Measurement of Time-Dependent $CP$ Violation Asymmetries in $B^0 \rightarrow \phi K^0_S$ Decay

Aleksey Chuvikov
© Copyright by Aleksey Chuvikov, 2006.

All rights reserved.
Abstract

We present measurements of time-dependent \( CP \) asymmetry parameters in \( B^0 \rightarrow \phi K_s \) transitions. We reconstruct one neutral \( B \) meson decaying into a \( CP \) eigenstate and use the inclusive information of the remaining particles to determine the flavor of the accompanying meson at its decay time. Using a data sample of 492 fb\(^{-1}\) collected at the \( \Upsilon(4S) \) resonance with the Belle detector at the asymmetric KEKB accelerator, we reconstruct 202 \( B^0 \rightarrow \phi K_s \) decay candidates with an estimated signal purity of 81.2% after vertexing, flavor tagging and other good vertex selection. From an unbinned maximum likelihood fit to the proper time interval distribution we obtain

\[
S = 0.55 \pm 0.31 (\text{statistical}) \pm 0.06 (\text{systematic})
\]

\[
A = 0.18 \pm 0.20 (\text{statistical}) \pm 0.05 (\text{systematic})
\]

The \( A \) term shows the absence of direct \( CP \) violation and the \( S \) term is consistent with the current world average of \( \sin 2\phi_1 \) as predicted in the Standard Model, where \( \phi_1 \) is one of the angles of the Unitarity Triangle corresponding to the Cabibbo-Kobayashi-Maskawa matrix.
Acknowledgments

There are many people I would like to thank whose contribution to this thesis was invaluable. First of all, I was very excited to join Princeton University back in 2001, where I had the chance to meet a lot of great people. It was an incredible honor for me to work under the supervision of Daniel Marlow. He was always recommending diverse and interesting topics of study from which I was able to learn various aspects of High Energy Physics. Many times his experience and valuable ideas rescued me from difficult problems arising along the way not letting me get stuck. At the same time, he offered me the complete freedom to be creative and not be concerned with less important aspects. Despite his busy schedule he managed not only to read my thesis, but also make a lot of corrections.

A great influence on me was exerted by Kirill Korotushenko. He was the individual who convinced me to start working with Dan immediately and pay more attention to research. He became my mentor from the beginning and was looking after me ever since. I greatly appreciate his opinions on the everyday problems. During my first two years I was also interacting a lot with another of Dan’s student - Sven Vahsen. Even though we were working on different projects, I enjoyed our late night conversations over a cup of tea about physics, technology and music. During my last year I had similar meetings with Srinivas Raghu. And when it was our turn to write a thesis our discussions were covering more serious and philosophical subjects.

I enjoyed working with the other members of Princeton’s Belle group: Tom Ziegler, Alexei Garmash, Bert Harrop and Stanley Chidzik. This analysis would be practically impossible to do without the tedious work done by CP Fit group. In particular, I thank M. Hazumi, Y. Sakai, K. Sumisawa, K. Hara, T. Higuchi, T. Tomura, H.
Kakuno for vertex fitting, flavor tagging and other studies related to this analysis.

I want to thank Christopher Tully, my second reader, for his time spent proofreading my thesis. Additional thanks to Igor Klebanov and Edward Groth for being committee members. I appreciate final comments from James Olsen.

I am highly in debt of Viktor Tkachuk, who was my first teacher of Physics. He instilled a passion for Physics and enabled me to compete with students from some of the best schools. I should also mention my previous advisor from ITEP Eugene Tarkovsky, who introduced me to High Energy Physics. His dedication and interest convinced me to continue working in this field.

I thank all of my friends I met through Princeton which made my stay enjoyable. Among them are Elena Loginova, Dima Sarkisov, Alexander Baitine, Artem Smirnov, Elena Nabieva, Pedro Goldbaum, Mike Leung, Nhan Tran, Slava Lukin, Alexey Kuritsyn.

Finally, I am thankful to my parents and my brothers Kolya and Sergej for their love and moral support throughout these long years of study. I especially thank Alfia for all great moments we spent together.
Contents

Abstract iii

Acknowledgments iv

1 Introduction and Overview 1

2 CP violation 4

2.1 Overview . . . . . . . . . . . . . . . . . . . . 4

2.2 Discrete Symmetries . . . . . . . . . . . . . . . . . . . . . . . . . . 5

2.3 Cabibbo-Kobayashi-Maskawa Mechanism . . . . . . . . . . . . . . . . 7

2.4 CKM Matrix Parametrization . . . . . . . . . . . . . . . . . . . . . . 9

2.5 Unitarity Triangles . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11

2.6 CP Violation in B Meson Decays . . . . . . . . . . . . . . . . . . . . . 13

2.6.1 Oscillations Of $B^0 - \bar{B}^0$ Mesons. Indirect CP Violation . . . . . . 14

2.6.2 CP Violation in Decay Rates . . . . . . . . . . . . . . . . . . . . . 16

2.6.3 Decay of Neutral B Mesons into CP Eigenstate . . . . . . . . . 17

2.7 CP Violation in $B^0 \to \phi K^0_S$ Decay . . . . . . . . . . . . . 18

2.8 Overview of the Measurement . . . . . . . . . . . . . . . . . . . . . . 20

3 Experimental Apparatus 23

3.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23

3.2 The KEKB accelerator and storage rings . . . . . . . . . . . . . . . . . 24
3.3 The Belle Detector

3.3.1 Silicon Vertex Detector (SVD)

3.3.2 Central Drift Chamber (CDC)

3.3.3 Aerogel Cherenkov Counter (ACC)

3.3.4 Time of Flight System (TOF)

3.3.5 Electromagnetic Calorimeter (ECL)

3.3.6 Extreme Forward Calorimeter (EFC)

3.3.7 $K_L$/Muon Detector (KLM)

4 Reconstruction of $B^0 \rightarrow \phi K_S$

4.1 Hadronic $B\bar{B}$ Selection

4.2 Event Reconstruction

4.2.1 Identification of Electrons and Muons

4.2.2 Particle Identification of Charged Hadrons

4.2.3 $K_S$ Reconstruction

4.2.4 $B^0$ Candidate Reconstruction

4.3 Background Study

4.3.1 Thrust Angle $\theta_T$ and Sphericity $S_\perp$

4.3.2 Fox-Wolfram Moments

4.3.3 $B$ Flight Direction and Helicity Angle of $\phi$ Meson

4.3.4 Likelihood Analysis

5 Measurement of the $CP$ Asymmetry

5.1 Unbinned Maximum Likelihood Fit

5.2 Probability Density Function

5.3 Flavor Tagging Algorithm

5.4 Proper-Time Difference Reconstruction

5.5 Resolution Functions

5.5.1 Detector Resolution Functions
5.5.2 Quality of Vertex Fit ............................... 79
5.5.3 Multiple-Track Vertex Reconstruction .................. 80
5.5.4 Single-Track Vertex Reconstruction .................... 81
5.5.5 Smearing due to Non-Primary Tracks .................. 84
5.6 Kinematic Approximation ............................... 89
5.7 Outliers ........................................... 91
5.8 Background Shape ................................... 92
5.9 Signal Probability .................................... 94
5.10 Validity Check ...................................... 96
  5.10.1 Lifetime Fit .................................... 97
  5.10.2 Linearity Check ................................ 98
  5.10.3 Asymmetry of the $B^\pm \to \phi K^\pm$ Control Sample . . . . . . . . . 99
5.11 Correction Due To The $B^0 \to K^+K^-K_s$ and $B^0 \to f_0K_s$ Contribution 100
5.12 Fit Result .......................................... 102
5.13 Systematic Uncertainties .............................. 103

6 Conclusions ............................................ 108

A The Special Properties of the UML Fit Applied To $CP$ Fitting 112

Bibliography ............................................. 116

viii
List of Figures

2.1 Hierarchy of transition amplitudes mediated by the charged weak current 8
2.2 Charge current interaction in SM. .................................................. 9
2.3 The CKM unitarity triangle. ............................................................ 12
2.4 $B^0 - \bar{B}^0$ mixing contribution to $B^0$ decay rate .................. 13
2.5 Dominant box diagrams for the $B^0 \rightarrow \bar{B}^0$ transition. ........ 14
2.6 QCD penguin contribution to the decay $B^0 \rightarrow \phi K^0$. ............ 19
2.7 Feynman diagrams for the $B^0 \rightarrow J/\psi K^0_S$ transition. ........ 20
2.8 Schematic of the time-dependent CP asymmetry measurement in $B^0 \rightarrow \phi K^0_S$. ........................................................... 21

3.1 The hadronic cross section for $e^+e^-$ annihilation around 10GeV/c^2. 25
3.2 The KEKB accelerator configuration. .............................................. 26
3.3 Side view of the Belle detector. ...................................................... 28
3.4 Momentum range coverage for $K/\pi$ separation provided by the CDC, ACC and TOF systems. ....................................................... 30
3.5 SVD1 layout .......................................................... 31
3.6 SVD2 layout .......................................................... 32
3.7 Overview of CDC ...................................................... 33
3.8 Barrel and endcap ACC modules .................................................. 35
3.9 The aerogel Cherenkov counter system (ACC). ............................. 36
3.10 (TOF) .......................................................... 37
3.11 $K^\pm/\pi^\pm$ separation performance of TOF system. .......................... 38
3.12 Configuration of the electromagnetic calorimeter (ECL). .................. 40
3.13 ECL counter module. .......................................................... 40
3.14 ECL energy resolution measured with Bhabha and $\gamma\gamma$ events. ... 41
3.15 Extreme Forward Calorimeter (EFC) ....................................... 42
3.16 Cross section of an RPC superlayer. ....................................... 43
3.17 Schematic diagram of KLM RPCs ........................................... 44

4.1 Invariant mass distribution of oppositely charged tracks after the $K_S$ selection criteria are applied. ........................................... 53
4.2 Invariant mass distribution of $K^+K^-$ .................................... 54
4.3 $\Delta E$ and $M_{bc}$ distributions of reconstructed $B^0 \rightarrow \phi K_S$ events for signal MC sample. ........................................... 55
4.4 A schematic of the topological difference between $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$ and $e^+e^- \rightarrow q\bar{q}$ continuum events. ........................................... 56
4.5 Distributions of (a) $|\cos(\theta_T)|$ and (b) $S_\perp$ for MC signal and MC $q\bar{q}$ background. ........................................... 57
4.6 Modified Fox-Wolfram moments distributions. ............................. 59
4.7 Super Fox Wolfram separation power. ....................................... 61
4.8 $R^S_2$ versus $|\cos(\theta_T)|$ distributions for (a) signal $B\bar{B}$ and (b) jet-like $q\bar{q}$ events. ........................................... 62
4.9 $R^S_2$ versus $S_\perp$ distributions for (a) signal $B\bar{B}$ and (b) jet-like $q\bar{q}$ events. 62
4.10 Distributions of the $B$ flight direction angle $\theta_B$ and the helicity of the $\phi$ meson for signal and continuum background MC samples. .... 63
4.11 Likelihood ratio distribution made with SFW, $|\cos\theta_B|$ and $|\cos\theta_H|$ variables. ........................................... 65
4.12 Figure of merit distribution depending on the likelihood ratio cut for different $r$ intervals. ........................................... 67
4.13 ΔE and $M_{bc}$ distributions of reconstructed $B^0 \rightarrow \phi K_S$ events. 67
4.14 Yield extraction for $r$ bins 1-3. 68
4.15 Yield extraction for $r$ bins 4-6. 69

5.1 Schematic overview of the flavor tagging algorithm. 73
5.2 Measured time-dependent $A(\Delta t)$ asymmetries between the SF and OF events for six tagging $r$ bins. 75
5.3 Reconstruction of $B$ decay vertices. 76
5.4 Residual $\delta z_q$ distributions and fits to the sum of two Gaussians. 79
5.5 Nominal $\chi^2$/n.d.f. vs $\xi$. 80
5.6 The pull distributions of $\delta z_q/\sigma^z_q$ divided by the goodness of vertex fit $\xi$. 82
5.7 Standard deviations of the pull distributions as a function of $\xi$. 83
5.8 Distributions of (a) $\delta z_{\text{rec}}$ and (b) $\delta z_{\text{asc}}$ for multiple-track vertices with superimposed event-by-event resolution functions. 83
5.9 Residual distributions of (a) $\delta z_{\text{rec}}$ and (b) $\delta z_{\text{asc}}$ for single-track vertices with $R_{\text{rec}}$ and $R_{\text{asc}}$ resolution functions. 84
5.10 Distribution of $z_{\text{asc}} - z_{\text{asc}}^{\text{noNP}}$ for multiple-track vertex. 85
5.11 Non-primary tracks shift $z_{\text{asc}} - z_{\text{asc}}^{\text{noNP}}$ versus (a) $\sigma^z_{\text{asc}}$ and (b) $\xi_{\text{asc}}$ for multiple-track vertices. 85
5.12 Distribution of $\delta z_{\text{asc}}$ for multiple-track vertices of (a) neutral and (b) charged $B$ mesons. 86
5.13 Non-primary track shift $z_{\text{asc}} - z_{\text{asc}}^{\text{noNP}}$ versus $\sigma^z_{\text{asc}}$ for single-track vertices. 87
5.14 Distribution of $\delta z_{\text{asc}}$ for single-track vertices of (a) neutral and (b) charged $B$ mesons. 87
5.15 Distribution of $x = \Delta t - \Delta t_{\text{true}}$ with $R_k(x)$ superimposed. 90
5.16 Background $\Delta t$ distribution for $B \rightarrow \phi K_s$ mode. 93
5.17 Fitted 2D probability density distribution. 96
5.18 The reconstructed $\Delta t$ distribution of $B^0 \rightarrow \phi K_s$ and $B^\pm \rightarrow \phi K^\pm$
candidates. ......................................................... 98

5.19 Linearity check plot using signal GEANT MC for $B^0 \rightarrow \phi K_s$ decay. . 99

5.20 Asymmetry plots for $B^0 \rightarrow \phi K_s$ GEANT MC sample with generated
$S = 0.7$. ............................................................... 100

5.21 Asymmetry plots for $B^\pm \rightarrow \phi K^\pm$ decay. ................................. 101

5.22 Distribution of the invariant mass of charged kaons after all selections. 102

5.23 $-2 \ln(\mathcal{L}/\mathcal{L}_{\text{max}})$ as a function of (a) $S$ and (b) $A$ ............................... 104

5.24 Asymmetry plots for $B^0 \rightarrow \phi K_s$. ................................................. 104

6.1 Comparisons of averages in different $b \rightarrow q\bar{q}s$ modes. ................. 109

6.2 $\bar{p}-\bar{\eta}$ plane with constraints from different measurements. .......... 111

A.1 Consistency check. 50 events per experiment. ................................. 117

A.2 Consistency check. 100 events per experiment. .............................. 118

A.3 Consistency check. 200 events per experiment. ............................. 119

A.4 Consistency check. 1000 events per experiment. .......................... 120

A.5 Consistency check. 10000 events per experiment. ......................... 121

A.6 Distribution of estimated MINUIT errors. ..................................... 122

A.7 $S - S_{\text{gen}}$ versus the MINUIT errors. ................................... 122

A.8 Dependence of $\mu_{\delta S}$ and $\sigma_{\delta S}$ on MINUIT errors. ........... 123

A.9 $S - S_{\text{gen}}$ for the same MINUIT errors slice. .............................. 124

A.10 $\sigma_l/\sigma_r$ distribution. ......................................................... 124

xii
# List of Tables

2.1 Three generations of quarks and leptons ........................................ 4
2.2 Effects of discrete symmetry transformations ................................... 6
2.3 Invariance of forces under the discrete symmetry transformations ... 7

3.1 Hadronic cross sections at the energy corresponding to the $\Upsilon(4S)$ . 24
3.2 KEKB parameters corresponding to the best peak luminosity $\mathcal{L} = 16.52$ nb$^{-1}$s$^{-1}$ recorded on June 30, 2006 with the design values in the parentheses. ............................................................... 27

4.1 Cross sections and efficiencies of HadronB skim for various processes . 47
4.2 $K_S$ selection criteria. ........................................................................ 52
4.3 Event feeds and correspondent purities for optimized $LR$ cuts and different $r$ intervals ............................................................... 66

5.1 Wrong tag fractions $w_l$ for different tagging $r$ bins. ................. 74
5.2 The list of fitted parameters of $R_{np}$ determined from SVD1 Monte Carlo. 88
5.3 The list of fitted parameters of $R_{np}$ determined from SVD2 Monte Carlo. 88
5.4 Background shape parameters for the $B^0 \rightarrow \phi K_s$ mode. ........ 93
5.5 Signal probability parameters for $B^0 \rightarrow \phi K_s$ and $B^\pm \rightarrow \phi K^\pm$ decay modes. ............................................................... 95
5.6 Detector resolution parameters for SVD1 and SVD2 data samples. . 97
5.7 Summary of the systematic uncertainties contributions .................. 107
Chapter 1

Introduction and Overview

In the first half of the last century particle physicists thought that all fundamental interactions were symmetric under the three discrete operations: charge conjugation ($C$), parity ($P$) and time reversal ($T$). Later on, this assumption started to fall apart.

In 1957 Lee and Yang [1] posited that parity was not conserved in weak interactions. This was done to resolve the $\theta - \tau$ paradox, wherein two particles, the $\theta^+$ and the $\tau^+$, were observed to have the same mass and lifetime, but to decay into final states with opposite parities. It was reasonable to speculate that they were in fact the same particle (presently known as the $K^+$) and that parity is not conserved in weak decays. Soon after this paper, a collaboration led by C.S. Wu demonstrated parity violation in an experiment involving $\beta$-decay of polarized nuclei[2]. It was also shown that there is no charge conjugation symmetry. Landau introduced a combined $CP$ symmetry as the product of charge conjugation and parity transformation[3]. He pointed out that although the $C$ and $P$ symmetries were separately violated by the weak Lagrangian, $CP$ symmetry was conserved.

But the era of “$CP$ conservation” did not last for long. In 1964 Christenson, Cronin and Fitch [4] unexpectedly discovered $CP$ violation in neutral $K^0_L$ kaons. In 1967 Sakharov increased interest in the topic by showing that $CP$ violation is a necessary ingredient in the baryon-number asymmetry of the universe[5]. Since then...
$CP$ violation has been one of the main subjects of interest in particle physics.

In 1963 before the existence of quarks was experimentally verified, Cabibbo investigated strangeness-changing decays. He found\[6\] that quarks did not interact weakly as mass eigenstates, but instead the down quark interacted as a superposition of down and strange flavor states, described by a mixing angle parameter. This phenomenon is now known as Cabibbo mixing. Glashow, Iliopoulos, and Maiani proposed\[7\] in 1970 the existence of a fourth quark as a partner of the strange quark. Even though the theory looked complete with two quark families and the Cabibbo mixing angle, it could not accommodate $CP$ violation.

In 1972 Kobayashi and Maskawa extended Cabibbo’s theory and introduced a third quark family\[8\]. This brought two more mixing angles and a phase in the mixing matrix. This mechanism (later known as the KM mechanism) results in $CP$ violation for a non-zero phase. Subsequent discoveries proved the existence of quarks and their three-generation arrangement. The mixing matrix become know as the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix. The Standard Model says that the CKM matrix must be unitary. The unitarity of the matrix can be visualized as a closed triangle in the complex plane. Experimental measurements of the sides and angles of the unitarity triangle can put constraints on the triangle and hence check unitarity. The Standard Model and the unitarity triangle will be discussed in detail in Chapter 2.

Carter, Sanda and Bigi showed that the KM mechanism predicts a large $CP$ violation in certain $B$ meson decays\[9, 10\]. This created a need for $B$-factories—accelerators producing a large number of $B$ mesons. Currently there are two such operating factories, one at KEK and one at SLAC, hosting the Belle and BaBar experiments, respectively. Both produce large numbers of $B_u$ and $B_d$ type of mesons and have asymmetric beams to facilitate the measurement of time dependent $CP$ asymmetry.

Soon after commencing operation, both experiments reported observation of $CP$
violation for $(c\bar{c})K^0$ states and the “gold-plated mode” $B^0 \rightarrow J/\psi K_S[11, 12]$. The latter got this name because it allows one to extract the $CP$ asymmetry parameter with a very high precision due to the large statistics and low level of background. This measurement puts a severe constraint on one angle of the unitarity triangle known as $\phi_1$ and also $\beta$. Further measurements of $B$ meson decays would put more constraints on the triangle and check the Standard Model for consistency.

This thesis presents a study of the $CP$ asymmetry in the charmless $B^0 \rightarrow \phi(s\bar{s})K_S$ decay mode. The relevant angle in this measurement is $\phi_1$, which is the same angle as for the $B^0 \rightarrow J/\psi K_S$ mode. Although the smaller signal and higher background associated with the $B^0 \rightarrow \phi(s\bar{s})K_S$ decay mode result in a less precise measurement of $\sin 2\phi_1$ than what can be achieved in the charmonium mode, a consistency check between the two modes is of considerable interest, since a difference in the value of $\phi_1$ obtained via these two different modes could be symptomatic of new physics, not accounted for in the Standard Model.
Chapter 2

CP violation

2.1 Overview

The Standard Model (SM) captures much of our current knowledge of particle physics. It is a $SU(3)_C \times SU(2)_L \times SU(1)_Y$ gauge theory with color, weak isospin and hypercharge being the parameter spaces of the $SU(3)_C$, $SU(2)_L$ and $SU(1)_Y$ gauge groups respectively[13]. It combines the theory of strong interactions, known as Quantum Chromodynamics (QCD) with the unified theory of weak interactions.

The SM arranges the quarks and leptons into three generations, see Table 2.1. There are three generations known at the moment, but there is the possibility to have more. The up-type quarks $u$, $c$ and $t$ have the electric charge of $+2/3$, and down-type quarks $d$, $s$, $b$ have the charge of $-1/3$. The families are ordered from the lightest to heaviest.

<table>
<thead>
<tr>
<th>Charge</th>
<th>Quarks</th>
<th>Charge</th>
<th>Leptons</th>
</tr>
</thead>
<tbody>
<tr>
<td>$+\frac{2}{3}$</td>
<td>$u$ $c$ $t$</td>
<td>-1</td>
<td>$e$ $\mu$ $\tau$</td>
</tr>
<tr>
<td>$-\frac{1}{3}$</td>
<td>$d$ $s$ $b$</td>
<td>0</td>
<td>$\nu_e$ $\nu_\mu$ $\nu_\tau$</td>
</tr>
</tbody>
</table>

Table 2.1: Three generations of quarks and leptons
QCD describes the strong interactions between colored quarks and gluons. There are three color charges denoted red, green and blue, needing 8 gauge bosons to mediate the transformations between these three colors. Quarks do not exist by themselves in nature. Instead, they are always confined with the other quarks in a color-neutral state. There is no analytic proof for that confinement, but it can be explained intuitively. When one tries to separate them, at some point it becomes more energetically favorable to produce a quark/anti-quark pair instead of the bound gluon. There are two typical known states: mesons, which are quark/anti-quark pairs; and baryons, which are quark triplets.

The combined theory of electroweak interactions describes the weak and electromagnetic interactions. The symmetry is spontaneously broken through the Higgs mechanism, which gives mass to the weak gauge bosons $W^\pm$ and $Z^0$, while leaving the electromagnetic gauge boson (photon) massless. The weak interactions distinguish between left handed and right handed particles.

### 2.2 Discrete Symmetries

A symmetry is a transformation, which when applied to a system, may or may not change some of its properties. If a system is unchanged after a certain transformation, it is said to be invariant under that transformation. Symmetry plays a dramatic role in physics. There is a well-known correspondence between invariance under transformation and conservation laws (Noether’s theorem). For example, invariance of a physical system under spatial translation leads to conservation of momentum, invariance under rotation leads to conservation of angular momentum, invariance under time evolution leads to conservation of energy.

All three symmetries mentioned above are continuous, meaning that they can be characterized by a continuous parameter of the transformation. For example, we
Symmetry | Action on  
---|---
| Position | Time | Velocity, Momentum | Spin |
| Parity \([P]\) | Reversed | Unchanged | Reversed | Unchanged |
| Charge Conjugation \([C]\) | Unchanged | Unchanged | Unchanged | Unchanged |
| Time Reversal \([T]\) | Unchanged | Reversed | Reversed | Reversed |
| \(CP\) | Reversed | Unchanged | Reversed | Unchanged |
| \(CPT\) | Reversed | Reversed | Unchanged | Reversed |

Table 2.2: Effects of discrete symmetry transformations

could shift the system coordinates by any distance. There are also discrete symmetries, which could simply flip a system from one state to another, such as reflecting a coordinate system in a mirror. These symmetries do not have corresponding conservation laws, however, they are important in particle physics to determine which particle interactions are possible.

There are three symmetries that are useful in particle physics:

- Charge conjugation \((C)\) is the transformation that turns particles into their corresponding antiparticles. It flips all charges of a particle such as electrical charge, lepton number, baryon number and so on.

- Parity transformation \((P)\) is the operator that reflects all three spatial coordinates, leaving the time coordinate intact.

- Time inversion \((T)\) is the transformation that reverses the direction of time. Invariance under this operator means that the particle reaction could also go from the final state back to the initial state.

There are also combinations of these operators, for example, \(CP\) is the operator where a parity transformation is followed by a charge conjugation, and a \(CPT\) is the combination of all three \(C\), \(P\) and \(T\) operators. The results of the \(C\), \(P\) and \(T\) operators and their combinations on the physical parameters are summarized in Table 2.2.
<table>
<thead>
<tr>
<th>Symmetry</th>
<th>Obeyed by</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strong forces</td>
</tr>
<tr>
<td>Parity[P]</td>
<td>yes</td>
</tr>
<tr>
<td>Charge Conjugation[C]</td>
<td>yes</td>
</tr>
<tr>
<td>Time Reversal[T]</td>
<td>yes</td>
</tr>
<tr>
<td>(CP)</td>
<td>yes</td>
</tr>
<tr>
<td>(CPT)</td>
<td>yes</td>
</tr>
</tbody>
</table>

Table 2.3: Invariance of forces under the discrete symmetry transformations

Although the weak interaction violates \(C\) and \(P\) separately, it was for some time assumed to be invariant under a combined \(CP\) transformation. The reason for that is that the \(CP\) transformation is almost conserved. A naive explanation is that the large violation of \(C\) and \(P\) almost cancel each other, making it difficult to observe violation in \(CP\). Finally, although it is now clear that the laws of physics are not invariant under \(C\), \(P\), or \(CP\), scientists are still optimistic that the ultimate discrete symmetry \(CPT\) is not broken. See the summary of invariance of physics laws under different symmetry transformation in Table 2.3.

### 2.3 Cabibbo-Kobayashi-Maskawa Mechanism

Within the Standard Model the only possible source of \(CP\) violation resides in the weak decay of quarks. The weak \((d'\ s'\ b')\) quark eigenstates are in general different from the mass \((d\ s\ b)\) eigenstates. The former could be presented as a linear combination of latter via the unitarity matrix \(\hat{V}_{\text{CKM}}\), or also know as the Cabibbo-Kobayashi-Maskawa (CKM) matrix[8],

\[
\begin{pmatrix}
  d \\
  s \\
  b
\end{pmatrix}
= \hat{V}_{\text{CKM}}
\begin{pmatrix}
  d' \\
  s' \\
  b'
\end{pmatrix};
\hat{V}_{\text{CKM}} = \begin{pmatrix}
  V_{ud} & V_{us} & V_{ub} \\
  V_{cd} & V_{cs} & V_{cb} \\
  V_{td} & V_{ts} & V_{tb}
\end{pmatrix}.
\]  

(2.1)

Specifically, the charged weak current mediating the weak interaction between quarks
The elements of the CKM matrix are in general complex and represent coupling constants that describe the strength of weak transitions from down-type quarks (charge $-1/3e$) to up-type quarks (charge $+2/3e$). Transitions between quarks with the same charge (flavor changing neutral currents) are not allowed and are only possible through the loop diagrams. The values of the matrix elements have been measured precisely in a set of experiments[14]. This matrix is almost diagonal, as the off-diagonal elements correspond to the transition between the quark in different families, which are much less favorable than transitions within a family. See the schematic of the hierarchy of transition amplitudes shown in Fig. 2.1. The only firm constraint, however, on the CKM matrix is that it should be unitary, meaning that the sum of transition probabilities of any quark into the quarks in the other generations is one.

Violation of $CP$ symmetry in the SM comes from the structure of the CKM matrix. Fig. 2.2 illustrates a weak transition of the $b$ to $u$ quark with the emission of a $W^-$ boson. The transition amplitude is equal to $V_{ub}$. If we take a look at this process under the $CP$ conjugate transformation, the $\bar{u}$ quark absorbs a $W^+$ boson, transforming it
into a $\bar{b}$ quark. The corresponding transition amplitude in this case would be $V_{ub}^*$—the complex conjugate of $V_{ub}$. Therefore, in order to have $CP$ violation, at least one element of the CKM matrix must be complex. In addition to that, there must be interference between two or more diagrams. The decay rate goes as the square of the transition amplitude, so one diagram by itself can not produce $CP$ violation.

### 2.4 CKM Matrix Parametrization

In the general case of $N$ quark generations the complex mixing matrix $N \times N$ consists of $2N^2$ real parameters. This number can be dramatically reduced given the invariance of quark fields with respect to an overall phase change and their unitarity.

The unitarity condition imposes several restriction on the matrix. First, the sum of the squares of the elements in any column must add up to one. This gives us $N$ equations. Second, any two different columns should be orthogonal, which gives us $\frac{N(N-1)}{2} \times 2 = N(N - 1)$ additional equations. We multiply the number of columns by 2 because each equation splits into parts, one real and the other imaginary. Thus, unitarity gives $N^2$ constraints on the mixing matrix.

The CKM matrix describes the coupling of $2N$ quarks. We can choose the complex phases of these quark states arbitrarily without affecting the laws of physics. But the redefinition of the complex phase of the quark state changes the corresponding

![Figure 2.2: Charge current interaction in SM.](image-url)
coupling element in the CKM matrix. We can therefore remove $2N - 1$ phases from the mixing matrix through a convenient choice of quark phases. This brings us to just $N^2 - (2N - 1) = (N - 1)^2$ independent elements in the CKM matrix.

In the case of two generations and a $2 \times 2$ mixing matrix, we get only one irreducible parameter—a single mixing angle $\theta_C$ and no mixing angles. The parametrization of the matrix was given by N. Cabibbo

$$\hat{V}_C = \begin{pmatrix}
\cos \theta_C & \sin \theta_C \\
-\sin \theta_C & \cos \theta_C
\end{pmatrix}$$

(2.3)

where $\sin \theta_C \approx 0.22$ is the Cabibbo angle, which can be measured in semileptonic $K \rightarrow \pi e^+\nu_e$ decays. The Cabibbo matrix is purely real, making $CP$ violation impossible.

In the case of three generations we get a total of four irreducible parameters—three Euler angles and one complex phase. This led Kobayashi and Maskawa to surmise the existence of a third generation of quarks to explain $CP$ violation observed experimentally.

The CKM matrix is almost never used in the form of 2.1. There are several representations that are more useful in particle physics. The “standard parametrization” utilizes the Euler angles $\theta_{12}$, $\theta_{23}$, $\theta_{13}$ and $\delta = \delta_{13}$, which is the $CP$-violating phase[15, 16]

$$\hat{V}_{\text{CKM}} = \begin{pmatrix}
{c}_{12}{c}_{13} & s_{12}{c}_{13} & s_{13}e^{-i\delta} \\
-s_{12}{c}_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}{c}_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}{c}_{13} \\
s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}{c}_{13}
\end{pmatrix},$$

(2.4)

where $c_{ij} = \cos \theta_{ij}$ and $s_{ij} = \sin \theta_{ij}$

This interpretation has a very distinct advantage. It is achieved by consequent rotations in planes (12), (13) and (23) and the rotation in plane (13) is accomplished by a phase rotation. The rotation angles are defined in a way that relates to the mixing of two specific generations. If one of these angles vanishes, so does the mixing between those two generations. In the limit of $\theta_{23} = \theta_{13} = 0$ the third generation...
decouples and the situation reduces to the usual Cabibbo mixing of the first two
generations with $\theta_{12}$ identified as the Cabibbo angle.

In practice, it is more convenient to use the CKM matrix in the form that reflects
the hierarchy of the transitional amplitudes

$$s_{12} \simeq 0.22 \gg s_{23} = \mathcal{O}(10^{-2}) \gg s_{13} = \mathcal{O}(10^{-3}) \quad (2.5)$$

Wolfenstein[17] defined a set of parameters, $\lambda$, $A$, $\rho$ and $\eta$ that captured the nearly
diagonal nature of the CKM matrix

$$s_{12} \equiv \lambda, \quad s_{23} \equiv A\lambda^2, \quad s_{13}e^{-i\theta_{13}} \equiv A\lambda^3 (\rho - i\eta), \quad (2.6)$$

then the CKM matrix could be approximated by neglecting terms of order $\lambda^4$ as

$$\hat{V}_{\text{CKM}} = \begin{pmatrix}
1 - \frac{1}{2} \lambda^2 & \lambda & A\lambda^3 (\rho - i\eta) \\
-\lambda & 1 - \frac{1}{2} \lambda^2 & A\lambda^2 \\
A\lambda^3 (1 - \rho - i\eta) & -A\lambda^2 & 1
\end{pmatrix} + \mathcal{O}(\lambda^4). \quad (2.7)$$

Although the Wolfenstein parametrization is more intuitive the standard one, it is
an approximation. For some phenomenological applications the $\mathcal{O}(\lambda^4)$ accuracy may
not be sufficient for extension of the Wolfenstein parametrization beyond the leading
order in $\lambda$ see[18].

### 2.5 Unitarity Triangles

The unitarity of the CKM matrix $VV^\dagger = V^\dagger V = 1$ implies that all columns are
orthogonal. Orthogonality of the first and the third columns gives

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad (2.8)$$

This equation can be visualized as a closed triangle in the complex plane. There are
additional triangles involving the other columns, but this one is of particular interest,
since it has sides of similar length, resulting in three relatively large angles. Usually
it is referred to as the *Unitarity Triangle*. Fig. 2.3 shows the Unitarity Triangle scaled such that the $V_{cd}V_{cb}^*$ side equals one. It was obtained by choosing the phase convention where $V_{cd}V_{cb}^*$ is real, which sets the base of the triangle between the points $(0,0)$ and $(1,0)$. The upper vertex has the coordinates $(\bar{\rho}, \bar{\eta})$ given by

$$\bar{\rho} \equiv (1 - \frac{\lambda^2}{2})\rho, \quad \bar{\eta} \equiv (1 - \frac{\lambda^2}{2})\eta$$

The internal angles correspond to the complex phase represented by the parameter $\eta$. In the absence of $CP$ violation ($\eta = 0$) the triangle would collapse to a straight line. Thus, an observation of $CP$ violation can be considered as a measurement of the parameters of the Unitarity Triangle.

The angles $\phi_1$, $\phi_2$, $\phi_3$ of the Unitarity Triangle are defined as

$$\phi_1 = \pi - \arg(-\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*})$$
$$\phi_2 = \arg(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*})$$
$$\phi_3 = \arg(\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*})$$

By making precise measurements of the sides and angles of the unitarity triangle we can check for consistency of the SM. The measurement of just the angles (which should, in principle, add up to $\phi_1 + \phi_2 + \phi_3 = \pi$) does not guarantee that we will
get a consistent triangle. In addition, constraints on the sides of the triangle should be imposed. It is also important to cross-check the parameters of the triangle using different decay modes.

2.6 \textit{CP} Violation in $B$ Meson Decays

Even before the $b$ quark was discovered, theorists speculated that $CP$ violation has a potentially large effect in decays of $B$ mesons. In 1980 Carter and Sanda\cite{9} proposed a new method for studying $CP$ violation in $B$ mesons. They suggested using the decay of neutral $B$ mesons into a final state $f$ common to both $B^0$ and $\bar{B}^0$. The decay amplitude of $B^0 \to f$ is the sum of the direct $B^0 \to f$ decay amplitude and the amplitude coming after a $B^0 \to \bar{B}_0 \to f$ transition as illustrated in Fig. 2.4. The same applies to the $\bar{B}^0 \to f$ decay amplitude. In the presence of a weak phase difference the sum of these two amplitudes would be different for $B^0$ and $\bar{B}^0$.

Large $B^0 - \bar{B}^0$ mixing is essential to the study of $CP$ violation. If the mixing were small, causing a low level of $B^0 - \bar{B}^0$ oscillation, $B$ mesons would decay before a noticeable difference in decay rates could develop. Fortunately, the ARGUS collaboration\cite{19} observed $B^0 - \bar{B}^0$ mixing at a level large enough to indicate that $CP$ violation in $B$ decays would be experimentally accessible. That triggered a lot of interest in $B$ physics and led to the building of a new generation of experiments specifically dedicated to the study of $CP$ violation in $B$-meson decay.
2.6.1 Oscillations Of $B^0 - \bar{B}^0$ Mesons. Indirect $CP$ Violation

We define $|B^0\rangle$ and $|\bar{B}^0\rangle$ as flavor eigenstates of the $B^0(\bar{b}d)$ and $\bar{B}^0(bd)$ mesons, respectively. The oscillation from one state to the other can occur through the weak interactions. The dominant box diagrams are shown in Fig. 2.5.

The evolution of the system is described by the Schrödinger equation.

$$i\hbar \frac{\partial}{\partial t} \Psi(t) = H \Psi(t) \quad (2.11)$$

where $H \equiv M - \frac{i}{2} \Gamma$ is the Hamiltonian, $M$ is the mass matrix and $\Gamma$ is the decay matrix. Given that both $M$ and $\Gamma$ are Hermitian we get

$$M_{21} = M_{12}^*, \quad \Gamma_{21} = \Gamma_{12}^* \quad (2.12)$$

$CPT$ invariance adds an additional constraint

$$M_{11} = M_{22}, \quad \Gamma_{11} = \Gamma_{22} \quad (2.13)$$

Thus, the Hamiltonian is given by

$$H = \begin{pmatrix} M_{11} - \frac{i}{2} \Gamma_{11} & M_{12} - \frac{i}{2} \Gamma_{12} \\ M_{12}^* - \frac{i}{2} \Gamma_{12}^* & M_{11} - \frac{i}{2} \Gamma_{11}^* \end{pmatrix} \quad (2.14)$$

Eigenstates of this Hamiltonian are the mass eigenstates $B_L$ and $B_H$. In the general case they are not the same as the flavor eigenstates, but can be expressed as
the superposition of the latter

$$|B_L\rangle = p|B^0\rangle + q|\bar{B}^0\rangle$$

(2.15)

$$|B_H\rangle = p|B^0\rangle - q|\bar{B}^0\rangle$$

(2.16)

with the solution of the equation for the eigenvalues

$$\frac{q}{p} = \frac{\sqrt{M_{12}^2 - \frac{i}{2} \Gamma_{12}}}{M_{12} - \frac{i}{2} \Gamma_{12}}$$

(2.17)

defining $\Delta m = m_H - m_L$ and $\Delta \Gamma = \Gamma_H - \Gamma_L$, the evolution of a pure $|B^0\rangle$ or $|\bar{B}^0\rangle$ state at $t = 0$ is given by

$$|B^0(t)\rangle = g_+(t)|B^0\rangle + \frac{q}{p} g_-(t)|\bar{B}^0\rangle$$

(2.18)

$$|\bar{B}^0(t)\rangle = g_+(t)|\bar{B}^0\rangle + \frac{p}{q} g_-(t)|B^0\rangle$$

(2.19)

which means that the flavor remains the same (+) or oscillates into the opposite flavor (−) with a time-dependent probability proportional to

$$|g_\pm(t)|^2 = \frac{e^{-\Gamma t}}{2} [\cosh(\frac{\Delta \Gamma}{2} t) \pm \cos(\Delta mt)]$$

(2.20)

where $\Gamma = (\Gamma_H + \Gamma_L)/2$

If $|p/q|$ differs from unity we observe $CP$ violation. This type of $CP$ violation is called indirect $CP$ violation. It results from the mass eigenstates being different from the $CP$ eigenstates. It could be observed through the asymmetries in semileptonic decays:

$$A_{sl} = \frac{\Gamma(B^0 \to l^+\nu X) - \Gamma(B^0 \to l^-\nu X)}{\Gamma(B^0 \to l^+\nu X) + \Gamma(B^0 \to l^-\nu X)} = \frac{1 - |p/q|^4}{1 + |p/q|^4} = 3 \left( \frac{\Gamma_{12}}{M_{12}} \right).$$

(2.21)

In the Standard Model indirect $CP$ violation in $B^0 - \bar{B}^0$ mixing is suppressed by a factor of $m_c^2/m_t^2$ and is expected to be of order $\mathcal{O}(10^{-3})$. Other types of $CP$ violation, to be explained further below, could be much larger. In the approximation $|q/p| \approx 1$ the mixing parameter $q/p$ can be approximated by a pure phase.
2.6.2 \textit{CP} Violation in Decay Rates

Since under \textit{CP} conjugation $V_{ij} \rightarrow V_{ij}^*$, the irreducible phase of the CKM matrix gives different decay amplitudes for weak transitions and their \textit{CP} conjugates. To avoid the complications associated with the mixing, let us consider decays of the charged $B^+$ and $B^-$ mesons into the final states $f$ and $\bar{f}$. If the amplitude of a weak transition is described by a single Feynman diagram, it can be written in the form

$$A(B^+ \rightarrow f) = |A|e^{i(\delta + \phi)}$$ \hspace{1cm} (2.22)

$$A(B^- \rightarrow \bar{f}) = |A|e^{i(\delta - \phi)}$$ \hspace{1cm} (2.23)

where $\delta$ is a \textit{CP}-even strong phase and $\phi$ is a \textit{CP}-odd weak phase. Since the decay rate is proportional to the square of the amplitude’s absolute value $|A|^2$, there is no difference in decay rates. We have a totally different situation if there is an interference between decay amplitudes, meaning two (or more) Feynman diagrams contribute to the $B^+(B^-) \rightarrow f(\bar{f})$ transition

$$A(B^+ \rightarrow f) = |A_1|e^{i(\delta_1 + \phi_1)} + |A_2|e^{i(\delta_2 + \phi_2)}$$ \hspace{1cm} (2.24)

$$A(B^- \rightarrow \bar{f}) = |A_1|e^{i(\delta_1 - \phi_1)} + |A_2|e^{i(\delta_2 - \phi_2)}$$ \hspace{1cm} (2.25)

defining $\Delta \delta = \delta_2 - \delta_1$ and $\Delta \phi = \phi_2 - \phi_1$, we obtain

$$\Gamma(B^+ \rightarrow f) \propto |A(B^+ \rightarrow f)|^2 = |A_1|^2 + |A_2|^2 + 2|A_1||A_2|\cos(\Delta \delta + \Delta \phi)$$ \hspace{1cm} (2.26)

$$\Gamma(B^- \rightarrow \bar{f}) \propto |A(B^- \rightarrow \bar{f})|^2 = |A_1|^2 + |A_2|^2 + 2|A_1||A_2|\cos(\Delta \delta - \Delta \phi)$$ \hspace{1cm} (2.27)

In the general case $\Gamma(B^+ \rightarrow f) \neq \Gamma(B^- \rightarrow \bar{f})$, the asymmetry in the decay rates between a process and its \textit{CP} conjugate is called \textit{direct \textit{CP} violation}. Substituting these decay rates into the equation for $A_{CP}$ asymmetry we get

$$A_{CP} = \frac{\Gamma(B^+ \rightarrow f) - \Gamma(B^- \rightarrow \bar{f})}{\Gamma(B^+ \rightarrow f) + \Gamma(B^- \rightarrow \bar{f})} = \frac{-2|A_1||A_2|\sin(\Delta \delta)\sin(\Delta \phi)}{|A_1|^2 + 2|A_1||A_2|\cos(\Delta \delta)\cos(\Delta \phi) + |A_2|^2}$$ \hspace{1cm} (2.28)
Note that in order to obtain measurable CP violation in decay rates there must be at least two contributing diagrams and that they must have different weak and strong phases. The extraction of the phase difference $\Delta \phi = \phi_2 - \phi_1$ from the experiment requires the knowledge of the strong amplitudes $A_{1,2}$ and strong phase difference $\Delta \delta$. Direct CP violation is expected to be fairly large in some $B$-meson decays. For example, it could be as large as 30% in $B^+ \to K^+ \bar{K}^0$ decay. Nevertheless, in such decay modes there will be two contributing diagrams with amplitudes of a similar magnitude. Since one is unlikely to find two large amplitudes contributing to a single process, it is generally true that modes with large asymmetries have small branching fractions. Therefore, it is very difficult to make a reliable measurement of the asymmetry due to the small numbers of events for these decay modes.

2.6.3 Decay of Neutral $B$ Mesons into CP Eigenstate

Consider the decay of neutral $B$ mesons into a common CP eigenstate $f_{CP}$. $B^0 - \bar{B}^0$ oscillation provides an elegant mechanism for studying CP violation. Define $A$ and $\bar{A}$ as the instantaneous amplitudes for processes $B^0 \to f$ and $\bar{B}^0 \to f$ assuming no mixing between $B^0$ and $\bar{B}^0$. The interference of $B^0 \to f_{CP}$ and of $B^0 \to \bar{B}^0 \to f_{CP}$ gives the following time dependent decay rates assuming that the $B$ meson is initially in a pure flavor eigenstate

$$\Gamma(B^0(t) \to f_{CP}) = e^{-\Gamma t} |A|^2 \left( \frac{1 + \cos(\Delta mt)}{2} + \right)$$

$$+ \left( \frac{1}{p} \right) ^2 |\bar{A}|^2 \left( \frac{1 - \cos(\Delta mt)}{2} - 3 \left( \frac{q}{p} \bar{A} \right) \sin(\Delta mt) \right) \quad (2.29)$$

$$\Gamma(\bar{B}^0(t) \to f_{CP}) = e^{-\Gamma t} |\bar{A}|^2 \left( \frac{1 + \cos(\Delta mt)}{2} + \right)$$

$$+ \left( \frac{1}{q} \right) ^2 |A|^2 \left( \frac{1 - \cos(\Delta mt)}{2} - 3 \left( \frac{p}{q} A \right) \sin(\Delta mt) \right) \quad (2.30)$$

In the neutral $B$ meson system $|q/p| = 1 + \mathcal{O}(10^{-2})$. Introducing $\lambda_f$ as

$$\lambda_f \equiv \frac{q \bar{A}}{p A}$$

17
the decay rates can be written as

\[
\Gamma(B^0(t) \rightarrow f_{CP}) = e^{-\Gamma t} |A|^2 \times \left( 1 - \frac{|\lambda_f|^2 - 1}{|\lambda_f|^2 + 1} \cos(\Delta mt) - \frac{23\lambda_f}{|\lambda_f|^2 + 1} \sin(\Delta mt) \right) \\
\Gamma(\bar{B}^0(t) \rightarrow f_{CP}) = e^{-\Gamma t} |A|^2 \times \left( 1 + \frac{|\lambda_f|^2 - 1}{|\lambda_f|^2 + 1} \cos(\Delta mt) + \frac{23\lambda_f}{|\lambda_f|^2 + 1} \sin(\Delta mt) \right)
\]

(2.31)

we see that the signs in front of the sine and cosine terms are opposite for the $B^0(t) \rightarrow f_{CP}$ and $\bar{B}^0(t) \rightarrow f_{CP}$ decay rates. By defining $A$ and $S$ as

\[
A \equiv \frac{|\lambda_f|^2 - 1}{|\lambda_f|^2 + 1} \\
S \equiv \frac{23\lambda_f}{|\lambda_f|^2 + 1}
\]

(2.32)

(2.33)

for the $CP$ violation asymmetry we obtain

\[
A_{CP} = \frac{\Gamma(\bar{B}^0 \rightarrow f_{CP}) - \Gamma(B^0 \rightarrow f_{CP})}{\Gamma(\bar{B}^0 \rightarrow f_{CP}) + \Gamma(B^0 \rightarrow f_{CP})} = A \cos(\Delta mt) + S \sin(\Delta mt)
\]

(2.34)

We see that even in the absence of indirect $CP$ violation when $|q/p| = 1$, we can still get a $CP$ asymmetry in the following cases

- $|\bar{A}/A| \neq 1$: $CP$ violation in the decay amplitudes (direct $CP$ violation)
- $\Re\lambda_f \neq 0$: mixing-induced $CP$ violation

The values of $CP$ asymmetry parameters $A$ and $S$ are limited to a circle $A^2 + S^2 \leq 1$ by definition of the parameters.

### 2.7 $CP$ Violation in $B^0 \rightarrow \phi K^0_S$ Decay

The decay $B^0 \rightarrow \phi K^0_S$ proceeds via the quark transition $b \rightarrow s \bar{s}s$, dominated by the QCD penguin transition[20] shown in Fig. 2.6. A small contribution from the electro-weak penguin transition is also possible[21, 22].

For a $b \rightarrow s \bar{s}s$ transition governed by a single decay amplitude, the ratio $\bar{A}/A$ is expressed in CKM matrix elements as

\[
\frac{\bar{A}}{A} = \xi_f \frac{V_{cb}V_{cs}^*}{V_{cb}^*V_{cs}}
\]

(2.35)
where $\xi_f$ is the $CP$ eigenvalue of the $\phi K^0_S$ state. In the calculation of $q/p$ for $B^0 \to \phi K^0_S$ one must also take into account $K^0 - \bar{K}^0$ mixing. Thus, $q/p$ is composed of two terms as

$$
\left( \frac{q}{p} \right)_{B \to \phi K_S} = \left( \frac{q}{p} \right)_{B} \left( \frac{q}{p} \right)_{K_S} \approx \frac{V_{tb}^* V_{td} V_{cs} V_{cd}^*}{V_{tb}^* V_{td}^* V_{cs} V_{cd}}
$$

Finally, we find the following expression for $\lambda_f$ defined in Equation 2.6.3

$$
\lambda_f = \xi_f \frac{V_{tb}^* V_{td} V_{cs} V_{cd}^*}{V_{tb} V_{td}^* V_{cs} V_{cd}} = \frac{V_{tb} V_{td}^* V_{cs} V_{cd}}{V_{tb} V_{td} V_{cs} V_{cd}^*} = \xi_f e^{-2\phi_1}
$$

Therefore the $CP$ asymmetry parameters $\mathcal{A}$ and $\mathcal{S}$ for $B^0 \to \phi K^0_S$ decay become

$$
\mathcal{A}_{\phi K} \simeq 0
$$

$$
\mathcal{S}_{\phi K} \simeq -\xi_{\phi K} \sin 2\phi_1 = \sin 2\phi_1
$$

we used the fact that $\phi K^0_S$ is a $CP$ odd state, hence $\xi_{\phi K} = -1$. The same angle could be measured precisely in $B^0 \to J/\psi K^0_S$ (the so-called gold-plated mode) and other $c\bar{c}K^0$ decays. The main difference is that $B^0 \to J/\psi K^0_S$ occurs through a tree diagram with a negligible contribution from penguin diagrams. Both diagrams are shown in Fig. 2.7. Therefore, this decay has very small theoretical hadronic uncertainties and
can be considered to be the primary measurement of the angle $\phi_1$. On the other hand, the decay $B^0 \rightarrow \phi K^0_S$ could be used as a probe of the Standard Model. Any deviation of the $\phi_1$ measured in the $\phi K^0_S$ channel could be clear evidence for new physics beyond the SM.

Figure 2.7: $B^0_d \rightarrow J/\psi K^0_S$ decay via (a) $b \rightarrow c\bar{c}s$ tree and (b) $b \rightarrow s\bar{c}c$ penguin transitions.

2.8 Overview of the Measurement

In this thesis we present a measurement of the time-dependent $CP$ violation parameters in $B^0 \rightarrow \phi K^0_S$ decays with data taken by the Belle experiment. In Belle, $B$ mesons are produced through the $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$ process. The $\Upsilon(4S)$ is an excited $b\bar{b}$ resonance with a mass of 10.58 GeV, which is just above the $BB$ production threshold. $\Upsilon(4S)$ nearly always decays into $B^0\bar{B}^0$ or $B^+B^-$ with almost the same branching fractions. These $B$ mesons are produced with a very low momentum $p_B^{CMS} \approx 300\text{MeV/c}$ in the $\Upsilon(4S)$ frame. To ensure that the vertices of the two $B$ mesons are separated, the KEKB accelerator works in the asymmetric mode. The energies of the electron and positron beams are 8 GeV and 3.5 GeV respectively, producing a Lorentz boost of $\beta\gamma \approx 0.425$ in the laboratory frame. This allows the $B$ mesons to travel about 200 $\mu$m in the direction of the electron beam before they decay.
In $\Upsilon(4S)$ decay, the $B^0 \bar{B}^0$ pair is produced in a $C$-odd state. Despite the flavor oscillations that occur during the time evolution of this state, the $B^0$ and the $\bar{B}^0$ must remain in opposite states until one of them decays. If we reconstruct one of the $B$ mesons in a flavor specific state, which we call $B_{\text{tag}}$, we can conclude that the other $B$ meson had the opposite flavor at the decay time $t_{\text{tag}}$ of the $B_{\text{tag}}$ meson. We are interested in the case where the other meson decays into the $CP$ eigenstate $K^0_S$. Then, we can approximate the time difference between decays of $B_{CP}$ and $B_{\text{tag}}$ mesons as

$$\Delta t \equiv t_{CP} - t_{\text{tag}} \approx \frac{\Delta z}{\gamma \beta_c} = \frac{z_{CP} - z_{\text{tag}}}{\gamma \beta_c} \quad (2.40)$$

where $z_{CP}$, $z_{\text{tag}}$ are the decay vertex positions of the mesons along the electron beam. Vertices are reconstructed from the charged decay products. The Standard Model predicts the following time-dependent decay rate

$$\Gamma_{B \rightarrow \phi K_S}(\Delta t) \propto e^{-\frac{\Delta z}{\gamma \beta_c}}[1 + q(A_{CP} \cos \Delta m \Delta t + S_{CP} \sin \Delta m \Delta t)] \quad (2.41)$$

where $q$ represents the flavor of the $B_{\text{tag}}$ meson; $q = 1$ corresponds to $B_{\text{tag}} = B^0$ and $q = -1$ corresponds to $B_{\text{tag}} = \bar{B}^0$. Although the explanation above was made assuming $\Delta t \geq 0$, Equation 2.41 also holds for $\Delta t < 0$. A schematic of the experimental measurement is shown in Fig. 2.8.
The observed $\Delta t$ distribution is affected by experimental effects such as background in the $\phi K_S$ sample, errors in assigning the correct flavor to the $B_{tag}$ meson, and finite vertex resolution. We construct the probability density function of the $\Delta t$ distribution taking into account these effects. Then, we perform a maximum likelihood optimization to extract the $A_{CP}$ and $S_{CP}$ parameters.

In keeping with the discussion above, the measurement of time-dependent $CP$ violation requires the following from the experimental apparatus

- Asymmetric collider
- Production of a large number of $B$ mesons
- Efficient $B^0 \rightarrow \phi K^0_S$ reconstruction
- Good decay vertex reconstruction
- Good flavor tagging system
Chapter 3

Experimental Apparatus

3.1 Introduction

The Belle experiment was specifically designed to make a precise measurement of the CP asymmetry in $B_d\bar{B}_d$ meson decays, yielding information on the CKM unitarity triangle angles $\phi_1$, $\phi_2$ and $\phi_3$. The branching fractions of the decays used to measure these asymmetries are of order $10^{-4} - 10^{-6}$. In order to reduce the statistical errors and make measurements competitive with other experiments, a large number of $B$ mesons is required. A good source of $B$ mesons is the $\Upsilon(4S)$ resonance, which is the first resonance with an energy above the $B_d\bar{B}_d$ production threshold. Cross sections of all hadronic reactions of $e^+e^-$ collisions at the energy corresponding to the $\Upsilon(4S)$ resonance are listed in Table 3.1. In 76% of the annihilation collisions that lead to an hadronic final state, jet-like $e^+e^- \rightarrow q\bar{q}$ continuum background is produced, where $q$ one of the light $u,d,c,s$ quarks. The remainder of the hadronic final state events proceed through the production of an $\Upsilon(4S)$ meson, which decays with almost 100% probability to either a $B^0\bar{B}^0$ or a $B^+B^-$ pair. The $e^+e^- \rightarrow \Upsilon(4S)$ cross section is $\approx 1.05 \text{ nb}^{-1}$ which gives roughly one million of pairs for each fb$^{-1}$ of data.

Measuring a time-dependent CP asymmetry requires that one be able to resolve flight length differences between the mesons. $B$ mesons are produced nearly at rest
Table 3.1: Hadronic cross sections at the energy corresponding to the $\Upsilon(4S)$

<table>
<thead>
<tr>
<th>Channel</th>
<th>$\sigma$(nb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Upsilon(4S)$</td>
<td>1.05</td>
</tr>
<tr>
<td>$u\bar{u}$</td>
<td>1.39</td>
</tr>
<tr>
<td>$d\bar{d}$</td>
<td>0.35</td>
</tr>
<tr>
<td>$s\bar{s}$</td>
<td>0.35</td>
</tr>
<tr>
<td>$c\bar{c}$</td>
<td>1.30</td>
</tr>
<tr>
<td>hadronic total</td>
<td>4.44</td>
</tr>
</tbody>
</table>

in the $\Upsilon(4S)$ frame along with the short life of $B$ mesons, making it practically impossible to study the time evolution at symmetric colliders, where the lab and CM frames are the same. To overcome this problem, an asymmetric $e^+e^-$ collider was constructed with a large Lorentz boost of $\beta\gamma \approx 0.42$ along the beam line. The idea of using asymmetric colliders to study the $CP$ violation was originally proposed by Pier Oddone[23]. Even so, a detector with excellent vertexing capability is required.

3.2 The KEKB accelerator and storage rings

The KEKB accelerator, shown schematically in Fig. 3.2, is located on the campus of the High Energy Accelerator Research Organization (KEK) in Tsukuba, Japan. KEKB consists of separate 3-km-long electron and positron rings built in the former TRISTAN tunnel. Electrons produced by an electron gun are accelerated in a linear accelerator (LINAC) to an energy of 7.996 GeV and then directly injected to the High Energy Ring (HER). In the middle of the LINAC part of the electron beam is diverted into a target to produce positrons, which are accelerated by the remaining part of the LINAC to 3.500 GeV and then injected into the Low Energy Ring (LER).

Since Belle is the only experiment that utilizes the KEKB accelerator, there is only one interaction point (IP) located underground in the Tsukuba experimental hall. The electron and positron beams cross at a finite angle of $\pm 11$ rad rather than colliding head-on. This strategy was chosen to eliminate parasitic collisions near
Figure 3.1: The hadronic cross section for $e^+e^-$ annihilation around 10GeV/$c^2$. The measurements are from the CLEO collaboration [24].

the IP due to the multi-bunch operation mode of KEKB. The center-of-mass energy collision of 10.58 GeV corresponds to the peak of the $\Upsilon(4S)$ resonance.

Full understanding of the $q\bar{q}$ background component is crucial for some measurements. For that purpose, about 10% of the time KEKB operates in off-resonance mode with the center of mass energy 60 MeV below the resonance energy to eliminate the production of $B^0\bar{B}^0$ events. KEKB is capable of tuning the energy to 10.869 GeV corresponding to the $\Upsilon(5S)$ resonance. A luminosity of 1.86 fb$^{-1}$ has been collected in June 2005 during a $\Upsilon(5S)$ test run.

The major parameters of the KEKB accelerator are listed in Table 3.2. Additional details can be found in [25].
Figure 3.2: The KEKB accelerator configuration.
<table>
<thead>
<tr>
<th>Description</th>
<th>LER</th>
<th>HER</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circumference, C</td>
<td>3016.26</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>Luminosity, ( \mathcal{L} )</td>
<td>16.52(10)</td>
<td>nb(^{-1})s(^{-1})</td>
<td></td>
</tr>
<tr>
<td>Crossing angle ( \theta_x )</td>
<td>±11 mrad</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particles</td>
<td>( e^+ )</td>
<td>( e^- )</td>
<td></td>
</tr>
<tr>
<td>Energy</td>
<td>3.5 GeV</td>
<td>8.0 GeV</td>
<td>GeV/c(^2)</td>
</tr>
<tr>
<td>Beam Current, I</td>
<td>1.616(2.6) A</td>
<td>1.210(1.1) A</td>
<td></td>
</tr>
<tr>
<td>Bunches, ( N_B )</td>
<td>1388(5000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bunch Current</td>
<td>1.16(0.52) mA</td>
<td>0.871(0.22) mA</td>
<td></td>
</tr>
<tr>
<td>Bunch spacing, ( s_b )</td>
<td>2.1 m</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emittance ( \epsilon_x/\epsilon_y )</td>
<td>18(18)</td>
<td>24(18)</td>
<td>nm</td>
</tr>
<tr>
<td>Beta ( \beta_x^* )</td>
<td>59(33) cm</td>
<td>56(33) cm</td>
<td></td>
</tr>
<tr>
<td>Beta ( \beta_y^* )</td>
<td>0.65(1.0) cm</td>
<td>0.59(1.0) cm</td>
<td></td>
</tr>
<tr>
<td>Hor. size at IP</td>
<td>103(77) ( \mu )m</td>
<td>116(77) ( \mu )m</td>
<td></td>
</tr>
<tr>
<td>Ver. size at IP</td>
<td>1.7(1.9) ( \mu )m</td>
<td>1.7(1.9) ( \mu )m</td>
<td></td>
</tr>
<tr>
<td>Beam-beam, ( \xi_x )</td>
<td>0.106(0.039) ( \mu )m</td>
<td>0.068(0.039) ( \mu )m</td>
<td></td>
</tr>
<tr>
<td>Beam-beam, ( \xi_y )</td>
<td>0.105(0.052) ( \mu )m</td>
<td>0.065(0.052) ( \mu )m</td>
<td></td>
</tr>
<tr>
<td>Beam lifetime at collision</td>
<td>148 at 1616 mA</td>
<td>204 at 1210 mA</td>
<td>min</td>
</tr>
</tbody>
</table>

Table 3.2: KEKB parameters corresponding to the best peak luminosity \( \mathcal{L} = 16.52 \text{ nb}^{-1}\text{s}^{-1} \) recorded on June 30, 2006 with the design values in the parentheses.
3.3 The Belle Detector

The Belle detector was specifically designed to study CP asymmetries in the decays of $B$ mesons produced by an asymmetric $e^+e^-$ collider operating at the $\Upsilon(4S)$ resonance. This requires good particle identification, vertex reconstruction of the charged particles, and precise measurement of the momenta and energies of the final state particles. In addition, the detector should provide good solid angle acceptance to allow reconstruction of modes with many final state particles. To accommodate the boost of the center-of-momentum (CM) frame of the $\Upsilon(4S)$, the detector is made backward/forward asymmetric, providing the coverage in polar angle $\theta$ from $17^\circ$ to $150^\circ$. 
Fig. 3.3 shows the configuration of the Belle detector. The inner part of the detector is inside a constant magnetic field of 1.5 T. The field, which runs parallel to the beam axis, is created by a superconducting solenoidal magnet. Surrounding the cylindrical beryllium beam pipe is the silicon vertex detector (SVD), which provides a very precise measurement of the vertex position of $B$ mesons and other short-lived particles. The central drift chamber (CDC) is used to do charged-particle tracking. A momentum measurement is obtained by measuring the deflection of the particles in the magnetic field and $dE/dx$ information is obtained by measuring the collected charge on the wires. The aerogel Cherenkov counter (ACC) and time of flight (TOF) systems are used for $K/\pi$ separation for different momentum ranges. Fig. 3.4 shows the contribution of each subsystem to the momentum coverage. The last subdetector inside the solenoid coil is the electromagnetic calorimeter (ECL). The electromagnetic showers produced by electrons and photons are detected in an array of CsI(Tl) crystals. Iron layers interspersed with resistive plate chambers (RPCs) form the KLM system. The iron layers provide the magnet’s flux return and the $K_L/\mu$ detection. The purpose of the extreme forward calorimeter (EFC) is to improve the coverage in the forward and backward directions. The EFC consists of a pair BGO crystal arrays placed at the front and at the end of the detector.

The Belle coordinate system is defined as follows:

- the $z$ axis points in the opposite direction of the positron (low energy) beam, aligned with the magnetic field.
- the $y$ axis points vertically up
- the $x$ axis points outward from the ring to form a right-hand system

The following sections give a brief description of each system. A detailed description including electronic layout can be found in [26]
3.3.1 Silicon Vertex Detector (SVD)

The silicon vertex detector plays a central role in the Belle experiment. It provides a precise measurement of the vertex position of charged particles and improves the tracking. In the measurement of time-dependent $CP$ asymmetries we need to know the proper-time $\Delta t$ difference, which is inferred from a measurement of the $z$-vertex displacement of the neutral $B$ mesons. In order to increase the separation of the vertices the $\Upsilon(4S)$ is boosted along the $z$ axis with a Lorentz factor $\beta\gamma \approx 0.425$. Still, due to the short lifetime of the $B$ meson, the average displacement of the $B$ mesons is just around 200$\mu$m. This requires the SVD resolution to be around 100 $\mu$m or better.

The SVD detector is based on double-sided silicon strip detectors (DSSDs) fabricated by Hamamatsu Photonics. Each DSSD consists of 1280 sense strips and 640 readout pads on opposite sides. The strips on one side make a measurement in the $z$ direction, while those on the other side are oriented in the $\phi$ direction. The $z$-strip pitch is 42 $\mu$m and the $\phi$-strip pitch is 25 $\mu$m. The induced current on the strips is readout using a hybrid card.

Two different designs were employed during the course of data taking: SVD1 and
SVD2. The original design configuration is shown in Fig. 3.5. Either one or two DSSDs connected to a hybrid unit form a short or long half ladder, respectively. A hybrid unit is produced by gluing two hybrids back-to-back to read both sides of the DSSD. Two electrically independent half ladders are then mechanically joined together to form a full ladder. Full ladders are arranged in three cylindrical layers placed around the beam pipe at 30 mm, 45.5 mm and 60.5 mm from the center. There are 8, 10, 14 ladders in the inner, middle and outer layers respectively. The SVD covers the polar angle range $23^\circ \leq \theta \leq 139^\circ$, which corresponds to 86% of the full solid angle.

In the summer of 2004 the three-layer SVD was replaced with a newly designed four-layer SVD, referred to as SVD2. Several improvements were made. SVD2 has a larger angular acceptance of $17^\circ \leq \theta \leq 150^\circ$. In addition, the first layer was moved closer to the primary interaction point at a distance of 20 mm instead of 30 mm as before. This was made possible by a significantly smaller beam pipe. The four layers
consist of 6, 12, 18 and 18 full ladders. The SVD2 geometry is shown in Fig. 3.6.

More information on SVD1 can be found in references[27, 28] and on SVD2 in [29, 30].

### 3.3.2 Central Drift Chamber (CDC)

The central drift chamber (CDC) is the core part of the tracking system. It provides reconstruction of charged particles, measurements of their momenta and also helps with particle identification by providing the ionization energy loss $dE/dx$.

The structure of the CDC is shown in Figure. 3.7. It is a cylindrical drift chamber filled with a mixture of 50% helium and 50% ethane gas, chosen to minimize multiple scattering and to provide good $dE/dx$ resolution. The length of the CDC is 2, 400 mm, and the inner and outer radii are 103.5 mm and 874 mm, respectively. It is asymmetric in the $z$ direction and provides an angular coverage $17^\circ \leq \theta \leq 150^\circ$.

The CDC consists of 50 anode sense wire layers and three cathode strip layers.
The 50 anode layers consist of 32 axial wire layers and 18 small-angle stereo wire layers. The axial wires are parallel to the $z$-direction and the crossing angles between stereo wires and $z$-direction reach from 42.5 mrad to 72.1 mrad. Axial layers measure the $r$-$\phi$ position. Stereo layers work in conjunction with axial layers to measure the $z$ position. The spatial resolution in $r$-$\phi$ is 130 $\mu$m, and is better than 2 mm in the $z$ direction. There are 8400 drift cells made of one sense wire surrounded by field wires. Reference [31] provides additional information on CDC.

The CDC is situated in a 1.5 T magnetic field supplied by a solenoid magnet. The field is nearly uniform and is parallel to the $z$ direction. The charged particles moving in the magnetic field follow a helical path: circular motion in the $r$-$\phi$ plane and motion in a straight line along the $z$ axis. There are five independent parameters characterizing the helix [32]. Initially, they are fitted from CDC information, then each track is matched with SVD information to improve the determination of these parameters.

In addition to track reconstruction, the CDC provides information on the energy
deposited in the gas by the charged particles. This measurement is taken from the ionization charge picked up by the sense wires along the particle’s trajectory. Since \(dE/dx\) mainly depends on the velocity of the particle \((\beta \equiv v/c = p/E)\), particles of different mass with the same momentum have a different ionization loss. The value of \(dE/dx\) is obtained using a truncated-mean method in order to remove Landau tails in the \(dE/dx\) distribution.

### 3.3.3 Aerogel Cherenkov Counter (ACC)

The aerogel Cherenkov counter system is designed for separation of \(K^\pm\) and \(\pi^\pm\) with momentum between 1.2 and 3.5 GeV/c. It uses the Cherenkov light that is emitted when a charged particle travels through a medium faster than the speed of light in that medium. Specifically, for a medium of refractive index \(n\), Cherenkov light is emitted if the velocity of the particle satisfies:

\[
n > \frac{1}{\beta} = \sqrt{1 + \left(\frac{m}{p}\right)^2}
\]

where \(m\) is the mass of the particle and \(p\) its momentum. The emitted light forms a cone with a half angle equal to:

\[
\cos \theta_c = \frac{1}{n\beta}
\]

Since kaons are heavier than pions, a kaon of a given momentum travels more slowly than a pion of the same momentum. So there is a momentum range for which pions will emit light in the aerogel and kaons will not.

An ACC module (see Fig. 3.8) is a threshold device which just determines the presence of Cherenkov light. It is made of five aerogel tiles stacked in a thin aluminum box. The Cherenkov light is detected by one or two fine mesh photomultiplier tubes (FM-PMTs) attached directly to the aerogel at the side of the box. If the index of refraction of the Cherenkov device is chosen appropriately, then at a given momentum, the lighter particles will trigger the ACC module, while heavier particles will
Figure 3.8: Barrel and endcap ACC modules
not. One of the advantages of aerogel is that its index of refraction can be varied in the production process.

The ACC system consists of a barrel part comprising 960 counter modules (around the CDC) and a forward endcap comprising 228 modules, as shown in Fig. 3.9. Due to the asymmetric KEKB beams, final state particles emitted at a small polar angle tend to have larger momenta than those at large polar angles. To compensate for that and to allow the optimum separation aerogel modules with different refraction indices are used.

Additional details on the ACC system can be found in [33].

3.3.4 Time of Flight System (TOF)

The time-of-flight (TOF) system provides additional information for $K/\pi$ separation of charged particles in the low momentum range $p \leq 1.2$ GeV/c. In addition to particle identification, it provides the fast timing signals for the trigger system.
A TOF module (see Fig. 3.10) consists of two adjacent TOF counters separated by a 1.5 cm gap from a trigger scintillation counter (TSC). The signal from a particle crossing the TSC is used in coincidence with the two adjacent TOF counters to create a trigger signal. In total, there are 64 TOF modules located at a radius of 1.2 m from the interaction point covering a polar angle range from 34° to 120°.

Bicron BC408 is used as scintillator material for the TOF system. BC408 is a fast plastic scintillator with a large attenuation length to meet the design goal of 100 ps time resolution. Since the TOF system is placed in a magnetic field of 1.5 T fine-mesh-dynode photo-multiplier tubes (FM-PMT) are used to collect the scintillation light from the counters. FM-PMTs are mounted directly to the TOF counters to avoid the use of light guides. This strategy was chosen to minimize the time dispersion of scintillation photons propagating in the counters.

The TOF measures the time of flight between a particle originating at the IP and passing through the scintillator. The mass of a particle of momentum $p$ traveling a distance $L$ in time $T$ is given by:

$$M = p\sqrt{\frac{1}{\beta^2} - 1} = p\sqrt{\frac{cT}{L}}$$

Fig. 3.11(a) shows the distribution of the calculated mass using the momentum
(a) Mass distribution from TOF measurements for particles with momentum below 1.2 GeV/c.

(b) $K^\pm/\pi^\pm$ separation by TOF

Figure 3.11: $K^\pm/\pi^\pm$ separation performance of TOF system for hadronic particles.

as measured by the CDC and time measured by the TOF for tracks with momentum below 1.2 GeV/c. The solid histogram represents a Monte Carlo prediction obtained assuming a time resolution of 100 ps. The three peaks correspond to pions, kaons and protons. The data points are in good agreement with the expected distribution. The level of $K^\pm/\pi^\pm$ separation as a function of particle momentum is plotted on Fig. 3.11(b). $\sigma_K$ and $\sigma_\pi$ are the time resolutions of kaons and pions respectively, $\mu_K$ and $\mu_\pi$ are the average flight time. The TOF system provides better than $2\sigma$ separation for particles with momenta up to 1.25 GeV/c.

More details on the TOF system can be found in [34].

3.3.5 Electromagnetic Calorimeter (ECL)

The main purpose of the electromagnetic calorimeter (ECL) is the detection of photons and electrons with good energy resolution and position. Most of the photons produced in the Belle experiment are the end products of cascade decays, and have energies below 500 GeV. Important two-body decay modes such as $B \rightarrow K^*\gamma$ and
$B \rightarrow \pi^0\pi^0$ produce photons with energies up to 4 GeV. In order to reduce background for these decay modes, high resolution is needed. Electron identification in Belle is mainly done by comparison of the particle momentum and deposited energy in the calorimeter. Reconstruction of $\pi^0$’s involves the separation of two nearby photons, separated by a small opening angle. Good performance of the ECL over a wide range of angles and energies is therefore important.

Calorimeters detect the electromagnetic showers produced by electrons and photons passing through dense media. Electromagnetic showers are the product of repeated cascades of pair production and brehmstrahlung. As the shower propagates, the number of particles produced increases exponentially, with a corresponding reduction in the energies of the particles in each succeeding generation. The cascade continues until ionization losses eventually terminate the shower. In a crystal scintillator such as CsI, electrons and positrons excite the bands in the crystal lattice, resulting in the subsequent emission of scintillation light in the visible spectrum.

In order to meet the resolution requirement, Belle employs a highly segmented array of CsI(Tl) crystals in its ECL system. The configuration of the ECL is shown in Fig. 3.12. It consists of a barrel section combined with two end caps, covering the azimuthal angle range $17^\circ \leq \theta \leq 150^\circ$, corresponding to a total solid-angle coverage of 91%. There are 1152, 6624 and 960 crystals in the forward end cap, barrel and backward end cap sections respectively. The individual crystals are oriented so that they point almost toward the interaction point. The crystals are tilted slightly to prevent photons from escaping through the gaps between the crystals.

Each CsI(Tl) crystal has a tower-like shape as shown in Fig. 3.13. The cross-sectional dimensions of the crystals vary with polar angle $\theta$, but the length of all crystals is the same and equal to 30 cm. This corresponds to 16.2 radiation lengths for electrons and 0.8 nuclear interaction lengths for $K_L$’s. Each crystal is wrapped in a 200 $\mu$m-thick porous goretex teflon diffusion material for better light collection.
Figure 3.12: Configuration of the electromagnetic calorimeter (ECL).

Figure 3.13: ECL counter module.
Figure 3.14: ECL energy resolution measured with Bhabha and $\gamma\gamma$ events.

The crystals are then covered by 25 $\mu$m-thick aluminum layers and 25 $\mu$m-thick mylar sheets for light and electrical shielding. A pair of independently read Hamamatsu S2744-08 photodiodes is glued at the center of crystal rear end surface via an 1 mm-thick acrylic plate.

The energy calibration of the ECL is based mainly on using $e^+e^- \rightarrow e^+e^-$ (Bhabha) and $e^+e^- \rightarrow \gamma\gamma$ events, for which the energy is known as function of the detection angle. The resulting energy resolution is plotted in Fig. 3.14. See [35] for more details.
3.3.6 Extreme Forward Calorimeter (EFC)

The extreme forward calorimeter (EFC) is designed to extend the polar angle coverage for the Belle detector. Photons and electrons are measured by the EFC in the extreme forward $6.4^\circ \leq \theta \leq 11.5^\circ$ and extreme backward angular ranges $163.3^\circ \leq \theta \leq 171.2^\circ$ can not be detected by the ECL. This additional angular coverage is useful for some physics processes such as $B \to \tau \nu$ and for two-photon analyses.

The EFC is placed around the beam pipe very close to the IP as shown in Fig. 3.15. Since the radiation level there is very high, bismuth germanate $\text{Bi}_4\text{Ge}_3\text{O}_{12}$ was chosen as a scintillating material. It sustains high radiation doses and still provides good energy resolution. Both crystal arrays are made of 160 bismuth germanate crystals arranged in 32 azimuthal segments and 5 polar segments. The forward and backward crystals correspond to 12 and 11 radiation lengths. The energy resolution of the forward EFC was measured to be 7.3\% at 8 GeV, while the backward EFC has 5.8\% resolution at 3.5 GeV.
In addition, the EFC is used for the protection of the CDC chamber from background and online luminosity monitoring based on the well known Bhabha scattering rate. See [36] for more details.

3.3.7 $K_L$/Muon Detector (KLM)

Muons and neutral long-lived kaons deposit only a small amount of their energy in the ECL. Therefore an additional massive detection sub-system was built as the outermost part of Belle detector.

The KLM detection system was designed to identify $K_L$’s and muons with high efficiency in the momentum range $0.6 - 5.0$ GeV/c. It consists of barrel, forward endcap and backward endcap sections. The barrel section covers an angular range from $45^\circ$ to $125^\circ$ in polar angle. The endcaps extend this range to $20^\circ \leq \theta \leq 155^\circ$. The basic detector element of the KLM is the Resistive Plate Counter (RPC). RPC modules are interspersed with 4.7 cm thick iron plates. There are 15 detector layers.
and 14 iron layers in each octagonal barrel section and 14 detector layers and 14 iron layers in each endcap. The iron plates provide a total of 3.9 hadron interaction lengths of material. In addition to that, 0.8 interaction lengths are provided by the ECL.

Each RPC module consists of two independent RPCs arranged back to back as depicted in Fig. 3.16. Each layer is made of two glass plates coated with conductive Koh-i-noor 3080F ink separated by extruded noryl spacers (Fig. 3.17). A 1.9 mm gap between the plates is filled with a mixture of the hydrofluorocarbon (HFC) gas R134a, argon and butane. This non-combustible and environmentally friendly mixture provides high detection efficiency and stable RPC operation. Each RPC is electrically insulated with a layer of 0.125 mm thick mylar. A high voltage 8000 V is applied to the conducting ink layers, which charge the glass plates. Two RPCs are sandwiched together between orthogonal θ and φ pickup-strips with ground planes for signal reference and proper impedance. The whole structure is placed into an aluminum box less than 3.7 cm thick.

Charged particles crossing the RPC leave a trail of ionization in the gas. This
results in a discharge between the plates along the ionization trail. Since glass is a poor electrical conductor, only the charge stored on the plates in the immediate vicinity of the track is available to the discharge, resulting in rapid quenching. The orthogonal strips sense the position of the discharge in two dimensions, giving a three-dimensional hit position. Since the total charge in the pulse depends on the amount of charge stored on the plates and is largely independent of the amount of primary ionization, there is no useful information on the energy deposition.

For momenta above 1.5 GeV/c, the muon identification efficiency is greater than 90%. The $K_L$ detection efficiency has been studied using $e^+e^- \rightarrow \gamma\phi(K_S K_L)$ events and was found to be 50% for $K_L$'s detected only in the KLM and 65% when KLM and ECL information were combined. The KLM angular resolution measured taking the IP as one end of the neutral track is better than 10 mrad. For more information on the KLM detector see [37, 38, 39].
Chapter 4

Reconstruction of $B^0 \rightarrow \phi K_S$

In this chapter we describe the experimental procedure for the reconstruction of $B^0 \rightarrow \phi K_S$ events from the data sample collected by the Belle experiment. We start with a description of the selection criteria used in the initial $e^+e^- \rightarrow B_0\bar{B}_0$ skim. We then explain the reconstruction of long lived particles like $K^\pm$ and $\pi^\pm$. Later we proceed to reconstruction of the composite particles $K_S$, $\phi$ and subsequently to the reconstruction of the $B^0$ meson. Finally, we discuss the combinatorial background contamination and describe techniques for its suppression.

4.1 Hadronic $B\bar{B}$ Selection

There are a number of other hadronic final state processes besides $e^+e^- \rightarrow B\bar{B}$ at an energy of 10.58 GeV, some with cross sections comparable to or even large than that for $B\bar{B}$ production. The main processes are $q\bar{q}$, QED processes like Bhabha or radiative Bhabha events, $e^+e^- \rightarrow \mu^+\mu^-$ and $e^+e^- \rightarrow \tau^+\tau^-$ pair production, two-photon events. There are also beam-gas backgrounds. The hadronic event skim “HadronB”[40] is used as a starting point for $B \rightarrow \phi K_S$ reconstruction. The HadronB selection criteria can be summarized as:

- At least three good charged tracks, satisfying $|r| < 2.0$ cm and $|z| < 4.0$ cm
<table>
<thead>
<tr>
<th>Process</th>
<th>$BB$</th>
<th>$q\bar{q}$</th>
<th>$\tau^+\tau^-$</th>
<th>QED</th>
<th>$\gamma\gamma$</th>
<th>Beam Gas</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{total}(nb)$</td>
<td>1.1</td>
<td>3.3</td>
<td>0.93</td>
<td>37.8</td>
<td>11.1</td>
<td>N/A</td>
</tr>
<tr>
<td>$\epsilon(%)$</td>
<td>0.991</td>
<td>0.795</td>
<td>0.049</td>
<td>0.00002</td>
<td>0.004</td>
<td>0.09$\epsilon_A$</td>
</tr>
<tr>
<td>$\sigma$(nb)</td>
<td>1.09</td>
<td>2.62</td>
<td>0.05</td>
<td>0.001</td>
<td>0.04</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Table 4.1: Cross sections and efficiencies of HadronB skim for various processes

at the closest approach to the beam axis and having transverse momentum $p_t > 0.1$ GeV/$c$.

- There are at least two good ECL clusters within the volume of $-0.7 < \cos \theta < 0.9$, where “good” means an energy deposition greater than 0.1 GeV.

- The total energy of the good ECL clusters in the $\Upsilon(4S)$ rest frame is required to be between 10% and 80% of the total energy $\sqrt{s}$.

- The total momenta of good charged tracks and good ECL clusters not associated with any track in CDC measured in the $\Upsilon(4S)$ rest frame has $z$ component smaller than 50% of the total energy $\sqrt{s}$.

- The total visible energy in the $\Upsilon(4S)$ rest frame $E_{vis}^{CMS}$ is greater than 20% of the total energy $\sqrt{s}$. Visible energy is defined as the sum of energies of good ECL clusters and good tracks under the pion mass assumption.

- The primary vertex formed by all good tracks must be within 1.5 cm and 3.5 cm of the interaction point in $r$ and $z$ directions respectively.

- The invariant mass of particles found in the hemispheres perpendicular to the event thrust axis is required to be greater than 1.8 GeV/$c^2$.

- The average energy deposition of good ECL clusters in the $\Upsilon(4S)$ rest frame $E_{sum}^{CMS}/N_{ECL}$ is smaller than 1 GeV.
• The ratio of the heavy-jet mass to the visible energy $M_{\text{HJ}}/E_{\text{vis}}^{\text{CMS}}$ is greater than 0.25 or the heavy-jet mass is greater than 1.8 GeV/c$^2$.

• The total energy deposition of good ECL clusters (with no polar angle requirement) in the $\Upsilon(4S)$ rest frame is greater than 18% of the total energy $\sqrt{s}$ or the heavy-jet mass $M_{\text{HJ}}$ is greater than 1.8 GeV/c$^2$.

The heavy-jet mass is calculated in the following way. The event is split into two hemispheres by the plane perpendicular to the thrust axis forming two jets from the tracks in the same hemisphere. The invariant mass then is calculated for both jets using good tracks under a pion mass assumption. The larger invariant mass is considered to be the heavy-jet mass of the event. The heavy-jet mass is essentially a $\tau$ invariant mass. Thus, the last two conditions pass continuum events if they are not consistent with a $\tau$-pair event.

Approximate cross sections and efficiencies for the HadronB hadronic skim are given in Table 4.1. $q\bar{q}$ denotes the continuum production of $e^+e^\rightarrow q\bar{q}$ quark pairs, where $q$ is one of the four lightest quarks $u, d, s$ or $c$. QED denotes Bhabha ($e^+e^- \rightarrow e^+e^-$) and radiative Bhabha scattering. The latter includes extra photons and leptons in the final state. The beam gas background results from imperfect vacuum in the beam pipe. Beam electrons and positrons interact with molecules of the gas in the beam pipe, producing charged tracks in the drift chamber. Despite a very large cross section for Bhabha type events, they are efficiently suppressed by the above criteria. HadronB skim retains 99% of $BB$ events, while keeping the contamination from non-hadronic processes smaller than 5%.

### 4.2 Event Reconstruction

We reconstruct $B^0 \rightarrow \phi K_S$ candidates using the following decay sub-channels $\phi \rightarrow K^+K^-$ and $K_S \rightarrow \pi^+\pi^-$. This requires us to be able to identify charged kaons and
pions. All good charged tracks are considered to be pion and kaon candidates and go through the PID algorithm identification.

### 4.2.1 Identification of Electrons and Muons

Electron and muon candidates are selected from the sample of reconstructed charged tracks through the imposition of selection criteria. In order to distinguish electrons from muons and hadrons we exploit the differences in the magnetic shower induced in the ECL calorimeter, energy loss $dE/dx$ measured in the central drift chamber, light yield in the Aerogel Cherenkov counters (ACC), and the hit pattern in the KLM detector. Information is then combined into a single variable using a likelihood method. We construct different likelihoods for electrons and muons, called EID and MuID, respectively. We identify the charged tracks as electrons or muons by requiring that the corresponding likelihood be above a certain threshold.

We use the following five discriminants in the EID likelihood

- Matching between the position of the charged track extrapolated to the ECL and the measured position of the energy cluster.

- Ratio of the energy measured by the ECL and the momentum measured by the CDC.

- $E_9/E_{25}$ transverse shower shape parameter: the ratio of the energy in the $3 \times 3$ and the $5 \times 5$ arrays of crystals surrounding the central crystal.

- Energy loss $dE/dx$ measured in the CDC.

- Light yield in the ACC.

Muons pass through the ECL with almost no interaction but leave a track in the KLM detector. Two discriminants are used for the MuID likelihood:
• $\Delta R$, which is the difference between the measured and expected penetration depth range of the track in the KLM, based on the momentum measured by the CDC.

• $\chi^2_r$, which corresponds to the difference between the extrapolated position of the CDC track into the KLM and the actual position of the associated RPC hits in the transverse direction. $\chi^2_r$ is normalized according to the number of hits associated with the charged track.

See detailed discussion of electron and muon identification in references [41] and [42] respectively.

4.2.2 Particle Identification of Charged Hadrons

Charged hadron identification in Belle is based on $dE/dx$ measurements in the central drift chamber, time-of-flight measurements and the response of the aerogel Cherenkov counters. These three nearly independent measurements provide good separation between particle species in different momentum and angular regions. Therefore, it is necessary to combine the sub-detector information in some way, to allow efficient particle identification for different charge track candidates. As is the case for electron and muon identification, we use a likelihood method to combine these measurements into a single cut variable

$$\mathcal{L}(h) = \mathcal{L}^{\text{ACC}}(h) \times \mathcal{L}^{\text{TOF}}(h) \times \mathcal{L}^{\text{CDC}}(h)$$  \hspace{1cm} (4.1)

where $h$ denotes the hadron type $K$, $\pi$ or $p$. To identify kaons and pions we impose a cut on the corresponding likelihood ratio $\text{PID}$

$$\text{PID}(K) = \frac{\mathcal{L}(K)}{\mathcal{L}(K) + \mathcal{L}(\pi)}$$ \hspace{1cm} (4.2)

$$\text{PID}(\pi) = \frac{\mathcal{L}(\pi)}{\mathcal{L}(K) + \mathcal{L}(\pi)} = 1 - \text{PID}(K)$$ \hspace{1cm} (4.3)

$$\text{PID}(p) = \frac{\mathcal{L}(p)}{\mathcal{L}(K) + \mathcal{L}(p)}$$ \hspace{1cm} (4.4)
In our analysis we impose only a very loose cut on kaons $PID(K) > 0.1$ to retain a good efficiency. In addition to the kaon-pion separation requirement, kaon track candidates must satisfy a proton veto: the proton likelihood ratio $PID(p)$ should be smaller than 0.95. We also reject charged tracks identified as electrons with corresponding likelihoods greater than 0.99.

We do not impose any PID requirement on pions used to reconstruct $K_S$. There the kinematic fit and vertex constraints are more efficient than generic particle identification. Refer to [43, 44] for more detailed information about PID in Belle.

### 4.2.3 $K_S$ Reconstruction

The main branching fractions of $K_S$ meson are

\begin{align*}
B(K_S \rightarrow \pi^+\pi^-) &= (68.95 \pm 0.14)\% \quad (4.5) \\
B(K_S \rightarrow \pi^0\pi^0) &= (31.05 \pm 0.14)\% \quad (4.6)
\end{align*}

We only use $K_S$ mesons reconstructed from $K_S \rightarrow \pi^+\pi^-$ decay. Neutral pions are reconstructed from pairs of photons $\pi^0 \rightarrow \gamma\gamma$, resulting in four relatively low-energy photons in the final state. There is also more than one $\pi^0$ meson on average coming from the other $B^0$ meson. This makes the use of $K_S$ reconstructed from $\pi^0$'s to be technically difficult due to a high level of combinatorial background.

In the reconstruction of $K_S$ we take advantage of its relatively large lifetime. In general, a $K_S$ travels several cm before it decays ($c\tau \approx 2.7$ cm). Since the majority of the charge tracks originate close to the IP, we can easily suppress them by requiring $K_S$ children to have a large distance from the IP. In addition to that, the reconstructed vertex of the $K_S$ decay should be displaced in the direction of the measured momentum of the $K_S$. These requirements turn out to be much more efficient in separation of the signal from background than PID, so we do not apply additional PID selection and take all charged tracks as $\pi^\pm$ candidates.

We use the following four variables to select $K_S$ candidates[45]:

---

51
<table>
<thead>
<tr>
<th>Momentum (GeV/c^2)</th>
<th>dr(cm)</th>
<th>d\phi(rad)</th>
<th>z_dist(cm)</th>
<th>fl(cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.5</td>
<td>&gt; 0.05</td>
<td>&lt; 0.3</td>
<td>&lt; 0.8</td>
<td>-</td>
</tr>
<tr>
<td>0.5 – 1.5</td>
<td>&gt; 0.03</td>
<td>&lt; 0.1</td>
<td>&lt; 1.8</td>
<td>&gt; 0.08</td>
</tr>
<tr>
<td>&gt; 1.5</td>
<td>&gt; 0.02</td>
<td>&lt; 0.03</td>
<td>&lt; 2.4</td>
<td>&gt; 0.22</td>
</tr>
</tbody>
</table>

Table 4.2: \(K_S\) selection criteria.

- *dr*: the smaller of \(dr_1\) and \(dr_2\), which are the closest approach of the charged tracks to the IP in the \(x-y\) plane.
- *d\(\phi\)*: the azimuthal angle between the momentum of the \(K_S\) and the vector from IP to decay vertex.
- *z_dist*: the mismatch in the \(z\) direction between two daughter tracks at the reconstructed vertex.
- *fl*: the flight length of the \(K_S\) candidate in the \(x-y\) plane.

We split \(K_S\) candidates into three categories depending on the momentum in the lab frame. The selection criteria for each momentum range \(< 0.5\), \(0.5-1.5\) and \(> 1.5\) are listed in Table 4.2. We define the mass of the \(K_S\) candidate as the invariant mass of the two charge tracks using the pion mass assumption:

\[
M_{K_S}^2 = \left( \sqrt{M_{\pi^+}^2 + |\vec{P}_{\pi^+}|^2} + \sqrt{M_{\pi^-}^2 + |\vec{P}_{\pi^-}|^2} \right)^2 - |\vec{P}_{\pi^+} + \vec{P}_{\pi^-}|^2. \tag{4.7}
\]

Fig. 4.1 illustrates the invariant mass distribution of \(K_S\) candidates after the vertex constraint. These selections reduce the background in the \(K_S\) sample to almost zero. We require \(M_{K_S}\) to be within 12 MeV/c^2 of the \(K_S\) mass, which corresponds to around 2.5\(\sigma\).

### 4.2.4 \(B^0\) Candidate Reconstruction

Given reconstructed \(K^+, K^-\) and \(K_S\) candidates, we can proceed to the reconstruction of \(B^0 \rightarrow \phi K_S\) decays. We first reconstruct a \(\phi\) meson from two charged kaons by
Figure 4.1: Invariant mass distribution of oppositely charged tracks after the $K_S$ selection criteria are applied. The dashed and solid lines represent the fitted curves of signal and signal plus background events respectively.

requiring the invariant mass of the $K^+$ and $K^-$ to be within 10 MeV of the $\phi$ mass:

$$M_{\phi}^2 = \left( \sqrt{M_{K^+}^2 + |\vec{P}_{K^+}|^2} + \sqrt{M_{K^-}^2 + |\vec{P}_{K^-}|^2} \right)^2 - |\vec{P}_{K^+} + \vec{P}_{K^-}|^2. \quad (4.8)$$

Fig. 4.2 shows the invariant mass distribution of the $\phi \rightarrow K^+K^-$ candidates. At this point, there is still a large level of combinatorial background. This background component is dramatically reduced after requiring the event to be consistent with $B^0 \rightarrow \phi K_S$ decay. A relatively aggressive cut on $M_\phi$ is used to reduce the contribution of $B^0 \rightarrow K^+K^-K_S$ and $B^0 \rightarrow f^0K_S$ decays, which have the opposite sign of $CP$ asymmetry.

$B^0$ candidates are reconstructed by combining $\phi$ and $K_S$ candidates. We use two variables to select signal candidates: the energy difference $\Delta E$ and the beam constraint mass $M_{bc}$ defined as:

$$\Delta E = E_{B}^{\text{cms}} - E_{\text{beam}}^{\text{cms}} \quad (4.9)$$

$$M_{bc} = \sqrt{(E_{\text{beam}}^{\text{cms}})^2 - (p_B^{\text{cms}})^2} \quad (4.10)$$
where $E_{\text{beam}}^{\text{cms}} = \sqrt{s}/2$ is the beam energy in the center-of-mass frame, and $E_B^{\text{cms}}$ and $p_B^{\text{cms}}$ are the cms energy and momentum respectively of the reconstructed $B$ candidate.

Since the $B$ meson pair is created by a beam collision, $\Delta E \approx 0$ for signal events. The $M_{bc}$ distribution peaks around the real mass of $B$ meson $M_B = 5.279$ GeV/$c^2$. We use the beam-constrained mass quantity because it takes into account the correction due to run-dependent beam energy. Fig. 4.3 shows the distributions of $M_{bc}$ and $\Delta E$ for the signal MC sample. We define the signal region as $|\Delta E| < 0.045$ GeV and $5.274 < M_{bc} < 5.286$ GeV/$c^2$.

If there are multiple candidates found in a single event, we select the candidate with the smallest vertex quality $\chi^2$ for the reconstructed $B_{\text{rec}}$ meson. If there are still two or more candidates remaining with the same $\chi^2$ (combination of the same $\phi$ meson candidate and different $K_s$ candidates), we use the most probable candidate based on the information on $\Delta E$, $M_{bc}$ and the invariant mass of the $\phi$ candidate $M_\phi$. The secondary $\chi^2_s$ is calculated as

$$\chi^2_s = \left( \frac{\Delta E}{\sigma_{\Delta E}} \right)^2 + \left( \frac{M_{bc} - M_B^s}{\sigma_{M_{bc}}} \right)^2 + \left( \frac{M_\phi - M_\phi^s}{\sigma_{M_\phi}} \right)^2$$

(4.11)
where $M_B^*$ and $M_\phi^*$ are the masses of $B$ and $\phi$ mesons, $\sigma_{\Delta E}$, $\sigma_{M_{bc}}$, $\sigma_{M_\phi}$ are the errors on $\Delta E$, $M_{bc}$ and $M_\phi$ respectively.

### 4.3 Background Study

The dominant background in the reconstruction of the $B^0 \rightarrow \phi K_S$ comes from continuum $e^+e^- \rightarrow q\bar{q}$ events, where $q\bar{q}$ is the one of the light quark-antiquark pairs with $q = u, d, c, s$. This type of process occurs in roughly three fourths of all hadronic events. There are several ways to distinguish signal $B\bar{B}$ from $q\bar{q}$ events, which take into account the event topology. Since both $B$ mesons are almost at rest in the $\Upsilon(4S)$ rest frame, there is no preferred direction and their decay products are isotropically distributed. By contrast, in the case of continuum events, quarks are produced back to back generating hadronic jets along a single axis. Fig. 4.4 illustrates the topological difference between $B\bar{B}$ and $q\bar{q}$ events. Several variables could be constructed to quantify the event topology - $\theta_T$, $S_\perp$ and Fox-Wolfram moments. We investigated
them in turn and tried to combine them together to achieve the best separation.

### 4.3.1 Thrust Angle $\theta_T$ and Sphericity $S_\perp$

For a given set of particles with momenta $\vec{p}_i$, the thrust is defined as the largest projected momentum along the axis referred as thrust axis:

$$T = \max_n \frac{\sum_i |\vec{p}_i \vec{n}|}{\sum_i |\vec{p}_i|}$$

(4.12)

where $\vec{n}$ is a unit vector characterizing the thrust axis. The thrust angle $\theta_T$ is defined as the angle between the thrust axis of the reconstructed $B_{\text{rec}}$ meson daughters and the thrust axis of the rest of the particles in the event. For the rest of the particles in the event we take the remaining charged tracks using a pion mass assumption and photons with energy greater than 50 MeV. In the case of a jet-like event topology
the distribution of $|\cos(\theta_T)|$ tends to peak near $|\cos(\theta_T)| = 1$, while for a spherical topology the distribution is flat as shown in Fig. 4.5 (a).

The sphericity $S_\perp$ is calculated as the scalar sum of the transverse momenta of all remaining particles outside of the 45° cone around the $B_{\text{rec}}$ thrust axis divided by the scalar sum of the momenta of all particles, including those outside of the cone

$$S_\perp = \frac{\sum_i |p_{T_i}|_{>45^\circ}}{\sum_i |p_i|} \quad (4.13)$$

Jet-like continuum events are likely to have a large percentage of particles within a 45° cone around the thrust axis of the reconstructed candidate and therefore tend to have lower values for $S_\perp$ than the signal $B\bar{B}$ events. Fig. 4.5 (b) illustrates the difference in $S_\perp$.  

Figure 4.5: Distributions of (a) $|\cos(\theta_T)|$ and (b) $S_\perp$ for MC signal and MC $q\bar{q}$ background.
4.3.2 Fox-Wolfram Moments

The original Fox-Wolfram moments are defined as

\[ H_l = \sum_{i,j} |\vec{p}_i||\vec{p}_j|P_l(\cos \theta_{ij}) \]  

(4.14)

where \( \vec{p}_i \) and \( \vec{p}_j \) are the momentum vectors of particles \( i \) and \( j \) respectively, \( \theta_{ij} \) is the angle between the momentum vectors of particles \( i \) and \( j \), \( P_l \) is the Legendre polynomial of the order \( l \) defined as

\[ P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l \]

Fox-Wolfram moments are usually normalized by dividing the normal Fox-Wolfram moments by the zeroth moment as

\[ R_l = H_l/H_0 \]

We divide all particles in the event into two categories: particles used in the reconstruction of the \( B_{\text{rec}} \) meson (signal group) and the remaining particles in the event (other group). Given these two groups we can calculate three sets of modified Fox-Wolfram parameters with different list of particles used in Eq. 4.14 as

\[ R_{nSS} = \frac{\sum_{i,j} |\vec{p}_i||\vec{p}_j|P_n(\cos \theta_{ij})}{\sum_{i,j} |\vec{p}_i||\vec{p}_j|} \]

\[ R_{nSO} = \frac{\sum_{i,k} |\vec{p}_i||\vec{p}_k|P_n(\cos \theta_{ik})}{\sum_{i,j} |\vec{p}_i||\vec{p}_j|} \]  

(4.15)

\[ R_{nOO} = \frac{\sum_{k,l} |\vec{p}_k||\vec{p}_l|P_n(\cos \theta_{kl})}{\sum_{k,l} |\vec{p}_k||\vec{p}_l|} \]

where \( i, j \) iterate through the signal particles and \( k, l \) iterate through the other particles. The momenta used in calculating the \( R_{nSS} \) moments are the same as those used in calculating the \( \Delta E \) and \( M_{bc} \) variables. We found that \( R_{nSS} \) is strongly correlated with \( \Delta E \) and \( M_{bc} \) and therefore is not used for the background separation. Fig. 4.6 shows the distribution of the \( R_{nSO} \) and \( R_{nOO} \) for \( n = 1 - 4 \) for signal \( B\bar{B} \) and background \( q\bar{q} \) Monte Carlo samples. \( R_{1SO} \), \( R_{3SO} \) and \( R_{1OO} \) were found to be somewhat correlated with \( M_{bc} \) and were also excluded from the separation analysis.
Figure 4.6: Modified Fox-Wolfram moments distributions.
The $R_2^{SO}$ moment is particularly useful since it provides the best separation power among all other moments. For a perfectly spherical event $R_2^{SO}$ should be equal to 0 and for an event completely collimated around the jet axis should be equal to 1. We combine the remaining modified moments into a linear Fisher discriminant\cite{47}. We also exclude $R_4^{OO}$ from consideration since it does not provide a noticeable mean separation, in fact, we found it even slightly dilutes the Fisher discriminant. The Fisher discriminant constructed from Fox-Wolfram moments is called Super Fox Wolfram (SFW) and defined as

$$SFW = \alpha_1 R_2^{SO} + \alpha_2 R_4^{SO} + \alpha_3 R_2^{OO} + \alpha_4 R_3^{OO}$$

(4.16)

where Fisher coefficients $\alpha_i$ are obtained by optimizing the separation between the signal and background training samples. The distributions of the SFW for MC signal and continuum background samples are shown in Fig. 4.7 (a). SFW is found to give better separation between signal and continuum background than using the original Fox-Wolfram moments. The Fisher discriminant combines the moments in an efficient way by taking into account the correlation between them. Fig. 4.7 (b) shows the power separation of the SFW parameter compared to $R_2^{SO}$. Each point on this plot represents the acceptance of the signal and background depending on the parameter threshold.

In some analyses the SFW is used in combination with $|\cos(\theta_T)|$ and the sphericity $S_\perp$ to construct the Fisher discriminant $F$

$$F = SFW + \alpha_5 |\cos(\theta_T)| + \alpha_6 S_\perp$$

The motivation for this is that using the additional variables in principle should improve the separation power. We found an extremely strong correlation between the $R_2^{SO}$ moment and both $|\cos(\theta_T)|$ and $S_\perp$ parameters, which is illustrated in Fig. 4.8 and Fig. 4.9. We compared the separation power in the same way as the comparison between the $R_2^{SO}$ and SFW and did not notice any improvement. Therefore, we
decided to discard the $|\cos(\theta_T)|$ and $S_\perp$ variables in our analysis and use only the SFW defined in Eq. 4.16 to characterize the event shape.

### 4.3.3 B Flight Direction and Helicity Angle of $\phi$ Meson

When a vector $J^P = 1^-$ $\Upsilon(4S)$ meson decays into two pseudo-scalar $0^-$ $B$ mesons, the $B\bar{B}$ pairs are produced as a P-wave state. This results in a differential production cross section that is proportional to $1 - \cos^2 \theta_B = \sin^2 \theta_B$, where $\theta_B$ is the angle of either $B$-meson momentum with respect to the beam axis in the $\Upsilon(4S)$ rest frame. The combinatorial background does not depend on $\theta_B$ and is practically flat. Therefore, $\cos \theta_B$ can be used to further suppress the background. The distribution of $|\cos \theta_B|$ is shown in Fig. 4.10 (a) for signal and continuum background MC samples.

Similarly, we can use the helicity angle $\theta_H$ of the $\phi$ meson, which is defined as the angle between the $B$ meson momentum and the momentum of either charged $K$ daughter in the $\phi$ meson rest frame. Since the vector $\phi$ meson decays into the kaons,
Figure 4.8: $R_2^{SO}$ versus $|\cos(\theta_T)|$ distributions for (a) signal $B\bar{B}$ and (b) jet-like $q\bar{q}$ events.

Figure 4.9: $R_2^{SO}$ versus $S_\perp$ distributions for (a) signal $B\bar{B}$ and (b) jet-like $q\bar{q}$ events.
which are pseudo-scalars, the angular distribution goes as $\cos^2 \theta_H$. The distribution of $|\cos \theta_H|$ is shown in Fig. 4.10 (b).

### 4.3.4 Likelihood Analysis

We have three uncorrelated variables: SFW, $|\cos \theta_B|$ and $|\cos \theta_H|$ which have different distributions for signal and background events. Instead of applying hard cuts on each one, we combine them to a single variable using a likelihood ratio technique. In this way, the level of suppression is controlled by only one parameter. In addition, the background suppression is done in a more efficient way than is the case for separate hard cuts.

Let us assume that we have an uncorrelated set of variables $\vec{v} = v^i$ with a given set of probability density functions (normalized distributions) $p^i(v^i)$. The N-dimensional probability density distribution is the product of all pdf’s. For a certain event, the
probability to have a certain set of values is proportional to

\[ P(\vec{v}) = \prod_{i} p^i(v^i) \quad (4.17) \]

Assuming that we have two categories of events—signal and background, we define likelihood functions in the similar way as the product of normalized distributions

\[ L_{\text{sig}} \equiv \prod_{i} p_{\text{sig}}^i(v^i) \]
\[ L_{\text{bkg}} \equiv \prod_{i} p_{\text{bkg}}^i(v^i) \]

there \( p_{\text{sig}}^i \) and \( p_{\text{bkg}}^i \) are the pdf’s for signal and background events respectively. The likelihood ratio defined as

\[ LR \equiv \frac{L_{\text{sig}}}{L_{\text{sig}} + L_{\text{bkg}}} = \frac{\prod_{i} p_{\text{sig}}^i(v^i)}{\prod_{i} p_{\text{sig}}^i(v^i) + \prod_{i} p_{\text{bkg}}^i(v^i)} \quad (4.18) \]

is then the probability that a certain event belongs to the signal category. In the general case, where there is some correlation between the variables, the likelihood ratio is not strictly the probability to be a signal event, but it behaves in the same way and can be used to classify events. The likelihood ratio has another advantage. If we have two different likelihood ratios \( LR_1 \) and \( LR_2 \) which are constructed from different sets of variables, a common likelihood ratio \( LR \) can be constructed using the following equation

\[ LR = \frac{LR_1 LR_2}{LR_1 LR_2 + (1 - LR_1)(1 - LR_2)} \quad (4.19) \]

giving only one variable to control the yield. This equation can be easily extended to three and more underlying likelihood ratios. Combining SFW, |\( \cos \theta_B \)| and |\( \cos \theta_H \)| variables gives us the following expression for signal and background likelihood functions

\[ L_{\text{sig}} = p_{\text{sig}}^{\text{SFW}} \times p_{\text{sig}}^{|\cos \theta_B|} \times p_{\text{sig}}^{|\cos \theta_H|} \]
\[ L_{\text{bkg}} = p_{\text{bkg}}^{\text{SFW}} \times p_{\text{bkg}}^{|\cos \theta_B|} \times p_{\text{bkg}}^{|\cos \theta_H|} \]
The combined likelihood ratio distribution for signal and continuum MC samples is shown in Fig. 4.11. It peaks around $LR = 1$ for signal and around $LR = 0$ for background. We accept events with a likelihood ratio above a certain threshold. We determine the likelihood ratio threshold from a figure of merit study, where we maximize the significance value

$$FOM = \frac{S}{\sqrt{S + B}}$$

where $S$ and $B$ are the numbers of signal and background events respectively after all selection criteria. That includes vertexing and flavor tagging selections used in the $CP$ fitting procedure described in Chapter 5. We use different $LR$ cuts depending on the flavor tagging quality of the associated $B$ meson characterized by the $r$ variable. The values of $r$ range from 0 for a completely ambiguous assignment to 1 when the assignment is certain to be correct. Events are divided into 6 intervals in the $r$ variable: 0.0-0.25, 0.25-0.5, 0.5-0.625, 0.625-0.75, 0.75-0.875, 0.875-1.0. There are different signal-to-background ratios in various $r$ bins. For bins with larger values of

![Figure 4.11: Likelihood ratio distribution made with SFW, $|\cos \theta_B|$ and $|\cos \theta_H|$ variables. The solid line shows the distribution of the signal $B^0 \to \phi K_S$ MC and the dotted line shows the continuum background MC sample.](image)
<table>
<thead>
<tr>
<th>$r$ bin</th>
<th>0.0-0.25</th>
<th>0.25-0.5</th>
<th>0.5-0.625</th>
<th>0.625-0.75</th>
<th>0.75-0.875</th>
<th>0.875-1.0</th>
<th>combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR cut</td>
<td>&gt; 0.48</td>
<td>&gt; 0.37</td>
<td>&gt; 0.37</td>
<td>&gt; 0.44</td>
<td>&gt; 0.36</td>
<td>&gt; 0.17</td>
<td>—</td>
</tr>
<tr>
<td>events</td>
<td>78</td>
<td>35</td>
<td>25</td>
<td>23</td>
<td>17</td>
<td>24</td>
<td>202</td>
</tr>
<tr>
<td>purity</td>
<td>0.78</td>
<td>0.77</td>
<td>0.79</td>
<td>0.86</td>
<td>0.86</td>
<td>0.92</td>
<td>0.81</td>
</tr>
<tr>
<td>$f_s$</td>
<td>0.08 ± 0.01</td>
<td>0.08 ± 0.02</td>
<td>0.09 ± 0.02</td>
<td>0.13 pm 0.03</td>
<td>0.13$^{+0.04}_{-0.03}$</td>
<td>0.23 ± 0.04</td>
<td>0.10 ± 0.01</td>
</tr>
</tbody>
</table>

Table 4.3: Event feeds and correspondent purities for optimized LR cuts and different $r$ intervals

We are more confident in our flavor assignment. Since the flavor tagging algorithm was trained on the products of decays of $B^0$ and $B^0$ mesons, there is a higher chance that events with large $r$ come from $\Upsilon(4S) \rightarrow B^0\bar{B}^0$ decay and not from continuum $q\bar{q}$.

We seek the largest significance $S/\sqrt{S + B}$ in each $r$ interval by varying the likelihood ratio cut. Fig. 4.12 shows the distributions of the significance depending on the LR cut for different $r$ bins. The combined purity is obtained from two-dimensional $M_{bc}$ - $\Delta E$ fits for each bin separately by integrating through the signal region. The functional forms of signal and background distributions are described in Section 5.9. The result of the fits and full summary are presented in Fig. 4.14, 4.15 and Table 4.3 respectively. The combined purity is calculated as

$$p = \frac{\sum_{i=1}^{6} S_i}{\sum_{i=1}^{6} (S_i + B_i)}$$

where $S_i$ and $B_i$ are the expected numbers of signal and background events in the correspondent bin. The $M_{bc}$ and $\Delta E$ distributions of the total yield are shown in Fig. 4.13.

In addition to providing a better yield for the same combined purity, usage of the $r$ dependent likelihood ratio cut has an another advantage in $CP$ fitting. It improves the event yield in intervals with higher $r$ intervals, which are more important in the calculation of $CP$ violation parameters.
Figure 4.12: Figure of merit distribution depending on the likelihood ratio cut for different $r$ intervals.

Figure 4.13: $\Delta E$ and $M_{bc}$ distributions of reconstructed $B^0 \rightarrow \phi K_S$ events. The histogram represents the data, the dashed line shows the distribution of the background component and the solid line shows the total distribution.
Figure 4.14: Yield extraction for \( r \) bins 1-3.
Figure 4.15: Yield extraction for $r$ bins 4-6.
Chapter 5

Measurement of the $CP$ Asymmetry

In this chapter we describe the analysis procedure for the measurement of the $CP$ violation parameters $S$ and $A$. We use one of the $B$ mesons reconstructed in the $CP$ eigenstate $\phi K_S$ (which we will call $B_{\text{rec}}$). We identify the flavor of the accompanying $B$ meson, $B_{\text{asc}}$, from its decay products. We determine the proper-time interval $\Delta t$ between the decay of the two $B$ mesons from the $z$ displacement of their decay vertices. Finally, we extract $S$ and the parameters $A$ from the distribution of the proper-time interval between $B_{\text{rec}}$ and $B_{\text{asc}}$ mesons decays using an unbinned maximum likelihood fit. We will use a modified theoretical $\Delta t$ distribution, which takes into account the finite resolution of the $B$ meson vertex measurements, possible wrong $B_{\text{asc}}$ flavor assignments, and the existence of the background in the data sample.

5.1 Unbinned Maximum Likelihood Fit

Suppose that we have an experiment that has collected $N$ independent events, consisting of a measurement of the quantity $x$, which may be multidimensional. Assume that we also know the probability density function (p.d.f) $P(x; \alpha)$ to observe $x$, where
\( \alpha \) is an unknown parameter and could be also multidimensional. \( P(x; \alpha) \) is the same for all \( x_i \) from the \( \{x_i\} \) set and is normalized to unity. We define the likelihood function as

\[
\mathcal{L}(\alpha) \equiv \prod_{i=1}^{N} P(x_i; \alpha)
\]  

(5.1)

The likelihood function can be considered a joint probability density function to observe this \( \{x_i\} \) set of measurements as a function of the parameter \( \alpha \). In the maximum likelihood method we obtain the estimator \( \hat{\alpha} \), which maximizes the \( \mathcal{L}(\alpha) \) likelihood function with respect to \( \alpha \). In practice, it is better to work with \( \ln \mathcal{L} \) instead of \( \mathcal{L} \), to avoid computational problems. Maximization of \( \ln \mathcal{L}(\alpha) \) is equivalent to the maximization of \( \mathcal{L}(\alpha) \). The usage of \( \ln \mathcal{L} \) has another advantage: in the limit of large numbers the likelihood becomes Gaussian, and the confidence interval for \( \alpha' \) can be determined from

\[
- \ln \mathcal{L}(\alpha') = - \ln \mathcal{L}(\hat{\alpha}) + 1/2
\]  

(5.2)

We use the MINUIT fitting package\[48\] in order to minimize \( - \ln \mathcal{L}(\alpha) \) and obtain the maximum likelihood estimator \( \hat{\alpha} \).

### 5.2 Probability Density Function

To use the maximum likelihood method in our measurement, we need to construct a probability density function for the proper-time interval \( \Delta t \). We know the theoretical PDF given by Eq. 2.41, but we can not use it directly in Eq. 5.1. The observed PDF will be smeared with experimental uncertainties. The combined PDF can be written as

\[
P = (1 - f_{\text{ol}})(f_{\text{sig}} P_{\text{sig}} + (1 - f_{\text{sig}}) P_{\text{bkg}}) + f_{\text{ol}} P_{\text{ol}}
\]  

(5.3)

where \( f_{\text{sig}} \) is the signal fraction, \( P_{\text{sig}} \) is the signal PDF smeared by wrong flavor tagging and the \( \Delta t \) resolution, and \( P_{\text{bkg}} \) is the PDF of the background. The small number of signal and background events that have a very large \( \Delta t \) are described by an outlier
component with outlier fraction of $f_{ol}$ and $P_{ol} \Delta t$ distribution. The observed signal probability function is described as the convolution of the theoretical PDF $P_{\text{sig}}$ with the resolution function $R_{\text{sig}}$:

$$P_{\text{sig}}(\Delta t) = \int_{-\infty}^{\infty} d(\Delta t') P_{\text{sig}}(\Delta t') R_{\text{sig}}(\Delta t - \Delta t')$$  

(5.4)

### 5.3 Flavor Tagging Algorithm

To measure the oscillation frequency $\Delta m_d$ and other $CP$ violation parameters we need to know the flavor of the associated $B$ meson at its decay time. The procedure to determine the flavor of the associated meson is called *flavor tagging*.

The flavor tagging algorithm is based on the analysis of the particles not belonging to the reconstructed $B_{\text{rec}}$ meson, under the assumption that we observe $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B_{\text{rec}}B_{\text{asc}}$ event. In order to get a good efficiency, we consider many possible decay channels of the $B$ mesons. For some flavor-specific decay channels the associated meson could be exclusively reconstructed, but in the majority of situations we have only inclusive information about the decay.

We determine the flavor of the $B_{\text{asc}}$ meson mainly based on the charge information of its final state particles. For example, the charge of the high-momentum leptons coming from semileptonic $B^0 \rightarrow Xl^+\nu$ or $\bar{B}^0 \rightarrow Xl^-\bar{\nu}$ decays provides a clean $B_{\text{asc}}$ flavor assignment. Similarly, the charge of the resulting kaons can be also used, since the majority of them come from $B^0 \rightarrow XK^+$ through the cascade transition $\bar{b} \rightarrow \bar{c} \rightarrow \bar{s}$. In addition, we can use other discriminant information like the charge of the medium-range momentum leptons coming from $c \rightarrow st^+\nu$ decays, high-momentum pions produced in $B^0 \rightarrow D^{(*)-}(\pi^+, \rho^+, a_1^+, \text{etc.})$ decays, slow pions coming from $D^{*-} \rightarrow D^0\pi^-$ decays and flavor of $\Lambda$ baryons from the cascade decays $b \rightarrow c \rightarrow s$.

Fig. 5.1 illustrates the principle of the flavor tagging algorithm. The process is essentially a look-up table for the combination of several likelihood quantities, which
was tuned on a large Monte Carlo sample. The algorithm returns two parameters to characterize the flavor of the $B_{\text{asc}}$ meson: $q$ and $r$. The value $q$ represents the assigned flavor of the $B_{\text{asc}}$ meson, $q = +1$ corresponds to $B^0$ and $q = -1$ corresponds to $\bar{B}^0$ meson. The parameter $r$ represents the reliability of the flavor assignment estimated from MC. It ranges from 0 for a completely ambiguous assignment to 1 when the assignment is certain to be correct. For each entry in the table the product of $q$ and $r$ is defined as

$$q \times r = \frac{N(B^0) - N(\bar{B}^0)}{N(B^0) + N(\bar{B}^0)}$$  \hspace{1cm} (5.5)$$

where $N(B^0)$ and $N(\bar{B}^0)$ are the number of $B^0$ and $\bar{B}^0$ decays respectively in MC with this assignment.

Since we use inclusive information for tagging, sometimes we incorrectly assign the flavor of the $B_{\text{asc}}$ meson. As a result of wrong tagging assignments, the $\Delta t$ time-dependent distribution is diluted. To avoid systematic bias in the measurement of the $CP$ violation parameters one must correctly estimate the wrong-tag probability $w$. For this reason we should estimate $w$ using experimental data. If our MC were perfect, the wrong tag probability would be related to $r$ as

$$r = 1 - 2w$$  \hspace{1cm} (5.6)$$

Figure 5.1: Schematic overview of the flavor tagging algorithm.
Table 5.1: Wrong tag fractions $w_l$ for different tagging $r$ bins. The errors on $w_l$ include both statistical and systematic errors.

\[ \begin{array}{|c|c|c|c|}
\hline
l & r \text{ range} & w_l \text{ SVD1} & w_l \text{ SVD2} \\
\hline
1 & 0.000 - 0.250 & 0.464 \pm 0.006 & 0.467 \pm 0.005 \\
2 & 0.250 - 0.500 & 0.331 \pm 0.008 & 0.324 \pm 0.007 \\
3 & 0.500 - 0.625 & 0.231 \pm 0.009 & 0.223 \pm 0.009 \\
4 & 0.625 - 0.750 & 0.163 \pm 0.008 & 0.160 \pm 0.009 \\
5 & 0.750 - 0.875 & 0.109 \pm 0.007 & 0.101 \pm 0.008 \\
6 & 0.875 - 1.000 & 0.020 \pm 0.005 & 0.020 \pm 0.005 \\
\hline
\end{array} \]

First, we distribute events into six bins in $r$. Then, we determine the wrong tag fraction $w$ for each $r$ bin directly from experimental data. We use a fully reconstructed data sample, where one of the $B$ mesons is reconstructed in the flavor-specific final state $D^{*\pm} l^\mp \nu$, $D^{(s)\pm} \pi^\mp$ and $D^{*\pm} \rho^\mp$. Then, we use our tagging algorithm to assign the flavor of the other $B$ meson. Due to the oscillations between $B^0$ and $\bar{B}^0$ mesons, the probabilities to observe the same flavor (SF) and the opposite flavor (OF) mesons for a given $\Delta t$ are given by

\[ P_{SF}(\Delta t) \sim 1 - (1 - 2w) \cos(\Delta m_d \Delta t) \tag{5.7} \]
\[ P_{OF}(\Delta t) \sim 1 + (1 - 2w) \cos(\Delta m_d \Delta t) \tag{5.8} \]

This results in the following expression for the time-dependent mixing asymmetry

\[ A(\Delta t) = \frac{P_{OF} - P_{SF}}{P_{OF} + P_{SF}} = (1 - 2w) \cos(\Delta m_d \Delta t) \tag{5.9} \]

By measuring the amplitude of this asymmetry we can estimate the wrong tag fraction for each $r$ bin. The result of the fits for all six $r$ bins is shown in Fig. 5.2.

The resulting wrong tag fractions $w_l$ are listed in Table 5.1. We have different sets of $w_l$ for SVD1 and SVD2 data samples.

Using $w$ obtained this way, the $CP$ analysis is not systematically biased due to the difference between MC simulation and real data. We use the errors on the wrong tag fractions to estimate the systematic uncertainties on measurement of the $CP$ violation parameters $A$ and $S$. 

74
Figure 5.2: Measured time-dependent $A(\Delta t)$ asymmetries between the SF and OF events for six tagging $r$ bins. Obtained using a sample where the flavor of one of the $B$ mesons is known and the flavor of the other is determined by flavor tagging algorithm[49].
5.4 Proper-Time Difference Reconstruction

Since the momenta of the $B$ mesons in the $\Upsilon(4S)$ rest frame are very small, we can use the difference of the decay vertex positions in the beam direction to extract the proper-time interval $\Delta t$. Disregarding the center-of-mass frame momenta of the $B$ mesons, we can approximate the interval between fully reconstructed and associated mesons decays $\Delta t \equiv t_{\text{rec}} - t_{\text{asc}}$ as

$$\Delta t \approx \frac{\Delta z}{c(\beta\gamma)_\Upsilon} = \frac{\Delta z}{c} \frac{m_\Upsilon}{p_\Upsilon} \quad (5.10)$$

where $\Delta z = z_{\text{rec}} - z_{\text{asc}}$ is the vertex displacement in the beam direction.

Fig. 5.3 illustrates the procedure for the reconstruction of decay vertices. We use the $K^+$ and $K^-$ charged tracks to reconstruct the vertex position of the $B_{\text{rec}}$ meson. We require tracks to have at least one associated SVD $r$-$\phi$ hit and at least two $z$ hits. We impose an IP profile[50] constraint to improve the vertex resolution and vertex reconstruction efficiency. The IP constraint also allows us to reconstruct a vertex position for an event with only one good charged track. The IP profile is run-dependent and calculated using hadronic events for each run. The typical size of the IP profile is 100 $\mu$m in $x$, 5 $\mu$m in $y$ and 3 mm in $z$. To take into account the transverse $B^0$ flight length in the $r$-$\phi$ plane, the IP profile is smeared out by a Gaussian with $\sigma = 21$ $\mu$m when it is used in the constraint vertex fit.

We use all remaining charged tracks with the same SVD hit requirement that was
used for the $B_{\text{rec}}$ side in the reconstruction of the vertex position of the associated $B_{\text{asc}}$ meson. We eliminate track pairs that are consistent with being a $K_S$ meson (within $\pm 15$ MeV/$c^2$) since they do not originate from the primary $B_{\text{asc}}$ vertex. We exclude poorly reconstructed tracks which have a position-measurement accuracy in the beam direction that is worse than 500 $\mu$m. We also exclude tracks that are more than 500 $\mu$m away from the fully reconstructed $B_{\text{rec}}$ vertex in the $r-\phi$ plane.

Next, we do an iterative IP constraint fit to determine the vertex position of the associated meson. We repeat the procedure by removing the track that gives the largest contribution to the reduced $\chi^2$ until the resulting $\chi^2$/n.d.f. is less than 20 or only one track is left. However, we do not remove the charged tracks associated with leptons having momentum greater than 1.1 GeV/$c$ in the center of mass frame, because they most likely come from primary semileptonic $B$ decays. In this case, we remove a track with the second largest contribution to $\chi^2$/n.d.f. This procedure cannot completely avoid including tracks coming from the secondary decay vertices, causing a possible shift of the reconstructed vertex in the flight direction of the parent particle.

We reject a small fraction $\approx 0.2\%$ of the events by requiring that $|\Delta t| < 70$ ps, which corresponds to $\approx 45\tau_B$. This affects the PDF normalization. In the likelihood we use the normalized versions of all PDFs defined as:

$$\tilde{P}(\Delta t) = \frac{P(\Delta t)}{\int_{-70}^{70} P(\Delta t)d(\Delta t)}$$  \hspace{1cm} (5.11)

### 5.5 Resolution Functions

We use an event-by-event resolution function that depends on the errors and goodness of fit of the reconstructed vertices. The resolution function is constructed as a convolution of four different components

$$R_{\text{sig}} = R_{\text{rec}} \otimes R_{\text{asc}} \otimes R_{\text{np}} \otimes R_k$$  \hspace{1cm} (5.12)
where $R_{\text{rec}}$ and $R_{\text{asc}}$ are the detector resolution functions for the determination of the $z_{\text{rec}}$ and $z_{\text{asc}}$, $R_{\text{np}}$ additional non-symmetrical smearing of $z_{\text{asc}}$ due to the secondary tracks not coming not from the associated vertex of the $B_{\text{asc}}$ meson, $R_k$ is the resolution function corresponding to the kinematic approximation of resulting from the non-zero momentum of the $B$ mesons in the $\Upsilon(4S)$ frame. The overall resolution $R_{\text{sig}}(\Delta t)$ is then expressed as

$$
R_{\text{sig}}(\Delta t) = \int \int \int_{-\infty}^{\infty} d(\Delta t')d(\Delta t'')d(\Delta t''') R_{\text{rec}}(\Delta t - \Delta t') \times
$$

$$
R_{\text{asc}}(\Delta t' - \Delta t'') R_{\text{np}}(\Delta t'' - \Delta t''') R_k(\Delta t''')
$$

(5.13)

We use a MC simulation to understand the resolution functions and to determine their forms. Almost all final parameters of the resolution functions are then fitted using experimental data.

### 5.5.1 Detector Resolution Functions

In order to separate the detector-resolution smearing from the smearing due to non-primary tracks, we use a Monte Carlo sample in which all secondary particles from $B_{\text{asc}}$ are generated with zero lifetime at the $B_{\text{asc}}$ decay vertex. The effects from non-primary smearing will be discussed in the next section. Fig. 5.4 shows the residual distributions of $z_{\text{rec}}$ and $z_{\text{asc}}$ defined as

$$
\delta z_q = z_{\text{rec}}^q - z_{\text{gen}}^q
$$

(5.14)

where rec superscript represents the measurement obtained by the vertex reconstruction and gen represents the position generated by MC, $q$ denotes fully reconstructed (rec) or associated (asc) $B$ mesons. Solid lines represent the fits to the sum of two Gaussians. The fitted curves only model residual distributions in the central region and fail completely for the tails. We also found that the sum of three or more Gaussians does not dramatically improve the picture.
Figure 5.4: Distribution of $\delta z_q$ for (a) fully reconstructed and (b) associated $B$ mesons. Superimposed are the fits to the sum of two Gaussians.

We therefore considered more elaborate resolution functions that take into account the event-by-event $z$-coordinate error $\sigma^z$ of the reconstructed vertex. The value of $\sigma^z$ is calculated from the error matrix of the tracks used in the vertex fit and the size of the IP profile. To construct the functional forms for $R_{rec}$ and $R_{asc}$ we studied the $\delta z_q/\sigma_q^z$ pull distributions. In particular, if the $\sigma_q^z$ estimation were correct, the distribution of the ratio of $\delta z_q/\sigma_q^z$ would be characterized by a single Gaussian with a variance equal to one.

Because of the IP constraint, we can reconstruct a decay vertex position even with a single track. We considered the single-track and multiple-track cases separately and came up with different forms of resolution functions.

### 5.5.2 Quality of Vertex Fit

From a MC study it was found that the residual distributions of $z_{rec}$ and $z_{asc}$ are highly depend on the $\chi^2$ of the vertex reconstruction. However, the normal $\chi^2$ of a vertex fit is found to be highly correlated with the $z$ position of the vertex due to a very tight constraint in the transverse IP plane. To avoid bias in the lifetime measurement we introduce the reduced assessment of the quality projected on the $z$ direction. We define a new parameter $\xi$ to characterize the fit quality as:
Figure 5.5: The average of (a) nominal $\chi^2$/n.d.f. and (b) newly defined goodness $\xi$ depending on the flight length. $\xi$ shows no dependence on the distance.

\[
\xi = \frac{1}{2n} \sum_{i=1}^{n} \left( \frac{z^i_{\text{after}} - z^i_{\text{before}}}{\varepsilon^i_{\text{before}}} \right)^2
\]

(5.15)

where $n$ is the number of tracks used in the fit, $z^i_{\text{before}}$ and $z^i_{\text{after}}$ are the $z$ positions of the $i$-th track at the closest approach to the origin before and after the fit, and $\varepsilon^i_{\text{before}}$ is the error on $z^i_{\text{before}}$. The Monte Carlo simulation in Fig. 5.5 shows that $\xi$ does not depend on the $B$ flight length. The goodness of vertex reconstruction parameter, $\xi$, is defined only when two or more charged tracks are used in the fit. If only one charged track is used, then the vertex position is just the crossing point of the track and IP profile.

To eliminate badly reconstructed vertices, we accept tracks with $\xi < 100$ for both reconstructed and associated vertex positions $z_{\text{rec}}$ and $z_{\text{asc}}$.

### 5.5.3 Multiple-Track Vertex Reconstruction

To study the detector resolution functions for multiple-track vertices, we studied the $\delta z_q/\sigma^z_q$ pull distributions with similar $\xi$ values defined in Eq.5.15. We found that for each goodness $\xi$, $\delta z_q/\sigma^z_q$ could be represented as a single Gaussian distribution. In addition, the standard deviation of each distribution has a linear dependence on $\xi$. Fig. 5.6 shows the pull distributions divided by the associated $\xi$ with superimposed
fits to a single Gaussian. The last histogram with $\xi > 6.0$ seems not to be fitted well to a single Gaussian. The reason is that the last histogram contains vertices with a very large $\xi$, which produces a long tail. Fig. 5.7 shows the standard deviation of the fitted Gaussian as a function of $\xi$.

Based on that finding, we introduce the detector resolution functions for multiple tracks $R_{\text{rec}}^{\text{mult}}$ and $R_{\text{asc}}^{\text{mult}}$ as

$$R_{\text{rec}}^{\text{mult}}(\delta z_{\text{rec}}) = G(\delta z_{\text{rec}}; s_{\text{rec}}\sigma_{\text{rec}})$$
$$R_{\text{asc}}^{\text{mult}}(\delta z_{\text{asc}}) = G(\delta z_{\text{asc}}; s_{\text{asc}}\sigma_{\text{asc}})$$

where $s_{\text{rec}}$ and $s_{\text{asc}}$ are first-order polynomials in vertex quality $\xi$

$$s_{\text{rec}} = s_{0\text{rec}} + s_{1\text{rec}}\xi_{\text{rec}}$$
$$s_{\text{asc}} = s_{0\text{asc}} + s_{1\text{asc}}\xi_{\text{asc}}$$

The scale factors $s_{0\text{q}}$ and $s_{1\text{q}}$ are determined from the lifetime fit to the experimental data. Fig. 5.8 shows the residual distributions of $\delta z_{\text{rec}}$ and $\delta z_{\text{asc}}$ with event-by-event resolution functions $R_{\text{rec}}^{\text{mult}}$ and $R_{\text{rec}}^{\text{mult}}$ superimposed.

In summary, we use event-by-event resolution functions with dependence on the vertex quality reconstruction $\sigma_z$, and, for the case of multiple tracks, the goodness of fit $\xi$. The resolution functions $R_{\text{rec}}^{\text{mult}}(\delta z_{\text{rec}})$ and $R_{\text{rec}}^{\text{mult}}(\delta z_{\text{asc}})$ can easily be converted into the $\Delta t$ resolution by replacing $\sigma_{\text{q}}$ with $\sigma_{\text{q}} = \sigma_{\text{q}}^{\text{z}}/(c(\beta\gamma)\tau)$.

### 5.5.4 Single-Track Vertex Reconstruction

The $\delta z_{\text{q}}/\sigma_{\text{q}}^{\text{z}}$ distribution for single-track vertices is well represented by a sum of two Gaussians. So we express the resolution $R_{\text{q}}^{\text{sgl}}(\delta z_{\text{q}})$ function as the main and tail Gaussian components

$$R_{\text{q}}^{\text{sgl}}(\delta z_{\text{q}}) = (1 - f_{\text{tail}})G(\delta z_{\text{q}}; s_{\text{main}}\sigma_{\text{q}}^{\text{z}}) + f_{\text{tail}}G(\delta z_{\text{q}}; s_{\text{tail}}\sigma_{\text{q}}^{\text{z}})$$

81
Figure 5.6: The pull distributions of $\delta z_q/\sigma^z_q$ divided by the goodness of vertex fit $\xi$. 

(a) Reconstructed $B_{\text{rec}}$ side

(b) Associated $B_{\text{ass}}$ side
Figure 5.7: Standard deviations of the pull distributions as a function of $\xi$ for (a) fully reconstructed and (b) associated $B$ meson vertices.

Figure 5.8: Distributions of (a) $\delta z_{\text{rec}}$ and (b) $\delta z_{\text{asc}}$ for multiple-track vertices with superimposed event-by-event resolution functions.
Figure 5.9: Residual distributions of (a) $\delta z_{\text{rec}}$ and (b) $\delta z_{\text{asc}}$ for single-track vertices with $R_{\text{rec}}^\text{sgl}$ and $R_{\text{asc}}^\text{sgl}$ resolution functions imposed.

where $s_{\text{main}}$ and $s_{\text{tail}}$ are the global scale coefficients common to both reconstructed and associated $B$ vertices with a single track, and $f_{\text{tail}}$ is the fractional tail component. The parameters $s_{\text{main}}$, $s_{\text{tail}}$ and $f_{\text{tail}}$ are determined from a lifetime fit to the experimental data. $G(x, \sigma)$ is a Gaussian with zero mean defined as

$$
G(x, \sigma) \equiv \frac{1}{\sqrt{2\pi\sigma}} \exp \left( -\frac{x^2}{2\sigma^2} \right)
$$

(5.21)

Fig. 5.9 illustrates the residual distributions of $\delta z_q$ for single-track vertices and the corresponding resolution functions $R_{\text{rec}}^\text{sgl}(\delta z_{\text{rec}})$ and $R_{\text{asc}}^\text{sgl}(\delta z_{\text{asc}})$.

### 5.5.5 Smearing due to Non-Primary Tracks

Since we use all charged tracks in the reconstruction of the $B_{\text{asc}}$ decay vertex, there is a bias due to charged tracks not coming from the primary vertex. These charged tracks are mainly produced in the decays of relatively long-lived charmed mesons. The vertex reconstruction algorithm rejects some of the non-primary tracks and the remaining tracks are taken into account by the resolution function.

To study the effects of non-primary-track smearing, we used two Monte Carlo samples. The first is a standard MC sample. The second sample has the non-primary-track smearing removed by generating secondary particles from the $B_{\text{asc}}$ meson with zero lifetime at the $B_{\text{asc}}$ decay vertex. We studied the difference between the standard
reconstructed vertex position $z_{\text{asc}}$ and the reconstructed vertex with the effect of non-primary tracks removed $z_{\text{asc}}^{\text{noNP}}$. The distribution of $\delta z_{\text{asc}}^{\text{NP}} \equiv z_{\text{asc}} - z_{\text{asc}}^{\text{noNP}}$ is shown in Fig. 5.10. The histogram can be considered as a representation of the resolution function $R_{\text{np}}$.

Since non-primary tracks mainly come from the finite lifetime of charm mesons, it is natural to introduce $R_{\text{np}}$ as the sum of the prompt component with no non-primary
track smearing and two lifetime components as

\[
R_{np}(\delta z_{asc}) = f_0 \delta(\delta z_{asc}) + (1 - f_0) \left[ f_p E_p(\delta z_{asc}; c(\beta \gamma) \tau_{np}^p) + \right. \\
\left. + (1 - f_p) E_n(\delta z_{asc}; c(\beta \gamma) \tau_{np}^n) \right]
\]  \hspace{1cm} (5.22)

where \( \delta(\delta z_{asc}) \) is the Dirac \( \delta \)-function corresponding to the prompt component, \( f_0 \) is the fraction of the prompt component, \( f_p \) is the fraction of the \( E_p \) component, \( \tau_{np}^p \) and \( \tau_{np}^n \) are parameters of the lifetime components. \( E_p \) and \( E_n \) are defined as follows

\[
E_p(x; \tau) \equiv \frac{1}{\tau} \exp \left( -\frac{x}{\tau} \right) \quad \text{for } x > 0, \text{ otherwise } 0
\]  \hspace{1cm} (5.23)

\[
E_n(x; \tau) \equiv \frac{1}{\tau} \exp \left( +\frac{x}{\tau} \right) \quad \text{for } x \leq 0, \text{ otherwise } 0
\]  \hspace{1cm} (5.24)

As in the case of the detector resolution functions, we studied the vertex position shift \( \delta z_{asc}^{NP} \) associated with non-primary tracks as a function of \( \sigma_{asc}^z \) and \( \xi_{asc} \). Fig. 5.11 shows that there is also a linear dependence of \( \delta z_{asc}^{NP} \) on vertex error and goodness of fit. Consequently, we define \( \tau_{np}^p \) and \( \tau_{np}^n \) as

\[
\tau_{np}^p = s_{asc}^3 |\tau_0^p + \tau_1^p (1.0 - s_{asc}^2 \xi_{asc}) \sigma_{asc}| \]  \hspace{1cm} (5.25)

\[
\tau_{np}^n = s_{asc}^3 |\tau_0^n + \tau_1^n (1.0 + s_{asc}^2 \xi_{asc}) \sigma_{asc}| \]  \hspace{1cm} (5.26)

where \( \tau_0^p, \tau_1^p, \tau_0^n \) and \( \tau_1^n \) are lifetime parameters, \( s_{asc}^2 \) and \( s_{asc}^3 \) are scaling factors. All of these parameters are obtained from the fit of \( \delta z_{asc} \) to the \( R_{asc} \otimes R_{np} \). First, we fit \( \delta z_{asc}^{NP} \)
Figure 5.13: Non-primary track shift $z_{\text{asc}} - z_{\text{asc}}^{\text{noNP}}$ versus $\sigma_{\text{asc}}$ for single-track vertices. The events where $z_{\text{asc}} < z_{\text{asc}}^{\text{noNP}}$ are excluded.

Figure 5.14: Distribution of $\delta z_{\text{asc}}$ for single-track vertices of (a) neutral and (b) charged $B$ mesons with superimposed $R_{\text{asc}} \otimes R_{\text{np}}$.

to the event-dependent $R_{\text{asc}}$ detector resolution to obtain the $s^0_{\text{asc}}$ and $s^0_{\text{asc}}$ parameters. Then, we fix the parameters of $R_{\text{asc}}$ and fit $\delta z_{\text{asc}}$ to the convolved resolution function $R_{\text{asc}} \otimes R_{\text{np}}$ to determine the parameters of $R_{\text{np}}$. The result of the fit and the residual distribution of $z_{\text{asc}}$ are shown in Fig. 5.12 for neutral $B^0$ and charged $B^-$ mesons. We use different $R_{\text{np}}$ parameters for neutral and charged $B$ mesons because the yields of each charmed meson $D^0$, $D^+$ and $D_s^+$ are different. This produces the difference in the effective resolution shape $R_{\text{np}}$.

For single-track vertices the goodness of fit $\xi_{\text{asc}}$ is not defined. We checked the correlation between the non-primary track displacement $\delta z_{\text{asc}}^{\text{NP}}$ and $\sigma_{\text{asc}}^z$ and also found
Table 5.2: The list of fitted parameters of $R_{np}$ determined from SVD1 Monte Carlo.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$B^0/B^0$</th>
<th>$B^+/B^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>single</td>
<td>multiple</td>
</tr>
<tr>
<td>$f_\delta$</td>
<td>0.692±0.035</td>
<td>0.557±0.040</td>
</tr>
<tr>
<td>$f_p$</td>
<td>0.791±0.009</td>
<td>0.940±0.002</td>
</tr>
<tr>
<td>$\tau_p^0$</td>
<td>0.458±0.047</td>
<td>0.050±0.006</td>
</tr>
<tr>
<td>$\tau_p^1$</td>
<td>1.326±0.064</td>
<td>0.704±0.014</td>
</tr>
<tr>
<td>$\tau_p^0$</td>
<td>0.293±0.047</td>
<td>0.444±0.013</td>
</tr>
<tr>
<td>$\tau_n$</td>
<td>1.498±0.124</td>
<td>1.287±0.068</td>
</tr>
</tbody>
</table>

Table 5.3: The list of fitted parameters of $R_{np}$ determined from SVD2 Monte Carlo.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$B^0/B^0$</th>
<th>$B^+/B^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>single</td>
<td>multiple</td>
</tr>
<tr>
<td>$f_\delta$</td>
<td>0.777±0.040</td>
<td>0.393±0.049</td>
</tr>
<tr>
<td>$f_p$</td>
<td>0.774±0.021</td>
<td>0.774±0.024</td>
</tr>
<tr>
<td>$\tau_p^0$</td>
<td>1.319±0.040</td>
<td>0.132±0.005</td>
</tr>
<tr>
<td>$\tau_p^1$</td>
<td>0.000±0.000</td>
<td>0.680±0.015</td>
</tr>
<tr>
<td>$\tau_p^0$</td>
<td>0.638±0.056</td>
<td>0.151±0.008</td>
</tr>
<tr>
<td>$\tau_n$</td>
<td>0.000±0.000</td>
<td>0.673±0.019</td>
</tr>
</tbody>
</table>

The linear dependence as illustrated in Fig. 5.13. We define $\tau_{np}^p$ and $\tau_{np}^n$ for the case of single-track vertex as

$$\tau_{np}^p = s_{asc}^3 |\tau_p^0 + \tau_p^1 \sigma_{asc}|$$  \hspace{1cm} (5.27)
$$\tau_{np}^n = s_{asc}^3 |\tau_n^0 - \tau_n^1 \sigma_{asc}|$$  \hspace{1cm} (5.28)

The residual distribution of $z_{asc}$ and the fitted distribution of $R_{np} \otimes R_{asc}$ for single-track vertices are shown for $B^0$ and $B^-$ mesons in Fig. 5.14. The fitted parameters of $R_{np}$ resolution function for SVD1 and SVD2 samples are listed in Table 5.2 and Table 5.3 respectively.
5.6 Kinematic Approximation

In the measurement of the proper-time difference $\Delta t$ we use a kinematic approximation, that does not take into account the non-zero momentum of $B$ mesons in the $\Upsilon(4S)$ frame

$$\Delta t \approx \frac{z_{\text{rec}} - z_{\text{asc}}}{c(\beta \gamma)_{\Upsilon}}$$

The resolution function $R_k$ associated with this approximation can be calculated analytically as a function of $E_{B_{\text{rec}}}^{\text{cms}}$ and $\cos \theta_{B_{\text{rec}}}^{\text{cms}}$ from the kinematics of the two-body decay, where $E_{B_{\text{rec}}}^{\text{cms}}$ and $\cos \theta_{B_{\text{rec}}}^{\text{cms}}$ are the energy and polar angle of the fully reconstructed $B_{\text{rec}}$ meson in the $\Upsilon(4S)$ rest frame. The difference between the approximation given above and the true proper-time interval $\Delta t_{\text{true}} = t_{\text{rec}} - t_{\text{asc}}$ can be expressed as

$$x \equiv \Delta t - \Delta t_{\text{true}} = \frac{z_{\text{rec}} - z_{\text{asc}}}{c(\beta \gamma)_{\Upsilon}} - (t_{\text{rec}} - t_{\text{asc}})$$

$$= \frac{c(\beta \gamma)_{\text{rec}} t_{\text{rec}} - c(\beta \gamma)_{\text{asc}} t_{\text{asc}}}{c(\beta \gamma)_{\Upsilon}} - (t_{\text{rec}} - t_{\text{asc}})$$

$$= \left[ \frac{(\beta \gamma)_{\text{rec}}}{(\beta \gamma)_{\Upsilon}} - 1 \right] t_{\text{rec}} - \left[ \frac{(\beta \gamma)_{\text{asc}}}{(\beta \gamma)_{\Upsilon}} - 1 \right] t_{\text{asc}}$$

(5.29)

where $(\beta \gamma)_{\text{rec}}$ and $(\beta \gamma)_{\text{asc}}$ are Lorentz boost factors of $B_{\text{rec}}$ and $B_{\text{asc}}$ respectively given by

$$(\beta \gamma)_{\text{rec}} = (\beta \gamma)_{\Upsilon} \left( \frac{E_{B_{\text{rec}}}^{\text{cms}}}{m_{B_{\text{rec}}} c^2} + \frac{p_{B_{\text{rec}}}^{\text{cms}} \cos \theta_{B_{\text{rec}}}^{\text{cms}}}{m_{B_{\text{rec}}} c \beta_{\Upsilon}} \right)$$

(5.30)

$$(\beta \gamma)_{\text{asc}} = (\beta \gamma)_{\Upsilon} \left( \frac{E_{B_{\text{asc}}}^{\text{cms}}}{m_{B_{\text{asc}}} c^2} - \frac{p_{B_{\text{asc}}}^{\text{cms}} \cos \theta_{B_{\text{asc}}}^{\text{cms}}}{m_{B_{\text{asc}}} c \beta_{\Upsilon}} \right)$$

(5.31)

Defining $a_k \equiv E_{B_{\text{rec}}}^{\text{cms}}/(m_{B_{\text{rec}}} c^2) \approx 1$ and $c_k \equiv p_{B_{\text{rec}}}^{\text{cms}} \cos \theta_{B_{\text{rec}}}^{\text{cms}}/(m_{B_{\text{rec}}} c \beta_{\Upsilon})$, $x$ can be written in terms of $a_k$ and $c_k$ as

$$x = (a_k + c_k - 1) t_{\text{rec}} - (a_k - c_k - 1) t_{\text{asc}}$$

(5.32)

The probability density function to observe $t_{\text{rec}}$ and $t_{\text{asc}}$ is given by

$$P_{\text{true}}(t_{\text{rec}}, t_{\text{asc}}) = P(t_{\text{rec}}) P(t_{\text{asc}}) = \frac{1}{t_B^2} \exp \left( -\frac{t_{\text{rec}} + t_{\text{asc}}}{\tau_B} \right)$$

(5.33)
Figure 5.15: Distribution of $x = \Delta t - \Delta t_{\text{true}}$ with $R_k(x)$ superimposed.

Then the probability of a proper-time interval $\Delta t_{\text{true}}$ is

$$F(\Delta t_{\text{true}}) = \int_0^{+\infty} dt_{\text{rec}} \int_0^{+\infty} dt_{\text{asc}} P_{\text{true}}(t_{\text{rec}}, t_{\text{asc}}) \delta(\Delta t_{\text{true}} - (t_{\text{rec}} - t_{\text{asc}}))$$

(5.34)

The simultaneous probability of getting $x$ and $\Delta t_{\text{true}}$ is

$$F(x, \Delta t_{\text{true}}) = \int_0^{+\infty} dt_{\text{rec}} \int_0^{+\infty} dt_{\text{asc}} P_{\text{true}}(t_{\text{rec}}, t_{\text{asc}}) \delta(\Delta t_{\text{true}} - (t_{\text{rec}} - t_{\text{asc}})) \times$$

$$\times \delta(x - [(a_k + c_k - 1)t_{\text{rec}} - (a_k - c_k - 1)t_{\text{asc}}])$$

(5.35)

Next, $R_k(x)$ can be analytically calculated as $R_k(x) = F(x, \Delta t_{\text{true}})/F(\Delta t_{\text{true}})$, which gives the following expression

$$R_k(x) = \begin{cases} 
E_p(x - [(a_k - 1)\Delta t_{\text{true}} + c_k|\Delta t_{\text{true}}|]; |c_k|\tau_B) & (c_k > 0) \\
\delta(x - (a_k - 1)\Delta t_{\text{true}}) & (c_k = 0) \\
E_n(x - [(a_k - 1)\Delta t_{\text{true}} + c_k|\Delta t_{\text{true}}|]; |c_k|\tau_B) & (c_k < 0)
\end{cases}$$

(5.36)

Fig. 5.15 shows the distribution of the difference between the approximate and true values of the proper-time interval $\Delta t - \Delta t_{\text{true}}$ with $R_k(x)$ superimposed. In practice, we deal with the convolution of the $R_k(x)$ resolution function and the theoretical distribution $P_{\text{sig}}(\Delta t_{\text{true}})$. For example, in the decay of two $B$ mesons observed the
$P(\Delta t)$ distribution is given by

$$P(\Delta t) = P_{\text{sig}} \otimes R_k = \begin{cases} \frac{1}{2a_k \tau_B} \exp \left( -\frac{|\Delta t|}{(a_k + c_k)\tau_B} \right) & (\Delta t \geq 0) \\ \frac{1}{2a_k \tau_B} \exp \left( -\frac{|\Delta t|}{(a_k - c_k)\tau_B} \right) & (\Delta t < 0) \end{cases} \quad (5.37)$$

More generally, when the probability to observe $t_{\text{rec}}$ and $t_{\text{asc}}$ is given by

$$P_{\text{true}}(t_{\text{rec}}, t_{\text{asc}}) = \frac{1}{t_B^2} \exp \left( -\frac{t_{\text{rec}} + t_{\text{asc}}}{\tau_B} \right) g(t_{\text{rec}} - t_{\text{asc}}) \quad (5.38)$$

we get the same expression for $R_k(x)$ as given by Eq. 5.36. Therefore, the $R_k$ resolution function thus obtained can be applied to $B^0 - \bar{B}^0$ mixing and $CP$ asymmetry distributions.

### 5.7 Outliers

Even after adding resolution smearing to the theoretical $\Delta t$ distribution, there still remains a very long tail that cannot be described by the resolution functions discussed above. We introduce an additional outlier term to the observed probability distribution function $P(\Delta t)$. An outlier term is represented by a single Gaussian with zero mean and a very large standard deviation

$$P_{\text{ol}}(\Delta t) = G(\Delta t, \sigma_{\text{ol}}) \quad (5.39)$$

Since this Gaussian outlier has a width of about 40 ps, we neglect the lifetime convolution and other resolution components. This long tail is thought to be caused by the mis-reconstruction of the tracks and attributed to both signal and background events. Therefore we introduce the outlier term as a third component of the $P(\Delta t)$ in the Eq. 5.3. We use different outlier fractions $f_{\text{ol}}$ for events where both $z_{\text{rec}}$ and $z_{\text{asc}}$ are reconstructed using two or more tracks, and events where at least one of the vertices is reconstructed using a single track.
5.8 Background Shape

For the $B^0 \rightarrow \phi K_S$ decay mode, the background is dominated by the $q\bar{q}$ continuum process and the contribution from feed across modes is negligible. We model the background PDF $P_{bkg}(\Delta t)$ distribution in a way similar to the signal. The background distribution is the convolution of the physical distribution $P_{bkg}$ and the background resolution function $R_{bkg}$

$$P_{bkg}(\Delta t) = \int_{-\infty}^{\infty} d(\Delta t') P_{bkg}(\Delta t') R_{bkg}(\Delta t - \Delta t')$$  \hspace{1cm} (5.40)

The physical distribution consists of a prompt component with non-zero mean and a shifted lifetime component

$$P_{bkg}(\Delta t) = f_{bkg}^\delta \delta(\Delta t - \mu_\delta) + (1 - f_{bkg}^\delta) \frac{1}{2\sigma_{bkg}} \exp \left(-\frac{\Delta t - \mu_\tau}{\tau_{bkg}}\right)$$  \hspace{1cm} (5.41)

where $f_\delta$ is the fraction of prompt component, $\mu_\delta$ and $\mu_\tau$ are the offsets of the prompt and lifetime components, respectively. We use a double-Gaussian for the background resolution function

$$R_{bkg}(\Delta t) = (1 - f_{tail}^{bkg}) G(\Delta t; s^{bkg}\sigma_{vtx}) + f_{tail}^{bkg} G(\Delta t; s_{tail}^{bkg}\sigma_{vtx})$$  \hspace{1cm} (5.42)

where $f_{tail}$ is the fraction of the Gaussian with a larger standard deviation, $\sigma_{vtx}$ is a common vertex error defined as $\sigma_{vtx} = \sqrt{\sigma_{rec}^2 + \sigma_{acc}^2}$. We used different values for $s^{bkg}_{main}, s^{bkg}_{tail}, f^{bkg}_{tail}, f^{bkg}_{\delta}$ depending on whether both vertices are reconstructed with two or more tracks. The parameters of the background distribution $P_{bkg}$ are determined from the unbinned maximum likelihood fit of the $\Delta t$ distribution of the events in the sideband region of real data. For this fit we used a modified background distribution with added outlier component

$$P_{bkg}^*(\Delta t) = (1 - f_{ol}^*) P_{bkg}(\Delta t) + f_{ol}^* P_{ol}(\Delta t)$$  \hspace{1cm} (5.43)

We use different sets of the parameters for SVD1 and SVD2 data. The result of the separate fits are listed in Table 5.4. Fig. 5.16 shows the background distribution $\Delta t$ with the fitted curves superimposed.
Figure 5.16: Background $\Delta t$ distribution for $B \rightarrow \phi K_s$ mode

![Graphs showing background distribution for SVD1 and SVD2](image)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SVD1</th>
<th>SVD2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(s_{\text{main}})_{\text{mult}}^\text{bkg}$</td>
<td>$0.86^{+0.39}_{-0.31}$</td>
<td>$1.38^{+0.06}_{-0.06}$</td>
</tr>
<tr>
<td>$(s_{\text{bkg}})_{\text{mult}}$</td>
<td>$1.86^{+0.90}_{-0.52}$</td>
<td>$3.38^{+0.39}_{-0.33}$</td>
</tr>
<tr>
<td>$(r_{\text{bkg}})_{\text{mult}}$</td>
<td>$0.80^{+0.14}_{-0.65}$</td>
<td>$0.11^{+0.04}_{-0.03}$</td>
</tr>
<tr>
<td>$(f_{\delta})_{\text{mult}}$</td>
<td>$0.77^{+0.06}_{-0.07}$</td>
<td>$0.55^{+0.08}_{-0.09}$</td>
</tr>
<tr>
<td>$(s_{\text{main}})_{\text{sgl}}^\text{bkg}$</td>
<td>$1.13^{+0.11}_{-0.18}$</td>
<td>$1.07^{+0.10}_{-0.10}$</td>
</tr>
<tr>
<td>$(s_{\text{tail}})_{\text{sgl}}^\text{bkg}$</td>
<td>$2.40^{+0.81}_{-0.66}$</td>
<td>$4.55^{+0.50}_{-0.43}$</td>
</tr>
<tr>
<td>$(f_{\text{tail}})_{\text{sgl}}^\text{bkg}$</td>
<td>$0.17^{+0.25}_{-0.10}$</td>
<td>$0.23^{+0.04}_{-0.04}$</td>
</tr>
<tr>
<td>$(f_{\delta})_{\text{sgl}}^\text{bkg}$</td>
<td>$0.78^{+0.14}_{-0.13}$</td>
<td>$0.42^{+0.14}_{-0.14}$</td>
</tr>
<tr>
<td>$\tau^\text{bkg (ps)}$</td>
<td>$1.27^{+0.20}_{-0.16}$</td>
<td>$0.75^{+0.10}_{-0.10}$</td>
</tr>
<tr>
<td>$\mu^\text{bkg (ps)}$</td>
<td>$-0.048^{+0.028}_{-0.028}$</td>
<td>$-0.039^{+0.017}_{-0.017}$</td>
</tr>
<tr>
<td>$\mu_\tau^\text{bkg (ps)}$</td>
<td>$-0.048^{+0.028}_{-0.028}$</td>
<td>$-0.032^{+0.017}_{-0.017}$</td>
</tr>
</tbody>
</table>

Table 5.4: Background shape parameters for the $B^0 \rightarrow \phi K_s$ mode.
5.9 Signal Probability

We introduce a signal fraction in the proper-time $\Delta t$ distribution given by Eq. 5.3. We use the event-dependent signal probability fraction $f_{\text{sig}}(\Delta E, M_{bc})$ based on the $\Delta E - M_{bc}$ distribution of signal and background and the measured values of $\Delta E$ and $M_{bc}$ in the event. To maximize the sensitivity of the $CP$ analysis, different fractions of signal distribution are used for each of the six $r$ bins, while the signal and background $\Delta E - M_{bc}$ shapes are assumed to be the same for all bins. In total, there are six parameters to represent the signal and two to represent the background shape distributions.

The two-dimensional probability density distribution in terms of $\Delta E$ and $M_{bc}$ is given by

$$p(\Delta E, M_{bc}) = f_{s} p_{\text{sig}}(\Delta E, M_{bc}) + (1 - f_{s}) p_{\text{bkg}}(\Delta E, M_{bc})$$

(5.44)

where $f_{s}$ is the signal pdf amplitude and $p_{\text{sig}}$ and $p_{\text{bkg}}$ are normalized probability density functions of signal and background distributions respectively. We model the signal probability density distribution as a double Gaussian in $\Delta E$ with the same non-zero mean and a single Gaussian in $M_{bc}$

$$p_{\text{sig}}(\Delta E, M_{bc}) = \left[ f_{s}^{\text{main}} G(\Delta E; \mu_{s}^{\Delta E}, \sigma_{s}^{\text{main}}) + (1 - f_{s}^{\text{main}}) G(\Delta E; \mu_{s}^{\Delta E}, \sigma_{s}^{\text{tail}}) \right] \times G(M_{bc}; \mu_{s}^{\Delta E}, M_{bc}^{\Delta E})$$

(5.45)

where $f_{s}^{\text{main}}$ is the fraction of the main Gaussian component, $\sigma_{s}^{\text{main}}$ and $\sigma_{s}^{\text{tail}}$ are the standard deviations of the main and tail components, respectively and $\mu_{s}^{\Delta E}$ is the common mean. The background distribution is modeled with a first-order polynomial in $\Delta E$ and the ARGUS function in $M_{bc}$

$$p_{\text{bkg}}(\Delta E, M_{bc}) = (1 + a_{0} \Delta E) \times \text{Ar}(M_{bc})$$

(5.46)

The ARGUS function $\text{Ar}(M_{bc})$ is defined as[51]

$$\text{Ar}(M_{bc}) = M_{bc} e^{a_{0} \left( \frac{M_{bc}}{E_{\text{beam}}} \right)^{2}} \sqrt{1 - \left( \frac{M_{bc}}{E_{\text{beam}}} \right)^{2}}$$

(5.47)
Parameter | $B^0 \rightarrow \phi K_s$ | $B^\pm \rightarrow \phi K^\pm$
--- | --- | ---
$f_s$ | 0.098 ± 0.008 | 0.068 ± 0.005
$\mu_s$ | 0.004 ± 0.001 | 0.001 ± 0.001
$\sigma_{s\text{main}}$ | 0.013 ± 0.001 | 0.015 ± 0.001
$\sigma_{s\text{tail}}$ | 0.041 ± 0.001 | 0.049 ± 0.002
$f_{s\text{main}}$ | 0.830 ± 0.003 | 0.917 ± 0.001
$\mu_{M_{bc}}$ | 5.278 ± 0.001 | 5.279 ± 0.001
$\sigma_{M_{bc}}$ | 0.0024 ± 0.0002 | 0.0024 ± 0.0001
$a_0$ | 1.12 ± 0.22 | 1.07 ± 0.13
$\alpha$ | $-28.3 \pm 2.8$ | $-29.9 \pm 1.8$

Table 5.5: Signal probability parameters for $B^0 \rightarrow \phi K_s$ and $B^\pm \rightarrow \phi K^\pm$ decay modes.

where $E_{\text{beam}}$ is the energy of the beams in the $\Upsilon(4S)$ center of mass frame, $\alpha$ is a floating parameter obtained from the fit. $E_{\text{beam}}$ is also a cut-off of the ARGUS distribution, which extends as $\text{Ar}(M_{bc}) \equiv 0$ for all $M_{bc} > E_{\text{beam}}$. We perform a two-dimensional $M_{bc}$-$\Delta E$ unbinned maximum likelihood fit to determine parameters of the signal and background shapes. Then, we fix all parameters except the signal amplitude $f_s$ and repeat the fit for six bins to determine all corresponding signal amplitudes.

Given the shapes of the signal and background distributions the sample purity can be calculated using the integrals over the signal region as

$$ P = \frac{\int \int_{\text{signal}} f_s p_{\text{sig}}(\Delta E, M_{bc}) d\Delta E dM_{bc}}{\int \int_{\text{signal}} p(\Delta E, M_{bc}) d\Delta E dM_{bc}} $$

The event-by-event signal fraction $f_{\text{sig}}$ is then calculated as a function of $\Delta E$ and $M_{bc}$ as

$$ f_{\text{sig}}(\Delta E, M_{bc}) = \frac{f_s p_{\text{sig}}(\Delta E, M_{bc})}{p(\Delta E, M_{bc})} = \frac{P p'_{\text{sig}}(\Delta E, M_{bc})}{P p'_{\text{sig}}(\Delta E, M_{bc}) + (1-P)p'_{\text{bkg}}(\Delta E, M_{bc})} $$

where $p'_{\text{sig}}(\Delta E, M_{bc})$ and $p'_{\text{bkg}}(\Delta E, M_{bc})$ are signal and background density functions which are normalized in the signal region. The result of the 2D fit for $B^0 \rightarrow \phi K_s$ is shown on Fig. 5.17. We use a different set of parameters for $B^0 \rightarrow \phi K_s$ and
Figure 5.17: Fitted 2D probability density distribution.

$B^\pm \to \phi K^\pm$ decay modes. The fitted signal probability parameters are presented in Table 5.5.

### 5.10 Validity Check

To avoid human bias in the measurement of the $CP$ asymmetry we used a blind analysis technique. The idea is to hide some part of the data on which the result is based. The final measurement of the $CP$ parameters is done as the last step of the analysis after determination of all procedures and parameters. No changes are allowed after opening the box. In our analysis we initially set the tag of the $B_{asc}$ meson at random. The $CP$ fitting of the data sample was done after performing of several consistency cross-checks of the proper time $\Delta t$ reconstruction and the $CP$ fitting algorithm.
### Table 5.6: Detector resolution parameters for SVD1 and SVD2 data samples.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SVD1</th>
<th>SVD2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s^0_{rec}$</td>
<td>$0.702^{+0.294}_{-0.262}$</td>
<td>$0.655^{+0.290}_{-0.114}$</td>
</tr>
<tr>
<td>$s^1_{rec}$</td>
<td>$0.132^{+0.027}_{-0.052}$</td>
<td>$0.129^{+0.031}_{-0.018}$</td>
</tr>
<tr>
<td>$s^0_{asc}$</td>
<td>$0.698^{+0.493}_{-0.063}$</td>
<td>$0.760^{+0.500}_{-0.105}$</td>
</tr>
<tr>
<td>$s^1_{asc}$</td>
<td>$0.050^{+0.037}_{-0.025}$</td>
<td>$0.045^{+0.021}_{-0.021}$</td>
</tr>
<tr>
<td>$s_{main}$</td>
<td>$1.11^{+0.04}_{-0.04}$</td>
<td>$1.02^{+0.05}_{-0.05}$</td>
</tr>
<tr>
<td>$s_{tail}$</td>
<td>$10.0^{+2.2}_{-2.8}$</td>
<td>$4.19^{+4.08}_{-0.83}$</td>
</tr>
<tr>
<td>$f_{tail}$</td>
<td>$0.035^{+0.026}_{-0.005}$</td>
<td>$0.087^{+0.028}_{-0.031}$</td>
</tr>
<tr>
<td>$\sigma_{ol}$</td>
<td>$44.5^{+6.6}_{-19.9}$</td>
<td>$35.5^{+7.8}_{-10.6}$</td>
</tr>
<tr>
<td>$f_{mult}^{ol}$</td>
<td>$0.00035^{+0.0004}_{-0.0001}$</td>
<td>$0.00036^{+0.0002}_{-0.0001}$</td>
</tr>
<tr>
<td>$f_{sng}^{ol}$</td>
<td>$0.012^{+0.004}_{-0.002}$</td>
<td>$0.018^{+0.002}_{-0.003}$</td>
</tr>
</tbody>
</table>

#### 5.10.1 Lifetime Fit

The detector resolution function is as broad as the lifetime distribution being fitted. Therefore, the uncertainty in the detector resolution parameters can be the dominant part of the systematic error. To reduce the dependence on the MC sample, we determine these parameters from a $B$ meson lifetime fit to real data. In this fit we fix the parameters of the $R_{np}$ resolution determined from the fit of the MC sample. Since we use the common resolution parameters for neutral and charged $B$ mesons we perform a simultaneous lifetime fit. The detector resolution parameters resulting from the fit to the $B^0 \to D^{(*)} \pi^+$, $D^{*0} \rho^+$, $D^{(*)} l^+ \nu$, $J/\psi K_s$, $J/\psi K^{*0}$ and $B^+ \to D^0 \pi^+$, $J/\psi K^+$ decays are presented in Table 5.6.

The parametrization of the resolution functions and background shapes are checked by the lifetime fit of the $B^0 \to \phi K_s$ and $B^{\pm} \to \phi K^{\pm}$ candidates. The reconstructed $\Delta t$ distributions with the fits imposed are shown in Fig. 5.18. The measured $B^0$ and $B^\pm$ lifetimes are

$$\tau_{B^0} = (1.63^{+0.18}_{-0.16}) \text{ ps}$$

$$\tau_{B^{\pm}} = (1.53^{+0.12}_{-0.11}) \text{ ps}$$

where the errors are statistical only. Both lifetimes are consistent with the world
average values[14] and suggests that the $\Delta t$ reconstruction is correct.

5.10.2 Linearity Check

To check the determination of the $CP$ violation parameters, we generated twenty one GEANT Monte Carlo samples with generated values of the $S$ parameter ranging from $-1.0$ to $1.0$ in steps of $0.1$. We examined the consistency between the input $S_{gen}$ and the fitted $S_{fit}$ values of the $CP$ asymmetry. In this test the corresponding MC detector resolution parameters were used. Fig. 5.19 illustrates the dependence of the fitted versus generated values for the $S$ parameter. The result of the linear fit gives $1.01 \pm 0.02$ for the slope term. We conclude that there is no noticeable bias in the $CP$ fitting procedure. Fig 5.20 illustrates the asymmetry plots for the sample with generated value $S_{gen} = 0.7$. The raw symmetry in each $\Delta t$ bin is defined as $(N_{q=+1} - N_{q=-1})/(N_{q=+1} + N_{q=-1})$, where $N_{q=+1}$ and $N_{q=-1}$ are the number of observed candidates with $q = +1$ and $q = -1$ respectively. The data points are in a
complete agreement with the expected distributions.

5.10.3 Asymmetry of the $B^\pm \rightarrow \phi K^\pm$ Control Sample

We performed a $CP$ fit on the accompanying control sample for charged $B^\pm \rightarrow \phi K^\pm$. The asymmetry term $S$ vanishes for this decay mode, but a non-zero value for the direct $CP$ asymmetry term $A$ is allowed. The same analysis procedure was applied for this decay mode with the set of parameters corresponding to the charged $B$ mesons. The measured $CP$ asymmetry found to be consistent with zero

$$S_{\phi K^\pm} = 0.02 \pm 0.24$$

$$A_{\phi K^\pm} = 0.24 \pm 0.17$$

Fig. 5.21 shows the corresponding asymmetry plots with results of the fits superimposed. Setting the tag of the associated $B_{asc}$ meson in $B^0 \rightarrow \phi K_s$ decay at random effectively produces a sample with no $CP$ asymmetry. The blinded $\phi K_s$ sample shows
(a) Proper time $\Delta t$ distribution where $q = 1$ corresponds to events where tagged $B$ meson identified as $B^0$ and $q = -1$ as $\bar{B}^0$.

(b) Raw asymmetry. The solid line represents the result of the fit.

Figure 5.20: Asymmetry plots for $B^0 \to \phi K_s$ GEANT MC sample with generated $S = 0.7$.

an absence of a $CP$ asymmetry as well.

5.11 Correction Due To The $B^0 \to K^+ K^- K_s$ and $B^0 \to f_0 K_s$ Contribution

So far we have considered $B^0 \to \phi K_s$ as the only peaking source. In actuality, there is some contribution from non-resonant $B^0 \to K^+ K^- K_s$ and $B^0 \to f_0 K_s$ events, which has the opposite $CP$ eigenvalue of $\xi = +1$. This effectively offsets the observed $CP$ parameters

\[ S_{\text{obs}} = (1 - 2f_{+1})S \]
\[ A_{\text{obs}} = (1 - 2f_{+1})A \]

where $S$ and $A$ are the actual $CP$ parameters corresponding to the $B^0 \to \phi K_s$ decay.
and $f_{+1}$ is the fraction of the $\xi = +1$ component. To estimate the non-$\phi$ contribution to the signal PDF we studied an extended region of invariant mass of the charged kaons. Fig. 5.22 shows the $M(K^+K^-)$ distribution after all selections except the $\phi$ mass cut. We estimated $B$ yields from the $M_{bc} - \Delta E$ fits of the regions just below the $\phi$ mass $M(K^+K^-) - M_{\phi} < -20 \text{ MeV}/c^2$ and above a $\phi$ mass $20 \text{ MeV}/c^2 < M(K^+K^-) - M_{\phi} < 60 \text{ MeV}/c^2$. We found $2.9 \pm 2.6$ and $35.0 \pm 8.0$ events after $B^0 \rightarrow \phi K_s$ subtraction in the lower and upper $\phi$ mass regions respectively.

Events in the region above the $\phi$ mass are dominated by non-resonant $B^0 \rightarrow K^+K^-K_s$ decays. The contribution from the $f_0$ in general is small and concentrated mainly in the low $K^+K^-$ and $\phi$ regions. To first order its contribution to the upper $\phi$ mass region is negligible. We assume that the yield in the upper region is accounted for only by the $K^+K^-K_s$ events. After the subtraction of the projected contribution of the $K^+K^-K_s$ events in the low $\phi$ region we calculated the feed across of the $f_0K_s$ to
be equal to $0.30^{+2.6}_{-0.3}$, consistent with zero. This confirms that our assumption that the contribution of $f_0$ events in the upper mass region is negligible. Finally, the estimated yields of $B^0 \rightarrow K^+ K^- K_s$ and $B^0 \rightarrow f_0 K_s$ events in the $\phi$ mass region are $8.4 \pm 1.9$ and $0.4^{+3.6}_{-0.4}$, respectively. The corresponding fractions are

\[ f_{KKK_s} = 5.1 \pm 1.2 \% \]
\[ f_{f_0 K_s} = 0.2^{+21}_{-0.2}\% \]

Since the contribution from $f_0 K_s$ decays is consistent with zero, we treated it as just an additional source of the systematic error. The errors of the obtained fractions were used for the systematic error study.

5.12 Fit Result

After determining all parameters and completing the cross-check studies we performed the final maximum likelihood fit of the $B^0 \rightarrow \phi K_s$ data sample. In this fit we fixed the $B^0$ lifetime to its world average value $\tau_{B^0} = 1.532 \pm 0.009$ ps. The resulting $CP$
asymmetry parameters $S^{\text{obs}}$ and $A^{\text{obs}}$ before $K^+K^-K_s$ correction are

$$S^{\text{obs}} = 0.50 \pm 0.31$$

$$A^{\text{obs}} = 0.18 \pm 0.20$$

The quoted errors are statistical only. We found that those errors are consistent with errors estimated by the toy MC study. In these pseudo MC experiments we generated $\Delta t$ and flavor tag values for the $B_{\text{asc}}$ meson according to the measured values of the $CP$ asymmetry and then smeared them with the detector resolution and flavor assignment errors. We did not notice the effects discussed in Appendix A mainly due to the fairly large statistics and the input $CP$ parameters being far from the physical boundary.

Fig. 5.23 shows the $-2\ln(L/L_{\text{max}})$ distribution as a function of the $S$ and $A$ around the fitted values. Fig. 5.24 (a) illustrates the difference in $\Delta t$ distribution for events tagged as $q = +1$ ($B^0$) and $q = -1$ ($\bar{B}^0$). The raw asymmetry between $B^0$ and $\bar{B}^0$ tagged mesons is shown in Fig. 5.24 (b) with the fitted curve superimposed.

The observed term $A$ is consistent with zero and indicates the absence of a direct $CP$ asymmetry in the current data set.

### 5.13 Systematic Uncertainties

The uncertainty in the parameters and shapes of distributions such as the signal fraction, the vertex resolution and the background results in a systematic error in the measurement of the $CP$ parameters. The effects of all parameters included in the fit are studied by varying their central values and repeating the $CP$ analysis. Below is the description of all sources of systematic error we considered and the procedure used to determine their contributions.

**Vertex Reconstruction**

Systematic uncertainties corresponding to the vertex reconstruction are studied
Figure 5.23: \(-2\ln(L/L_{max})\) as a function of (a) \(S\) and (b) \(A\).

(a) Proper time \(\Delta t\) distribution where \(q = 1\) corresponds to events where tagged \(B\) meson identified as \(B^0\) and \(q = -1\) as \(\bar{B}^0\).

(b) Raw asymmetry. The solid line represents the result of the fit.

Figure 5.24: Asymmetry plots for \(B^0 \rightarrow \phi K_s\).
by varying the track and vertex selection criteria. The following components were considered:

- We use an IP constrained fit in the determination of the position of both $B$ mesons. The IP profile is smeared due to the finite $B$ meson lifetime. We varied the transverse $x$-$y$ smearing by $\pm 10$ $\mu$m to determine the change in measured $CP$ parameters.

- Only tracks with $|d\tau| < 0.05$ cm and $\sigma_z < 0.05$ cm are used by the tag side vertex algorithm. We varied these two parameters by $\pm 0.01$ cm.

- The vertex quality parameter $\xi$ depends on the track reconstruction errors. The measured errors are corrected by a tracking error matrix determined from cosmic muon data. The $CP$ fit is repeated with this correction removed.

- Small biases in the $\Delta z$ measurement of the reconstructed $e^+e^- \rightarrow \mu^+\mu^-$ have been observed. The contribution to the systematic error is determined by repeating the $CP$ analysis after applying a special correction to the reconstructed charged tracks.

- We select events with vertex quality parameter $\xi < 100$ for reconstructed and tagging sides in the $CP$ fit. This requirement is varied from $\xi < 50$ to $\xi < 200$.

- The reconstructed proper time $\Delta t$ is required to satisfy $|\Delta t| < 70$ ps. This selection also affects the PDF normalization. We varied it by 30 ps to determine its systematic error.

**Flavor Tagging**

The systematic uncertainty in the flavor tagging is determined by varying the wrong-tag fraction of each $r$ bin by its error, given in Table 5.1. The combined deviations are added in quadrature.

**Resolution Function**
The resolution function plays an important role in the time-dependent $CP$ analysis. All parameters of the resolution function are in turn varied by $\pm \sigma$ from their nominal value except for the parameters of $R_{np}$ component, which are determined from Monte Carlo. They are varied by $\pm 2\sigma$ to take into account the difference in MC and real data.

**Physics Parameters**

In the nominal $CP$ fit physics parameters such as the $B^0$ lifetime $\tau_{B^0}$ and $\Delta m_d$ are fixed to the world average values $\tau_{B^0} = 1.532 \pm 0.009$ ps, $\Delta m_d = 0.505 \pm 0.005$ ps$^{-1}$. Each parameter is varied by its error.

**Possible Fit Bias**

Large statistics MC samples were generated to check if there is any possible bias in the $CP$ analysis procedure. We found that reconstructed $CP$ parameters are distributed around their input values. Some bias can occur if the measured $CP$ asymmetry is close to the physical boundary.

**Background Fraction**

The signal probability fraction depends not only on the uncertainty in the determination of the signal PDF fraction, but also on the uncertainty in the signal and background $M_{bc}-\Delta E$ shapes. All parameters are varied by their statistical errors obtained from simultaneous $M_{bc}-\Delta E$ 2D fit which are summarized in Table 5.5. The effects of contributions from $B^0 \rightarrow K^+K^-K_s$ and $B^0 \rightarrow f^0K_s$ decays have been discussed in Section 5.11. The errors in fractions of these decay modes affect the correction coefficient.

**Background $\Delta t$ Distribution**

The PDF of $\Delta t$ distribution has been discussed in Section 5.8. The systematic error due its uncertainty is estimated the same way: all parameters are individually varied by their errors.

**Tag-side Interference**

The effect of the interference between CKM-favored and CKM-suppressed $B \rightarrow D$
transitions in the tag side final state\cite{52} has been studied. There is a small correction to the signal $\Delta t$ distribution from the interference term. The size of the correction is estimated using the $B^0 \rightarrow D^{*-} l^+ \nu$ sample. The systematic biases in the $CP$ asymmetry are analyzed from the MC sample with the corrected PDFs. The systematic shift in $S$ is found to be negligibly small but sizable in the $A$ term.

The result of the systematic study for each source of the systematic error is summarized in Table 5.7. By adding all components in quadrature we obtain the following combined systematic errors:

$$\sigma_{S}^{\text{syst}} = 0.06$$

$$\sigma_{A}^{\text{syst}} = 0.05$$

<table>
<thead>
<tr>
<th>Source</th>
<th>$\delta S$</th>
<th>$\delta A$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertex Reconstruction</td>
<td>$\pm 0.01$</td>
<td>$\pm 0.02$</td>
</tr>
<tr>
<td>Flavor Tagging</td>
<td>$\pm 0.01$</td>
<td>$\pm 0.01$</td>
</tr>
<tr>
<td>Resolution Function</td>
<td>$\pm 0.05$</td>
<td>$\pm 0.02$</td>
</tr>
<tr>
<td>Physics Parameters</td>
<td>$\pm 0.01$</td>
<td>$\pm 0.01$</td>
</tr>
<tr>
<td>Possible Fit Bias</td>
<td>$\pm 0.01$</td>
<td>$\pm 0.01$</td>
</tr>
<tr>
<td>Background Fraction</td>
<td>$\pm 0.03$</td>
<td>$\pm 0.01$</td>
</tr>
<tr>
<td>Background $\Delta t$ Distribution</td>
<td>$\pm 0.01$</td>
<td>$\pm 0.02$</td>
</tr>
<tr>
<td>Tag Side Interference</td>
<td>$\pm 0.01$</td>
<td>$\pm 0.03$</td>
</tr>
<tr>
<td><strong>total</strong></td>
<td>$\pm 0.06$</td>
<td>$\pm 0.05$</td>
</tr>
</tbody>
</table>

Table 5.7: Summary of the systematic uncertainties contributions on the time-dependent parameters $S$ and $A$. Each entry represents the cumulative uncertainty by adding each contribution in quadrature.
Chapter 6

Conclusions

Using the data sample of 492 fb\(^{-1}\) (which corresponds to 535 million \(B\bar{B}\) pairs) collected with the Belle detector at the asymmetric KEKB accelerator working at the \(\Upsilon(4S)\) production energy, we reconstruct 202 good \(B^0 \rightarrow \phi K_s\) decay candidates used in the \(CP\) fit with an estimated purity of 81.2\%. The time-dependent asymmetry parameters \(S\) and \(A\) are extracted from a maximum likelihood fit to the proper-time interval between the two \(B\) candidates. The observed \(CP\) asymmetry after correction due to the \(B^0 \rightarrow K^+ K^- K_s\) contribution is:

\[
S = 0.55 \pm 0.31(\text{statistical}) \pm 0.06(\text{systematic}) \quad (6.1)
\]

\[
A = 0.18 \pm 0.20(\text{statistical}) \pm 0.05(\text{systematic}) \quad (6.2)
\]

In the Standard Model the \(S\) parameter is expected to be equal to the \(\sin 2\phi_1\) parameter of the Unitarity Triangle corresponding to the Cabibbo-Kobayashi-Maskawa matrix. This parameter has been measured precisely in \(B^0 \rightarrow J/\psi K_s\) and other \(c\bar{c}K^{(*)0}\) transitions and is considered to be the primary measurement of the \(\sin 2\phi_1\) parameter. Our result is consistent with its current average of all charmonium modes \(\sin 2\phi_1 = 0.685 \pm 0.032[53]\). Fig. 6.1 illustrates the averaged results of the \(\sin 2\phi_1\) measurements with various \(b \rightarrow q\bar{q}s\) modes. For each of these modes the expected \(S_f\) is given by \(-\xi_f \times \sin 2\phi_1\), where \(\xi_f\) is the \(CP\) eigenvalue of the final state \(f\). Averaging
over all $b \rightarrow s$-penguin modes is done assuming that there are no contributions to the decay amplitudes with non-zero weak phases. Although this assumption is valid to a very good approximation for $B^0 \rightarrow \phi K_s$ decay, it may be severely violated in some other modes, especially in non-$s\bar{s}$-resonant states, such as $B^0 \rightarrow \pi^0 K_s, \omega K_s$ and $K^+ K^- K_s$ decays.

The parameter $\mathcal{A}$ corresponds to $(|\lambda|^2 - 1)/(|\lambda|^2 + 1)$ in the SM and is consistent with zero. This shows that there is no observed direct $CP$ violation in $B^0 \rightarrow \phi K_s$ decay mode in the current data sample.

There are other measurements of Unitarity Triangle parameters beside $\sin 2\phi_1$, which allow us to check the consistency of the KM mechanism. This is done by the CKM fitter group[54]. Fig. 6.2 shows the rescaled version of the triangle under constraints from the current measurements. The apex of this triangle is located at $(\bar{\rho} + i\bar{\eta})$ in a $\bar{\rho}$-$\bar{\eta}$ complex plane. The other constraints are obtained from the following measurements:

Figure 6.1: Comparisons of averages in different $b \rightarrow q\bar{q} s$ modes.
\( \phi_2 \) is measured using time-dependent CP analyses in \( b \to u \) transitions like \( B \to \pi^+\pi^-, \rho^0\phi^0, \rho^+\rho^- \).

\( \phi_3 \) can be measured using time-dependent CP analyses in \( b \to c\bar{u}d \) transitions like \( B \to D^\pm\pi^\mp \), which provides the measurement of \( \sin(2\phi_1 + \phi_3) \). It can be directly obtained from the rate and asymmetry measurements of \( B^- \to D_{\text{CP}}^{(*)}K^{(*)}- \), where the \( D^{(*)} \) meson decays to CP even \( (D_{\text{CP}}^{(*)}) \) and CP odd \( (D_{\text{CP}}^{(*)}) \) eigenstates[55, 56].

- \( |V_{ub}| \) is measured using inclusive and exclusive semileptonic \( B \) decays, involving \( b \to u \) transition.

- \( |V_{cb}| \) is measured using inclusive and exclusive semileptonic \( B \) decays, involving \( b \to c \) transition.

- \( |V_{td}| \) can be obtained from the measurement of \( B \) meson mixing. For example, \( B_d^0 \) mixing is governed through the box diagram and \( \Delta m_d \propto |V_{ub}^*V_{td}|. \) The uncertainty in \( \Delta m_d \) can be translated into an uncertainty in \( V_{td} \). The recent measurement of the mixing in \( B_s^0 \) mesons dramatically improves the uncertainty in \( V_{td} \). Fig 6.2 shows the constraint of the \( V_{td} \) based on the measurement of the \( \Delta m_d \) alone together with the combined constraint using both \( \Delta m_d \) and \( \Delta m_s \) measurements.

- Another constraint can be obtained from the measurement of indirect CP violation in the neutral \( K \) meson system. The \( \epsilon \) constrains the apex of the Unitary triangle to lie on a hyperbola. The width of this hyperbola comes mainly from the uncertainty of the hadronic matrix elements.

At this point there seems to be no clear evidence of discrepancies between the experimental measurements and the Standard Model. Our measurement of \( \sin 2\phi_1 \) lies only 0.44\( \sigma \) away from the average from the charmonium modes[57, 58]. It is
also in a good agreement with the result obtained in BaBar experiment[59]. Due to the low branching fraction of the $B^0 \rightarrow \phi K_s$ decay mode, the uncertainty of our measurement is dominated by its statistical error. To reduce it to the level of the systematic error a significant increase in the integrated luminosity is needed, which requires a major upgrade of the accelerator. Such an upgrade is planned in the near future. A data set corresponding to an integrated luminosity of 5 ab$^{-1}$ is expected to be collected within one year by the Super KEKB, which should bring the statistical error down by the factor of 3, assuming that the statistical error scales as $1/\sqrt{L_{\text{int}}}$ of the integrated luminosity.
Appendix A

The Special Properties of the UML Fit Applied To CP Fitting

In this appendix we discuss some special features of the unbinned maximum likelihood fit applied to the CP fitting procedure. Recent studies of low-statistics CP samples have exhibited sample-to-sample variations larger than what one would naively expect from simple by statistical fluctuations. This leads to the question of whether the quoted statistical errors obtained the in the usual way are underestimated.

The CP violation parameters $S$ and $A$ are extracted from the proper time difference $\Delta t$ fit to the decay rate given by

$$
\Gamma_{B-\bar{B}}(\Delta t) \propto e^{-\frac{|\Delta t|}{\tau}} [1 + q(A_{CP} \cos \Delta m \Delta t + S_{CP} \sin \Delta m \Delta t)] \quad (A.1)
$$

where $q$ is the flavor of the tagging $B$. For a given data set of independent measurements $\Delta t_i, q_i$ characterized by the probability density function $P(\Delta t, q, S, A)$ the likelihood function is constructed as

$$
\mathcal{L}(S, A) = \prod_i P_i(\Delta t_i, q_i, S, A) \quad (A.2)
$$

To estimate the CP fit parameters $S$ and $A$ we need to maximize the likelihood function $\mathcal{L}$. In practice, it is easier to deal with $\ln \mathcal{L}$ instead of the likelihood function...
itself. Since the logarithm is a monotonic function, maximization of \( \ln L \) is equivalent to the maximization of \( L \). Using the logarithm also solves the problem of reaching the computer limit for a very small likelihood \( L \). There is an additional advantage of using logarithm: in the limit of large number of events in the sample \(-2 \ln L\) behaves like a \( \chi^2 \) and the standard deviation errors \( s \) can be determined from the contour given by

\[
-2 \ln L(S', A') = -2 \ln L(S, A) + s^2
\]  

(A.3)

The minimization software most commonly used by high energy physicists is MINUIT. It carries out the minimization (MIGRAD method) of a given function using a stable variation of the Davidson-Fletcher-Powell variable-metric algorithm. This algorithm converges to the correct error matrix as it converges to the function minimum. In practice, MIGRAD usually yields good estimates of the error matrix, but it is not always reliable. MINUIT can also estimate the standard deviation errors given by Eq. A.3 by calling the MINOS method. Let us emphasize that the errors returned by MINUIT are only approximate errors. It is worthwhile to perform independent calculations to cross check them.

We have examined the \( CP \)-fit algorithm for systematic consistency. To do that we generated twenty one samples with generated values of the \( S \) parameter ranging from \(-1.0 \) to \( 1.0 \) in steps of \( 0.1 \). Each sample consisted of 10000 experiments with 50 events in each experiment. These experiments were generated assuming an ideal situation with no background and without experimental errors in the measurement of the proper time difference and flavor of the tagging \( B \). Instead of plotting the more common linearity plot of the fitted value \( S_{\text{fit}} \) versus \( S_{\text{gen}} \), we plot \( S_{\text{fit}} - S_{\text{gen}} \) versus \( S_{\text{gen}} \) to magnify the difference. Each point represents an average of all 10000 psuedo experiments.

\[
S_{\text{fit}} = \frac{1}{n} \sum_{i=1}^{n} S_{\text{fit}}^i
\]

(A.4)

We introduce \( \chi^2/n \) to characterize the difference of the estimated errors in MINUIT
with the actual spreads of the fitted values. $\chi^2/n$ is defined as

$$\chi^2/n = \frac{1}{n} \sum_{i=1}^{n} \frac{(S_{\text{fit}} - S_{\text{gen}})^2}{(\sigma_S)^2}$$  \hspace{1cm} (A.5)$$

where $n = 10000$ is the number of experiments and $\sigma_S$ is the error estimated by MINUIT. For these cases we use the symmetric error returned by the MIGRAD method. If the estimated errors are correct the $\chi^2/n$ should be consistent with unity. The upper plots in Fig. A.1 show the average values of $S_{\text{fit}} - S_{\text{gen}}$ and $\chi^2/n$. In the actual experiment, the physical distribution is diluted by incorrect flavor assignments, finite detector resolution and the presence of background. To determine the impact of these physical effects on the $CP$ fit, we generate samples with proper-time values smeared by Gaussian resolution functions and a Gaussian background. The number of events in each experiment, which represents the total number of signal and background events, is kept the same. The values of $S_{\text{fit}} - S_{\text{gen}}$ and $\chi^2/n$ for these samples are shown in the lower plots of Fig. A.1.

As one can see, there is a bias in the determination of the $S$ value which increases as the absolute value of the generated parameter $S$ increases. Adding resolution and background effects makes the situation noticeably worse. $\chi^2/n$ has a similar tendency. For small $S_{\text{gen}}$ the estimated errors describe the spread of fitted values quite well, but for large values of $S_{\text{gen}}$ the difference can be substantial. Adding the experimental smearing and background increases the estimated errors and the difference becomes less apparent.

The existence of a non-vanishing bias with the increasing of the number of experiments used for averaging does not mean the absence of convergence of the algorithm. The term convergence here should not be confused with the convergence in the limit of the infinite number of the events in the sample. $CP$ fitting is not a linear operation and increasing the statistics by averaging over more experiments can not remove this bias. Figures A.2-A.5 show that with increasing numbers of events not only do the fitted values of $S_{\text{fit}}$ converge to the generated values, but also the errors estimated in
MINUIT converge to the observed statistical errors.

Since in the case of small statistics the observed errors are much larger than the errors returned by MINUIT, the question arises if the estimated errors have any meaning and whether we should use them at all or estimate them differently. To study that, we generated two high statistics MC samples with $S_{\text{gen}} = 0.9$ and 50 events per experiment. The first sample consists of idealized experiments and the second takes into account the physical smearing described above. Later in the discussion we will do side-by-side comparison of these two samples to check if the discrepancies get magnified or smeared due to the experimental dilutions. We chose this extreme value for $S_{\text{gen}}$ to maximize the discrepancy effects and still be relatively far from the physical boundary.

Fig. A.6 shows the distribution of the MINUIT errors. The distribution of the smeared sample is wider due to the larger uncertainty in the measurements of the parameters. We found that the MINUIT errors are highly correlated with the statistical spread of the fitted $S_{\text{fit}}$ values. See Fig. A.7 for details. This means that errors returned by MINUIT carry some information about the fit and should not be simply discarded. We split our experiments into different MINUIT error bins and checked the distribution of the $\delta S = S_{\text{fit}} - S_{\text{gen}}$, which in the general case happened not to be Gaussian. The $\delta S$ distribution for the slice in MINUIT error around the most probable value is shown in Fig. A.9. For each bin we calculate the mean $\mu_{\delta S}$ and the standard deviation $\sigma_{\delta S}$ of the $\delta S$ distribution. As one can see from Fig. A.8, the errors obtained by the MINUIT routine should be corrected depending on the estimated errors themselves.

Another thing that should be kept in mind is the difference between the mean and the most probable value (mode) of the distribution. An unbinned maximum likelihood algorithm finds the most probable value with estimated asymmetric errors. Assuming that the distribution can be described as an asymmetric Gaussian, the expectation would be biased by the value $\sqrt{2/\pi}(\sigma_r - \sigma_l)$, where $\sigma_r$ and $\sigma_l$ are the right
and left errors respectively. For positive $S$ the left errors are systematically larger than the right errors as shown on Fig. A.10. Still, this correction is too small to account for the positive bias of the mean in $\delta S$. The mode of the distributions also shows a dependence on the MINUIT error but does not have such a large bias.

Given these findings, we conclude that the best way to estimate the statistical errors is by using a toy Monte Carlo with a generated value of $S_{\text{gen}}$ that is equal to the measured value of $S$ from the experimental data. In this MC sample the generated $\Delta t$ values should be smeared with the same experimental resolution functions to correctly model the experiment. Then, using pseudo experiments with the MINUIT errors close to the MINUIT error obtained by fitting the data sample, we can determine the expected profile $\delta S$ of the statistical errors. Finally, the statistical asymmetric errors can be estimated from this distribution based on a desired confidence interval. If the statistics are very small, a correction of the measured value might also be required.
Figure A.1: Consistency check. 50 events per experiment.
Figure A.2: Consistency check. 100 events per experiment.
Figure A.3: Consistency check. 200 events per experiment.
Figure A.4: Consistency check. 1000 events per experiment.
Figure A.5: Consistency check. 10000 events per experiment.
Figure A.6: Distribution of estimated MINUIT errors.

(a) Signal events with idealistic measurements.  
(b) With experimental smearings and background.

Figure A.7: $S - S_{\text{gen}}$ versus the MINUIT errors.

(a) Signal events with idealistic measurements.  
(b) With experimental smearings and background.
Figure A.8: Dependence of $\mu_{S}$ and $\sigma_{S}$ on MINUIT errors.
(a) Signal events with idealistic measurements. Slice is taken around MINUIT error = 0.18 with the width 0.01.

(b) With experimental smearing and background. Slice is taken around MINUIT error = 0.37 with the width 0.01.

Figure A.9: $S - S_{\text{gen}}$ for the same MINUIT errors slice.

(a) Signal events with idealistic measurements.  

(b) With experimental smearing and background.

Figure A.10: $\sigma_l/\sigma_r$ distribution.
Bibliography


[56] M. Gronau and D. London., “How to determine all the angles of the unitarity triangle from $B(d)0 \to D K(s)$ and $B(s)0 \to D0$,” Phys. Lett. B253, 483 (1991).

[57] B. Aubert et al., “Improved measurement of CP asymmetries in $B^0 \to (c\bar{c})K^{(*)0}$ decays,” Phys. Rev. Lett. 94, 161803 (2005).

[58] K. Abe et al., “Time-dependent CP asymmetries in $b \to s$ anti-q q transitions and $\sin(2\phi(1))$ in $B0 \to J/\psi K0$ decays with 386 million $B$ anti-$B$ pairs,” (2005).

[59] B. Aubert et al., “Measurement of CP asymmetries in $B^0 \to \phi K^0$ and $B0 \to K^+K^-K^0_s$ decays,” Phys. Rev. D71, 091102 (2005).