BABAR Collaboration, B. AUBERT ET AL. Measurements of time-dependent CP asymmetries in $B^0 \to D^{(*)+}D^{(*)-}$ decays. *arXiv:0808:1866*.

[66] BABAR Collaboration, B. AUBERT ET AL. Measurement of the Branching Fraction and Time-Dependent CP Asymmetry in the Decay $B^0 \to D^{*+}D^{*-}K^0_{\rm S}$. *Phys. Rev. D*, 74:091101, 2006.

[67] Belle Collaboration, J. DALSENO ET AL. Measurement of Branching Fraction and Time-Dependent CP Asymmetry Parameters in $B^0 \to D^{*+}D^{*-}K^0_{\rm S}$ Decays. *Phys. Rev. D.*, 76:072004, 2007.

[68] Belle Collaboration, T. AUSHEV ET AL. Search for CP violation in the decay $B^0 \to D^{*\pm}D^{\mp}$. *Phys. Rev. Lett.*, 93:201802, 2004.

[69] BABAR COLLABORATION. The BABAR detector. *Nucl. Instr. Meth. A*, 479:1, 2002.

[70] BABAR Collaboration, V. RE ET AL. BaBar Silicon Vertex Tracker: Status and Prospects. *Nucl. Instr. Meth. A*, 569:1–4, 2006.

[71] M. SATPHPATHY K. ABE AND H. YAMAMOTO. Time dependent angular analysis of B decays. *Belle Note 419*, 2001.

[72] HEAVY FLAVOR AVERAGING GROUP (HFAG). `http://www.slac.stanford.edu/xorg/hfag/osc/PDG_2008/`. 2008.

[73] S. FRATINA. Measurement of time dependent CP violation in $B \to D^+D^-$ decays. 2006.

[74] `http://belle.kek.jp/secured/indirectcp/cpfit/ichep06/BAM20060523-updated.ppt`.

[75] T. TOMURA T. HIGUTCHI AND H. TAJIMA. Precise measurement of B meson lifetimes with hadronic decay final states. *Belle Note 470*, 2002.

[76] Belle Collaboration, K. ABE ET ALL. Precise measurement of B meson lifetimes with hadronic decay final states. *Phys. Rev. Lett.*, 88:171801, 2002.

[77] `http://belle.kek.jp/secured/indirectcp/cpfit/ichep06/wtag_svd1_ichep06`, `http://belle.kek.jp/secured/indirectcp/cpfit/ichep06/wtag_svd2_ichep06`.

[78] F. JAMES AND M. WINKLER. Minuit user's guide. *http://www.cern.ch/minuit, CERN*, 2004.

Kim Vervink

Address: 22, Chemin du Mottey
1020 Renens, VD
Switzerland

Phone: +41 79 763 54 13
E-mail: kvervink@hotmail.com

Education and degrees

2004-present	**Ecole Polytechnique Fédérale de Lausanne (EPFL)** Laboratoire de physique des Hautes Energies Teaching assistant for introductory course on particle physics	Lausanne, Switzerland
Feb 2007	**CERN-CLAF School of High-Energy Physics**	Vina del Mar Chile
Since 2006	**Member of the Belle Collaboration**	Tsukuba Japan
Since 2004	**Member of the LHCb Collaboration** Design and production of the Inner Tracker Software development: alignment and detector material studies	Geneva Switzerland
2000-2004	**University of Antwerp (UA)** Master in Physics Dissertation: Particle Physics, “Calibration of the VFPS detector of H1 through elastic ρ-meson production at HERA”	Antwerp, Belgium
Jun-Aug 2003	**CERN summer student program** NA48 Collaboration	Geneva Switzerland
2002-2003	**Universidad de Granada (UGR)** Erasmus exchange program	Granada, Spain
1999-2000	**Mount Rainier high school** AFS exchange program on inter-cultural learning	Seattle, USA
1997-1999	**Onze-Lieve-Vrouw-van-Lourdes College** High school (ASO): Math-Science Section Diploma of General Secondary Education	Edegem, Belgium

Publications directly related to the thesis work

Public:

- K. Vervink, "CP violation in B decays to charm and charmonium at Belle", to appear in the proceedings of the International Conference on High Energy Physics (ICHEP), Philadelphia, July 2008
- K. Vervink, "Measurements of the CKM angle ϕ_1/β from Belle and BaBar", arXiv:0807.0496. Appeared in the proceedings of the International Conference of Flavor Physics and CP violation (FPCP), Taipei, May 2008
- K. Vervink et al., "Compton scattering in scalar electrodynamics ", CERN2008004. Appeared in the yellow book of the "4th CERN - CLAF School of High-Energy Physics: Latin American School of High Energy Physics 2007", Vina del Mar, March 2007
- K. Vervink et al, "Design and performance of the LHCb Silicon Tracker", Nucl. Instr. Meth. A 566:170-173, 2006. Appeared in the proceedings of the International Workshop on Tracking In High Multiplicity Environments (TIME05), Zurich, October 2005
- W. Baldini et al, "LHCb alignment strategy", LHCb public note 2006-063, Geneva, 2006
- A. Perrin and K. Vervink, "The Inner Tracker detector description and its implementation in the XML database", LHCb public note 2006-018, Geneva, 2006

Internal:

- K. Vervink and T. Aushev, "Measurement of polarization fraction and time-dependent CP asymmetries in $B^0 \to D^{*+}D^{*-}$", Belle note 1007, Tsukuba, 2008
- K. Vervink, "A first study of the Inner Tracker alignment", LHCb note 2006-060, Geneva, 2006