Measurement of Branching Fraction and Time-dependent CP Asymmetry Parameters in $B^0 \to K^0\pi^0$ Decays

January 2009

School of Interdisciplinary Research of Scientific Phenomena and Information
Graduate School of Humanities and Sciences
Nara Women's University

Miyuki Fujikawa

Measurement of Branching Fraction and Time-dependent CP Asymmetry Parameters in $B^0 \to K^0\pi^0$ Decays

January 2009

School of Interdisciplinary Research of Scientific Phenomena and Information
Graduate School of Humanities and Sciences
Nara Women's University

Miyuki Fujikawa

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Abstract

We present a measurement on the branching fraction and time-dependent CP violation parameters of $B^0 \to K^0\pi^0$ decays. These results are obtained from a data sample that contains 657×10^6 $B\bar{B}$ pairs collected near the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB asymmetric e^+e^- collider.

We obtain the branching fraction,

$$\mathcal{B}(B^0 \to K^0\pi^0) = [8.72\ ^{+0.51}_{-0.50}\ (\text{stat})\ ^{+0.46}_{-0.40}\ (\text{syst})] \times 10^{-6},$$

which is in agreement with the current world average.

We measure the CP parameters,

$$\begin{aligned} \mathcal{A}_{CP} &= +0.14 \pm 0.13\ (\text{stat})\ \pm 0.06\ (\text{syst}), \\ \mathcal{S}^{\text{eff}}_{CP} &= +0.67 \pm 0.31\ (\text{stat})\ \pm 0.08\ (\text{syst}), \end{aligned}$$

where no evidence for direct CP violation is found and the mixing-induced component is consistent with the expectation from the Kobayashi-Maskawa theory of CP violation. We find the measurement of $\mathcal{A}_{CP}$ to deviate from the $K - \pi$ sum rule expectation by 1.9σ.

Acknowledgements

First and foremost, I wish to express my gratitude to my supervisor, Prof. H. Hayashii. Thank you for your constant encouragement beginning from those early tau days when we would discuss physics anytime, anywhere.

I appreciate the care of my "physics father", Prof. S. Noguchi, who broadened my horizons. I'd also like to thank Assoc. Prof. K. Miyabayashi for his support.

I am grateful to my seniors, Dr. S. Kataoka and Dr. A. Sekiya, who treated me as a sister and taught me many things. I especially enjoyed all those extracurricular activities.

Of course, where would I be without the love and support of my family and friends, especially Ms. Y. Nakahama and Ms. J. Sekihara.

Moving on to KEK, I'd like to thank Prof. Y. Sakai, Dr. K. Hara, Dr. T. Higuchi and all the other members of the ICPV group for their input into this analysis. Service work would have been impossible without the guidance of Assoc. Prof. Y. Ushiroda, Dr. I. Nakamura and the other members of the ECL group.

To my analysis partner, Dr. Y. Yusa, thank you for considering my doctoral schedule in performing the $B^0 \to K^0_L \pi^0$ analysis.

The cross-check of my analysis was kindly performed by Dr. K. Sumisawa who made time for me even when he was busy.

I would like to express my special thanks to Dr. J. Dalseno. Without his great help and precious advice, this thesis work would not have finished.

Last but not least, my supervisor at KEK, Prof. M. Hazumi, suggested this topic and taught me the necessary theory with regular meetings. Thank you for your continued encouragement and acceptance.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

One of the most important discoveries in cosmology and particle physics is that the universe started with the Big Bang (Fig. 1.1 [1]). It is supported by the discovery of 3K microwave background radiation, the abundance of light nuclei and the observation of distant galaxies receding from us. At present, there is no theory that can account for the origins of the universe, but from 10^{-10} s onwards after the Big Bang, the universe can be described by cosmology and particle physics. These predict that as the universe cools, particle-antiparticle pairs are converted to photons, and therein lies the paradox. Either the particles and antiparticles annihilate one another ultimately creating a sea of photons in which matter could not exist, otherwise an anti-universe must exist. However, there is no evidence for the universal antimatter counterpart.

In 1967, Sakharov postulated a mechanism to generate this cosmological charge asymmetry based on three conditions [2],

non-conservation of baryonic number,

breaking of symmetry between particles and anti-particles, both C and CP, and

deviation from thermal equilibrium.

Parity, P, charge conjugation, C, and time reversal, T, are discrete transformation operators. The P operator sends a particle from $(t, x) \rightarrow (t, -x)$, C interchanges a particle with its anti-particle and T transforms a particle from $(t, x) \rightarrow (-t, x)$. These three discrete operators form, in combination, a single operator called the CPT operator which is conserved for any Lorentz invariant quantum field theory.

The high energy particle physics community was particularly concerned with Sakharov's second condition, the essential difference between matter and antimatter, called CP violation.

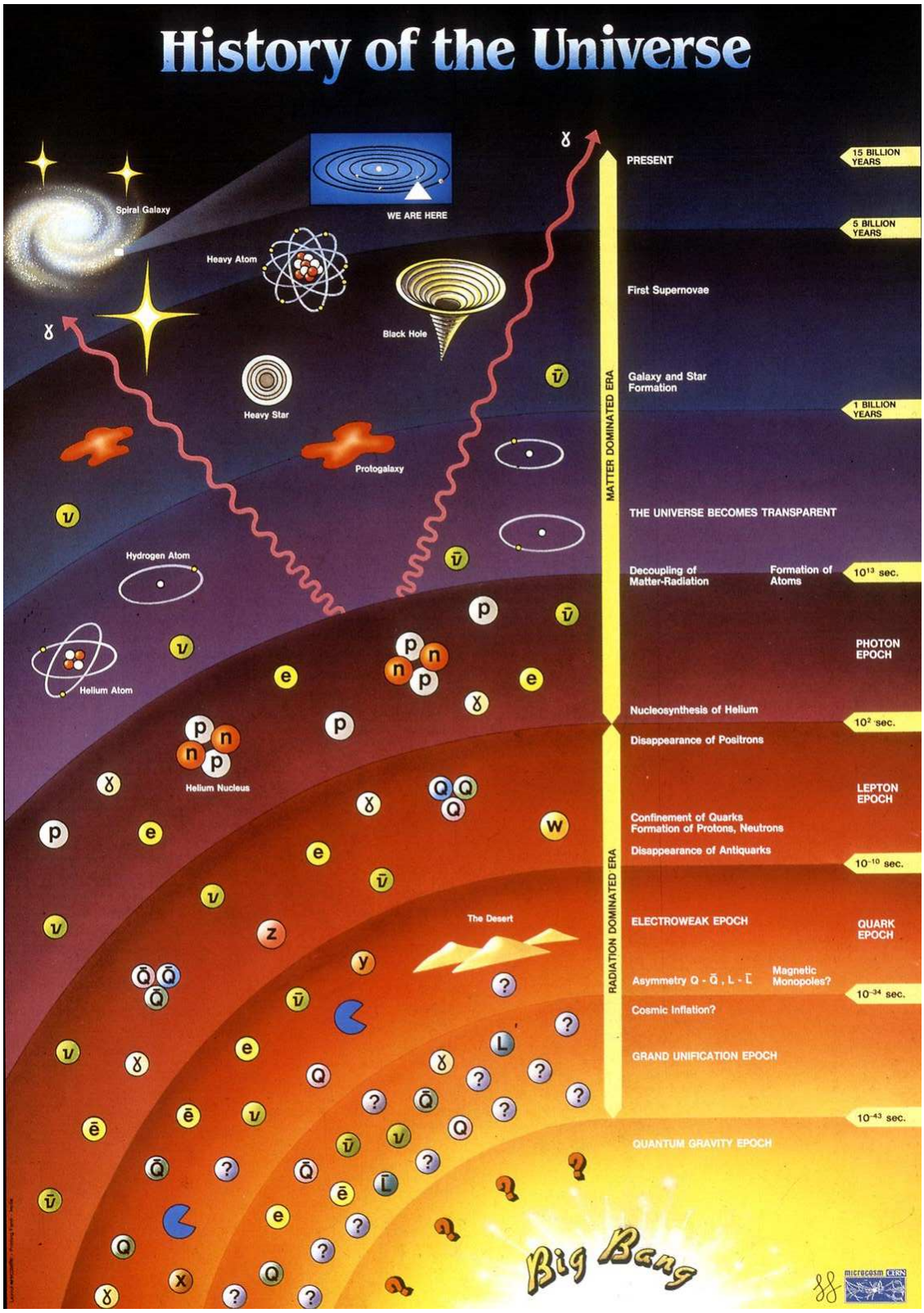


Figure 1.1: History of the universe.

Interestingly enough, before the proposal of the Sakharov conditions, CP violation was already discovered in 1964 completely by accident. While investigating an anomaly in neutral kaon regeneration, the CP violating decay, $K_L^0 \to \pi^+ \pi^-$, was found [3].

Then in 1973, Kobayashi and Maskawa incorporated into the Glashow-Weinberg-Salam (GWS) framework [4, 5, 6] a mechanism that allowed CP violation for three quark families [7]. In their model, CP violation arises from a single irreducible complex phase in the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix. Though the group of GWS, CKM and Quantum Chromodynamics (QCD) theories, which we call the Standard Model (SM), was very successful, CP violation had only been observed in neutral kaons.

In 1980, Bigi, Carter and Sanda pointed out that $B^0\bar{B}^0$ mixing may induce a large time-dependent CP asymmetry [8, 9]. Two B factories, Belle and BaBar, were commissioned in Japan and USA respectively, and CP violation in the B meson sector was confirmed by 2001 in $B^0 \to J/\psi\ K_S^0$ decays [10, 11].

However, the amount of CP violation in the Standard Model is not enough to generate our matter-dominated universe [12]. In the future, we must look for physics beyond the Standard Model to find new sources of CP violation besides the single KM phase of the Standard Model. Theories such as Supersymmetry and Grand Unified Theories, which expect more than one source of CP violation, may provide the breakthrough that is needed to explain our matter-dominated universe.

In this chapter, we introduce CP violation in the Standard Model, phenomenologies of CP violation in the B meson system. After that, we discuss the special cases of CP violation in $b \to c\bar{c}s$ and $b \to q\bar{q}s$ transitions. The $B^0 \to J/\psi K_S^0$ decay is the most typical mode that contains the $b \to c\bar{c}s$ transitions, while the $b \to q\bar{q}s$ transitions can be studied with several decay modes such as $B^0 \to K^0\pi^0, B^0 \to \phi K^0, B^0 \to \eta' K^0$ and $B^0 \to \omega K^0$. Among these, $B^0 \to K^0\pi^0$ decays are the subject of this thesis. As is discussed later in detail, decays proceeding through $b \to q\bar{q}s$ transitions are expected to be sensitive to possible sources of CP violation beyond the Standard Model, since additional contributions to their decay amplitudes in the loop diagrams (penguin diagrams) can be large.

1.2 *CP* Violation in the Standard Model

CP violation arises from the GWS $SU(2) \times U(1)$ model for weak and electromagnetic interactions. twelve fields are needed to describe the quarks. Following Ref. [13], we assign the left-handed quarks to $SU(2)$ doublets and the right-handed quarks to $SU(2)$ singlets.

$SU(2)$ doublets, Q_L^i:

$$Q_L^1 = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \quad Q_L^2 = \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \quad Q_L^3 = \begin{pmatrix} t_L \\ b_L \end{pmatrix}.$$

$SU(2)$ singlets, u_R^i, d_R^i:

$$u_R^1 = u_R, \quad u_R^2 = c_R, \quad u_R^3 = t_R$$
$$d_R^1 = d_R, \quad d_R^2 = s_R, \quad d_R^3 = b_R.$$

The gauge invariant Lagrangian is given by

$$\mathcal{L}_1 = \overline{Q}_L^i i\not{D} Q_L^i + \overline{u}_R^i i\not{D} u_R^i + \overline{d}_R^i i\not{D} d_R^i - \frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{1.1}$$

where $\not{D} \equiv \gamma^\mu D_\mu$. The symbol, D_μ, represents the gauge covariant derivative and takes the form,

$$D_\mu = \partial_\mu - igA_\mu^a T^a - ig' Y B_\mu, \tag{1.2}$$

where g is the $SU(2)$ coupling constant, A_μ^a $(a = 1, 2, 3)$ are the vector gauge fields associated with $SU(2)$, T^a $(a = 1, 2, 3)$ are the generators of $SU(2)$ in the fundamental representation, g' is the $U(1)$ coupling constant, B_μ is the vector gauge field associated with $U(1)$ and Y is the generator of $U(1)$.

The $SU(2)$ field strength is given by

$$F_{\mu\nu}^a \equiv \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g\epsilon_{abc} A_\mu^b A_\nu^c, \tag{1.3}$$

and the $U(1)$ field strength is given by

$$B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{1.4}$$

Now, we generate mass terms for the quarks. This requires an $SU(2) \times U(1)$ multiplet of spin-less fields, ϕ, that can couple Q_L^i to u_R^i and Q_L^i to d_R^i using a Yukawa coupling, λ_{ij}. The mass terms are given by

$$\mathcal{L}_2 = -\lambda_{ij}^d \overline{Q}_L^i \cdot \phi d_R^j - \lambda_{ij}^u \overline{Q}_L^i \cdot \tilde{\phi} u_R^j + h.c., \tag{1.5}$$

where $h.c.$ denotes the Hermitian conjugate.

In accordance with the Nambu-Goldstone theorem, gauge bosons are massless in a gauge theory where spontaneous breaking of global symmetry has occurred. At this point, there are 4 massless vector gauge bosons, but only one in nature is observed to be massless, that is, the photon. So, we introduce the Higgs mechanism and spontaneously break local symmetry to generate masses for three of the four gauge fields. By choosing the appropriate vacuum expectation value, $\langle\phi\rangle$, local $SU(2) \times U(1)$ symmetry is broken and results in just local $U(1)$ symmetry. Thus, the electromagnetic gauge invariance is not broken by the vacuum state and as a result, the vector boson associated with $U(1)$, identified as the photon, remains massless.

In the "unitarity gauge", which is chosen to make ϕ locally invariant, we rewrite $\mathcal{L}_2$ as

$$\mathcal{L}_2 = -\lambda_{ij}^d \begin{pmatrix} \bar{u}_L^i & \bar{d}_L^i \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} d_R^j - \lambda_{ij}^u \begin{pmatrix} \bar{u}_L^i & \bar{d}_L^i \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} r + h(x) \\ 0 \end{pmatrix} u_R^j + h.c., \tag{1.6}$$

where $h(x)$ is called the Higgs field.

As we have introduced the Higgs mechanism, it is also necessary to include the Higgs Lagrangian. From the Higgs Lagrangian, we can define the three massive vector bosons whose mass we have only just generated,

$$W_\mu^+ = \frac{1}{\sqrt{2}}(A_\mu^1 - iA_\mu^2), \tag{1.7}$$

$$W_\mu^+ = \frac{1}{\sqrt{2}}(A_\mu^1 - iA_\mu^2), \tag{1.8}$$

$$Z_\mu^0 = \frac{1}{\sqrt{g^2+g'^2}}(gA_\mu^3 - g'B_\mu), \tag{1.9}$$

and the massless vector boson,

$$A_\mu = \frac{1}{\sqrt{g^2+g'^2}}(g'A_\mu^3 + gB_\mu). \tag{1.10}$$

Using these definitions, one can introduce the weak mixing angle, θ_W, and rewrite the gauge covariant derivative, putting $\mathcal{L}_1$ into the new form,

$$\mathcal{L}_1 = \overline{u}^i i\partial\!\!\!/\, u^i + \overline{d}^i i\partial\!\!\!/\, d^i + g(W_\mu^+ J_W^{\mu+} + W_\mu^- J_W^{\mu-} + Z_\mu^0 J_Z^\mu) + eA_\mu J_{EM}^\mu - \frac{1}{4}F_a^{\mu\nu}F_{\mu\nu}^a - \frac{1}{4}B^{\mu\nu}B_{\mu\nu}, \tag{1.11}$$

where

$$J_W^{\mu+} = \frac{1}{\sqrt{2}}\overline{u}_L^i \gamma^\mu d_L^i \tag{1.12}$$

$$J_W^{\mu-} = \frac{1}{\sqrt{2}}\overline{d}_L^i \gamma^\mu u_L^i \tag{1.13}$$

$$\begin{aligned} J_Z^\mu = \frac{1}{\cos\theta_W}\Bigg[&\overline{u}_L^i\gamma^\mu\left(\frac{1}{2} - \frac{2}{3}\sin^2\theta_W\right)u_L^i + \overline{u}_R^i\gamma^\mu\left(-\frac{2}{3}\sin^2\theta_W\right)u_R^i \\ &+\overline{d}_L^i\gamma^\mu\left(-\frac{1}{2} + \frac{1}{3}\sin^2\theta_W\right)d_L^i + \overline{d}_R^i\gamma^\mu\left(\frac{1}{3}\sin^2\theta_W\right)d_R^i\Bigg] \end{aligned} \tag{1.14}$$

$$J_{EM}^\mu = \overline{u}^i\gamma^\mu\left(+\frac{2}{3}\right)u^i + \overline{d}^i\gamma^\mu\left(-\frac{1}{3}\right)d^i. \tag{1.15}$$

$J_W^{\mu+}$ and $J_W^{\mu-}$ are called the charged weak currents, J_Z^μ, the neutral weal current and J_{EM}^μ, the Electromagnetic Current.

It is possible to diagonalise the mass matrices of $\mathcal{L}_2$ through use of unitary matrices. This diagonalisation performs a change of basis from the weak eigenstates to the physical mass eigenstates, denoted by the primed fields, by the following transformations:

$$\begin{aligned} u_L^i = U_{ij}^u u_L'^j, \quad u_R^i = W_{ij}^u u_R'^j \\ d_L^i = U_{ij}^d u_L'^j, \quad d_R^i = W_{ij}^d d_R'^j \end{aligned} \tag{1.16}$$

We can now write the Lagrangian in the mass eigenstate basis. In particular, we are especially interested in the Charged Weak Currents as they are the only pieces of the Lagrangian whose terms do not maintain their form under transformation to the mass eigenstate basis. Under this transformation,

$$J_W^{\mu+} \to J_W'^{\mu+} = \frac{1}{\sqrt{2}} \overline{u}_L'^i \gamma^\mu (U_u^\dagger U_d)_{ij} d_L'^j \tag{1.17}$$

$$J_W^{\mu-} \to J_W'^{\mu-} = \frac{1}{\sqrt{2}} \overline{d}_L'^i \gamma^\mu (U_d^\dagger U_u)_{ij} u_L'^j \tag{1.18}$$

We define $V_{ij} \equiv (U_u^\dagger U_d)_{ij}$, where V_{ij} is a unitary matrix known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix, which allows for mixing between the quark families. It is possible for this $SU(2) \times U(1)$ gauge theory to violate CP conservation. The full electroweak Lagrangian transforms into itself under CP except for those terms describing charged current weak interactions for quarks,

$$\mathcal{L}_{cc}^q = \frac{g}{2\sqrt{2}} [V_{ij} W_\mu^+ \overline{u}'^i \gamma^\mu (1-\gamma^5) d'^j + V_{ij}^* W_\mu^- \overline{d}'^j \gamma^\mu (1-\gamma^5) u'^i]. \tag{1.19}$$

Under CP, these two terms transform as

$$W_\mu^+ \overline{u}'^i \gamma^\mu (1-\gamma^5) d'^j \xrightarrow{CP} W_\mu^- \overline{d}'^j \gamma^\mu (1-\gamma^5) u'^i, \tag{1.20}$$

$$W_\mu^- \overline{d}'^j \gamma^\mu (1-\gamma^5) u'^i \xrightarrow{CP} W_\mu^+ \overline{u}'^i \gamma^\mu (1-\gamma^5) d'^j, \tag{1.21}$$

and thus,

$$\begin{aligned} \mathcal{L}_{cc}^q &\to CP \mathcal{L}_{cc}^q \\ &= \frac{g}{2\sqrt{2}} [V_{ij} W_\mu^- \overline{d}'^j \gamma^\mu (1-\gamma^5) u'^i + V_{ij}^* W_\mu^+ \overline{u}'^i \gamma^\mu (1-\gamma^5) d'^j]. \end{aligned} \tag{1.22}$$

After acting on the Lagrangian with the CP operator, we can see by comparison with Eq. (1.19), that CP is only conserved if, $V_{ij}^* = V_{ij}$, that is, there is no CP violation if V_{ij} is real. However, we know from the unitarity of V_{ij}, that the CKM matrix can be complex for three families of quarks. Therefore, this $SU(2) \times U(1)$ gauge theory has the potential to violate CP.

1.3 The CKM Matrix

The CKM quark mixing matrix, $V_{ij} \equiv V_{\text{CKM}}$, arises from the transformation of the full electroweak GWS Lagrangian from the weak eigenstate into the physical mass eigenstate,

$$\begin{pmatrix} d \\ s \\ b \end{pmatrix}_{\text{weak}} = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{\text{mass}} \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{\text{mass}}. \tag{1.23}$$

The CKM matrix is required to be unitary to conserve probability, and for three or more families of quarks, it can be shown that $V_{\rm CKM}$ may be complex in general. The Lagrangian does not maintain its form under CP transformation if $V_{\rm CKM}$ is indeed complex, so the electroweak gauge theory has the potential to violate CP symmetry.

The CKM matrix contains information on the strength of flavour changing weak decays, so it is natural to parametrise $V_{\rm CKM}$ in a way that conveys their relative strengths more intuitively. This is usually realised with the Wolfenstein parametrisation [14],

$$V_{\rm CKM} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta + \frac{i}{2}\eta\lambda^2) \\ -\lambda & 1-\frac{1}{2}\lambda^2 - i\eta A^2\lambda^4 & A\lambda^2(1+i\eta\lambda^2) \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^6), \quad (1.24)$$

which expands $V_{\rm CKM}$ in powers of $\lambda \equiv \sin\theta_C = 0.22$, where θ_C is the Cabibbo angle [15]. The real parameters, A, ρ and η, are of order unity. From unitarity of the CKM matrix, $V_{\rm CKM}V_{\rm CKM}^\dagger = \mathbf{1}$, we may obtain nine relations. One of them is

$$\begin{array}{ccccccc} V_{ud}V_{ub}^* & + & V_{cd}V_{cb}^* & + & V_{td}V_{tb}^* & = & 0, \\ \mathcal{O}(\lambda^3) & & \mathcal{O}(\lambda^3) & & \mathcal{O}(\lambda^3) & & \end{array} \quad (1.25)$$

which is relevant to B decays. Note that since the relation is the sum of three complex terms, it may be represented graphically as a triangle in the complex plane and is known as the CKM Unitarity Triangle. It is interesting to see that the lengths of each side are of the same order, $\mathcal{O}(\lambda^3)$. This means that all the internal angles are naturally large which implies that sizable CP violation can be observed in the B meson sector.

It is convenient to normalise Eq. (1.25) by $|V_{cd}V_{cb}^*|$, so that one side becomes aligned to the real axis with length of order unity. The rescaled CKM Unitarity Triangle is shown in Fig. 1.2 where

$$\overline{\rho} \equiv \left(1 - \frac{\lambda^2}{2}\right)\rho, \quad \overline{\eta} \equiv \left(1 - \frac{\lambda^2}{2}\right)\eta \quad (1.26)$$

and

$$\phi_1 \equiv \pi - \arg\left(\frac{-V_{td}V_{tb}^*}{-V_{cd}V_{cb}^*}\right), \quad \phi_2 \equiv \arg\left(\frac{V_{td}V_{tb}^*}{-V_{ud}V_{ub}^*}\right), \quad \phi_3 \equiv \arg\left(\frac{V_{ud}V_{ub}^*}{-V_{cd}V_{cb}^*}\right). \quad (1.27)$$

1.4 Phenomenology of Time-Dependent CP Violation in B Meson Decays

1.4.1 Time Evolution of Neutral B Mesons

The neutral pseudo-scalar meson, $B^0(\overline{b}d)$, and its CP conjugate, $\bar{B}^0(b\overline{d})$, produced in the strong interaction is a pure flavour eigenstate. In the Standard Model, time-dependent

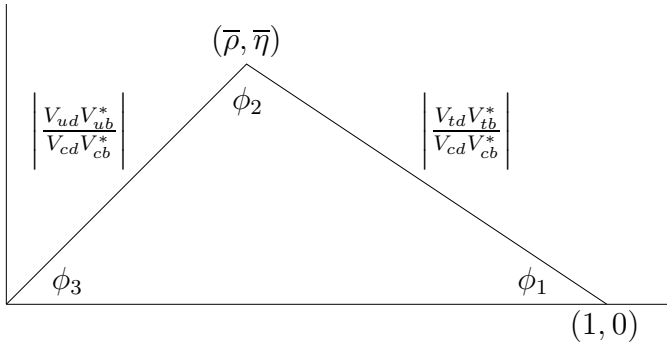


Figure 1.2: Rescaled CKM Unitarity Triangle.

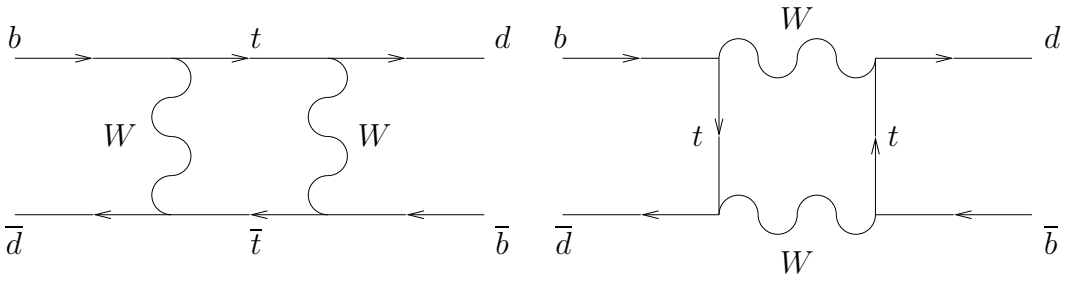


Figure 1.3: Box diagrams of $B^0\bar{B}^0$ mixing.

oscillation, or neutral meson mixing, may occur if its mass and flavour eigenstates differ. $B^0\bar{B}^0$ mixing dominantly proceeds via the second order weak interaction where d and b quarks couple to W bosons and t quarks as shown in Fig. 1.3.

The time evolution of the neutral B meson state can be written as a linear combination of pure B^0 and $\bar{B}^0$ flavour eigenstates,

$$|B(t)\rangle = a(t)|B^0\rangle + b(t)|\bar{B}^0\rangle, \tag{1.28}$$

and is governed by the Schrödinger equation,

$$i\frac{d}{dt}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \mathbf{H}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix}. \tag{1.29}$$

The weak effective Hamiltonian in the flavour basis, $\mathbf{H}$, is a 2×2 matrix which takes the form,

$$\mathbf{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{21} - \frac{i}{2}\Gamma_{21} & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}, \tag{1.30}$$

where $\mathbf{M}$ and $\mathbf{\Gamma}$ are the mass and decay matrices, respectively. From hermiticity of the Hamiltonian, $H_{21} = H_{12}^*$ holds. The off-diagonal elements are associated with flavour changing transitions, $B^0 \leftrightarrow \bar{B}^0$, so CP violation occurs if $H_{12} \neq H_{12}^*$. The diagonal elements are associated with the flavour conserving transitions, $B^0 \to B^0$ and $\bar{B}^0 \to \bar{B}^0$. Assuming CPT invariance, or in other words, the B^0 couples to itself in the same way a $\bar{B}^0$ couples to itself, the equality $H_{11} = H_{22} \equiv M - i/2\Gamma$ holds. Thus, from diagonalising $\mathbf{H}$, the eigenvalues, $\mu_\pm$,

are

$$\mu_\pm \equiv M - \frac{i}{2}\Gamma \pm \sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)}. \tag{1.31}$$

We now acquire the corresponding eigenstates of this Hamiltonian for each eigenvalue,

$$\begin{aligned} |B_{\rm L}\rangle &\equiv p|B^0\rangle + q|\bar{B}^0\rangle \quad \text{for } \mu_-, \\ |B_{\rm H}\rangle &\equiv p|B^0\rangle - q|\bar{B}^0\rangle \quad \text{for } \mu_+, \end{aligned} \tag{1.32}$$

where

$$\frac{q}{p} \equiv \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} \quad \text{and} \quad \sqrt{p^2+q^2} = 1. \tag{1.33}$$

$|B_{\rm L,H}\rangle$ represent the physical mass eigenstates, where L and H denote the light and heavy states, respectively. The mass and decay rates of $|B_{\rm L,H}\rangle$ are

$$\begin{aligned} m_{\rm L} &= \Re(\mu_-), \quad \Gamma_{\rm L} = -2\Im(\mu_-), \\ m_{\rm H} &= \Re(\mu_+), \quad \Gamma_{\rm H} = -2\Im(\mu_+). \end{aligned} \tag{1.34}$$

For convenience, the mass and lifetime differences and the mass and lifetime averages are introduced as

$$\Delta m_d \equiv m_{\rm H} - m_{\rm L}, \quad \Delta\Gamma \equiv \Gamma_{\rm L} - \Gamma_{\rm H}, \quad \frac{m_{\rm H} + m_{\rm L}}{2} = M, \quad \equiv \frac{\Gamma_{\rm H} + \Gamma_{\rm L}}{2} = \Gamma. \tag{1.35}$$

The mass eigenstates have exponential evolution as a function of time,

$$\begin{aligned} |B_{\rm L}(t)\rangle &= e^{i\mu_+ t}|B_{\rm L}\rangle = e^{-im_{\rm L}t}e^{-\Gamma_{\rm L}t/2}(p|B^0\rangle + q|\bar{B}^0\rangle), \\ |B_{\rm H}(t)\rangle &= e^{i\mu_+ t}|B_{\rm H}\rangle = e^{-im_{\rm H}t}e^{-\Gamma_{\rm H}t/2}(p|B^0\rangle - q|\bar{B}^0\rangle), \end{aligned} \tag{1.36}$$

using Eq. (1.34) and Eq. (1.32). The time evolution of the mass eigenstates can also be expressed with the time-dependent form of Eq. (1.32),

$$\begin{aligned} |B_{\rm L}(t)\rangle &= p|B^0(t)\rangle + q|\bar{B}^0(t)\rangle, \\ |B_{\rm H}(t)\rangle &= p|B^0(t)\rangle - q|\bar{B}^0(t)\rangle. \end{aligned} \tag{1.37}$$

With Eq. (1.36), Eq. (1.37) and Eq. (1.35), we solve for the time evolution of B^0 and $\bar{B}^0$,

$$\begin{aligned} |B^0(t)\rangle &= g_+(t)|B^0\rangle + \frac{q}{p}g_-(t)|\bar{B}^0\rangle, \\ |\bar{B}^0(t)\rangle &= g_+(t)|\bar{B}^0\rangle - \frac{p}{q}g_-(t)|B^0\rangle, \end{aligned} \tag{1.38}$$

where

$$g_\pm(t) \equiv \frac{1}{2}e^{iMt}e^{-\Gamma t/2}\left(e^{i\Delta m_d t/2}e^{-\Delta\Gamma t/4} \pm e^{-i\Delta m_d t/2}e^{\Delta\Gamma t/4}\right). \tag{1.39}$$

Since $\Delta\Gamma/\Gamma \sim \mathcal{O}(10^{-3})$ in the B^0 system, the lifetime difference between the light and heavy mass eigenstates can be ignored, i.e. $\Delta\Gamma = 0$. The decay rate is redefined as $\Gamma_{\rm L} = \Gamma_{\rm H} = \Gamma$, and therefore, Eq. (1.39) becomes

$$g_\pm(t) = e^{-\Gamma t/2}\left(\frac{e^{i\Delta m_d t/2} \pm e^{-i\Delta m_d t/2}}{2}\right), \tag{1.40}$$

where e^{iMt} is removed by a phase convention. Thus, the time evolution of the B^0 and $\bar{B}^0$is

$$|B^0(t)\rangle = e^{-\Gamma t/2}\left(\cos\frac{\Delta m_d t}{2}|B^0\rangle + i\frac{q}{p}\sin\frac{\Delta m_d t}{2}|\bar{B}^0\rangle\right),$$
$$|\bar{B}^0(t)\rangle = e^{-\Gamma t/2}\left(\cos\frac{\Delta m_d t}{2}|\bar{B}^0\rangle - i\frac{p}{q}\sin\frac{\Delta m_d t}{2}|B^0\rangle\right). \quad (1.41)$$

1.4.2 Time-Dependent *CP* Violation

Consider the case where both B^0 and $\bar{B}^0$ decay into the same final state, otherwise known as a CP eigenstate, f_{CP}. The decay amplitudes are written as

$$A_{CP} = \langle f_{CP}|\mathbf{H}|B^0\rangle,$$
$$\overline{A}_{CP} = \langle f_{CP}|\mathbf{H}|\bar{B}^0\rangle. \quad (1.42)$$

With the time evolution of the B mesons determined in the previous section as Eq. (1.41), the time-dependent decay amplitudes are

$$A_{CP}(t) = A_{CP}e^{-\Gamma t/2}\left(\cos\frac{\Delta m_d t}{2} + i\lambda_{CP}\sin\frac{\Delta m_d t}{2}\right),$$
$$\overline{A}_{CP}(t) = \overline{A}_{CP}e^{-\Gamma t/2}\left(\cos\frac{\Delta m_d t}{2} - \frac{i}{\lambda_{CP}}\sin\frac{\Delta m_d t}{2}\right), \quad (1.43)$$

where

$$\lambda_{CP} \equiv \frac{q}{p}\frac{\overline{A}_{CP}}{A_{CP}}. \quad (1.44)$$

Recall the box diagram for the $B^0\bar{B}^0$ mixing in Fig. 1.3. The b and d quarks couple to W bosons and u, c, t quarks. However, the intermediate mixing state is dominated by the virtual t quark because of $V_{tb} \gg V_{cb} \gg V_{ub}$ and Glashow, Iliopoulos, and Maiani (GIM) cancellations. Theory indicates that [16, 17, 18, 19],

$$M_{12} \propto (V_{tb}V_{td}^*)^2 m_t^2,$$
$$\Gamma_{12} \propto (V_{tb}V_{td}^*)^2 m_b^2, \quad (1.45)$$

where $m_{t,b}$ is the mass of the t and b quark. If we expand Eq. (1.33),

$$\begin{aligned}
\frac{q}{p} &\simeq \sqrt{\frac{M_{12}^*}{M_{12}}} + \mathcal{O}\left(\frac{\Gamma_{12}}{M_{12}}\right) \\
\Rightarrow \left|\frac{q}{p}\right| &= \left|\frac{V_{td}V_{tb}^*}{V_{td}^*V_{tb}}\right| + \mathcal{O}\left(\frac{m_b^2}{m_t^2}\right), \quad \text{but } m_b^2/m_t^2 \ll 1, \\
\Rightarrow \left|\frac{q}{p}\right| &\simeq 1. \quad (1.46)
\end{aligned}$$

This allows the relation from Eq. (1.44) to also be written as

$$|\lambda_{CP}|^2 = \frac{|\overline{A}_{CP}|^2}{|A_{CP}|^2}, \quad (1.47)$$

and so the time-dependent decay rates are,

$$\begin{aligned}
\Gamma[B^0(t) \to f_{CP}] &= |\langle f_{CP}|B^0(t)\rangle|^2 \\
&= \frac{e^{-\Gamma t}}{2}|A_{CP}|^2\Big[(|\lambda_{CP}|^2+1) - (|\lambda_{CP}|^2-1)\cos\Delta m_d t - 2\Im(\lambda_{CP})\sin\Delta m_d t\Big], \\
\Gamma[\bar{B}^0(t) \to f_{CP}] &= |\langle f_{CP}|\bar{B}^0(t)\rangle|^2 \\
&= \frac{e^{-\Gamma t}}{2}|A_{CP}|^2\Big[(|\lambda_{CP}|^2+1) + (|\lambda_{CP}|^2-1)\cos\Delta m_d t + 2\Im(\lambda_{CP})\sin\Delta m_d t\Big].
\end{aligned} \tag{1.48}$$

The time-dependent CP rate asymmetry is defined as,

$$\begin{aligned}
a_{CP}(t) &\equiv \frac{\Gamma(\bar{B}^0(t) \to f_{CP}) - \Gamma(B^0(t) \to f_{CP})}{\Gamma(\bar{B}^0(t) \to f_{CP}) + \Gamma(B^0(t) \to f_{CP})} \\
&= \frac{(|\lambda_{CP}|^2-1)\cos\Delta m_d t + 2\Im(\lambda_{CP})\sin\Delta m_d t}{|\lambda_{CP}|^2+1} \\
&= \mathcal{A}_{CP}\cos\Delta m_d t + \mathcal{S}_{CP}\sin\Delta m_d t,
\end{aligned} \tag{1.49}$$

where the CP parameters are defined as,

$$\mathcal{A}_{CP} \equiv \frac{|\lambda_{CP}|^2-1}{|\lambda_{CP}|^2+1}, \quad \mathcal{S}_{CP} \equiv \frac{2\Im(\lambda_{CP})}{|\lambda_{CP}|^2+1}. \tag{1.50}$$

The interpretation of $\mathcal{A}_{CP}$ and $\mathcal{S}_{CP}$ is as follows. Recall that the definition of λ_{CP} in Eq. (1.44) contains two distinct parts. The q/p part describes mixing while $\overline{A}_{CP}/A_{CP}$ describes the decay amplitudes. This implies that λ_{CP} can be written as

$$\lambda_{CP} \propto e^{-i\phi_{\rm M}}\frac{|\overline{A}_{CP}|}{|A_{CP}|}e^{-i\phi_{\rm D}} = \frac{|\overline{A}_{CP}|}{|A_{CP}|}e^{-i(\phi_{\rm M}+\phi_{\rm D})}, \tag{1.51}$$

where $\phi_{\rm M}$ represents a weak phase difference in the $B^0\bar{B}^0$ mixing and $\phi_{\rm D}$ represents a weak phase difference in the decay amplitudes.

Consider the case where there is a difference in the decay rates, that is, $\Gamma(B^0 \to f) \neq \Gamma(\bar{B}^0 \to f)$. Then, $|\overline{A}_{CP}|^2/|A_{CP}|^2 \neq 1 \Rightarrow \mathcal{A}_{CP} \neq 0$. This is called direct CP violation.

Now consider the case where there is interference between the mixing and decay phase differences, $\phi_{\rm M} + \phi_{\rm D} \neq 0$. Thus, $\Im(\lambda_{CP}) \neq 0 \Rightarrow \mathcal{S}_{CP} \neq 0$. This is called mixing-induced CP violation.

1.4.3 Coherent $B^0\bar{B}^0$ Mixing

So far, the time-dependent decay rate asymmetry in Eq. (1.49) requires that the flavour of the B meson is known at the time of production. For a CP eigenstate where both B^0 and

$\bar{B}^0$ may decay to the same final state, and may also have mixed in the decay time, t, after production, flavour identification is not possible experimentally.

Now, B mesons produced in the decay, $\Upsilon(4S) \to B^0\ \bar{B}^0$, inherit the quantum numbers, $J^{PC} = 1^{--}$, because they are conserved in the strong interaction. Since B^0 is a pseudoscalar particle, the $B^0\bar{B}^0$ system should have an orbital angular momentum of $L = 1$, also known as a p-wave configuration. The state must also be antisymmetric. Therefore, by Bose-Einstein statistics, the states $B^0 - B^0$ or $\bar{B}^0 - \bar{B}^0$ are forbidden and thus the $B^0\bar{B}^0$ pair must oscillate coherently,

$$|B_1(t_1), B_2(t_2)\rangle = \frac{1}{\sqrt{2}}\bigg(|B_1^0(t_1)\rangle|\bar{B}_2^0(t_2)\rangle - |\bar{B}_1^0(t_1)\rangle|B_2^0(t_2)\rangle\bigg). \tag{1.52}$$

By substitution of Eq. (1.41), the time evolution of the $B^0\bar{B}^0$ system becomes,

$$\begin{aligned}|B_1(t_1), B_2(t_2)\rangle = \frac{1}{\sqrt{2}}e^{-\Gamma(t_1+t_2)/2}\Bigg[&\cos\frac{\Delta m_d\Delta t}{2}\bigg(|B_1^0\rangle|\bar{B}_2^0\rangle - |\bar{B}_1^0\rangle|B_2^0\rangle\bigg)\\ &+i\sin\frac{\Delta m_d\Delta t}{2}\bigg(\frac{q}{p}|B_1^0\rangle|\bar{B}_2^0\rangle - \frac{p}{q}|\bar{B}_1^0\rangle|B_2^0\rangle\bigg)\Bigg],\end{aligned} \tag{1.53}$$

where $\Delta t = t_2 - t_1$. Let t_i be the decay time of each B meson. If we can measure Δt and determine the flavour of one B meson, then the flavour and time evolution of the accompanying B meson is known. Take again the specific case where one B decays to a CP eigenstate, B_{CP}, and the other B meson decays to a flavour specific state from which the flavour can be determined, $B_{\rm Tag}$. A schematic of this scenario is shown in Fig. 1.4. Then, the time-dependent decay rates become

$$\begin{aligned}\Gamma(f_{CP}, f_{\rm Tag}) &= |\langle f_{CP}, f_{\rm Tag}|B_{CP}(t_{CP}), B^0_{\rm Tag}(t_{\rm Tag})\rangle|^2\\ &= \frac{1}{4}e^{-\Gamma(t_{CP}+t_{\rm Tag})}A_{CP}^2A_{\rm Tag}^2\\ &\quad\bigg[(|\lambda_{CP}|^2+1) - (|\lambda_{CP}|^2-1)\cos\Delta m_d\Delta t - 2\Im(\lambda_{CP})\sin\Delta m_d\Delta t\bigg],\\ \Gamma(f_{CP}, \overline{f}_{\rm Tag}) &= |\langle f_{CP}, \overline{f}_{\rm Tag}|B_{CP}(t_{CP}), \bar{B}^0_{\rm Tag}(t_{\rm Tag})\rangle|^2\\ &= \frac{1}{4}e^{-\Gamma(t_{CP}+t_{\rm Tag})}A_{CP}^2A_{\rm Tag}^2\left|\frac{p}{q}\right|^2\\ &\quad\bigg[(|\lambda_{CP}|^2+1) + (|\lambda_{CP}|^2-1)\cos\Delta m_d\Delta t + 2\Im(\lambda_{CP})\sin\Delta m_d\Delta t\bigg],\end{aligned} \tag{1.54}$$

where the decay amplitude, $A_{\rm Tag} \equiv \langle f_{\rm Tag}|B^0\rangle = \langle\overline{f}_{\rm Tag}|\bar{B}^0\rangle$ and λ_{CP} retain their usual definitions. It is interesting to see that the resulting time-dependent CP asymmetry,

$$\begin{aligned}a_{CP}(\Delta t) &\equiv \frac{\Gamma(f_{CP}, \overline{f}_{\rm Tag}) - \Gamma(f_{CP}, f_{\rm Tag})}{\Gamma(f_{CP}, \overline{f}_{\rm Tag}) + \Gamma(f_{CP}, f_{\rm Tag})}\\ &= \mathcal{A}_{CP}\cos\Delta m_d\Delta t + \mathcal{S}_{CP}\sin\Delta m_d\Delta t,\end{aligned} \tag{1.55}$$

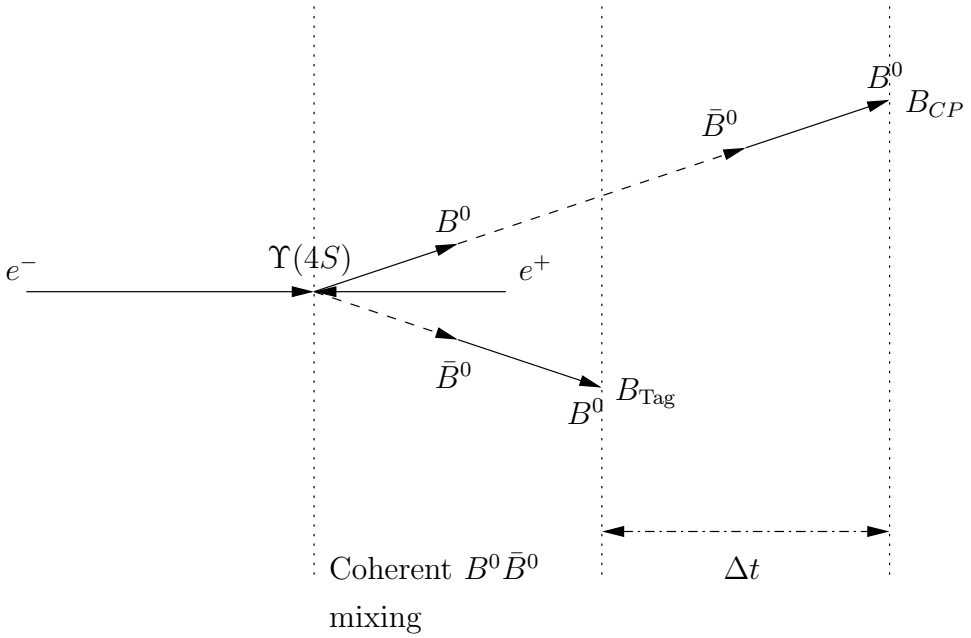


Figure 1.4: Scenario in which two B mesons are produced in a coherent state from the $\Upsilon(4S)$ decay. One B meson decays to a CP eigenstate, B_{CP}, while the other B meson decays to a flavour specific state, $B_{\rm Tag}$.

retains its form in Δt, when compared to Eq. (1.49). Thus, CP violation parameters can be extracted as a physical observable if Δt and the flavour of $B_{\rm Tag}$ can be determined. Notice how the right hand of Eq. (1.55) resembles the equation for a circle while the asymmetry of left hand side may only take on values between $-1 \leq a_{CP}(\Delta t) \leq +1$ by definition. So, $(\mathcal{A}_{CP}, \mathcal{S}_{CP})$ is physically restricted to lie within the unit circle in CP violation space.

Now, if Eq. (1.54) is integrated over the unmeasurable $t_{CP} + t_{\rm Tag}$ direction and normalised in the region, $-\infty < \Delta t < +\infty$, we obtain the probability of finding f_{CP} at time, Δt, for a given flavour tag, q,

$$\mathcal{P}(\Delta t, q) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}} \left[1 + q\left(\mathcal{A}_{CP}\cos\Delta m_d \Delta t + \mathcal{S}_{CP}\sin\Delta m_d \Delta t\right)\right], \tag{1.56}$$

where $q = +1(-1)$ for $B_{\rm Tag} = B^0(\bar{B}^0)$. Using this probability density function, Fig. 1.5 demonstrates the manifestation of different types of CP violation in Δt.

1.5 CP Violation in $b \to c\bar{c}s$ Transitions

CP violation has already been confirmed in the B meson sector and was first observed in $b \to c\overline{c}s$ transitions such as $B^0 \to J/\psi K^0_S$ [10, 11]. The Feynman diagrams of the leading order processes that contribute to the amplitude are shown in Fig. 1.6. The first order tree process proceeds by internal emission of a spectator W boson. The internal vertex cancels

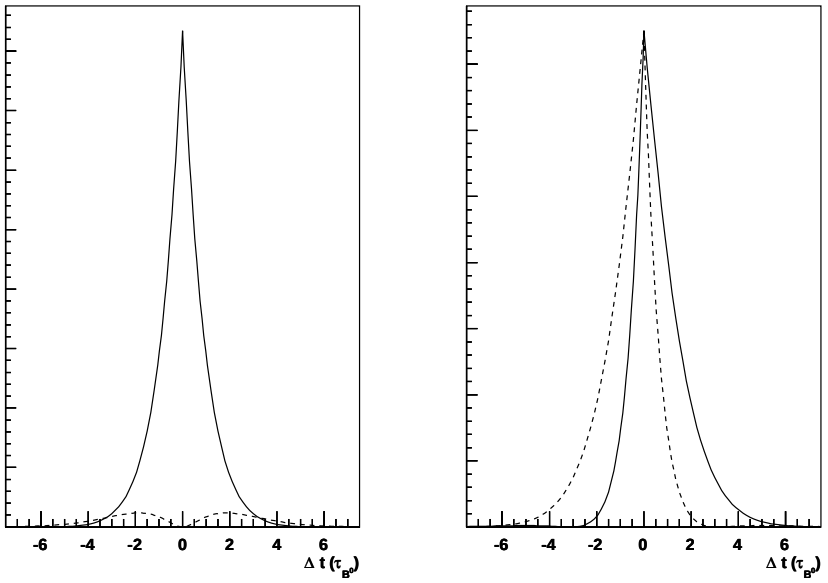


Figure 1.5: CP violation effects on Δt in units of the B lifetime on an arbitrary scale. The left plot shows $(\mathcal{A}_{CP}, \mathcal{S}_{CP}) = (1, 0)$ and the right shows $(\mathcal{A}_{CP}, \mathcal{S}_{CP}) = (0, 1)$. The solid curve represents $q = +1$ ($B_{\rm Tag} = B^0$) and the dashed curve represents $q = -1$ ($B_{\rm Tag} = \bar{B}^0$).

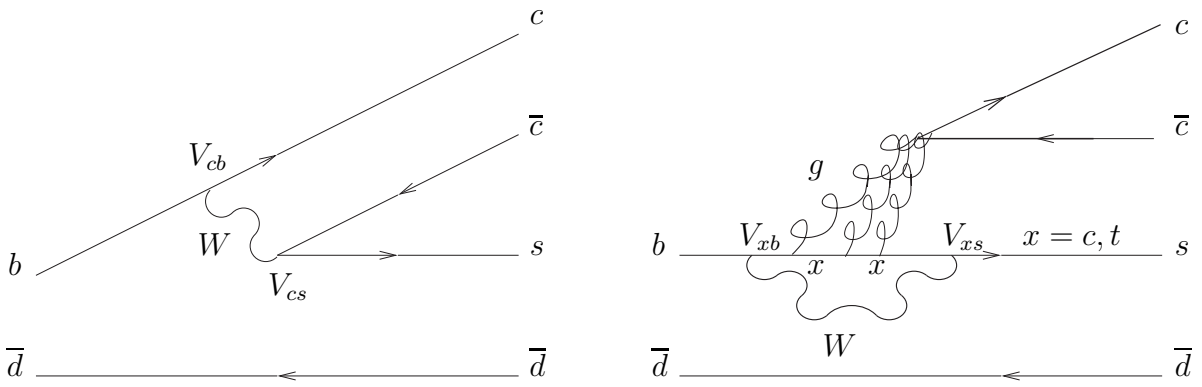


Figure 1.6: Feynman diagram for tree (left) and penguin (right) amplitudes in the $B^0 \to J/\psi K^0_S$ decay.

the colour of the spectator and so the decay is colour suppressed which has the effect of reducing the branching fraction.

The second order penguin process contains loops with virtual t and c quarks that have the CKM elements of the same order (Eq. (1.23)) as the tree diagram, so are not suppressed relative to the tree. The suppression in this process relative to the tree diagram arises from generating the massive $c\overline{c}$ colour singlet from a gluon to produce a charmonium. Loops containing the virtual u quark would receive an $\mathcal{O}(\lambda^2)$ (Eq. (1.24)) suppression. Very little penguin pollution is thus expected for $b \to c\bar{c}s$ transitions. This means that only one weak phase will contribute to CP violation, also implying that no direct CP violation is expected. For $b \to c\bar{c}s$ eigenstates, we obtain the CP violating parameter,

$$\lambda_{CP} = \left(\frac{q}{p}\right)\left(\frac{\overline{A}_{CP}}{A_{CP}}\right)\left(\frac{q}{p}\right)_{K^0} \tag{1.57}$$

where Eq. (1.44) has been adapted to include the phase from neutral kaon mixing as there is a K^0 in the final state. Neglecting the penguin contribution, Eq. (1.46) and the box diagram analogous to Fig. 1.3 for $K^0 - \overline{K}^0$ mixing, can be used to express λ_{CP} in terms of CKM elements,

$$\begin{aligned} \lambda_{CP} &= \frac{V_{td}V_{tb}^*}{V_{td}^*V_{tb}}\,\frac{V_{cb}V_{cs}^*}{V_{cb}^*V_{cs}}\,\frac{V_{cs}V_{cd}^*}{V_{cs}^*V_{cd}} \\ &= \eta_{CP}e^{-i2\phi_1}, \end{aligned} \tag{1.58}$$

where η_{CP} is the CP eigenvalue of the $b \to c\bar{c}s$ transition and ϕ_1 (Eq. (1.27)) is an internal angle in the CKM unitarity triangle. Therefore, we obtain

$$\mathcal{S}_{CP} = -\eta_{CP}\sin 2\phi_1, \tag{1.59}$$

if the penguin contribution is ignored. The current world average for $\sin 2\phi_1$ from all $b \to c\bar{c}s$ transitions is,

$$\sin 2\phi_1 = 0.672 \pm 0.024 \text{ [20]}, \tag{1.60}$$

which provides a tight constraint on other CP violating parameters as can be seen in Fig. 1.7 [21]. There is a two-fold ($\phi_1 \leftrightarrow \pi/2 - \phi_1$) ambiguity associated with $\sin 2\phi_1$. The solution preferred from other measurements [22, 23, 24, 25, 26, 27] is shown in Fig. 1.7.

1.6 CP Violation in $b \to q\bar{q}s$ Transitions

In contrast to $b \to c\bar{c}s$ transitions that are induced by the charged current, the neutral current is flavour-conserving, which is ensured by the unitarity of the CKM matrix. Thus, Flavor Changing Neutral Currents (FCNC), such as those found in $b \to q\bar{q}s$ transitions, proceed via second order decay processes that can be represented by a loop diagram.

The total amplitude for second order processes is

$$A_{CP} = \sum_{q=u,c,t}(V_{qb}^*V_{qs})P_{CP}^q, \tag{1.61}$$

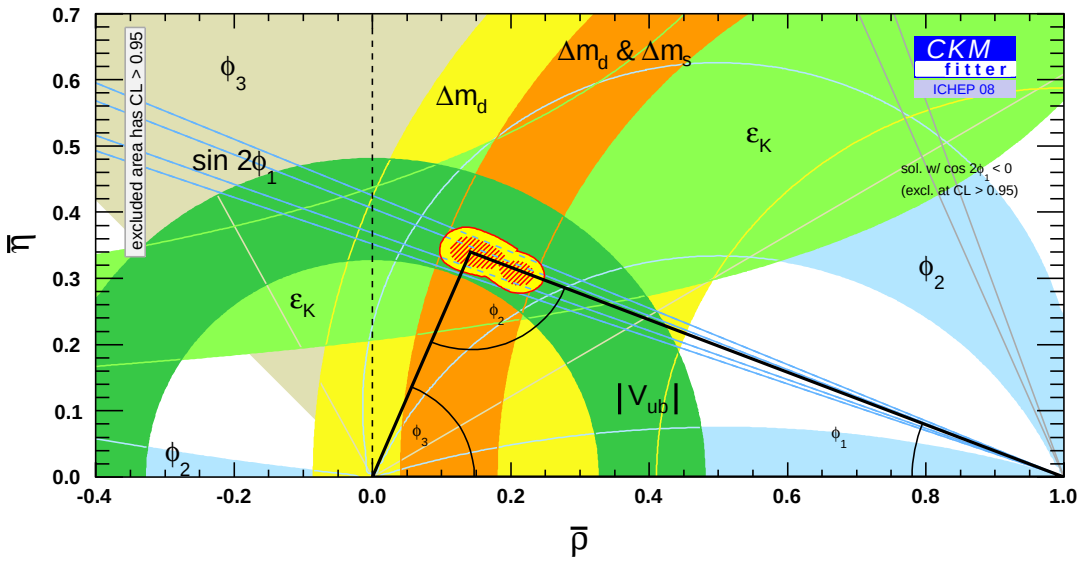


Figure 1.7: Current experimental constraints to the CKM unitarity triangle [19]. Previously unmentioned parameters contained in this figure include the CP violation parameter in the neutral kaon system, ϵ_K, and the mass difference between the B_s mass eigenstates, Δm_s.

where P^q_{CP} are the penguin amplitudes for the decay. Using the unitary constraint, $\sum_{q=u,c,t}(V^*_{qb}V_{qs}) = 0$, we obtain

$$A_{CP} = (V^*_{cb}V_{cs})(P^c_{CP} - P^t_{CP}) + (V^*_{ub}V_{us})(P^u_{CP} - P^t_{CP}). \tag{1.62}$$

Since $V^*_{ub}V_{us} \ll V^*_{cb}V_{cs}$, the ratio of amplitudes for penguin decays is given by

$$\frac{\bar{A}_{CP}}{A_{CP}} \sim \frac{V_{cb}V^*_{cs}}{V^*_{cb}V_{cs}}, \tag{1.63}$$

which is the same as that of $b \to c\bar{c}s$ induced decays ignoring higher-order effects.

Considering the smaller effects, the CP asymmetries in $b \to q\bar{q}s$ transitions are expected in the SM to be slightly higher than those observed in $b \to c\bar{c}s$ transitions [28, 29, 30, 31, 32, 33, 34, 35, 36]. However, current experimental measurements shown in Fig. 1.8 tend to be lower than those for $b \to c\bar{c}s$ transitions motivating more precise experimental determinations.

Within the loop diagram, a heavy particle unknown in the SM may appear due to the uncertainty principle. This amplitude may interfere with the SM amplitude. If the CP phases of these decay amplitudes are different, the measured CP asymmetry may contain a sizable contribution from the new weak phase. In some models, the value of $\sin 2\phi_1$ may differ from the SM expectation by more than ten percent and thus, $b \to s$ modes are an excellent probe for New Physics (NP). Taking into account NP effects, the CP phase is sometimes written as $\phi_1^{\rm eff}$. Figure 1.9 shows the loop diagram for $b \to q\bar{q}s$ transitions and where NP may enter the loop.

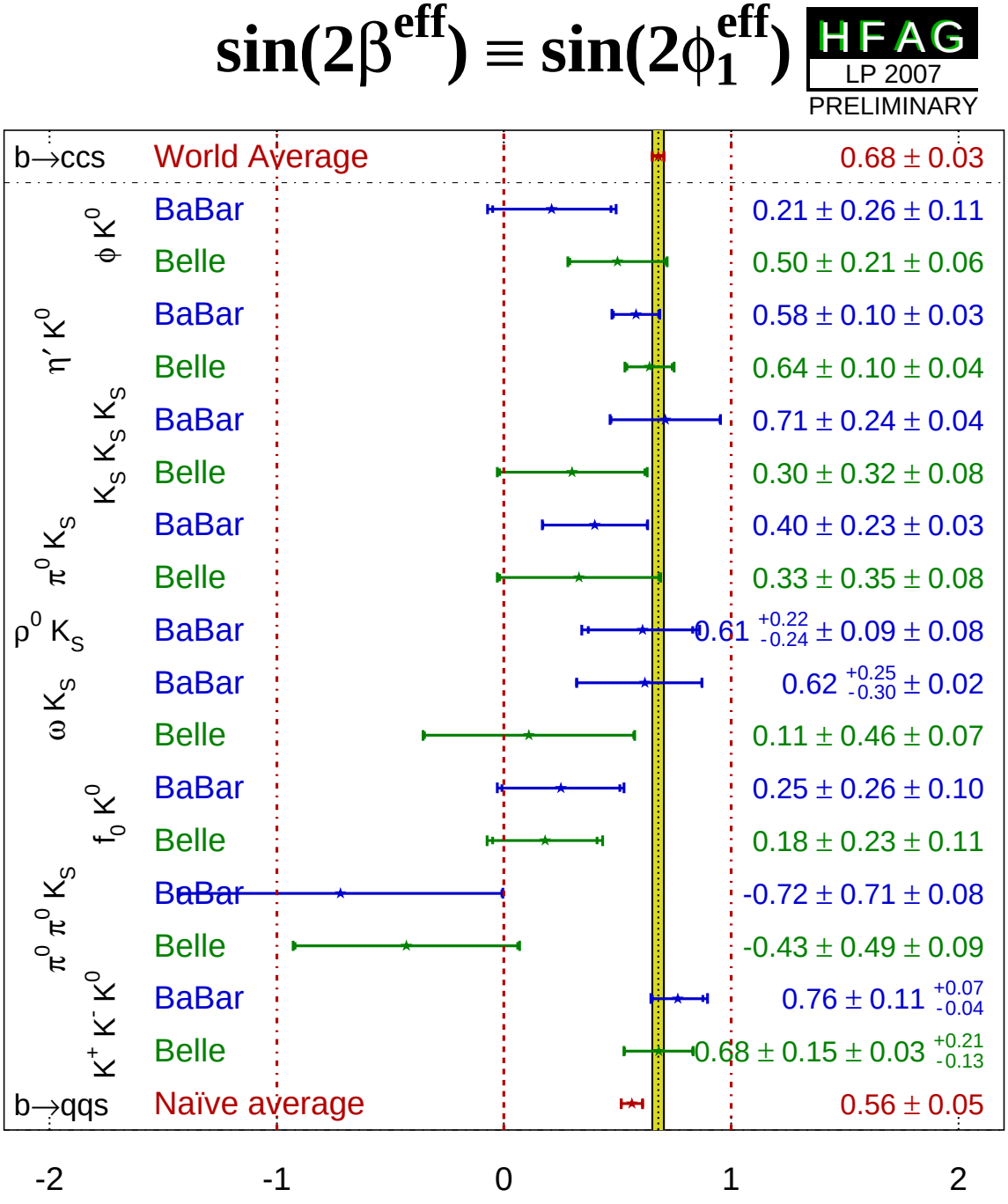


Figure 1.8: Summary of $\mathcal{S}_{CP}^{\text{eff}}$ for all $b \to s$ modes at the end of 2007.

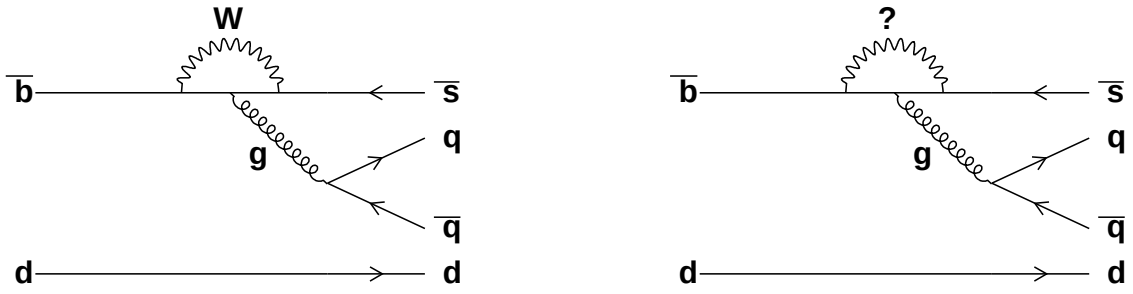


Figure 1.9: Penguin diagram of $b \to q\bar{q}s$ transitions including the SM phase (left) and NP phase (right).

1.7 CP Violation in $B^0 \to K^0\pi^0$ Decays

A recent unexpected result was that direct CP violation in the flavour-specific, $B^0 \to K^+\pi^-$ differs from direct CP violation in $B^+ \to K^+\pi^0$ [37, 38]. The result is against the naive expectation from the spectator model that they should be equal. There were several proposals that suggested this was due to new physics effects. However, such arguments suffered from hadronic uncertainties. Recently authors in Ref. [39] proposed to test the following sum rule that is held more precisely within the SM,

$$\mathcal{A}(K^+\pi^-)+\mathcal{A}(K^0\pi^+)\frac{\mathcal{B}(K^0\pi^+)\tau_{B^0}}{\mathcal{B}(K^+\pi^-)\tau_{B^+}} = \mathcal{A}(K^+\pi^0)\frac{2\mathcal{B}(K^+\pi^0)\tau_{B^0}}{\mathcal{B}(K^+\pi^-)\tau_{B^+}}+\mathcal{A}(K^0\pi^0)\frac{2\mathcal{B}(K^0\pi^0)}{\mathcal{B}(K^+\pi^-)}, \quad (1.64)$$

where τ_{B^0} and τ_{B^+} are the B^0 and B^+ lifetimes, respectively, and $\mathcal{B}$ represents the branching fraction. Leading terms are identical in the isospin symmetry limit, while sub-leading terms are equal in the flavour $SU(3)$ and heavy quark limits.

As $\mathcal{A}(K^0\pi^0)$ is experimentally the least known quantity, the sum rule can give a precise prediction of its expected value, $\mathcal{A}(K^0\pi^0) = -0.148 \pm 0.044$ [40]. The SM expectation of $\mathcal{A}(K^0\pi^0)$ can be expressed graphically (Fig. 1.10) as a function of $\mathcal{A}(K^+\pi^-) - \mathcal{A}(K^+\pi^0)$ since the experimental value of $\mathcal{A}(K^0\pi^+)$ is consistent with zero. A violation of the sum rule would be clear evidence for new physics in $b \to q\bar{q}s$ transitions. Therefore, a precise measurement of $\mathcal{A}(K^0\pi^0)$ is a powerful method to address the $K-\pi$ puzzle mentioned above, using the sum rule.

The measurement of $\mathcal{S}(K^0\pi^0)$ is also as useful as other $\mathcal{S}$ measurements in $b \to s$ penguin modes. The shift in $\mathcal{S}(K^0\pi^0)$ can be evaluated in the $1/m_b$ expansion and/or using SU(3) flavour symmetry to deviate upwards from the SM expectation by a few percent [41, 42].

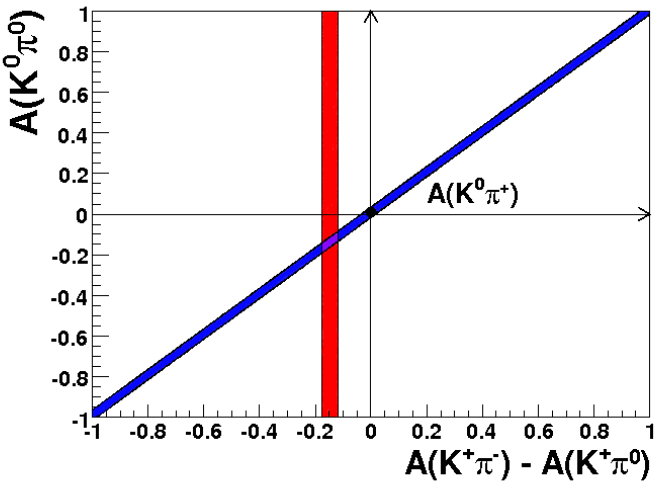


Figure 1.10: The blue region shows the sum rule expectation when fixing $\mathcal{A}(K^0\pi^+)$ as the y-intercept shown as the black dot. The red region shows the current experimental range of $\mathcal{A}(K^+\pi^-) - \mathcal{A}(K^+\pi^0)$. The resulting overlapping region shown in purple gives the expected range for $\mathcal{A}(K^0\pi^0)$.

1.8 Previous Measurements

The decay mode, $B^0 \to K^0\pi^0$, has previously been studied by the BaBar and Belle Collaborations. BaBar measures the branching fraction,

$$\mathcal{B}(B^0 \to K^0\pi^0) = (10.1 \pm 0.6\,(\text{stat}) \pm 0.4\,(\text{syst})) \times 10^{-6}, \tag{1.65}$$

from 467 million $B\bar{B}$ pairs [43] while Belle measures the branching fraction,

$$\mathcal{B}(B^0 \to K^0\pi^0) = (9.2 \pm 0.7\,(\text{stat})\,^{+0.6}_{-0.7}\,(\text{syst})) \times 10^{-6}, \tag{1.66}$$

from 449 million $B\bar{B}$ pairs [44]. With their same sample, BaBar measures the CP parameters [45],

$$\begin{aligned} \mathcal{A}_{CP} &= -0.13 \pm 0.13\ (\text{stat}) \pm 0.03\ (\text{syst}), \\ \mathcal{S}^{\text{eff}}_{CP} &= +0.55 \pm 0.20\ (\text{stat}) \pm 0.03\ (\text{syst}). \end{aligned} \tag{1.67}$$

while Belle measures,

$$\begin{aligned} \mathcal{A}_{CP} &= -0.05 \pm 0.14\ (\text{stat}) \pm 0.05\ (\text{syst}), \\ \mathcal{S}^{\text{eff}}_{CP} &= +0.33 \pm 0.35\ (\text{stat}) \pm 0.08\ (\text{syst}), \end{aligned} \tag{1.68}$$

with an increased data sample of 535 million $B\bar{B}$ pairs [46].

1.9 Thesis Outline

The goal of this thesis is to present a measurement of CP asymmetries in the $B^0 \to K^0\pi^0$ decay.

We have introduced the theoretical framework and physics motivation in this chapter. In Chapter 2, we describe the experimental apparatus of the KEK B-factory including the accelerator and each component of the Belle detector. The event selection procedure is explained in Chapter 3. The measurement of the branching fraction of $B^0 \to K^0\pi^0$ decays are described in Chapter 4. In Chapter 5, we explain in detail the method of extracting the CP asymmetries and show the results. Chapter 6 is devoted to the discussion of our results. Finally, the conclusion of this thesis is given in Chapter 7.

Chapter 2

The Belle Experiment

2.1 Overview

The Belle experiment is one of two dedicated B physics experiments in the world and is located at the High Energy Accelerator Research Organization (KEK), Tsukuba, Japan. The BaBar experiment is performed at the Stanford Linear Accelerator Center (SLAC), California, USA. The B-factory at KEK consists of two separate entities, the KEKB e^+e^- collider and the Belle detector. We describe both of them in this chapter2.

2.2 The KEKB Collider

The KEKB accelerator [47], which was commissioned in December 1998, is designed to produce large numbers of $B\bar{B}$ pairs by colliding electrons and positrons of asymmetric energy. It consists of two storage rings, an 8 GeV electron high energy ring (HER) and a 3.5 GeV positron low energy ring (LER). The two storage rings are side by side, 11 m underground and have a circumference of roughly 3 km. There is only one crossing point where the e^+e^- collision takes place, known as the interaction point (IP). A linear accelerator (Linac) accelerates the electrons and positrons to their required energies and inject them into their respective storage ring. Figure 2.1 shows a schematic of the KEKB collider.

Electrons and positrons are kept in bunches around the storage ring. There are approximately 1000 bunches in each ring which corresponds to a bunch separation of around 3.0 m. The electrons and positrons collide with a finite crossing angle of ± 11 mrad to avoid the parasitic interactions of bunches near but not at the IP. A depiction of the finite crossing angle is shown in Fig. 2.2.

The centre-of-mass energy at the collision point is $\sqrt{s} = 10.58$ GeV, which coincides with the mass of the $\Upsilon(4S)$ resonance. This is a convenient choice because the $\Upsilon(4S)$ decays to a

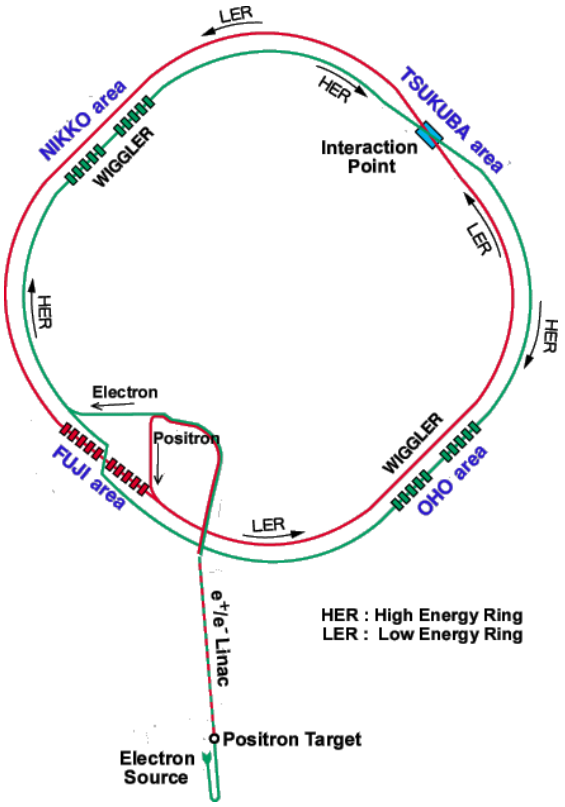


Figure 2.1: The KEKB collider.

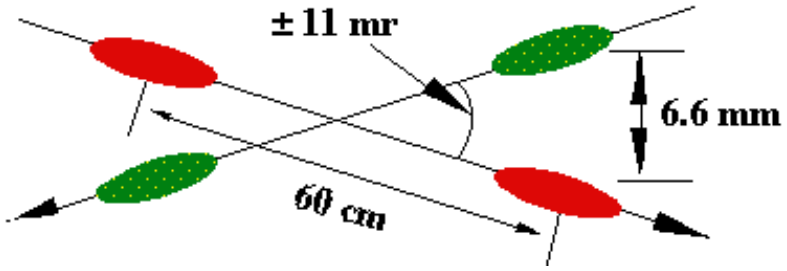


Figure 2.2: The finite crossing angle.

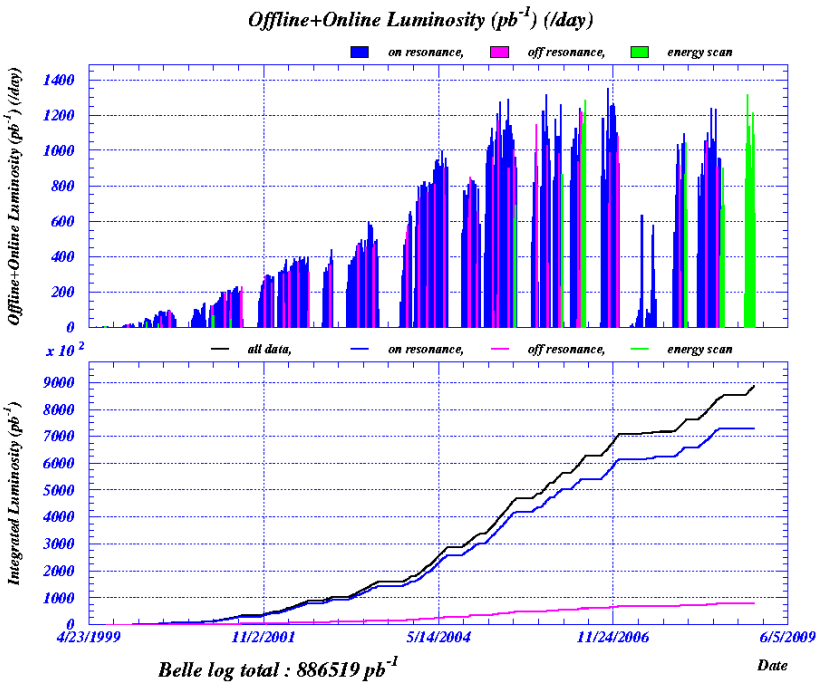


Figure 2.3: The KEKB luminosity history.

$B\bar{B}$ pair more than 96% of the time [48]. The mass of the B meson is almost half the $\Upsilon(4S)$ mass, meaning that it will be produced roughly at rest in the $\Upsilon(4S)$ rest frame.

The advantage of the asymmetric collider is that the B mesons receive a Lorentz boost, $\beta\gamma = 0.425$, in the laboratory frame along the direction of the beam line, thus simplifying the kinematics of the $B\bar{B}$ pair to one-dimension. The boost also gives the B mesons a mean flight length of about 200 μm in the laboratory frame making it feasible with current technology to measure a separation between the decay points of the $B\bar{B}$ pair. This allows the time difference between B mesons decays to be calculated, $\Delta t \sim \Delta z/\beta\gamma c$, which was shown in §1.4.3 to be useful in extracting time-dependent CP information.

The design luminosity of the KEKB collider was $\mathcal{L} = 1.0 \times 10^{34}$ cm^{-2}s^{-1}, which corresponds to an approximate production rate of 10 $B\bar{B}$ pairs per second. As of November 2006, we have well exceeded this design luminosity with the world record of $\mathcal{L} = 1.7118 \times 10^{34}$ cm^{-2}s^{-1}, and our accumulated luminosity is now greater than 800 fb^{-1} and the number of $B\bar{B}$ pairs recorded at Belle has exceeded 9×10^8. The KEKB luminosity history is shown in Fig. 2.3.

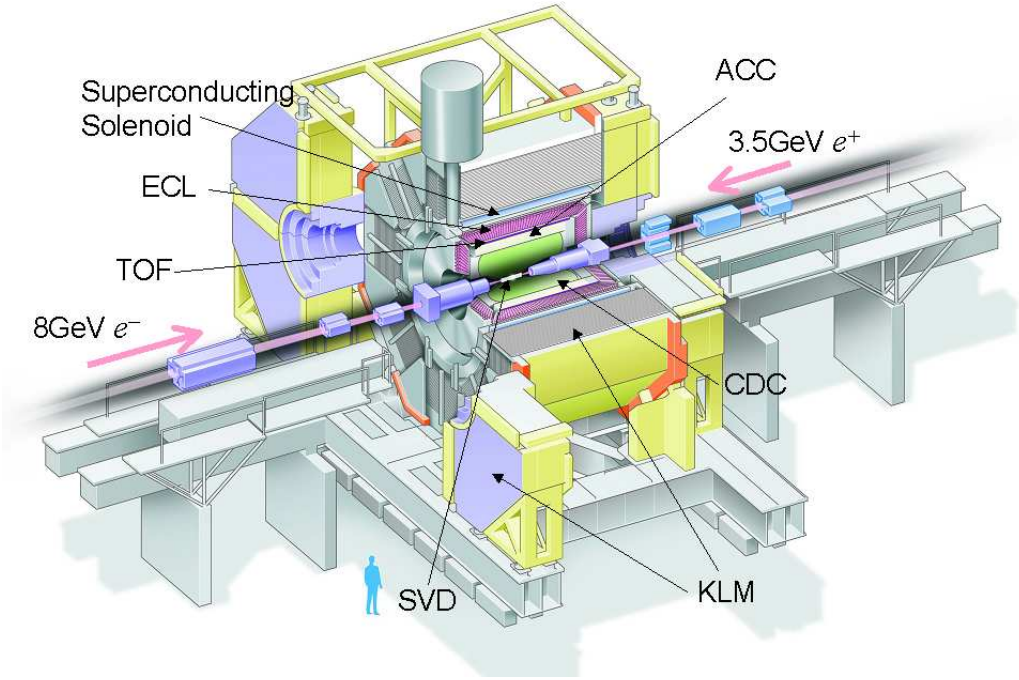


Figure 2.4: Cut-away view of the Belle detector.

2.3 The Belle Detector

The Belle detector is a large solid-angle magnetic spectrometer that can detect the following final state particles,

Charged particles: $K^{\pm}$, $\pi^{\pm}$, $e^{\pm}$, $p^{\pm}$, $\mu^{\pm}$

Neutral particles: γ, K_L^0.

Since a B meson will ultimately decay to some combination of these final state particles most of the time, it is important that these particles can be detected and identified with high efficiency. For a time-dependent CP analysis, it is also imperative that the tracks they leave can be used to measure a vertex position with great precision.

The Belle detector was designed to achieve this task with many sub-detectors as shown in Fig. 2.4, and is placed asymmetrically around the IP because of the direction of the boost which is better seen in Fig. 2.5. The following sections describe the various sub-detectors and further details can be found in Ref. [49].

Now we describe the Belle coordinate systems used in this thesis. In Cartesian coordinates, $\hat{z}$ is defined as being the direction opposite the positron beam line, $\hat{y}$ is vertically upwards,

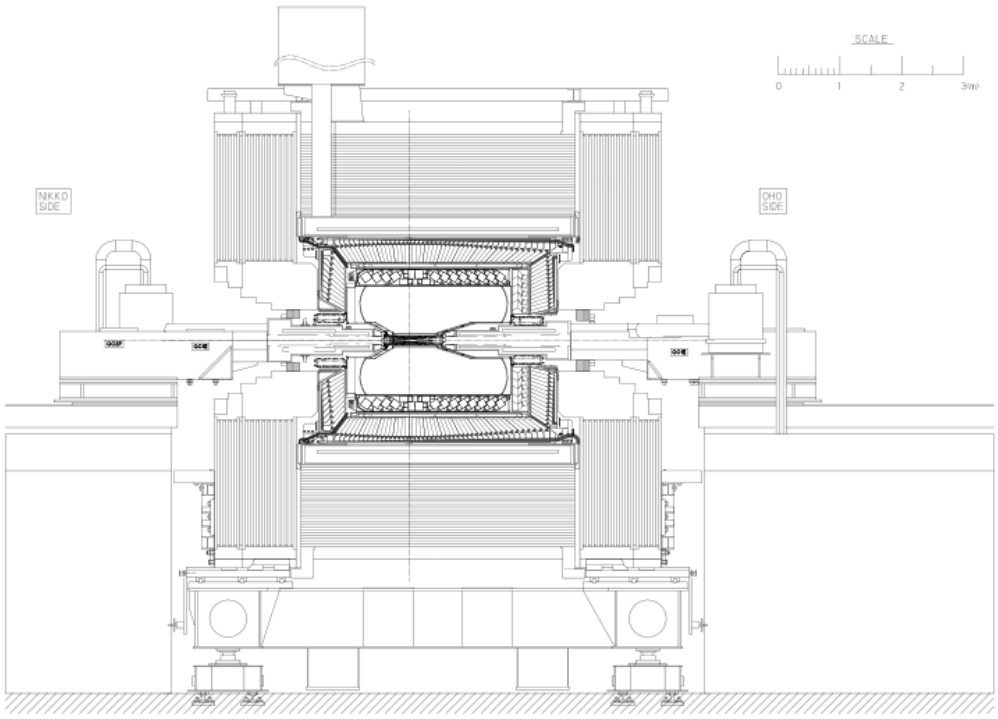


Figure 2.5: Side view of the Belle detector.

and $\hat{x}$ is in the direction of the cross product, $\hat{x} = \hat{y} \times \hat{z}$. Cylindrical coordinates are also used. The radius, $r = \sqrt{x^2 + y^2}$, is defined in the $x - y$ plane, θ is the polar angle from the z-axis, and ϕ is the azimuthal angle around the z-axis. This definition of the coordinate system is illustrated in Fig. 2.6.

2.3.1 Beampipe

Although not part of the Belle detector, the beampipe around the IP is the first piece of material through which particles must traverse before reaching the detector. Since Coulomb

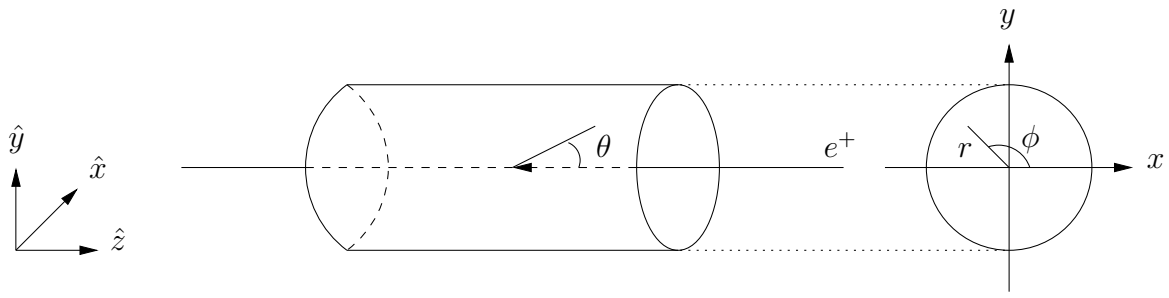


Figure 2.6: Coordinate systems of the Belle detector.

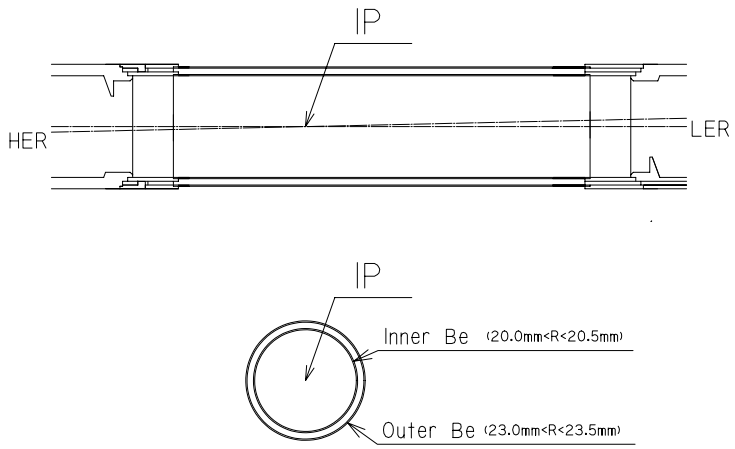


Figure 2.7: Side view and cross section of the beampipe used for accumulation the first 152×10^6 $B\bar{B}$ pairs.

scattering affects track resolution, it is important to minimise the impact of the beampipe on particle trajectories by choosing a thin material with low atomic number. In addition, track quality is reduced as the detector is further removed from the IP. So the diameter of the beampipe must also be as small as possible.

With these considerations in mind, a beryllium beampipe was chosen. The beampipe is a dual layer cylinder whose radii are 20.0 mm and 23.0 mm respectively, where each cylinder has a 0.5 mm thickness. The 2.5 mm gap between the cylinders is used as a helium gas channel which acts as a coolant. This prevents overheating induced by the beam and minimally interferes with tracks due to its low Z number. The outer layer is covered with a 20 μm thick gold sheet to reduce low energy X-ray background from the HER. The total thickness of the beampipe corresponds to 0.9% of a radiation length.

After the first 152×10^6 $B\bar{B}$ pairs were accumulated, the innermost detector was redesigned, placing it closer to the IP. The radius of the inner cylinder of the current beampipe is now 15.0 mm. Fig. 2.7 shows the design of the original beampipe.

2.3.2 Silicon Vertex Detector

The Silicon Vertex Detector (SVD) is the most important detector in the time-dependent CP analysis as it is responsible for precise measurement of B vertex positions. This is achieved via the high spatial resolution of the SVD.

The SVD is the innermost detector and consists of three concentric cylindrical layers of silicon sensors which cover the polar angle, $23° < \theta < 139°$. This corresponds to 86% of the solid angle. The radii of the innermost, middle, and outermost layers are 30.0, 45.5 and 60.5

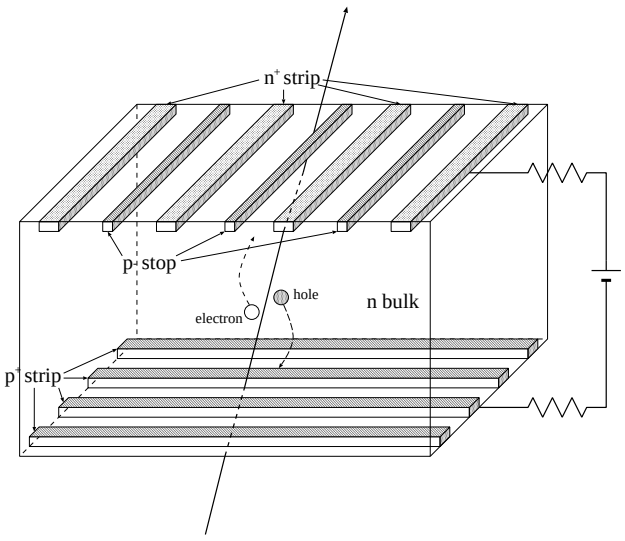


Figure 2.8: DSSD schematic.

mm, consisting of 8, 10 and 14 ladders, respectively.

Each ladder is made up of two half-ladders that are joined by a support structure but are electrically independent of each other. A long half-ladder contains two double sided silicon strip detectors (DSSD) and a hybrid unit which processes signals from the DSSD. A short half-ladder contains a DSSD and a hybrid unit. The innermost layer consists of two short half-ladders, the middle layer consists of a short and long half-ladder and the outermost-layer ladder consists of two long half-ladders.

The DSSD, manufactured by Hamamatsu Photonics, was originally developed for the DELPHI detector [50]. Its dimensions are 57.5×33.5 mm^2, with a thickness of 300 μm. Each DSSD consists of 1280 sense strips and 640 readout pads on both sides. One side of the DSSD (n-side) has its n^+ sense strips, each separated by 42 μm, oriented perpendicular to the beam direction to measure the z coordinate. The p^+ sense strips, each separated by 25 μm, on the other side (p-side) are oriented longitudinally which allows the $r - \phi$ position to be measured.

A DSSD is basically a pn junction. A bias of 75V is supplied to the n-side, while the p-side is grounded. The n^+ strips are interleaved by p^+ implants (p-stops) to electrically separate the consecutive strips. A charged particle passing through the n bulk silicon creates electron-hole pairs. The electrons and holes drift to their corresponding biased side of the DSSD potentially making a 2-dimensional hit signal. On the n-side, adjacent strips are read out by a single channel, giving an effective strip separation of 84 μm. On the p-side, every other strip is connected to a readout channel. Charge collected by the floating strips in between is read from adjacent strips by capacitive charge division. A schematic of the DSSD is shown in Fig. 2.8.

As mentioned in the previous section, §2.3.1, a redesigned SVD was installed after 152×10^6 $B\bar{B}$ pairs were accumulated. The SVD already described is called SVD1 and was replaced

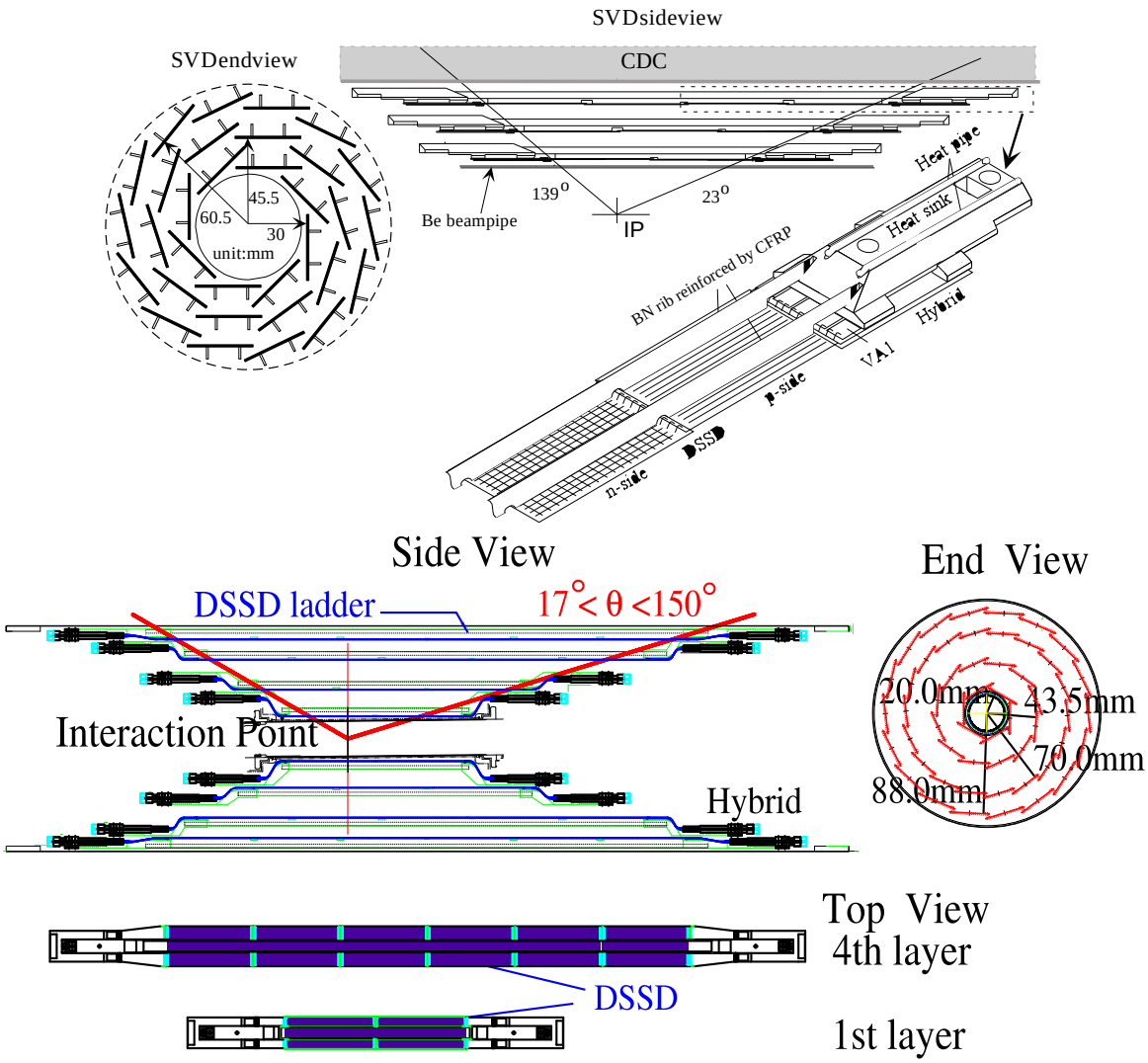


Figure 2.9: The top (bottom) schematic shows the SVD1 (SVD2) geometry.

with the 4-layer SVD2. Among the several improvements were that the angular acceptance of SVD2 was increased to $17° < \theta < 150°$, and the innermost layer was moved 1.0 cm closer to the IP to a radius of 2.0 cm. The fourth layer is accommodated by redesigning the inner region of the CDC. Fig. 2.9 shows the geometry of the SVD and Table 2.1 shows the number of ladders in each layer and the number of DSSDs in each half-ladder.

We estimate the performance of the SVD with two quantities. One is the SVD-CDC track matching efficiency which is defined as the probability that a CDC track passing through the SVD acceptance has associated SVD hits in at least two layers, and at least one 2-D hit. The CDC is the next innermost detector after the SVD and will be described later. The average matching efficiency is better than 98.7%.

The second estimate of SVD performance is the impact parameter resolution of tracks with associated SVD hits. The impact parameter is the distance of closest approach to the IP

Layer	# Ladders	# DSSDs in forward/backward half-ladder
1	6	1/1
2	12	1/2
3	18	2/3
4	18	3/3

Table 2.1: Number of ladders in each layer and number of DSSDs in each half-ladder of SVD2.

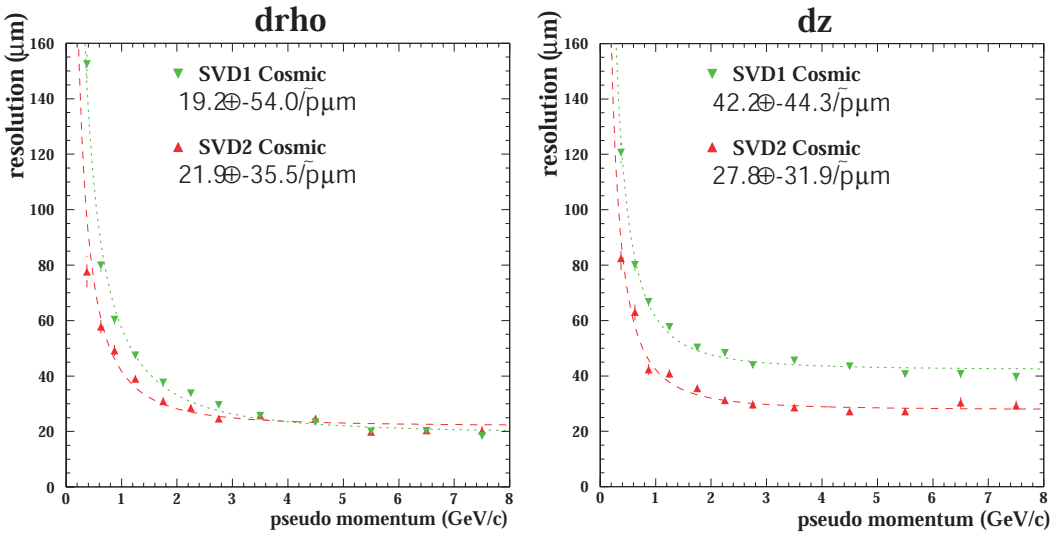


Figure 2.10: Impact parameter resolution of charged tracks with associated SVD hits from cosmic ray data. In the left plot, $\tilde{p} \equiv p\beta \sin^{3/2}\theta$, and in the right plot, $\tilde{p} \equiv p\beta \sin^{5/2}\theta$.

and its resolution is measured as a function of the track's momentum, p, and polar angle, θ. For SVD1,

$$\sigma_{xy} = 19.2 \oplus \frac{54.0}{p\beta \sin^{3/2}\theta}\ \mu\text{m}, \quad \sigma_z = 42.2 \oplus \frac{44.3}{p\beta \sin^{5/2}\theta}\ \mu\text{m}, \tag{2.1}$$

while for SVD2,

$$\sigma_{xy} = 21.9 \oplus \frac{35.5}{p\beta \sin^{3/2}\theta}\ \mu\text{m}, \quad \sigma_z = 27.8 \oplus \frac{31.9}{p\beta \sin^{5/2}\theta}\ \mu\text{m}, \tag{2.2}$$

where $\oplus$ indicates a quadratic sum. For a time-dependent CP analysis, the z-separation between B vertices must be measured with a precision of about 100 μm. Figure 2.10 shows the momentum and angular dependence of the impact parameter resolution.

2.3.3 Central Drift Chamber

The Central Drift Chamber (CDC) is a charged particle tracking system that measures track momenta from their curvature in the magnetic field induced by the solenoid magnet. The CDC also measures dE/dx of charged tracks to provide particle identification information.

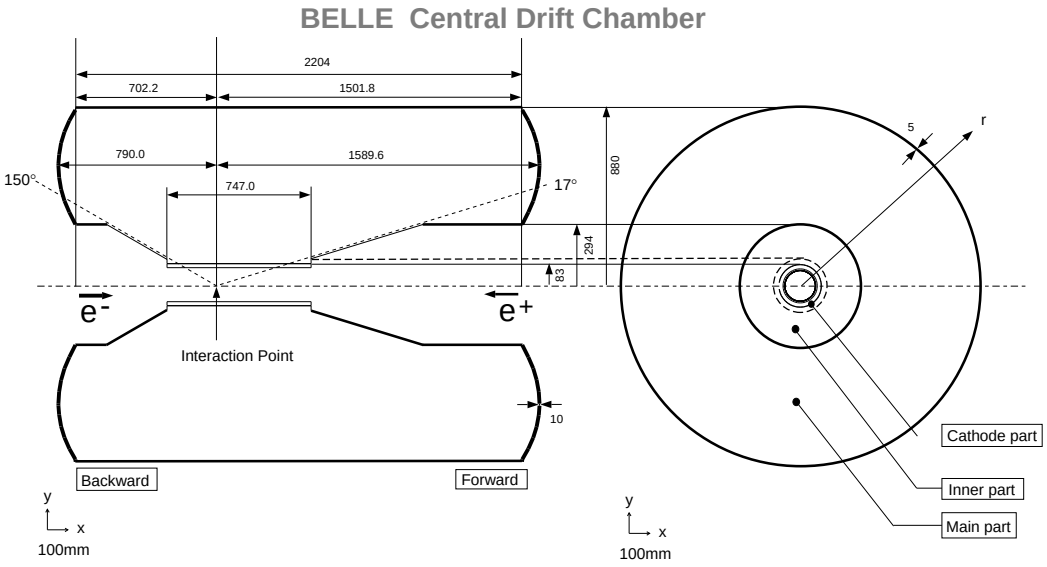


Figure 2.11: CDC geometry.

A charged particle moving through a constant magnetic field will follow the path of a helix which is defined by five parameters. The first is the curvature of the helix which gives the transverse momentum, p_T, and the second gives the pitch which is related to the longitudinal momentum, p_L. The other three give the impact parameters.

The inner side of the CDC is not encased with aluminium like the outer edges. By minimising the amount of material in the way, good tracking efficiency can be maintained even for low-p_T tracks. The CDC coverage is $17° < \theta < 150°$ which corresponds to 92% of the solid-angle. Figure 2.11 shows the geometry of the CDC.

The CDC is filled with a gas consisting of 50% helium and 50% ethane. The low-Z gas mixture is useful for minimising multiple Coulomb scattering to ensure a good momentum resolution, especially for low momentum tracks. The CDC contains a total of 8400 drift cells. A drift cell is made up of eight negatively biased field wires providing an electric field which surrounds a positively biased sense wire. The CDC cell structure is shown in Fig. 2.12.

When a charged particle passes through a drift cell, electrons are dislodged from the gas and drift towards the sensor wire. In the final 1 mm, the electric field increases as $1/r$. This accelerates the electrons sufficiently to cause secondary ionisation which, in turn, cause further ionisation resulting in a cascade of charge. This process, called gas amplification, increases the signal by a factor in excess of 10^6. Before amplification, the electrons have a specific drift velocity, so the measured pulse height and drift time are related to the energy deposit, dE/dx, and distance from the sense wire. Roughly half the wires are oriented parallel to the z-direction (axial) to provide p_T information while the remaining wires are oriented at a small angle, ± 50 mrad, to the z-direction (stereo) to give p_L.

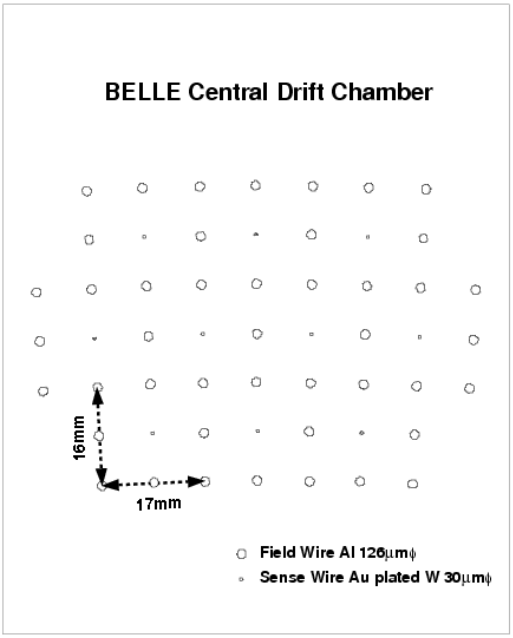


Figure 2.12: CDC cell structure.

The p_T resolution of the CDC is given by

$$\sigma(p_T) = \left(0.28 p_T \oplus \frac{0.35}{\beta}\right)\%, \tag{2.3}$$

and this is improved with SVD information,

$$\sigma(p_T) = \left(0.19 p_T \oplus \frac{0.30}{\beta}\right)\%. \tag{2.4}$$

The typical p_T resolution is shown in Fig. 2.13.

Figure 2.14 shows a scatter plot of the measured dE/dx as a function of the particle momentum. Separation between pions, kaons, protons and electrons can clearly be seen with particle momenta below 1 GeV/c. Now that the behaviour of particles with different mass have shown distinct characteristics in dE/dx, we can perform particle identification on an unknown track that traverses the CDC. The χ^2 for a given track, with each mass hypotheses, i, is constructed as

$$\chi^2_i \equiv \frac{(dE/dx)^{\text{Meas}} - (dE/dx)^{\text{Exp}}_i}{\sigma(p_T)^{\text{Exp}}_i}, \tag{2.5}$$

which compares measured with expected values. The likelihood that the charged track is of mass, i, is calculated assuming a Gaussian distribution,

$$\mathcal{L}_i = \frac{e^{-\frac{1}{2}\chi^2_i}}{\sqrt{2\pi}\sigma(p_T)_i}. \tag{2.6}$$

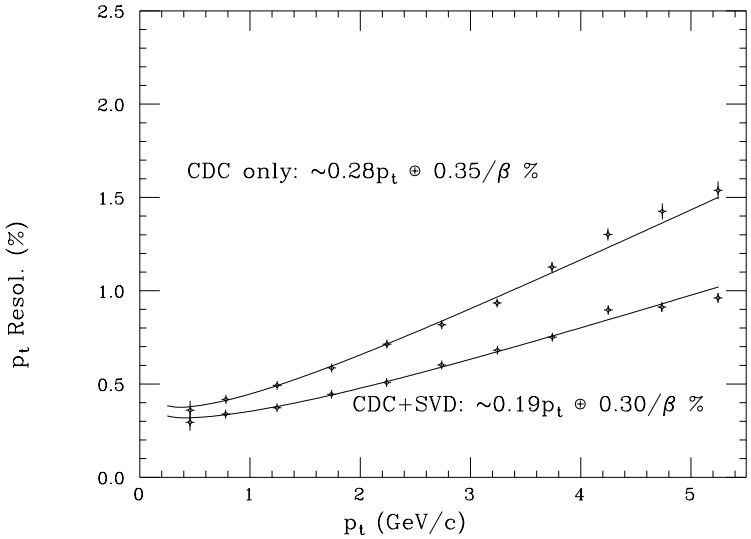


Figure 2.13: The p_T dependence of the p_T resolution using cosmic ray data.

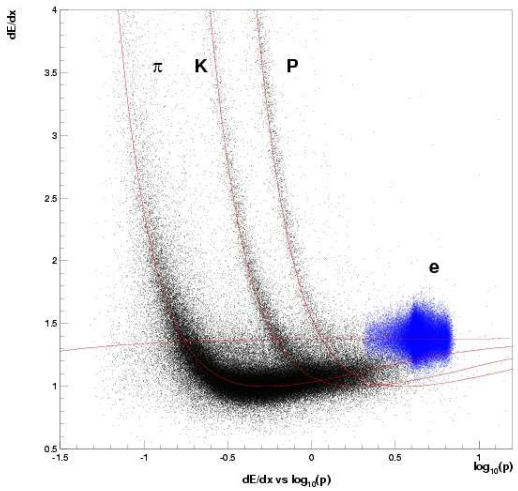


Figure 2.14: dE/dx vs momentum taken from collision data.

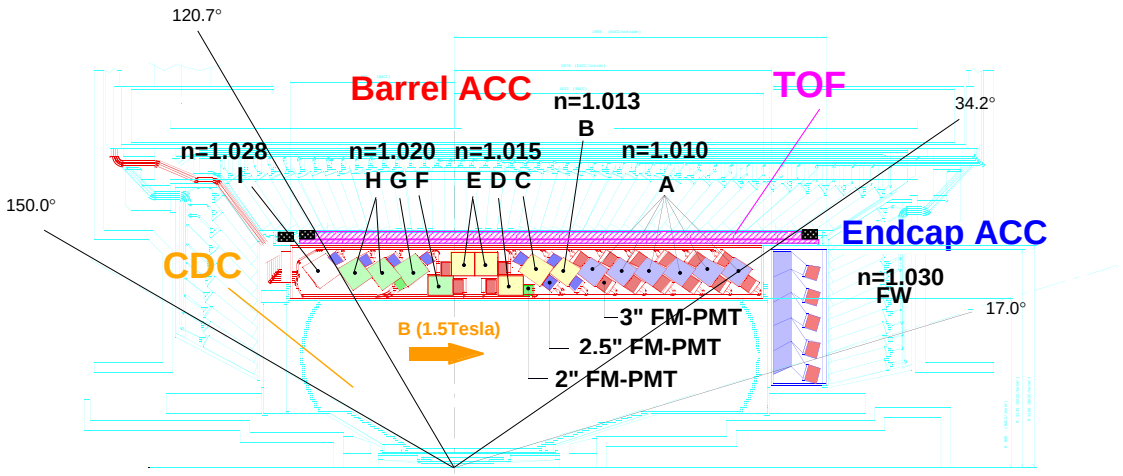


Figure 2.15: ACC geometry.

2.3.4 Aerogel Čerenkov Counter

The function of the Aerogel Čerenkov Counter (ACC) is to provide particle identification information to distinguish charged kaons from charged pions in the high momentum range of 1.0 to 4.0 GeV/c. Čerenkov radiation is emitted when the velocity of a charged particle in a medium exceeds the speed of light in the same medium,

$$n > \frac{1}{\beta} = \sqrt{1 + \left(\frac{m}{p}\right)^2}, \tag{2.7}$$

where m and p are the mass and momentum of the particle, and n is the refractive index of the material. Thus, we can distinguish kaons from pions and electrons by selecting a material in which pions will emit Čerenkov light, but heavier kaons will not.

The ACC can be divided into two regions, the barrel and the forward endcap. The barrel consists of 960 counter modules separated into sets of 60 in the ϕ direction. There are 228 modules that occupy the forward endcap in 5 concentric layers. Each counter is arranged in such a way that it points towards the direction of the IP. Fig. 2.15 shows the geometry of the ACC.

A counter module consists of silica aerogel encased in an aluminium box of roughly $12 \times 12 \times 12$ cm^3 in size. One or two photomultiplier tubes capable of operating in the 1.5 T magnetic field are attached to the sides of each box to detect light pulses. The refractive indices of the silica aerogel blocks are selected to maintain good kaon/pion separation over the kinematic range stated earlier. For the barrel modules, silica aerogel with five different refractive indices, $n = 1.010, 1.013, 1.015, 1.020$ and 1.028, are used depending on the polar angle. Silica aerogel with $n = 1.030$ is used in the forward endcap modules, to encompass lower momentum particles. This is because the TOF, designed for particle identification with low momentum particles and the next sub-detector in line, is not present in the endcap.

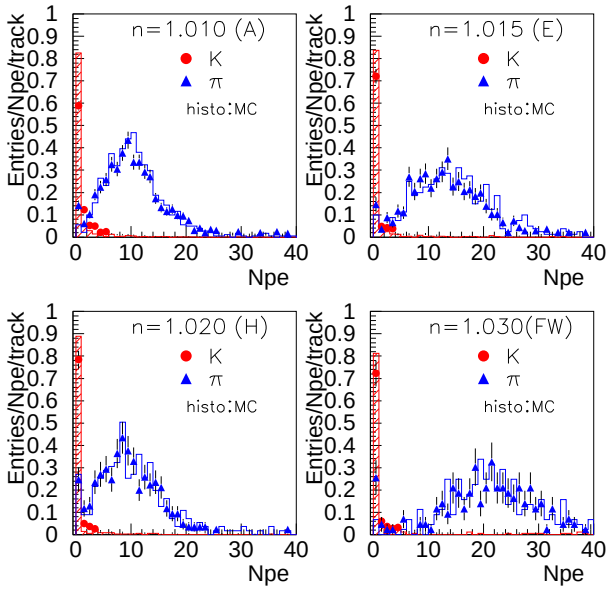


Figure 2.16: Pulse height in units of photo-electrons observed in the ACC barrel for kaons and pions in different regions of the ACC as labelled in Fig. 2.15.

Figure 2.16 shows the measured pulse height distributions in the ACC barrel for $\pi^{\pm}$ and $K^{\pm}$ candidates from $D^{*\pm}$ decays. There is good agreement between data and Monte Carlo simulation and a clear separation between kaons and pions can be seen. Of course, the ability of the ACC to distinguish between particles is not perfect. This can also be seen where the kaon and pion pulses slightly overlap.

The ACC is a threshold device, so basically acts as an on/off detector. The observed number of photo-electrons, $N(\mathrm{PE})^{\mathrm{Meas}}$, is compared with N_i^{Thr}, the momentum-dependent expected minimum number of photo-electrons for each particle type. The likelihood that an unknown track is of type, i, is given by,

$$\mathcal{L}_i = \begin{cases} \epsilon_i^{\mathrm{Exp}} & \text{if } N(\mathrm{PE})^{\mathrm{Meas}} \geq \mathrm{N(PE)_i^{Thr}} \\ 1 - \epsilon_i^{\mathrm{Exp}} & \text{if } N(\mathrm{PE})^{\mathrm{Meas}} < \mathrm{N(PE)_i^{Thr}} \end{cases} \tag{2.8}$$

where the expected efficiencies, $\epsilon_i^{\mathrm{Exp}}$, are determined from simulation studies.

2.3.5 Time of Flight Counter

The Time of Flight Counter (TOF) gives particle identification information to distinguish charged kaons from pions in the low momentum region, below 1.2 GeV/c. The particle mass,

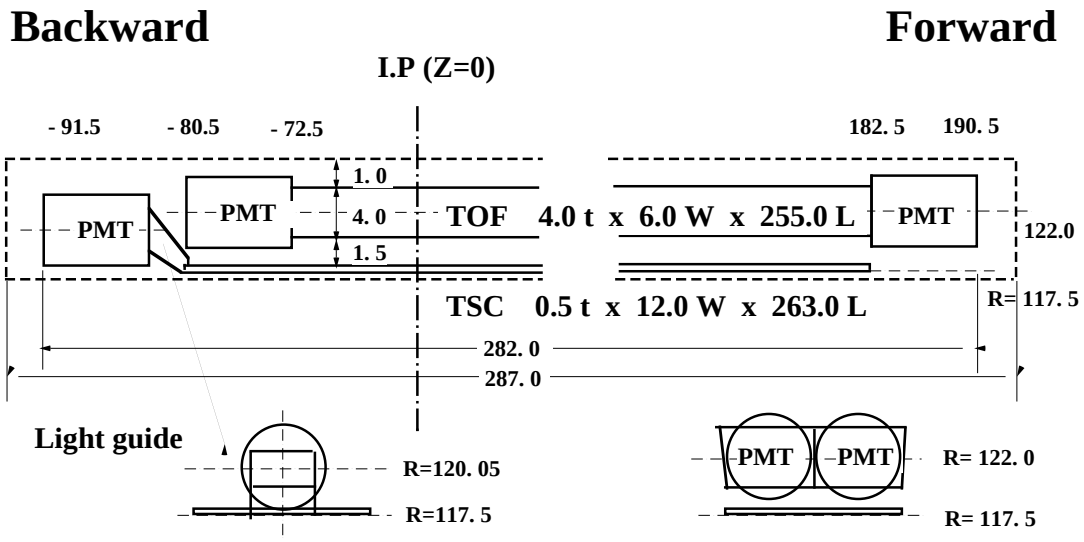


Figure 2.17: TOF module geometry.

m, can be determined from its relation to the measured elapsed time from collision at the IP, and is expressed as,

$$T = \frac{L}{c\beta} = \frac{L}{c}\sqrt{1 + \left(\frac{m}{p}\right)^2} \tag{2.9}$$

where p is the momentum of the track and L is the flight length.

The TOF sub-detector consists of plastic scintillation counters attached to photomultiplier tubes and has a timing resolution of 100 ps. The TOF also provides fast timing signals for the data acquisition trigger system. To sustain a fast trigger rate in any beam background condition, thin trigger scintillation counters (TSC) are placed just before the TOF counter. A TOF module consists of two TOF counters and one TSC counter, a schematic of which is shown in Fig. 2.17.

In total, there are 64 TOF modules located in the barrel region at a radius of 1.2 m from the IP, covering a polar angle of $34° < \theta < 120°$. The TOF provides separation between kaons, pions and protons in the momentum region below 1.2 GeV/c. Figure 2.18 shows the mass distribution obtained from TOF measurements. The data points are consistent with the simulation prediction that assumes a time resolution of 100 ps. Figure 2.19 shows the K/π separation performance as a function of the particle's momentum.

A χ^2 value from TOF hit information is calculated by measuring the time interval between hits in the TSC and TOF counters and comparing this with the expected time a particle of type, i, with known momentum would take to travel through the detector. The likelihood is given by

$$\mathcal{L}_i = \frac{e^{-\frac{1}{2}\chi_i^2}}{\prod_{j=1}^{\mathrm{ndf}} \sqrt{2\pi}\sigma_j}, \tag{2.10}$$

where ndf is the number of photomultiplier time intervals included in the calculation of χ_i^2,

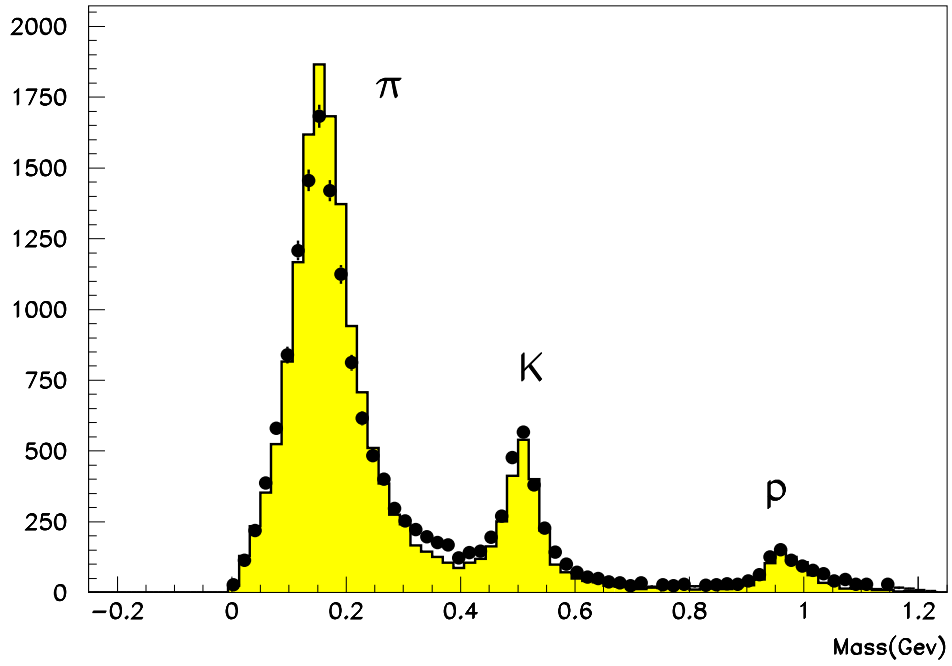


Figure 2.18: Mass distribution from TOF measurements for particle momenta below 1.2 GeV/c.

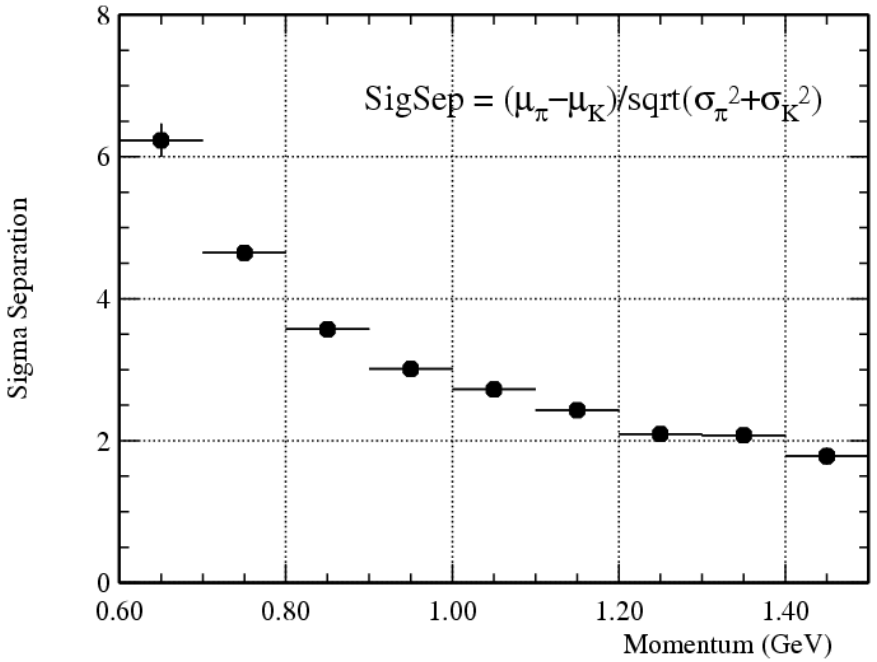


Figure 2.19: K/π separation performance of the TOF as a function of the particle's momentum.

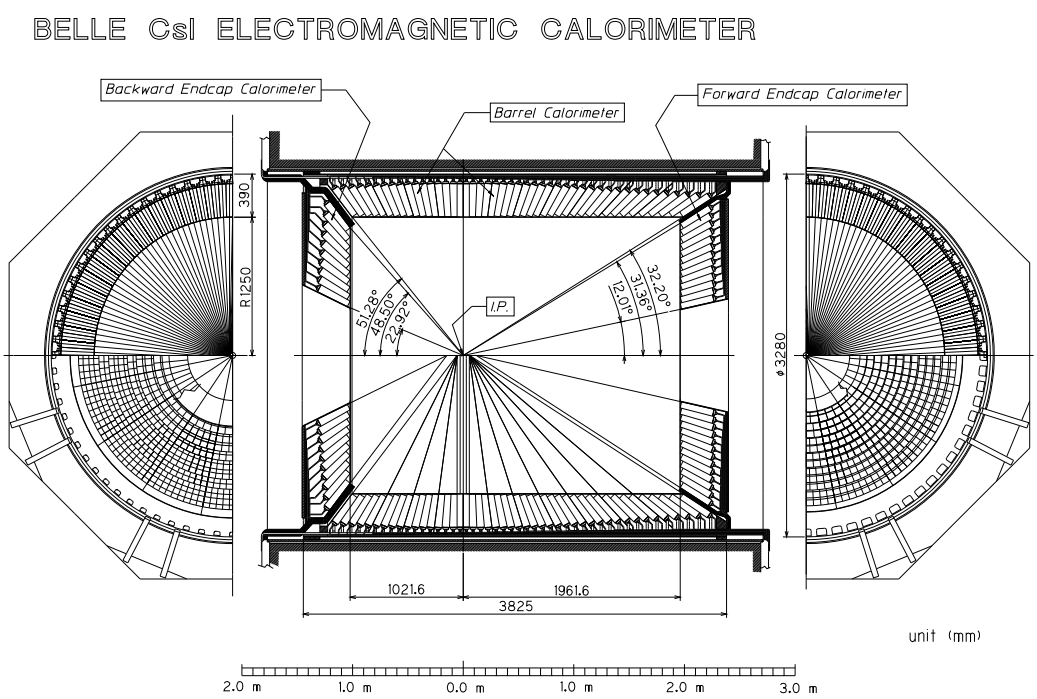


Figure 2.20: ECL geometry.

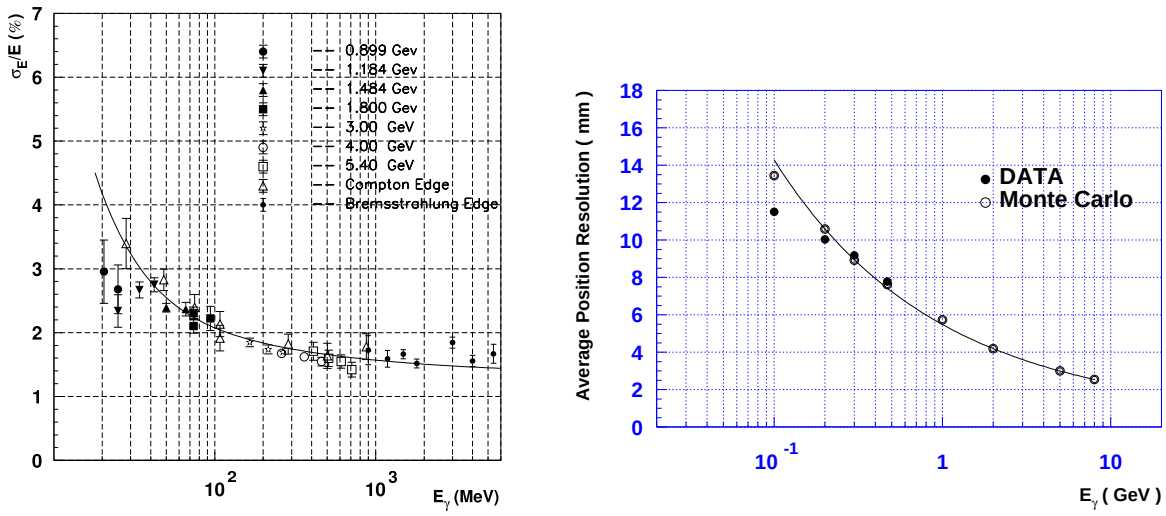
and σ_j is the timing resolution.

2.3.6 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECL) serves to identify electrons and photons by measuring electromagnetic showers. Electrons in this analysis are required for flavour tagging while photons are needed for the reconstruction of π^0.

The ECL consists of 8736 thallium-doped (Tl) CsI crystal counters. The CsI(Tl) crystal has a tower shape and is 30 cm long, which corresponds to 16.2 radiation lengths. Each CsI(Tl) crystal is arranged so that it points towards the IP. The barrel component has 6624 crystals divided into 46 in θ and 144 in ϕ. The forward (backward) endcap has 1152 (960) crystals divided into 13 (10) in θ and $48 - 144$ $(64 - 144)$ in ϕ depending on θ. The geometry of the ECL is shown in Fig. 2.20.

When an electron or photon hits a crystal, its energy is deposited in electromagnetic showers produced by bremsstrahlung and pair production. Other charged particles deposit a small amount of energy by ionisation. Therefore, the ratio of the cluster energy measured by the ECL to the momentum of the charged track momentum as measured by the CDC, E/p, is close to unity for electrons and lower for other particles. In this way, electron identification can be performed.



(a) Energy resolution

(b) Position resolution

Figure 2.21: Energy and position resolutions of the ECL as a function of the incident photon energy deposit.

The energy resolution of the ECL is given by

$$\frac{\sigma_E}{E} = \left(1.34 \oplus \frac{0.066}{E} \oplus \frac{0.81}{E^{1/4}}\right)\%, \tag{2.11}$$

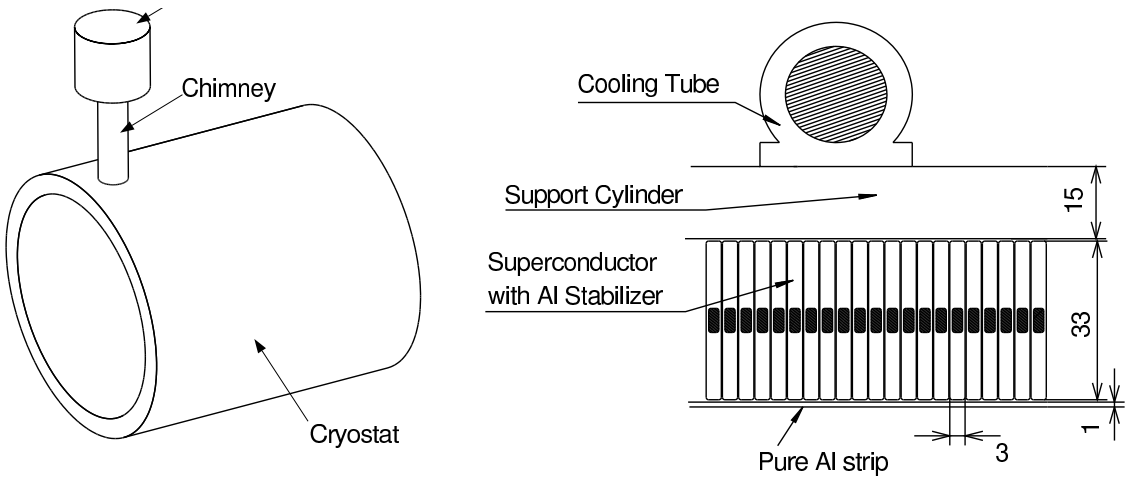
and the position resolution is

$$\sigma_{\mathrm{Pos}} = \left(0.27 + \frac{3.4}{E^{1/2}} + \frac{1.8}{E^{1/4}}\right) \mathrm{mm}, \tag{2.12}$$

where E is in units of GeV. Figure 2.21 shows the energy and position resolution of the ECL.

2.3.7 Solenoid Magnet

The superconducting solenoid provides a magnetic field of 1.5 T, that bends charged particles in a helix from which track momentum can be measured in the CDC. The superconducting coil consists of a single layer niobium-titanium-copper alloy embedded in a high purity aluminium stabiliser. The coil is wound around the inner surface of an aluminium support cylinder of 3.4 m in diameter and 4.4 m in length. Cooling is provided by circulating liquid helium through a tube on the inner surface of the aluminium cylinder. Figure 2.22 shows the layout of the superconducting solenoid.



(a) Outlook of the magnet

(b) Cross sectional view of the coil

Figure 2.22: The left figure shows outlook of the solenoid and the right shows a cross sectional view of the coil in units of mm.

2.3.8 K_L^0 and Muon Detector

The K_L^0 and muon detector (KLM) provides muon identification for charged particles with enough momentum to reach the KLM, $P_T > 0.6$ GeV/c. It can also detect the neutral K_L^0. Since these are highly penetrative particles, a lot of material is needed to detect them efficiently.

The KLM consists of alternating layers of charged particle detectors and 4.7 cm thick iron plates. There are 15 resistive plate counter (RPC) superlayers and 14 iron layers in the barrel region and 14 RPC superlayers in each endcap, covering the polar region, $20° < \theta < 155°$. The iron layers also serve as a return yoke for the magnetic flux provided by the superconducting solenoid. The KLM barrel part of the iron yoke is shown in Fig. 2.23. Each RPC superlayer consists of two RPC modules to provide 2-dimensional $\theta - \phi$ information. The cross section of an RPC superlayer is shown in Fig. 2.24.

Hadrons interacting with the iron plates produce a shower of ionising particles that are detected by the RPC layers. The result is a cluster of hits deposited in the KLM. A K_L^0 candidate can be distinguished from another charged hadron because it will not leave an associated track in the CDC. A muon, on the other hand, does leave a charged track in the CDC. However, muons can still be distinguished from charged and neutral hadrons because they do not feel the strong interaction. Hadrons are more rapidly absorbed and deflected by strong interactions with iron resulting in wide clusters and are stopped within a few layers of iron. Muons only experience electromagnetic multiple scattering and energy loss, so their clusters tend to be thinner and they have far greater penetration depth.

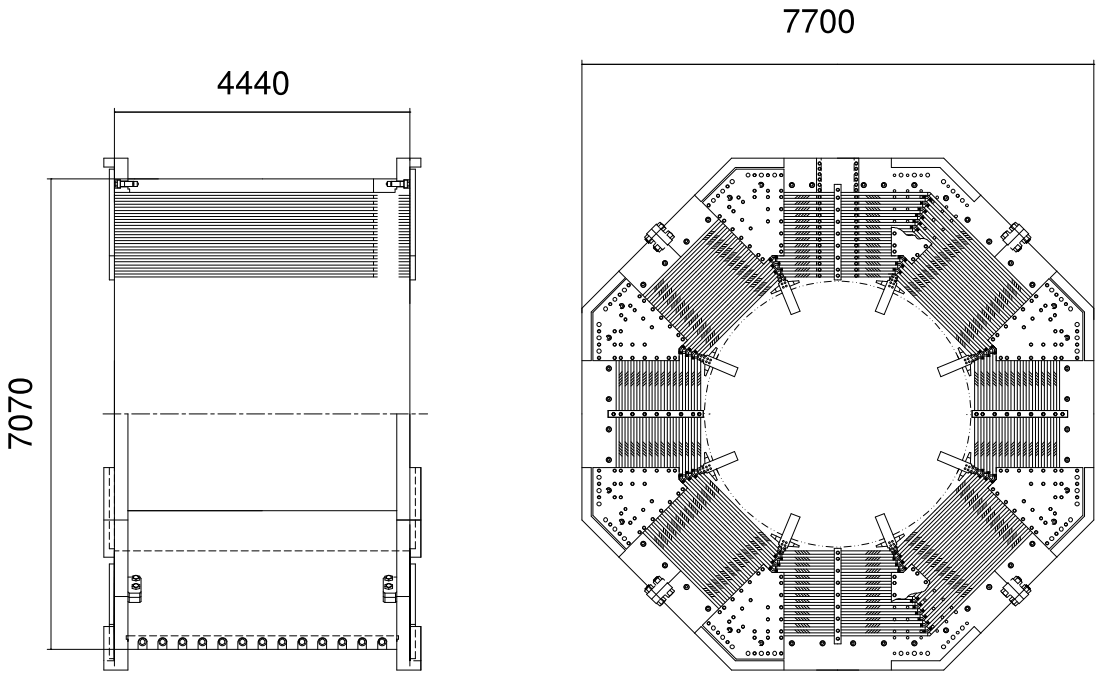


Figure 2.23: KLM geometry in units of mm.

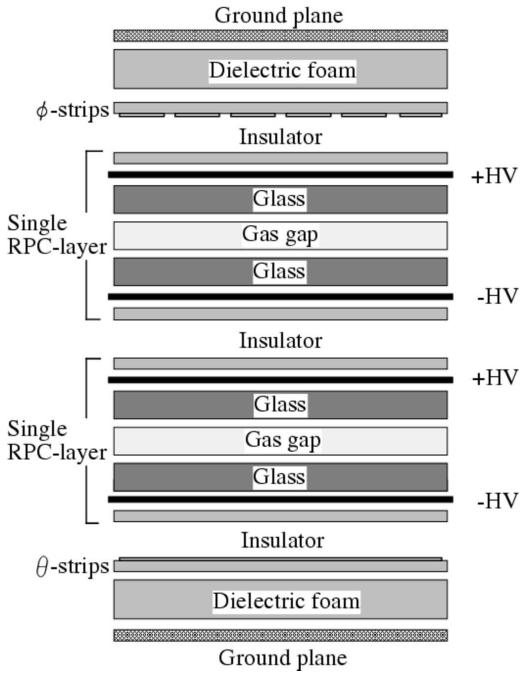


Figure 2.24: Cross section of an RPC superlayer.

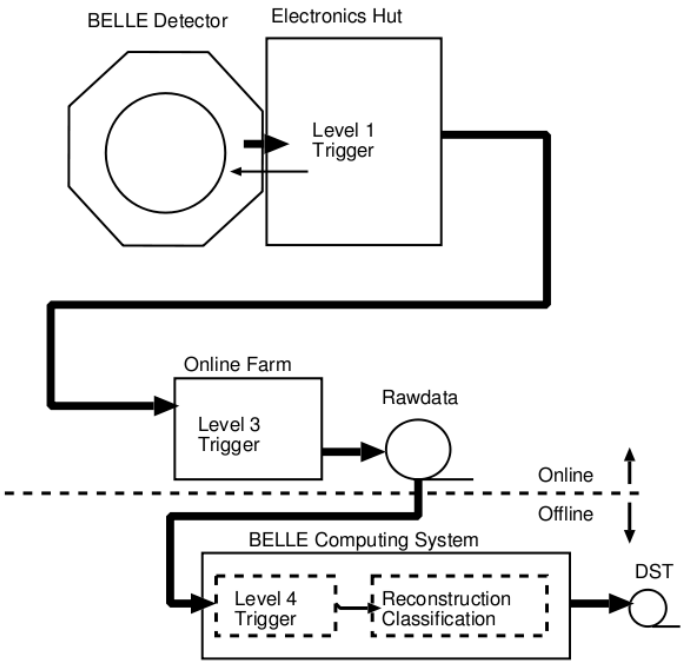


Figure 2.25: An overview of the Belle trigger system.

2.4 Trigger and Data Acquisition

The role of the trigger system is to distinguish and record the interesting physics events from the enormous amount of background that enters the detector, using fast signals from the sub-detectors. Because of the high beam current of the KEKB accelerator, the trigger suffers severe beam background, and since the rates of beam background are very sensitive to real-time accelerator conditions, it is difficult to estimate reliably. For this reason, the trigger system is required to be flexible so that background rates can be kept within the tolerance of the data acquisition system. Redundant triggers are also needed to keep the high trigger efficiency for interesting physics events.

The Belle trigger system consists of the Level-1 hardware trigger and the Level-3 software trigger which operate in real-time. The Level-4 trigger and event reconstruction and classification is performed offline. A flow diagram of the Belle trigger system is shown in Fig. 2.25.

The Level-1 trigger consists of the sub-detector trigger systems and the central trigger system called the Global Decision Logic (GDL). The sub-detector trigger systems are categorised into two: track triggers and energy triggers. The CDC and TOF are responsible for charged track trigger signals while the ECL provides triggers based on the total energy deposit. The KLM provides additional trigger signals for muons. A schematic of the Belle Level-1 trigger system is shown in Fig. 2.26.

These are four main sub-triggers in the GDL, the two track trigger, the three track trigger, the cluster number trigger and the energy sum trigger. The two track trigger requires two

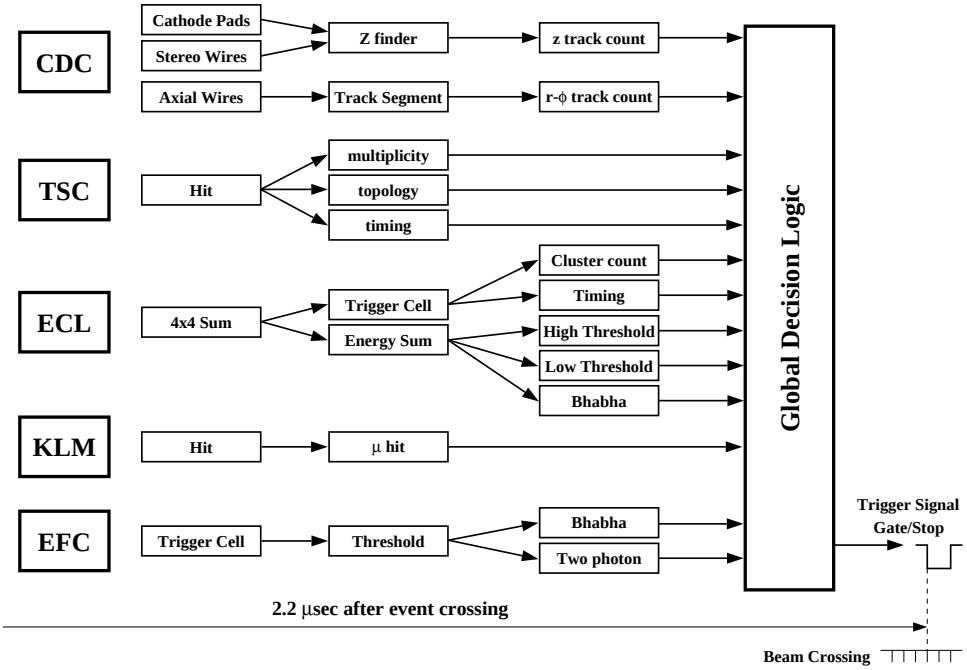


Figure 2.26: The Belle Level-1 trigger system.

tracks with an opening angle greater than 135°. At least one track must have z information and a minimum of 2 hits in the TOF. The three track trigger is similar to the two track trigger, but no opening angle condition is required. At least three hits in the inner tracking sub-detectors are required. The cluster number trigger requires at least 4 isolated clusters in the ECL. The energy sum trigger requires at least 1 GeV of energy deposited in the ECL.

The GDL combines the sub-detector trigger signals and makes the final decision to initiate a Belle-wide data acquisition (DAQ) within 2.2 μs of the beam crossing. The typical trigger rate is 200-250 Hz. With redundant triggers in place, the trigger efficiency for $B\bar{B}$ events is greater than 99.5%.

The performance goal of the DAQ is to be operational at a maximum trigger rate of 500 Hz, while keeping the dead time fraction to less than 10%. In order to achieve this, the entire DAQ system is divided into seven sub-systems running in parallel. An event builder combines the signals from sub-detectors into a single event and passes it to an online computer farm. The online computer performs a basic track and cluster reconstruction and further rejects unwanted events with the Level-3 trigger. The remaining events are sent to a computer centre for offline processing. A schematic of the Belle DAQ system is shown in Fig. 2.27.

The Level-4 trigger is applied first during offline processing. Events must have at least one track originating from the IP, $dr < 1.0$ cm and $|dz| < 4.0$ cm, with $p_T > 300$ MeV/c. Events passing the Level-4 trigger undergo full event reconstruction. Tracks are reconstructed with hits in the CDC, then these tracks are extrapolated towards the IP to search for associated SVD hits. They are also extrapolated outwards to search for hits in the outer detectors. Thus, 4-vectors and particle identification likelihoods can be assigned. Clusters in the ECL

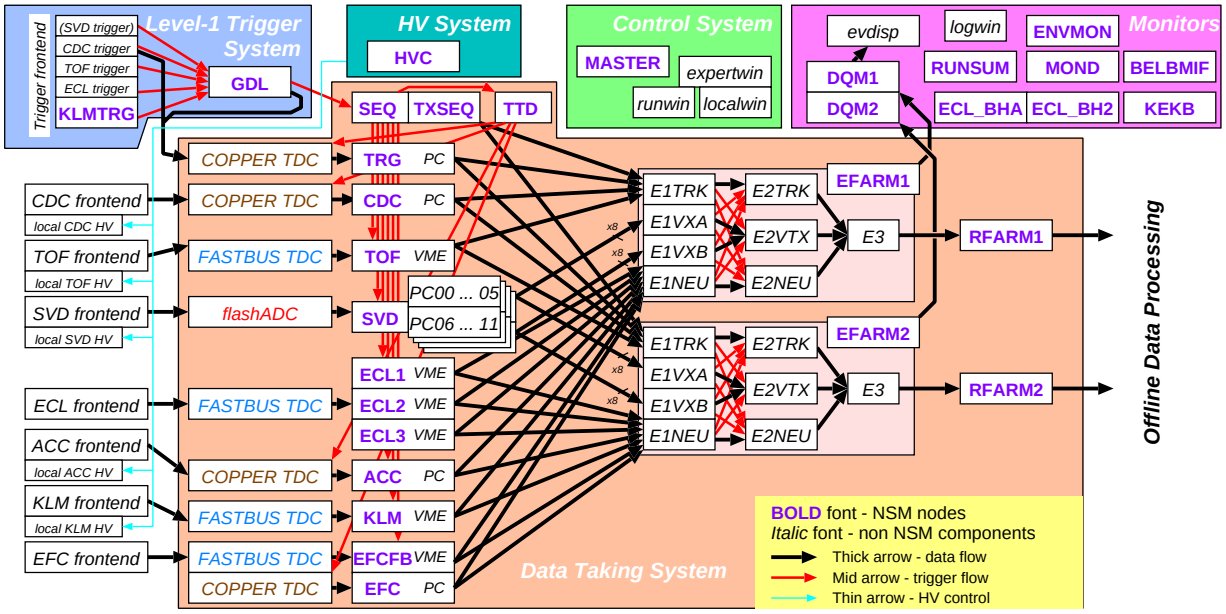


Figure 2.27: The Belle DAQ system.

are also reconstructed to form photons. This information along with many other calculated variables are stored in a format available to collaborators for analysis.

We close with a picture of a $B^0 \to J/\psi K^0_S$ candidate in the Belle detector, a culmination of all the detectors and procedures described in this chapter shown in Fig. 2.28.

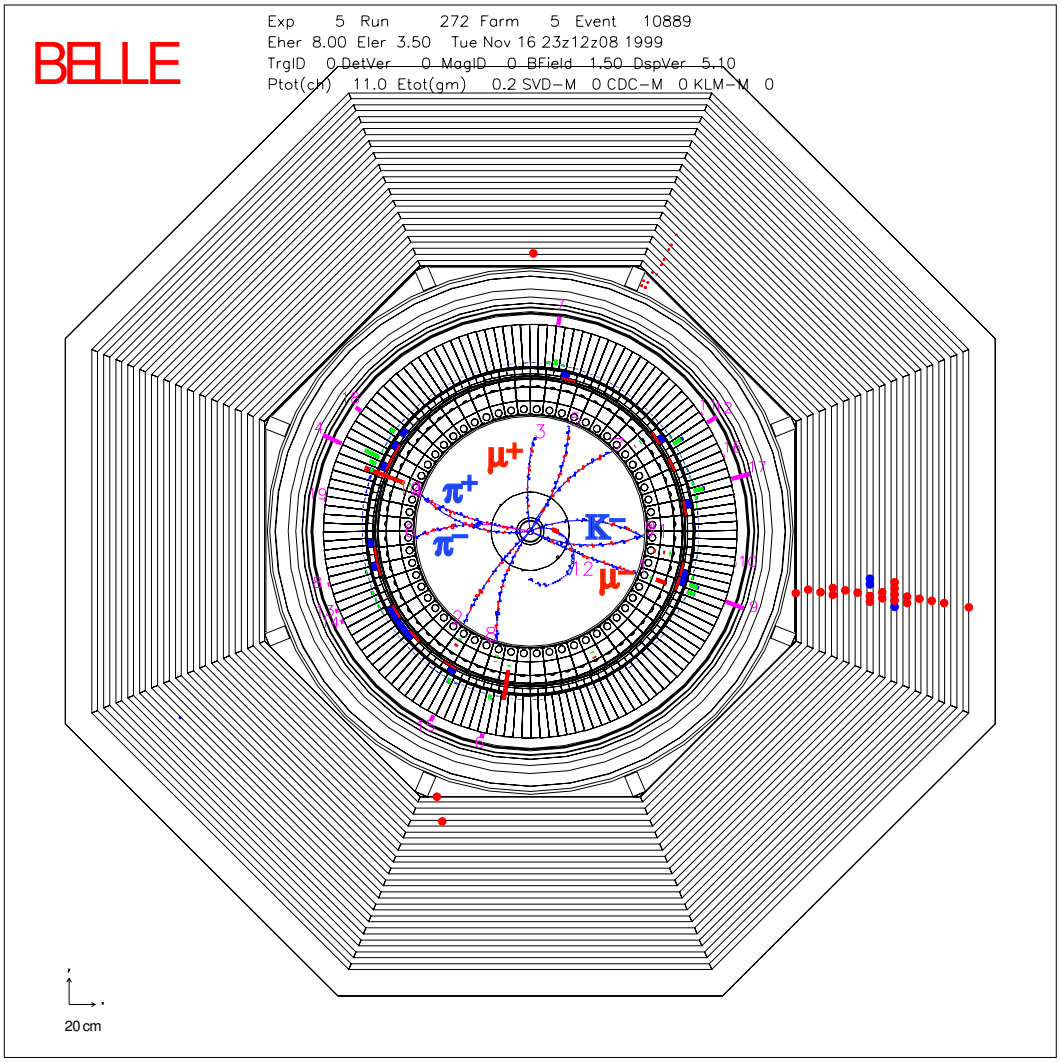


Figure 2.28: A $B^0 \to J/\psi K^0_S$ candidate in the $r - \phi$ plane.

Chapter 3

Event Selection

3.1 Overview

In this chapter, we first describe the event pre-selection procedure to isolate $B\bar{B}$ events produced in the chain, $e^+e^- \to \Upsilon(4S) \to B^0\overline{B}^0$. Then we explain the reconstruction procedure for the $B_{\rm Rec} \to K^0\pi^0$ decay.

In this analysis, we reconstruct the signal mode $B^0 \to K^0_S\pi^0$ and two control samples $B^+ \to K^0_S\pi^+$ and $B^+ \to K^+\pi^0$.

Control samples provides a valuable cross checks for the signal mode of interest and serve for two purposes. If the analysis of a control sample can produce expected results, this inspires confidence in the analysis procedure of the signal mode. In addition, since the control sample has many similarities to the signal mode, it may even aid in the extraction of physics parameters. The branching fraction measurement will be checked with the $B^+ \to K^+\pi^0$ control sample while the CP measurement will be checked against the $B^+ \to K^0_S\pi^+$ control sample.

3.2 Data Sample

This analysis is performed on a data sample containing an integrated luminosity of $605\,\mathrm{fb}^{-1}$ which corresponds to 657×10^6 $B\bar{B}$ pairs accumulated at the $\Upsilon(4S)$ resonance by the Belle detector. SVD1 was used for the first sample of 152×10^6 $B\bar{B}$ pairs, while SVD2 was used to record the remaining 505×10^6 $B\bar{B}$ pairs.

3.3 Event Pre-selection

3.3.1 $B\bar{B}$ Event Selection

Electron-positron colliders operating at the $\Upsilon(4S)$ resonance do not always produce $B\bar{B}$ events. In fact, the cross section is dominated by non-$B\bar{B}$ events, the most common of which is e^+e^- (Bhabha) scattering. Also present are radiative Bhabha, other lepton pair production, $e^+\ e^- \to l^+\ l^-$, where $l = \mu, \tau$, two-photon events, $e^+\ e^- \to \gamma\gamma$, and beam-gas interactions. Non-$B\bar{B}$ hadronic events may also be produced and occur when, $e^+\ e^- \to q\overline{q}$ where $q = u, d, s, c$ represents the light quarks and is known as continuum. This section discusses the suppression of these backgrounds.

3.3.2 Non-Hadronic Event Suppression

A standard set of selection criteria was devised to suppress non-hadronic non-$B\bar{B}$ events [51].

At least three "good" charged tracks must exist where a "good" charged track is defined as having $p_{\rm track} > 0.1$ GeV/c, $dr < 2.0$ cm and $|dz| < 4.0$ cm, where $p_{\rm track}$ is the transverse momentum and dr $|dz|$ are impact parameters which describe the distance of the closest approach to the e^+e^- interaction point (IP).

At least two "good" neutral clusters must be observed in the barrel region of the ECL where a "good" neutral cluster has $E_{\rm cluster} > 0.1$ GeV and $-0.7 < \cos\theta < 0.9$, where $E_{\rm cluster}$ is the energy deposit and θ is the polar angle.

The sum of momentum magnitudes in the z-direction calculated in the $\Upsilon(4S)$ rest frame, $p_z^{\rm CMS}$, must be less than half the total available energy, $\sqrt{s}$, $\sum |p_z^{\rm CMS}| \leq 0.5\ \sqrt{s}$ and the primary event vertex calculated from the "good" charged tracks must satisfy $dr < 1.5$cm and $|dz| < 3.5$ cm.

Assuming the pion mass for "good" charged tracks, the total visible energy in the $\Upsilon(4S)$ rest frame, $E_{\rm vis}^{\rm CMS}$, should be $E_{\rm vis}^{\rm CMS} > 0.18\ \sqrt{s}$. The total energy of "good" neutral clusters in the $\Upsilon(4S)$ rest frame, $E_{\rm cluster}^{\rm CMS}$, has to be inside $0.1\ \sqrt{s} < \sum E_{\rm cluster}^{\rm CMS} < 0.8\ \sqrt{s}$. The invariant mass of particles in each hemisphere defined as perpendicular to the boost, $M_{\rm jet}$, must satisfy, $M_{\rm jet} > 1.8$ GeV/c^2.

These selection criteria retain more than 99 % of $B\bar{B}$ events while reducing the contamination from non-hadronic processes to less than 5 %.

3.4 Particle Identification

Pions and kaons cannot be distinguished unambiguously by the Belle detector. However, the sub-detectors are able to provide a likelihood that a charged track is a kaon given that the track is hypothesised to be a pion,

$$\mathcal{P}(K:\pi) = \frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi}, \tag{3.1}$$

where $\mathcal{P}_i$ $(i = K, \pi)$ is the product of the likelihoods that the particle is of type i and is calculated as,

$$\mathcal{L}_i = \mathcal{L}_i^{\mathrm{CDC}} \times \mathcal{L}_i^{\mathrm{TOF}} \times \mathcal{L}_i^{\mathrm{ACC}}, \tag{3.2}$$

where each component, $\mathcal{L}_i^{\mathrm{DET}}$, is the likelihood that the particle is of type i in their respective sub-detectors defined in Eq. (2.6), Eq. (2.10) and Eq. (2.8).

Figure 3.1 shows an example distribution of the likelihood ratio, $\mathcal{P}(K:\pi)$, for charged tracks. The left half contains the region where the track is most likely to be a pion and on the right side, the track is most likely a kaon. The peaks at the extremities demonstrate that most tracks can be clearly identified. A non-intuitive feature is the spike that occurs at $\mathcal{P}(K:\pi) = 0.5$ which arises because the likelihood, $\mathcal{L}_i^{\mathrm{DET}}$, is set to 0.5 when a sub-detector cannot identify a track.

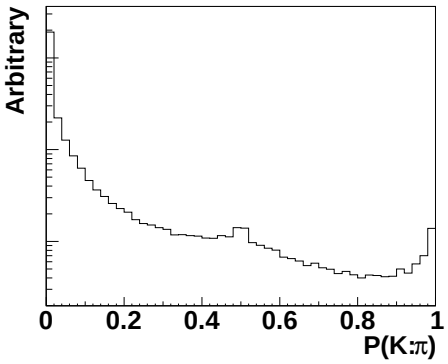


Figure 3.1: PID likelihood ratio, $\mathcal{P}(K:\pi)$, on a logarithmic scale.

3.5 Kinematic Fit

Kinematic fitting reduces the effects of detector resolution which can be used to improve the mass resolution and determine the decay vertex of a candidate particle. There are two types of kinematic fitting used in this analysis: mass-constrained fitting and vertex-constrained fitting.

In the mass-constrained fitting, the invariant mass of the candidate, M_{cand}, is set to the known mass and the momentum of its daughters are re-calculated. The underlying motivation for

vertex-constrained fitting is that the tracks used to reconstruct a candidate may not pass through the same point. The $M_{\rm cand}$ decay vertex is obtained by tuning the momentum and position of each daughter according to its measurement errors so that all tracks pass through a single point. An illustration of this procedure is shown in Fig. 3.2.

The fitting technique is based on the least square method using the Lagrange multiplier method and is described at length in [52]. A measure of the goodness-of-fit is the χ^2 of the vertex reconstruction which is defined as,

$$\chi^2 \equiv \vec{\lambda}^T V_D^{-1} \vec{\lambda}, \tag{3.3}$$

where $\vec{\lambda}$ is difference in track parameters before and after the vertex reconstruction, and V_D is the error matrix of $\vec{\lambda}$.

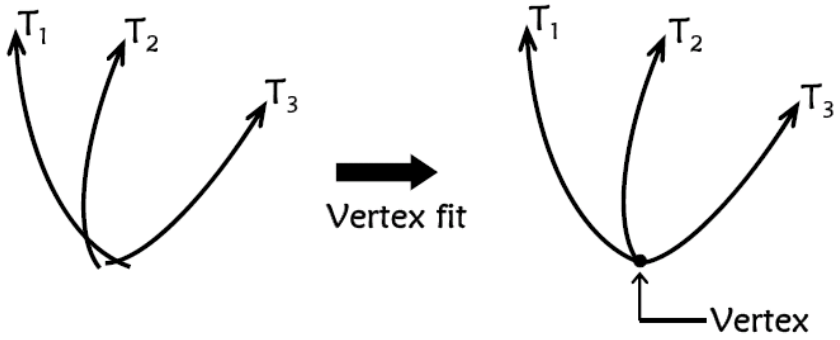


Figure 3.2: Schematic of a vertex-constrained fit.

In this analysis, we use mass-constrained kinematic fitting for the π^0 mass reconstruction (§3.6) and vertex-constrained fitting to determine the vertex of the K^0_S (§3.7) and both B mesons (§5.2).

3.6 π^0 Reconstruction

The π^0 is reconstructed from two photons as $\pi^0 \rightarrow \gamma\gamma$ and undergoes mass-constrained fitting. The variables which discriminate real π^0 mesons from combinatorial background are the daughter photon energies, E_γ, the invariant mass of the π^0 candidate before mass-constrained fitting, $m(\gamma\gamma)$, the mass-constrained goodness-of-fit χ^2 and the angle between the π^0 momentum in the B meson rest frame and the photon momentum in the π^0 rest frame, θ^*. Figure 3.3 shows a typical $m(\gamma\gamma)$ distribution.

We select π^0 candidates requiring the following selection criteria: $E_\gamma > 0.05$ GeV in the barrel region, $E_\gamma > 0.10$ GeV in the endcaps, $\chi^2 < 50$, $\cos\theta^* < 0.95$ and $0.115\ \text{GeV}/c^2 < m(\gamma\gamma) < 0.152\ \text{GeV}/c^2$.

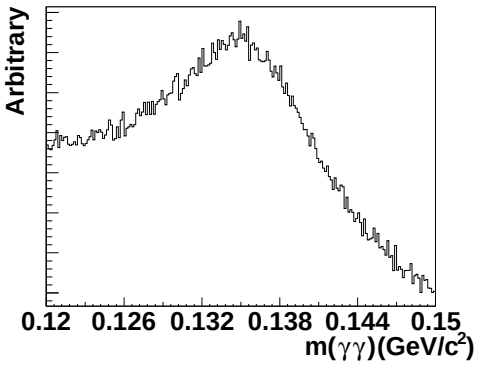


Figure 3.3: An invariant mass distribution of π^0.

3.7 K^0_S Reconstruction

The K^0_S is reconstructed from a charged pion pair, $K^0_S \rightarrow \pi^+\pi^-$, and a vertex-constrained fit is performed. Besides the reconstructed K^0_S mass, $m(\pi^+\pi^-)$, which is shown in Fig. 3.4, a set of four variables described in Table 3.1 are used to optimise K^0_S quality. A standard

Table 3.1: Variables for K^0_S optimisation.

dr	The smaller of dr_1 and dr_2, the shortest distance between the two K^0_S daughter tracks and the IP
$d\phi$	The azimuthal angle between the momentum vector and decay vertex vector of a K^0_S candidate
z_dist	The distance between the two K^0_S daughter tracks at their point of interception
fl	The flight length of a K^0_S candidate in the $x-y$ plane

set of selection criteria based on these four variables, called the "goodKs", were created to significantly enhance K^0_S purity and are summarised in Table 3.2. The effects of these requirements are discussed further in [53].

The $K^0_S \rightarrow \pi^+\pi^-$ channel is reconstructed and subject to the "goodKs" selection and the mass window, $0.480\ \text{GeV}/c^2 < m(\pi^+\pi^-) < 0.516\ \text{GeV}/c^2$.

Table 3.2: The "goodKs" selection criteria.

Momentum (GeV/c)	dr (cm)	$d\phi$ (rad)	z_dist (cm)	fl (cm)
< 0.5	> 0.05	< 0.30	< 0.8	–
$0.5 - 1.5$	> 0.03	< 0.10	< 1.8	> 0.08
> 1.5	> 0.02	< 0.03	< 2.4	> 0.22

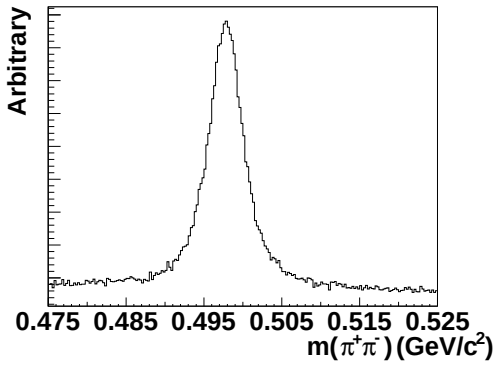


Figure 3.4: An invariant mass distribution of K^0_S.

3.8 Reconstruction of B mesons

The reconstructed B meson, $B_{\rm Rec}$, is not characterised by its invariant mass unlike other reconstructed particles. From 2-body kinematics, the energy of $B_{\rm Rec}$ in the centre-of-mass system (CMS), $E_B^{\rm CMS}$, must be half the total energy of the e^+e^- system in the same frame, $E_{\rm beam}^{\rm CMS}$. With this constraint, the B meson is described by two kinematic variables defined as,

$$\begin{aligned} M_{\rm bc} &\equiv \sqrt{(E_{\rm beam}^{\rm CMS})^2 - (p_B^{\rm CMS})^2}, \\ \Delta E &\equiv E_B^{\rm CMS} - E_{\rm beam}^{\rm CMS}. \end{aligned} \tag{3.4}$$

The variable, $M_{\rm bc}$, is known as the beam-constrained mass and has much better resolution than the reconstructed B mass, thus providing better separation of signal and background. The energy difference, ΔE, should peak around zero for correctly reconstructed B mesons. Examples of the distributions of $M_{\rm bc}$ and ΔE using simulated $B^0 \to K^0_S\pi^0$ events are shown in Fig. 3.5.

3.9 Continuum Background Suppression

Continuum background can be suppressed with criteria that utilise the event shape differences between the $B\bar{B}$ and continuum background. Continuum events produce light mesons with high momentum which tend towards a jet-like shape while $B\bar{B}$ events are more spherical in shape as shown in Fig. 3.6. Event shape variables can be constructed from modified Fox-Wolfram moments divided into three components: a component including only particles belonging to the reconstructed candidate, a component including particles other than the

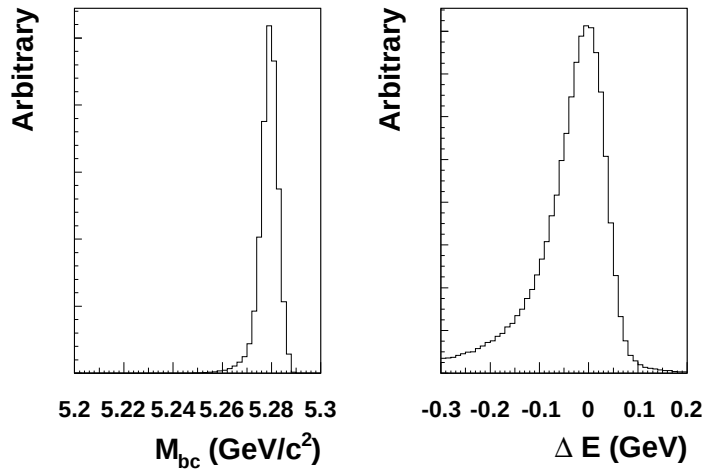


Figure 3.5: $M_{\rm bc}$ and ΔE distributions using simulated $B^0 \to K^0_S\pi^0$ events.

reconstructed candidate, and a component including all particles. They are defined as,

$$
\begin{aligned}
R_n^{\rm Rec} &\equiv \frac{\sum_{i,j} |\vec{p}_i||\vec{p}_j| P_n(\cos\theta_{ij})}{\sum_{i,j} |\vec{p}_i||\vec{p}_j|}, \\
R_n^{\rm Tag} &\equiv \frac{\sum_{i,j} |\vec{p}_i||\vec{p}_j| P_n(\cos\theta_{ij})}{\sum_{i,j} |\vec{p}_i||\vec{p}_j|}, \\
R_n^{\rm All} &\equiv \frac{\sum_{i,j} |\vec{p}_i||\vec{p}_j| P_n(\cos\theta_{ij})}{\sum_{i,j} |\vec{p}_i||\vec{p}_j|},
\end{aligned}
\tag{3.5}
$$

where $|\vec{p}|$ indicates the particle momentum evaluated in the CMS, P_n is an n-th order Legendre polynomial, the indices i, j enumerate over all particles in their respective categories and θ_{ij} is the angle between the momentum vectors. We do not use $R_n^{\rm Rec}$ because they are strongly correlated with $M_{\rm bc}$ and ΔE. The moments $R_1^{\rm Tag}$, $R_3^{\rm Tag}$ and $R_1^{\rm All}$ are also not used because they are correlated with $M_{\rm bc}$ [54].

We then introduce the Fisher discriminant [55], F, that is made of the following five variables $R_2^{\rm Tag}$, $R_4^{\rm Tag}$, $R_2^{\rm All}$, $R_3^{\rm All}$ and $R_4^{\rm All}$ defined as,

$$
F \equiv \sum_{n=2,4} \alpha_n R_n^{\rm Tag} + \sum_{n=2,3,4} \beta_n R_n^{\rm All}, \tag{3.6}
$$

where the coefficients α_n and β_n are determined by training the separation between the $B\bar{B}$ and continuum events. In addition, we include the cosine of the polar angle of the B candidate in the CMS, $\cos\theta_B$ which follows a $(1-\cos^2\theta_B)$ distribution for $B\bar{B}$ events, while $q\bar{q}$ events distribute uniformly. Figure 3.7 shows the different distributions of $\cos\theta_B$. These

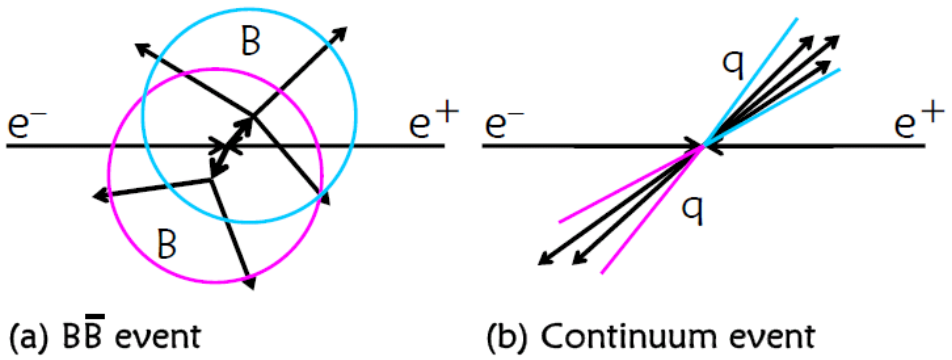


Figure 3.6: Schematically highlighting the topological differences between a $B\bar{B}$ event and a continuum event.

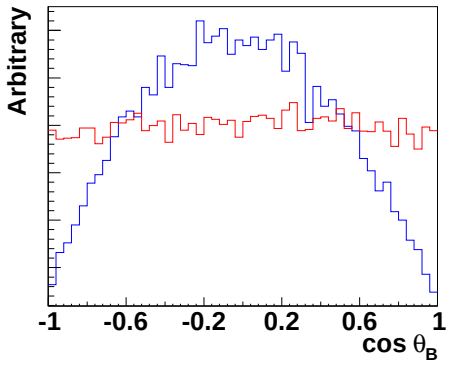


Figure 3.7: An example of the $\cos\theta_B$ distribution for $B\bar{B}$ events (blue) and continuum events (red).

two variables, F and $\cos\theta_B$, are combined to form a likelihood ratio, $\mathcal{L}_{\mathrm{S/B}}$ defined as

$$\begin{aligned} \mathcal{L}_{\mathrm{S/B}} &\equiv \frac{L_{B\bar{B}}}{L_{B\bar{B}} + L_{q\bar{q}}}, \\ L_{B\bar{B}} &\equiv P^{F}_{B\bar{B}} \times P^{\cos\theta_B}_{B\bar{B}}, \\ L_{q\bar{q}} &\equiv P^{F}_{q\bar{q}} \times P^{\cos\theta_B}_{q\bar{q}}, \end{aligned} \tag{3.7}$$

where $P_{B\bar{B}\,(q\bar{q})}$ is a probability density function of $B\bar{B}$ ($q\bar{q}$) events for F and $\cos\theta_B$. Figure 3.8 shows an example distribution of $\mathcal{L}_{\mathrm{S/B}}$.

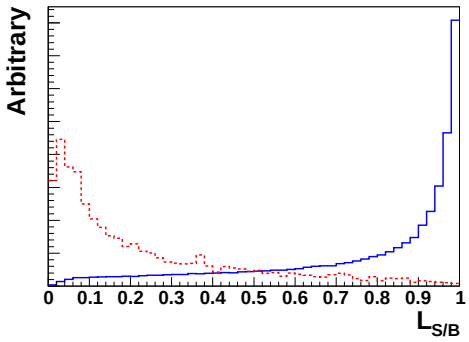


Figure 3.8: An example $\mathcal{L}_{\mathrm{S/B}}$ distribution. The blue curve represents $B\bar{B}$ events and the red curve shows $q\bar{q}$ events.

3.10 B meson Selection Criteria

3.10.1 $B^0 \to K^0_S\pi^0$

We reconstruct $B^0 \to K^0_S\pi^0$ events using reconstructed K^0_S and π^0 candidates. We select B_{Rec} candidates requiring $0.3 < \mathcal{L}_{\mathrm{S/B}}$, which retains $\sim 90.7\%$ of the signal while rejecting $\sim 70.5\%$ of the continuum background.

Signal events are described by the variables, M_{bc}, ΔE and $\mathcal{L}_{\mathrm{S/B}}$ for which the following spaces are defined:

the Variable space,

$5.20\ \mathrm{GeV}/c^2 < M_{\mathrm{bc}} < 5.30\ \mathrm{GeV}/c^2$,

$-0.30\ \mathrm{GeV} < \Delta E < 0.20\ \mathrm{GeV}$, and

$0.3 < \mathcal{L}_{\mathrm{S/B}}$;

the Signal region,

$5.27\ \mathrm{GeV}/c^2 < M_{\mathrm{bc}} < 5.29\ \mathrm{GeV}/c^2$,

$-0.15\ \mathrm{GeV} < \Delta E < 0.1\ \mathrm{GeV}$, and

$0.3 < \mathcal{L}_{\mathrm{S/B}}$;

the Sideband region,

$5.20\ \mathrm{GeV}/c^2 < M_{\mathrm{bc}} < 5.26\ \mathrm{GeV}/c^2$,

$0.05\ \text{GeV} < \Delta E < 0.2\ \text{GeV}$, and

$0.3 < \mathcal{L}_{\text{S/B}}$.

Due to the combinatorial nature of reconstruction, more than one B_{Rec} candidate per event is a possibility. The multiplicity of B_{Rec} candidates is ~ 1.007 events in B_{Rec} variable space and is illustrated in Fig. 3.9. However, as there can be only one reconstructed $B^0 \to K^0_S\pi^0$ per event, it is necessary to devise best B candidate selection criteria. For the case where B candidates were reconstructed from different π^0's, the B_{Rec} candidate with the lowest χ^2 for the π^0 mass-constrained fit is chosen. Otherwise, in the case of different K^0_S candidates, the B_{Rec} candidate is selected at random. After the best candidate selection, the fraction of mis-reconstructed events is estimated to be $\sim 20\%$, which is estimated from Monte Carlo simulation, and is almost completely dominated by fake π^0.

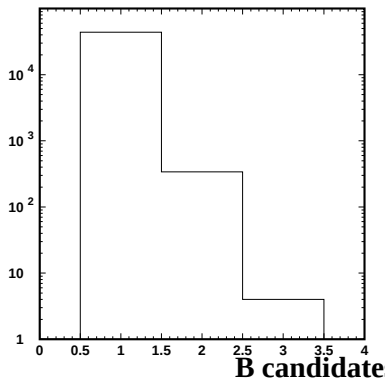


Figure 3.9: The number of B_{Rec} candidates per event.

Figure 3.10 shows the distribution of $\mathcal{L}_{\text{S/B}}$ vs M_{bc}, in which clear separation between signal and background events can be seen. The next chapter will describe how the different contributions to this plot are modelled in order to extract the signal yield.

3.10.2 $B^+ \to K^+\pi^0$

Since the extraction of the signal yield of $B^0 \to K^0_S\pi^0$ involves the kinematic variables, M_{bc} and ΔE, whose shapes are largely influenced by the presence of the π^0. A control sample with higher statistics, similar topology and a prompt π^0 is needed for which $B^+ \to K^+\pi^0$ is a natural choice.

The selection criteria for $B^+ \to K^+\pi^0$ are kept as similar as possible to that of $B^0 \to K^0_S\pi^0$. The main difference is that the charged kaon is selected with a PID requirement, $\mathcal{P}(K : \pi) > 0.6$.

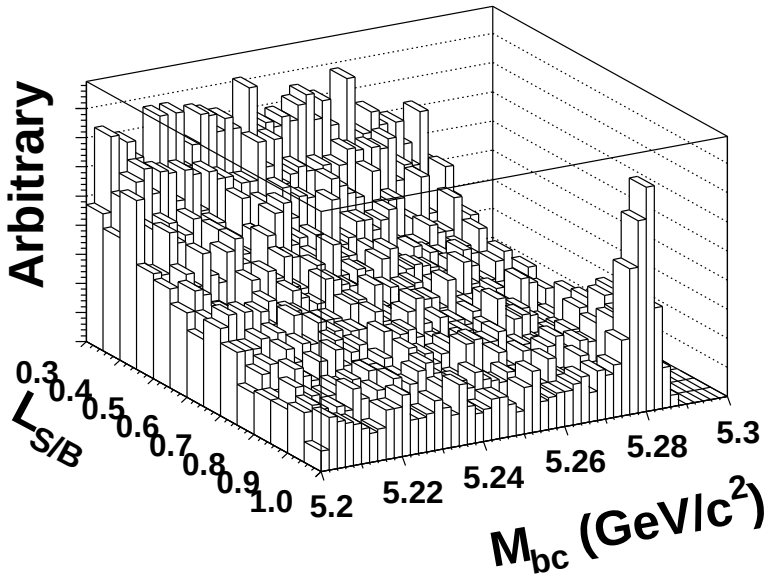


Figure 3.10: $\mathcal{L}_{\mathrm{S/B}}$ vs M_{bc}.

3.10.3 $B^+ \to K^0_S \pi^+$

A second control sample will be needed for the time-dependent CP analysis which uses the variable, Δt. As Δt is determined from a vertex using a K^0_S pseudo-track, $B^+ \to K^0_S \pi^+$ is chosen as a higher statistics control sample with similar topology. Only the K^0_S is used to reconstruct the vertex of the control sample.

The main difference for the reconstruction of $B^+ \to K^0_S \pi^+$ is that the prompt pion must satisfy $\mathcal{P}(K : \pi) < 0.4$ and that the best B candidate is selected at random if there are more than one candidates.

Chapter 4

Branching Fraction Measurement

4.1 Basic Formula

The branching fraction of a fully reconstructed B decay is given by

$$\mathcal{B} = \frac{N_{\rm Sig}}{\epsilon N(B\bar{B})}, \tag{4.1}$$

where $N_{\rm Sig}$ is the signal yield, ϵ is the reconstruction efficiency and $N(B\bar{B})$ is the number of $B\bar{B}$ events given in §3.2. Here we assume $N(B^+B^-) = N(B^0\bar{B}^0)$ and $N(B\bar{B}) = N(B^+B^-) + N(B^0\bar{B}^0)$. This chapter describes the procedures to obtain these parameters required to calculate the branching fraction of $B^0 \to K^0_S\pi^0$.

4.2 Calculation of $N(B\bar{B})$

The number of $B\bar{B}$ pairs passing the event pre-selection criteria of §3.3.2, $N(B\bar{B})$ is given by

$$N(B\bar{B}) = N(\text{on}) - acN(\text{off}), \tag{4.2}$$

where $N(\text{on})$ is the number of events collected at the $\Upsilon(4S)$ resonance, $N(\text{off})$ is the number of events collected off resonance, a is the scaling factor for on resonance to off resonance data and c is the ratio of the continuum efficiency for on resonance and off resonance [56].

We calculate a using both barrel bhabha and dimuon events, $a = N_{ll}(\text{on})/N_{ll}(\text{off})$, and take the average of the two. We calculate c as $c = \epsilon(\text{on})/\epsilon(\text{off})$, using simulated events.

4.3 Reconstruction Efficiency

Simulated events, or Monte Carlo (MC), play an important role in forming an expectation on how a decay will behave inside the Belle detector. This process of MC production occurs in two stages. Firstly, the physical decay vectors are simulated with the EvtGen [57] package according to some prescribed decay models. Secondly, the response of the Belle detector is modelled with GEANT [58], a package designed to imitate the interactions between particles and matter. GEANT takes each particle from EvtGen and traces its behaviour as it traverses the detector and returns an event intended to resemble real detector output.

The probability of detecting a given decay is known as the reconstruction efficiency. In its simplest form, the reconstruction efficiency can be determined from signal MC, ϵ_{MC}, as the number of reconstructed events in variable space out of the total number of simulated events. We generate samples of MC containing 1 000 000 signal events for SVD1 and SVD2. The reconstruction efficiencies of each decay are summarised in Table 4.1 where correct reconstruction was not a requirement.

Table 4.1: Reconstruction efficiency obtained from signal MC.

	$B^0 \to K^0_S\pi^0$	$B^+ \to K^+\pi^0$
SVD1	0.2182 ± 0.0005	0.3310 ± 0.0006
SVD2	0.2286 ± 0.0005	0.3362 ± 0.0006

Sometimes, a given reconstruction procedure may affect the detection efficiency of MC in a different way to data. In such cases, it becomes necessary to introduce a correction factor that scales a MC determined efficiency into one more appropriate for data. In general, the correction factor for a reconstruction procedure, i, is the ratio,

$$\xi_{\mathrm{CF}}(i) \equiv \frac{\epsilon_{\mathrm{Data}}(i)}{\epsilon_{\mathrm{MC}}(i)} \tag{4.3}$$

and is usually obtained by a separate study. The corrected efficiency, ϵ_{Data}, is simply the efficiency taken from MC scaled by the product of each correction factor,

$$\epsilon_{\mathrm{Data}} = \epsilon_{\mathrm{MC}} \prod_i \xi_{\mathrm{CF}}(i) \tag{4.4}$$

Two methods used in $B^0 \to K^0_S\pi^0$ reconstruction have associated correction factors, π^0 reconstruction [59] and K^0_S reconstruction [53]. These references provide lookup tables with correction factors that depend on the momentum of the K^0_S or π^0.

The correction factors for π^0 are obtained using the signal yields of η decays and calculating the ratios,

$$\frac{\epsilon_{\mathrm{Data}}(2\pi^0)}{\epsilon_{\mathrm{MC}}(2\pi^0)} = \frac{N_{\mathrm{Data}}(\eta \to 3\pi^0)/N_{\mathrm{MC}}(\eta \to 3\pi^0)}{N_{\mathrm{Data}}(\eta \to \gamma\gamma)/N_{\mathrm{MC}}(\eta \to \gamma\gamma)}, \tag{4.5}$$

Table 4.2: Efficiency correction factors

Category	Correction Factor
π^0 reconstruction	0.979 ± 0.027
K^0_S reconstruction	1.000 ± 0.018

and

$$\frac{\epsilon_{\rm Data}(2\pi^0)}{\epsilon_{\rm MC}(2\pi^0)} = \frac{N_{\rm Data}(\eta \to 3\pi^0)/N_{\rm MC}(\eta \to 3\pi^0)}{N_{\rm Data}(\eta \to \pi^+\pi^-\pi^0)/N_{\rm MC}(\eta \to \pi^+\pi^-\pi^0)}, \tag{4.6}$$

where the correction factor can be calculated as

$$\xi_{\rm CF}(\pi^0) = \sqrt{\frac{\epsilon_{\rm Data}(2\pi^0)}{\epsilon_{\rm MC}(2\pi^0)}}. \tag{4.7}$$

The K^0_S correction factors are obtained by comparing K^0_S signal yields in data and MC with and without the "goodKs" selection criteria. The correction factors for $B^0 \to K^0_S\pi^0$ are determined from signal MC by taking an average of the correction factors over all events and are shown in Table 4.2.

4.4 Branching Fraction PDF

The branching fractions of $B^0 \to K^0_S\pi^0$ and the control samples are determined from an extended three-dimensional unbinned maximum likelihood fit to the $M_{\rm bc}$-ΔE-$\mathcal{L}_{\rm S/B}$ distribution. The unbinned maximum likelihood method is discussed in Appendix A.

4.4.1 Signal PDF

In order to extract the signal yield from data, it is important to understand the signal shape in $M_{\rm bc}$, ΔE and $\mathcal{L}_{\rm S/B}$. Signal MC provides a way to form an expectation of the shape the signal will take in data. The signal shape is described by a probability density function (PDF) which is a three-dimensional histogram of all the reconstructed events in signal MC,

$$\mathcal{P}_{\rm Sig} \equiv H_{\rm Sig}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B}). \tag{4.8}$$

Usually, analytical models are preferred to describe signal, however, these were found to be insufficient due to strong correlations between $M_{\rm bc}$ and ΔE. As this PDF is determined from MC, it may not be an accurate portrayal of the signal shape in data. For modes with a π^0 in the final state, we introduce a smearing function, $S(M_{\rm bc}, \Delta E; \mu_i, \sigma_i)$, with parameters, μ_i, which shifts the histogram in $M_{\rm bc}$ and ΔE, respectively, and σ_i, which scales the histogram about the nominal B mass. For these cases, the signal PDF becomes,

$$\mathcal{P}_{\rm Sig} \equiv H_{\rm Sig}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B}) \otimes S(M_{\rm bc}, \Delta E). \tag{4.9}$$

4.4.2 Continuum Background PDF

The continuum background shape is modelled with the product,

$$\mathcal{P}_{q\bar{q}} \equiv \mathcal{P}_{q\bar{q}}(M_{\rm bc})\mathcal{P}_{q\bar{q}}(\Delta E)\mathcal{P}_{q\bar{q}}(\mathcal{L}_{\rm S/B}), \tag{4.10}$$

where $\mathcal{P}_{q\bar{q}}(M_{\rm bc})$ is an ARGUS function [60],

$$\mathcal{P}_{q\bar{q}}(M_{\rm bc}) \equiv M_{\rm bc}\sqrt{1-\left(\frac{M_{\rm bc}}{E^{\rm CMS}_{\rm beam}}\right)^2}\exp\left\{a\left[1-\left(\frac{M_{\rm bc}}{E^{\rm CMS}_{\rm beam}}\right)^2\right]\right\}, \tag{4.11}$$

with the shape parameter, a, which is a free parameter in the branching fraction measurement. The $\mathcal{P}_{q\bar{q}}(\Delta E)$ component is a first-order polynomial,

$$\mathcal{P}_{q\bar{q}}(\Delta E) \equiv 1 + c(\mathcal{L}_{\rm S/B})\Delta E. \tag{4.12}$$

As there is a correlation between the ΔE slope, c and $\mathcal{L}_{\rm S/B}$, the parameter, c, will be determined in three $\mathcal{L}_{\rm S/B}$ bins, $0.3 < \mathcal{L}_{\rm S/B} \leq 0.5$, $0.5 < \mathcal{L}_{\rm S/B} \leq 0.7$ and $0.7 < \mathcal{L}_{\rm S/B} \leq 1.0$, in the branching fraction measurement. The last component, $\mathcal{P}_{q\bar{q}}(\mathcal{L}_{\rm S/B})$, is a one-dimensional histogram,

$$\mathcal{P}_{q\bar{q}}(\mathcal{L}_{\rm S/B}) \equiv H(\mathcal{L}_{\rm S/B}), \tag{4.13}$$

determined from reconstructed events in the sideband region.

4.4.3 $B\bar{B}$ Background PDF

The $B\bar{B}$ background is determined from a large sample of generic MC B decays corresponding to an integrated luminosity roughly an order of magnitude larger than the data set used in this analysis. The dominant contribution towards the background comes from charmless B decays. As the contribution of charm events to the total background is negligible, it is treated together with the charmless events. The background PDF is modelled with a three-dimensional histogram,

$$\mathcal{P}_{B\bar{B}} \equiv H_{B\bar{B}}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B}). \tag{4.14}$$

4.4.4 Total PDF

The total extended likelihood is given by the product over all N events,

$$\mathcal{L} = \frac{e^{-(N_{\rm Sig}+N_{\rm Bkg})}}{N!}\prod_{i=1}^{N}\{N_{\rm Sig}\mathcal{P}^i_{\rm Sig} + N_{\rm Bkg}[f_{q\bar{q}}\mathcal{P}^i_{q\bar{q}} + (1-f_{q\bar{q}})\mathcal{P}^i_{B\bar{B}}]\}, \tag{4.15}$$

where $f_{q\bar{q}}$ is the $q\bar{q}$ fraction and is a free parameter, $N_{\rm Sig}$ is the extracted signal yield and $N_{\rm Bkg}$ is the fitted number of background events. Equation (4.1) is substituted into the extended

likelihood so that the branching fraction becomes the free parameter rather than the signal yield. This substitution is slightly complicated by the different SVD configurations which lead to different efficiencies for SVD1 and SVD2. To account for this, the signal yield can be written as

$$N_{\rm Sig} = \mathcal{B} \sum_{i=\rm SVD1}^{\rm SVD2} N(B\bar{B})_i \, \epsilon_i. \tag{4.16}$$

4.5 Correction Factors

We perform a fit to extract the branching fraction of the control sample, $B^+ \to K^+\pi^0$, with the signal PDF that includes factors which account for the possible differences between data and MC (Eq. (4.9)) as free parameters. However, the signal efficiency of $B^+ \to K^+\pi^0$ does not contain these corrections as this measurement is not intended to be a result of this analysis. The fit result is summarised in Table 4.3 and shown in Fig. 4.1.

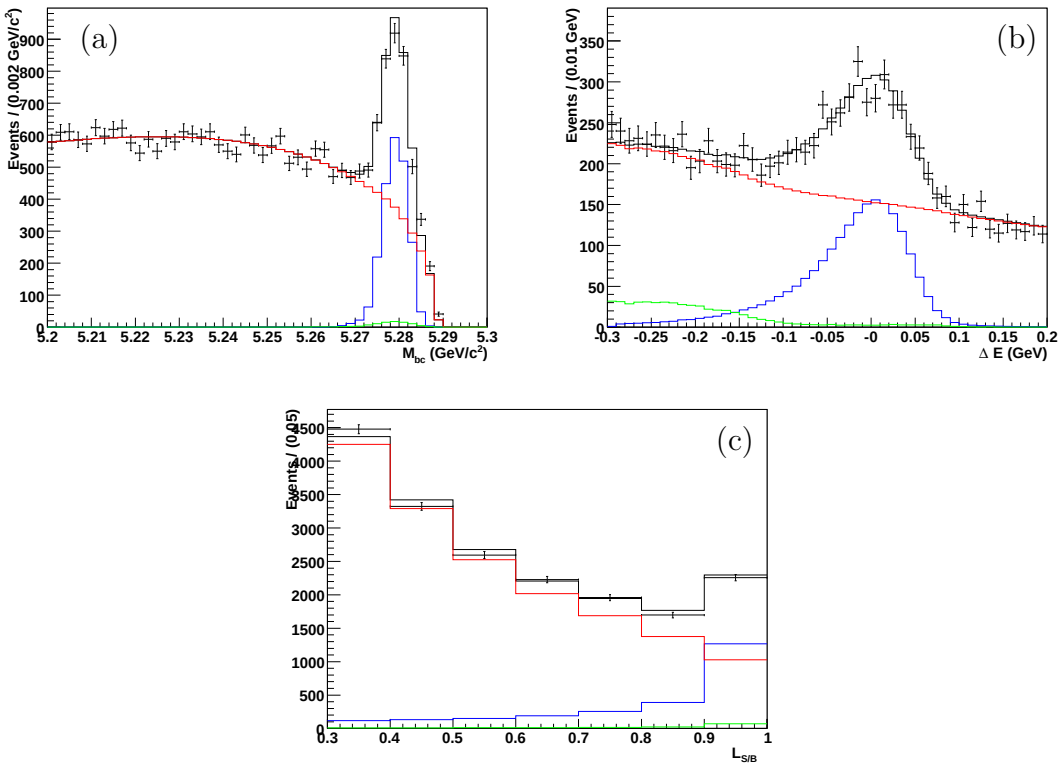


Figure 4.1: $M_{\rm bc}$-ΔE-$\mathcal{L}_{\rm S/B}$ fit projections of $B^+ \to K^+\pi^0$. The black histograms show the fit results. The blue histograms show the signal contribution, the green, the $B\bar{B}$ background and the red, the sum of continuum and $B\bar{B}$ background. (a) shows the $M_{\rm bc}$ projection within the ΔE signal region and enhanced $\mathcal{L}_{\rm S/B}$ region($\mathcal{L}_{\rm S/B} > 0.7$), (b) shows the ΔE projection within the $M_{\rm bc}$ signal region and enhanced $\mathcal{L}_{\rm S/B}$ region, and (c) shows the $\mathcal{L}_{\rm S/B}$ projection within the signal region.

Table 4.3: $B^+ \to K^+\pi^0$ fit result.

Parameter	Fit Result
$\mu_{M_{\rm bc}}$	-0.0006 ± 0.0004
$\sigma_{M_{\rm bc}}$	0.904 ± 0.013
$\mu_{\Delta E}$	0.007 ± 0.001
$\sigma_{\Delta E}$	0.999 ± 0.089
a	-21.83 ± 0.36
$c(0.3-0.5)$	-1.47 ± 0.03
$c(0.5-0.7)$	-1.33 ± 0.04
$c(0.7-1.0)$	-1.20 ± 0.05
$f_{q\bar{q}}$	0.991 ± 0.001
$\mathcal{B}(B^+ \to K^+\pi^0)$ (units 10^{-6})	9.966 ± 0.257
$N_{\rm Bkg}$ (units 10^3)	108.8 ± 0.3

4.6 Branching Fraction Measurement of $B^0 \to K^0_S\pi^0$

Finally, a fit is performed to extract the branching fraction of the signal mode, $B^0 \to K^0_S\pi^0$, using the signal PDF, Eq. (4.9), with correction factors fixed to the values found in the control sample $B^+ \to K^+\pi^0$. The signal detection efficiency is also corrected with the factors given in Table 4.2. The fit result is summarised in Table 4.4 and shown in Fig. 4.2. The branching

Table 4.4: $B^0 \to K^0_S\pi^0$ fit result.

Parameter	Fit Result
a	-23.29 ± 0.61
$c(0.3-0.5)$	-1.63 ± 0.06
$c(0.5-0.7)$	-1.30 ± 0.08
$c(0.7-1.0)$	-1.10 ± 0.08
$f_{q\bar{q}}$	0.993 ± 0.001
$\mathcal{B}(B^0 \to K^0_S\pi^0)$ (units 10^{-6})	4.362 ± 0.251
$N_{\rm Bkg}$ (units 10^3)	37.0 ± 0.2

fraction of the $B^0 \to K^0\pi^0$ decay which contains a scale factor of 2 to convert K^0_S to K^0, is found to be,

$$\boxed{\mathcal{B}(B^0 \to K^0\pi^0) = (8.72\,^{+0.51}_{-0.50}\,(\text{stat})\,^{+0.46}_{-0.40}\,(\text{syst})) \times 10^{-6}} \tag{4.17}$$

which corresponds to a signal yield of 634 ± 37 $B^0 \to K^0_S\pi^0$ events. This measurement is in agreement with the current world average, $\mathcal{B}(B^0 \to K^0\pi^0) = (9.8 \pm 0.6) \times 10^{-6}$ within 2 standard deviations.

The next section deals with the systematic uncertainties that originated from possible effects of various assumptions and fixed parameters used in the fit to obtain the branching fraction.

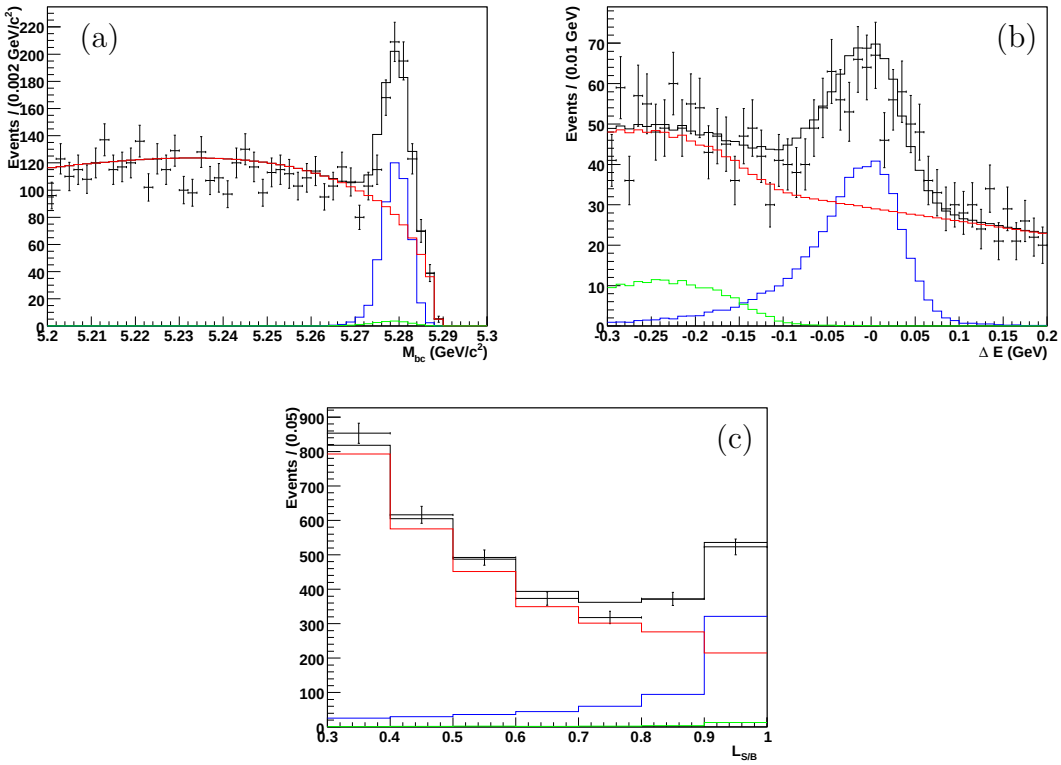


Figure 4.2: $M_{\rm bc}$-ΔE-$\mathcal{L}_{\rm S/B}$ fit projections of $B^0 \to K^0_S\pi^0$. The black histograms show the fit results. The blue histograms show the signal contribution, the green, the $B\bar{B}$ background and the red, the sum of continuum and $B\bar{B}$ background. (a) shows the $M_{\rm bc}$ projection within the ΔE signal region and enhanced $\mathcal{L}_{\rm S/B}$ region($\mathcal{L}_{\rm S/B} > 0.7$), (b) shows the ΔE projection within the $M_{\rm bc}$ signal region and enhanced $\mathcal{L}_{\rm S/B}$ region, and (c) shows the $\mathcal{L}_{\rm S/B}$ projection within the signal region.

4.7 Systematic Uncertainties

4.7.1 $N(B\overline{B})$

The number of $B\bar{B}$ pairs which appears in Eq. (4.1) has an associated error recorded at [56], which may affect the branching fraction measurement. The error is taken as a systematic uncertainty.

4.7.2 Efficiency

The signal detection efficiency also appears in Eq. (4.1) and its error depends on the amount of MC statistics in the histogram. This error is calculated from the values given in Table 4.1 by taking an average of the SVD1 and SVD2 efficiency errors for $B^0 \to K^0_S\pi^0$ weighted by

their relative amounts of data collected with each detector configuration.

4.7.3 K_S^0 Reconstruction

As described in §4.3, the K_S^0 selection criteria requires a correction factor for the efficiency to account for the difference between data and MC. The error on this correction factor given in Table 4.2 is taken as the systematic uncertainty of K_S^0 reconstruction.

4.7.4 π^0 Reconstruction

The systematic uncertainty of π^0 reconstruction is also due to its associated correction factor and is given in Table 4.2.

4.7.5 Histogram Binning

Histograms were used to represent certain components in the fit for the branching fraction, where the sizes of the histogram bins were fixed. To account for uncertainties in the shape, each bin size is varied by twice its bin error. Then a fit is performed noting the percentage difference between this fit result and the central value obtained in Eq. (4.17). This procedure is repeated for all bins of each histogram and the square root of the quadratic sum of the differences between each fit and the central value is taken as the systematic uncertainty.

4.7.6 Correction Factors

Correction factors obtained in §4.5 from the branching fraction measurement of $B^+ \to K^+\pi^0$ were used to smear the shape of $B^0 \to K_S^0\pi^0$ obtained from MC and were fixed in the fit to obtain the $B^0 \to K_S^0\pi^0$ branching fraction. The procedure to estimate the systematic error from these factors is similar to the case for the histogram bin sizes except that in each fit, the correction factors are varied by $\pm 1\sigma$.

4.7.7 Total Systematic Uncertainty

The systematic errors for each category are summarised in Table 4.5. The total systematic uncertainty of the branching fraction is taken as the quadratic sum of the uncertainties of each category.

Table 4.5: Branching fraction systematics.

Category	$\delta\mathcal{B}(\%)$
Number of $B\bar{B}$ pairs	± 1.36
Efficiency	± 0.21
K^0_S reconstruction	± 1.84
π^0 reconstruction	± 2.75
Histogram binning	$^{+1.47}_{-1.63}$
Correction factors	$^{+3.59}_{-2.37}$
Total	$^{+5.28}_{-4.60}$

4.8 Validity Test

We check branching fraction of the data sample collected with each SVD configuration. The fit results are,

$$\begin{aligned} \text{SVD1} &: \mathcal{B}(B^0 \to K^0\pi^0) = (9.00\,^{+1.07}_{-1.01}\,(\text{stat})) \times 10^{-6}, \\ \text{SVD2} &: \mathcal{B}(B^0 \to K^0\pi^0) = (8.63\,^{+0.58}_{-0.57}\,(\text{stat})) \times 10^{-6}, \end{aligned} \tag{4.18}$$

which are in agreement with each other and the final fit result.

Chapter 5

CP Violation Measurement

5.1 Overview

In this chapter, we describe the measurement of CP violation in $B^0 \to K^0\pi^0$ decays. Recall that CP violation parameters can be extracted using the physical observables Δt and the flavour of the tag-side B meson, using Eq. (1.56). Although this thesis is predominantly about $B^0 \to K^0_S\pi^0$, this part of the analysis was performed in conjunction with $B^0 \to K^0_L\pi^0$. Details of the $B^0 \to K^0_L\pi^0$ analysis will not be described here, but can be found in Ref. [61] and Appendix E.

5.2 Δt Reconstruction

5.2.1 Vertex Reconstruction of $B_{\rm Rec}$

In this analysis, there are no primary tracks originating from the B meson. In such cases, the vertex of $B_{\rm Rec}$, $z_{\rm Rec}$, has to be determined by extrapolating the K^0_S pseudo-track back to the IP. However, as the K^0_S has a relatively long lifetime, its charged daughters do not always register hits in the SVD which leads to a lower vertex reconstruction efficiency.

We can improve the $B_{\rm Rec}$ vertex resolution further by including the IP profile in the vertex fit. The shape of the IP profile is represented by a three-dimensional Gaussian distribution. Since the actual IP position and its spread varies with accelerator conditions, the mean and standard deviation of each Gaussian is determined frequently using pre-scaled hadronic events. The IP profile, which is typically 100 μm in x, 5 μm in y and 3 mm in z, is smeared in the $x-y$ plane to account for the finite flight length of the B meson in this plane, ${\rm IP}_{\rm smear} = 21$ μm.

5.2.2 Vertex Reconstruction of $B_{\rm Tag}$

After fully reconstructing $B_{\rm Rec}$, all remaining tracks should belong to the other B meson, $B_{\rm Tag}$. The following algorithm is employed to determine the vertex of $B_{\rm Tag}$, $z_{\rm Tag}$. Initially, all tracks with poor position measurement, $\sigma_z < 0.5$ mm, are excluded. Another requirement is placed on the impact parameter of the track with respect to the $B_{\rm Rec}$ vertex in the $r-\phi$ plane, $dr < 0.5$ mm.

Finally, the vertex of $B_{\rm Tag}$ is calculated with the remaining tracks and the IP constraint. To further minimise the effect of tracks with poor resolution and tracks with a relatively large displacement from the majority, the goodness-of-fit, χ^2, is checked. If $\chi^2 > 20$, the track providing the largest contribution to the χ^2 is removed and the vertex recalculated. However, we always keep high momentum leptons, $p_l^{\rm CMS} > 1.1$ GeV/c, as they are likely to originate from semi-leptonic $b \to cl\nu_l$ decays. The cycle of removing the worst track and re-fitting continues until $\chi^2 < 20$, whereupon the vertex is accepted as that of $B_{\rm Tag}$.

5.2.3 Δt Reconstruction

Time-dependent CP parameters are extracted from the Δt distribution which is calculated from the $B_{\rm Rec}$ and $B_{\rm Tag}$ vertices. As the $B\bar{B}$ pair is practically at rest in the$\Upsilon(4S)$ frame, Δt can be determined using a kinematic approximation,

$$\Delta t \sim \frac{\Delta z}{\beta\gamma c} \equiv \frac{z_{\rm Rec} - z_{\rm Tag}}{\beta\gamma c}, \tag{5.1}$$

where $\beta\gamma = 0.425$ is the Lorentz boost of the $\Upsilon(4S)$.

The quality of the vertex is usually determined by the χ^2 of the vertex fit. However, it is not a good index in this case because it is correlated with the vertex z-position due to the IP constraint in the $x-y$ plane. This would bias the calculated Δt distribution, so we therefore define a new measure of vertex quality,

$$\xi \equiv \frac{1}{2n}\sum_{i=1}^{n}\left[\frac{z^i_{\rm before} - z^i_{\rm after}}{\epsilon^i_{\rm before}}\right]^2, \tag{5.2}$$

which is essentially the reduced χ^2 projected along the z-axis. The parameters $z^i_{\rm before}$ and $z^i_{\rm after}$ are the vertex positions of each track before and after the fit, respectively, and $\epsilon^i_{\rm before}$ is the error of $z^i_{\rm before}$. Note that ξ cannot be defined for single-track vertices.

For Δt reconstruction, the following selection criteria are applied: the vertex goodness-of-fit, $\xi < 250$ for the reconstructed and tag-side vertices and $|\Delta t| < 70$ ps. For a good quality vertex, at least one $r-\phi$ hit and two z hits for each K^0_S daughter is required. The vertex reconstruction efficiency for SVD1 is found from MC to be 24% and 36% for SVD2. The vertex position resolution for the z direction with SVD1 (SVD2) is 141 μm (172 μm). The fraction of misreconstructed events in Δt is estimated to be $\sim 0.4\%$.

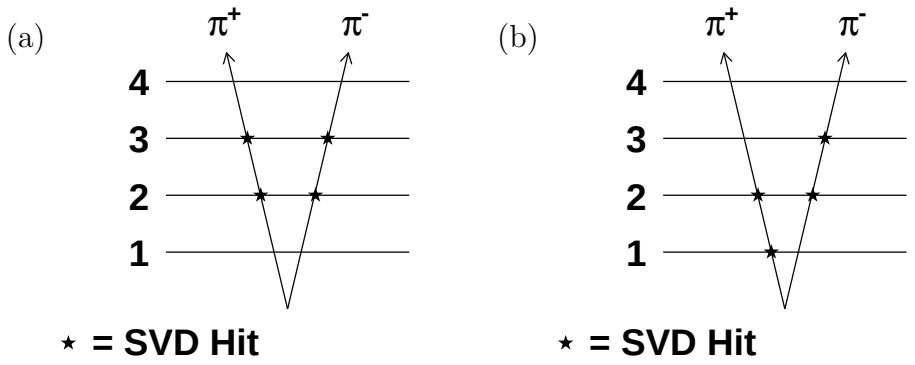


Figure 5.1: SVD configurations of K^0_S vertexing. (a) shows the "On-diagonal" configuration while (b) shows the "Off-diagonal".

Cosmic ray data taken with the Belle detector provide a better understanding of the track parametrisation. This is due to the fact that cosmic muon tracks have an entry and exit point through the SVD as opposed to a track which originated from inside the SVD which can only have an exit point. Thus the track resolution is further improved by correcting the track errors with global scaling factors obtained from these cosmic ray events.

To account for Δt smearing of signal events caused by our experimental environment, a resolution function, $R_{\rm Sig}$, is incorporated which considers three sources of Δt smearing,

$$R_{\rm Sig} \equiv R_{\rm Det} \otimes R_{\rm NP} \otimes R_{\rm K}. \tag{5.3}$$

The resolution function, $R_{\rm Det}$ represents the detector resolution, $R_{\rm NP}$ describes the smearing due to reconstructing the $B_{\rm Tag}$ vertex from non-primary tracks and $R_{\rm K}$ models the effects of the kinematic approximation that neglects the transverse momentum of the B mesons in the CMS. The details of the resolution function are contained in Appendix C.

The default resolution function is modified depending on the configuration of SVD hits of the K^0_S daughter tracks. Figure 5.1(a) shows the case where the innermost SVD hits of each track occur in the same layer and Figure 5.1(b) shows the case where they are not.

5.3 Flavour Tagging

Time-dependent CP asymmetry manifests only in CP eigenstates. As a consequence, the flavour of $B_{\rm Rec}$, or in other words, whether it was a B^0 or $\bar{B}^0$ that decayed, is unknown. By harnessing the property that the $B^0\bar{B}^0$ pair is produced in a coherent state, it is possible to infer the flavour of $B_{\rm Rec}$ at the same time $B_{\rm Tag}$ decayed if the flavour of $B_{\rm Tag}$ can be determined. This is called flavour tagging and the procedure is detailed in Appendix B. Two parameters are used to represent flavour information, q and r. The discrete variable, q, is the determined flavour of $B_{\rm Tag}$ and takes the value $q = +1$ when $B_{\rm Tag} = B^0$ and $q = -1$

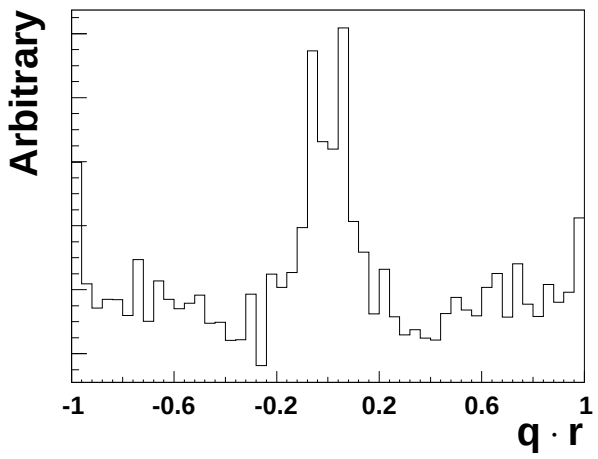


Figure 5.2: Distribution of $q \cdot r$. The edges show where $B_{\rm Tag}$ has been identified with high quality. The majority of events are in the middle with weak flavour discrimination.

when $B_{\rm Tag} = \bar{B}^0$. The continuous MC-determined r is the expected flavour dilution of q and can range from $r = 0$ for no flavour information, to $r = 1$ for unambiguous flavour assignment. Another two variables are used to describe incorrect flavour tagging, w, which is wrong tag probability and the difference in wrong tag probabilities between B^0 and $\bar{B}^0$ decays, Δw. Figure 5.2 shows a distribution of $q \cdot r$ obtained from signal MC.

5.4 Probability Density Function

5.4.1 Signal PDF

The time-dependent signal PDF which accounts for dilution from flavour tagging and resolution effects is given by

$$\mathcal{P}_{\rm Sig}(\Delta t, q) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}}\Big[1 - q\Delta w + q(1-2w)(\mathcal{A}_{CP}\cos\Delta m_d\Delta t - \mathcal{S}_{CP}^{\rm eff}\sin\Delta m_d\Delta t)\Big] \otimes R_{\rm Sig}, \tag{5.4}$$

where τ_{B^0} is the B^0 lifetime and Δm_d is the mass difference between the B mass eigenstates. Even though the vertex reconstruction efficiency is not perfect, events that lack Δt information or fail the Δt selection criteria can still be used to evaluate the $\mathcal{A}_{CP}$ component. This is achieved by integrating Eq. (5.4) over Δt,

$$\mathcal{P}_{\rm Sig}(q) = \frac{1}{2}\Big[1 - q\Delta w + q(1-2w_l)(1-2\chi_m)\mathcal{A}_{CP}\Big], \tag{5.5}$$

where,

$$\chi_m \equiv \frac{(\Delta m_d \Delta t)^2}{2[1+(\Delta m_d \Delta t)^2]}. \tag{5.6}$$

5.4.2 Continuum Background PDF

The continuum Δt distribution is empirically described with an effective lifetime and prompt component which is motivated by the finite lifetime of D mesons and the presence of short-lived particles in continuum,

$$\mathcal{P}_{q\bar{q}}(\Delta t) = (1-f_\delta)\frac{e^{-|\Delta t|/\tau_{q\bar{q}}}}{2\tau_{q\bar{q}}} + f_\delta \delta(\Delta t - \mu_\delta). \tag{5.7}$$

This PDF is convolved with a double Gaussian to describe Δt resolution effects,

$$R_{q\bar{q}}(\Delta t) = (1-f_{\rm tail})G(\Delta t; \mu_{\rm mean}, S_{\rm main}\sigma) + f_{\rm tail}G(\Delta t; \mu_{\rm mean}, S_{\rm main}S_{\rm tail}), \tag{5.8}$$

which uses the event-dependent Δt error constructed from the vertex resolution, $\sigma \equiv (\sqrt{\sigma^2_{\rm Rec}+\sigma^2_{\rm Tag}})/\beta\gamma c$, as a scale factor. The fit parameters are determined from the sideband region for the control sample, $B^+ \to K^0_S\pi^+$ and the signal mode. They can be found in Table 5.1 with the fit result shown in Fig. 5.3.

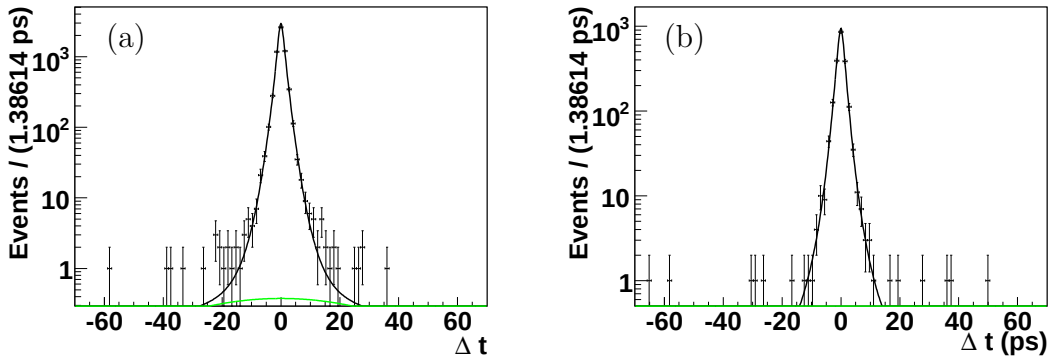


Figure 5.3: Fit results to determine the continuum background shape in Δt. The black curves show the fit results, and the green curve shows the contribution from an additional Gaussian similar to the one used in the lifetime measurement of signal MC. (a) shows the results for $B^+ \to K^0_S\pi^+$ and (b) shows $B^0 \to K^0_S\pi^0$.

5.4.3 Charged B Background PDF

The charged B background in Δt is modelled with a lifetime function,

$$\mathcal{P}_{B^+B^-}(\Delta t) = \frac{e^{-|\Delta t|/\tau_{B^+_{\rm eff}}}}{2\tau_{B^+_{\rm eff}}} \otimes R_{B^+B^-}, \tag{5.9}$$

Table 5.1: Fit parameters for continuum background in Δt.

Fit Parameter	$B^+ \to K^0_S\pi^+$	$B^0 \to K^0_S\pi^0$
$\tau_{\rm eff}$	0.875 ± 0.118	0.968 ± 0.086
μ_δ	0.013 ± 0.068	-0.031 ± 0.031
f_δ	0.475 ± 0.098	0.567 ± 0.063
$\mu_{\rm mean}$	-0.091 ± 0.092	0.135 ± 0.065
$S_{\rm main}$	1.126 ± 0.054	1.105 ± 0.034
$S_{\rm tail}$	3.520 ± 0.520	2.895 ± 0.291
$f_{\rm tail}$	0.073 ± 0.028	0.100 ± 0.026

where $R_{B^+B^-}$ is the Δt resolution function for charged B decays. As the source of this background involves particle exchange with the tag-side which may shift the vertex position, the lifetime may also be affected in consequence. We extract an effective lifetime from generic charged MC events to determine its shape in Δt. The fit result gives $\tau_{B^+_{\rm eff}} = 1.70 \pm 0.05$ ps for $B^+ \to K^0_S\pi^+$ and $\tau_{B^+_{\rm eff}} = 1.66 \pm 0.08$ ps for $B^0 \to K^0_S\pi^0$, and is shown in Fig. 5.4.

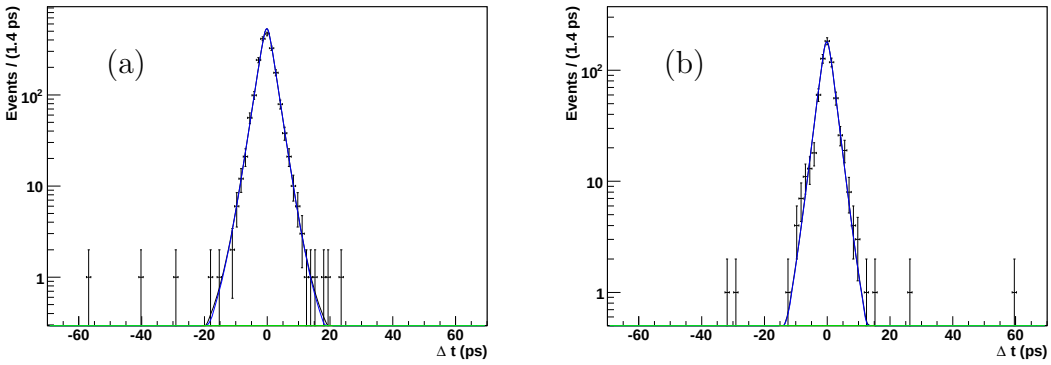


Figure 5.4: Fit results to determine the charged B background shape in Δt. The black curves show the fit results, and the green curves show the contribution from an additional Gaussian similar to the one used in the lifetime measurement of signal MC. (a) shows the results for $B^+ \to K^0_S\pi^+$ and (b) shows $B^0 \to K^0_S\pi^0$.

5.4.4 Neutral B Background PDF

Due to $B^0\bar{B}^0$ mixing, the neutral B background is described with the same PDF as signal, Eq. (5.4), however the CP parameters are fixed to zero and the effect of this choice will be accounted for in the systematic errors. For similar reasons to charged B events, we extract effective lifetimes to describe the neutral B background shape: $\tau_{B^0_{\rm eff}} = 1.68 \pm 0.11$ ps for $B^+ \to K^0_S\pi^+$ and $\tau_{B^0_{\rm eff}} = 1.21 \pm 0.21$ ps for $B^0 \to K^0_S\pi^0$. The fit result is shown in Fig. 5.5.

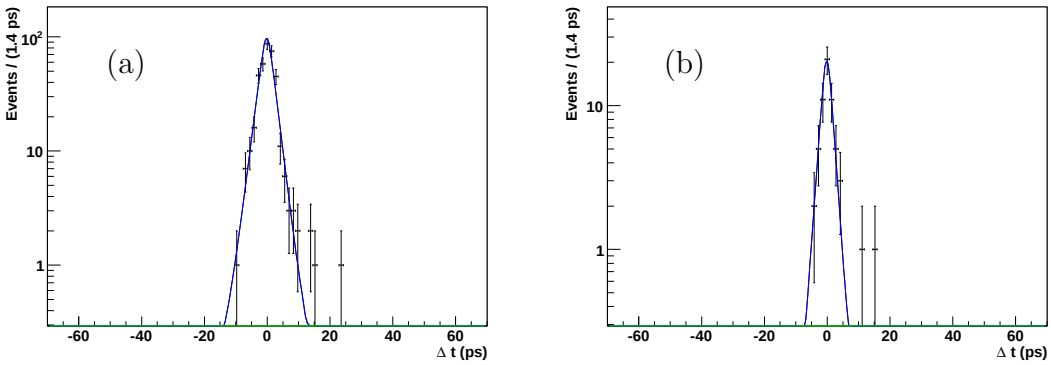


Figure 5.5: Fit results to determine the neutral B background shape in Δt. The black curves show the fit results, and the green curves show the contribution from an additional Gaussian similar to the one used in the lifetime measurement of signal MC. (a) shows the results for $B^+ \to K^0_S\pi^+$ and (b) shows $B^0 \to K^0_S\pi^0$.

5.4.5 Outlier PDF

At this point, there are a few remaining events that cannot be accounted for by the Δt resolution function or physics background PDFs. These so-called outlier events form a broad shape in Δt and are thought to originate from mis-reconstruction. They are modelled with a Gaussian centred around zero,

$$\mathcal{P}_{\rm Out}(\Delta t) = G(\Delta t; 0, \sigma_{\rm Out}), \tag{5.10}$$

where $\sigma_{\rm Out}$ is the width of the Gaussian.

5.4.6 Δt Component Probabilities

The relative contributions of each component in Δt is calculated in the following way. The ratio of B^+B^- decays to the total number of $B\bar{B}$ background events in the signal box, $f_{B^+B^-}$, is determined from MC to be $f_{B^+B^-} = 0.815 \pm 0.008$ for $B^+ \to K^0_S\pi^+$ and $f_{B^+B^-} = 0.915 \pm 0.018$ for $B^0 \to K^0_S\pi^0$. Discrimination between continuum and $B\bar{B}$ background is calculated event-by-event from their shapes in $M_{\rm bc}$, ΔE and $\mathcal{L}_{\rm S/B}$ as given in Eq. (4.10) and Eq. (4.14),

$$f_{q\bar{q}}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B}) = \frac{p_{q\bar{q}}\mathcal{P}_{q\bar{q}}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B})}{p_{q\bar{q}}\mathcal{P}_{q\bar{q}}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B}) + (1 - p_{q\bar{q}})\mathcal{P}_{B\bar{B}}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B})}. \tag{5.11}$$

The shapes are fixed from the signal yield extraction of the control sample, $B^+ \to K^0_S\pi^+$, and $B^0 \to K^0_S\pi^0$[†], described later in this section, where $p_{q\bar{q}}$ is the $q\bar{q}$ fraction of the total

[†]The signal yield extraction of $B^0 \to K^0_S\pi^0$ occurred chronologically before its branching fraction measurement in §4.6 and did not contain correction factors from $B^+ \to K^+\pi^0$.

background events in the signal box. The values of $p_{q\bar{q}}$ are calculated to be $p_{q\bar{q}} = 0.985 \pm 0.001$ for $B^+ \to K^0_S\pi^+$ and $p_{q\bar{q}} = 0.990 \pm 0.001$ for $B^0 \to K^0_S\pi^0$.

To maximise sensitivity to signal, the signal probability is calculated per event from the event shape in each r-bin,

$$f_{\rm Sig}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B}, r) = \frac{p_{\rm Sig}(r)\mathcal{P}_{\rm Sig}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B})}{p_{\rm Sig}(r)\mathcal{P}_{\rm Sig}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B}) + (1 - p_{\rm Sig}(r))\mathcal{P}_{\rm Bkg}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B})}, \tag{5.12}$$

where $\mathcal{P}_{\rm Sig}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B})$ is given in Eq. (4.8) and

$$\mathcal{P}_{\rm Bkg}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B}) = p_{q\bar{q}}\mathcal{P}_{q\bar{q}}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B}) + (1 - p_{q\bar{q}})\mathcal{P}_{B\bar{B}}(M_{\rm bc}, \Delta E, \mathcal{L}_{\rm S/B}). \tag{5.13}$$

The r-bin dependent signal purity, $p_{\rm Sig}(r)$, is calculated from signal MC and sideband scaled using the average purities in the signal region, $p_{\rm Sig}$. These values are given in Table 5.2.

	$B^+ \to K^0_S\pi^+$		$B^0 \to K^0_S\pi^0$	
r-bin	SVD1	SVD2	SVD1	SVD2
0	0.181 ± 0.011	0.180 ± 0.005	0.147 ± 0.018	0.120 ± 0.008
1	0.190 ± 0.011	0.181 ± 0.005	0.175 ± 0.021	0.154 ± 0.010
2	0.239 ± 0.013	0.222 ± 0.006	0.240 ± 0.026	0.195 ± 0.012
3	0.250 ± 0.013	0.232 ± 0.007	0.209 ± 0.024	0.187 ± 0.012
4	0.250 ± 0.013	0.246 ± 0.007	0.243 ± 0.026	0.176 ± 0.011
5	0.298 ± 0.015	0.315 ± 0.008	0.289 ± 0.030	0.247 ± 0.015
6	0.633 ± 0.017	0.661 ± 0.008	0.704 ± 0.032	0.613 ± 0.019
$p_{\rm Sig}$	0.241 ± 0.013	0.235 ± 0.007	0.221 ± 0.025	0.184 ± 0.012

Table 5.2: Summary of r-bin dependent purities.

Signal Yield Extraction of $B^+ \to K^0_S\pi^+$

We perform a fit to extract the signal yield of the control sample, $B^+ \to K^0_S\pi^+$, with the uncorrected signal PDF, Eq. (4.8), to be consistent with $B^0 \to K^0_S\pi^0$. The fit result is summarised in Table 5.3 and shown in Fig. 5.6. The signal yield of $B^+ \to K^0_S\pi^+$ is found to be,

$$N_{\rm Sig}(B^+ \to K^0_S\pi^+) = 2297 \pm 58 \text{ events}. \tag{5.14}$$

Signal Yield Extraction of $B^0 \to K^0_S\pi^0$

We perform a three-dimensional unbinned maximum likelihood fit to $B^0 \to K^0_S\pi^0$ events to extract the signal yield without fudge factors obtained from $B^+ \to K^+\pi^0$. The fit result gives $N_{\rm Sig} = 657 \pm 37$ events with the remaining parameters also summarised in Table 5.4. These parameters are in statistical agreement with the branching fraction measurement given in §4.6.

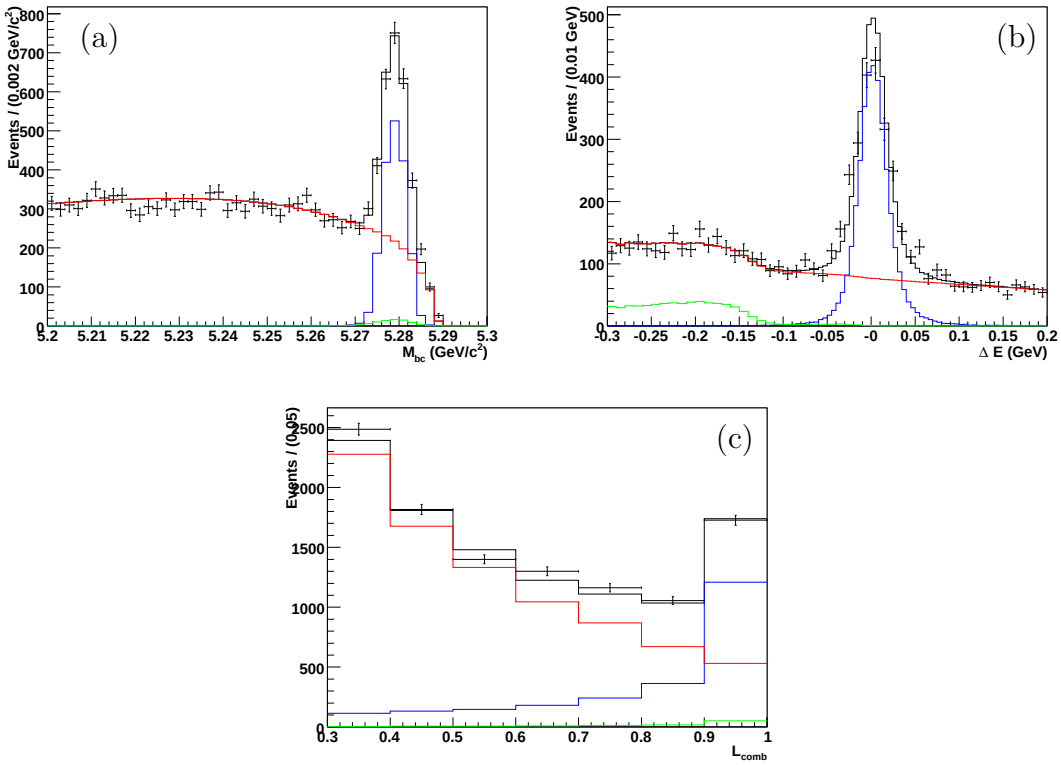


Figure 5.6: $M_{\rm bc}$-ΔE-$\mathcal{L}_{\rm S/B}$ fit projections of $B^+ \to K^0_S\pi^+$. The black histograms show the fit results. The blue histograms show the signal contribution, the green, the $B\bar{B}$ background and the red, the sum of continuum and $B\bar{B}$ background. (a) shows the $M_{\rm bc}$ projection within the ΔE signal region and enhanced $\mathcal{L}_{\rm S/B}$ region($\mathcal{L}_{\rm S/B} > 0.7$), (b) shows the ΔE projection within the $M_{\rm bc}$ signal region and enhanced $\mathcal{L}_{\rm S/B}$ region, and (c) shows the $\mathcal{L}_{\rm S/B}$ projection within the signal region.

5.4.7 Total CP Violation PDF

Combining Eq. (5.4), Eq. (5.7), Eq. (5.9) and Eq. (5.10) with their relative contributions given in §5.4.6, the full time-dependent CP violation PDF becomes

$$\begin{aligned}\mathcal{P}(\Delta t, q) = (1 - f_{\rm Out}) \Big[& f_{\rm Sig}\mathcal{P}_{\rm Sig} \\ & + \frac{1}{2}(1 - f_{\rm Sig}) f_{q\bar{q}} \mathcal{P}_{q\bar{q}} \\ & + \frac{1}{2}(1 - f_{\rm Sig})(1 - f_{q\bar{q}}) f_{B^+B^-} \mathcal{P}_{B^+B^-} \\ & + (1 - f_{\rm Sig})(1 - f_{q\bar{q}})(1 - f_{B^+B^-}) \mathcal{P}_{B^0\bar{B}^0} \Big] + \frac{1}{2} f_{\rm Out} \mathcal{P}_{\rm Out}. \quad (5.15)\end{aligned}$$

Table 5.3: $B^+ \to K^0_S\pi^+$ fit result.

Parameter	Fit Result
a	-21.66 ± 0.36
$c(0.3-0.5)$	-1.47 ± 0.03
$c(0.5-0.7)$	-1.33 ± 0.04
$c(0.7-1.0)$	-1.20 ± 0.05
$f_{q\bar{q}}$	0.991 ± 0.001
$N_{\rm Sig}$	2297 ± 58
$N_{\rm Bkg}$ (units 10^3)	108.6 ± 0.3

Table 5.4: $B^0 \to K^0_S\pi^0$ fit result.

Parameter	Fit Result
a	-23.3 ± 0.6
$c(0.3-0.5)$	-1.63 ± 0.06
$c(0.5-0.7)$	-1.30 ± 0.08
$c(0.7-1.0)$	-1.09 ± 0.08
$f_{q\bar{q}}$	0.993 ± 0.001
$N_{\rm Sig}$	657 ± 37
$N_{\rm Bkg}$ (units 10^3)	37.0 ± 0.2

The factor of 1/2 normalises over flavour for PDFs that are not flavour-dependent. Similarly, the CP violating PDF for events without Δt information is constructed from Eq. (5.5),

$$\begin{aligned}\mathcal{P}(q) = & f_{\rm Sig}\mathcal{P}_{\rm Sig} \\ & + \frac{1}{2}(1-f_{\rm Sig})f_{q\bar{q}} \\ & + \frac{1}{2}(1-f_{\rm Sig})(1-f_{q\bar{q}})f_{B^+B^-} \\ & + (1-f_{\rm Sig})(1-f_{q\bar{q}})(1-f_{B^+B^-})\mathcal{P}_{B^0\bar{B}^0}. \end{aligned} \tag{5.16}$$

5.5 Time-dependent CP Violation in $B^0 \to K^0\pi^0$

We perform a time-dependent CP violation fit to $B^0 \to K^0_S\pi^0$ and $B^0 \to K^0_L\pi^0$ events and measure the CP parameters,

$$\boxed{\begin{array}{l}\mathcal{A}_{CP} = +0.138 \pm 0.126\ (\text{stat}) \pm 0.064\ (\text{syst}), \\ \mathcal{S}^{\rm eff}_{CP} = +0.667 \pm 0.312\ (\text{stat}) \pm 0.088\ (\text{syst}),\end{array}} \tag{5.17}$$

where the fit results are shown in Fig. 5.7. There is no evidence for direct CP violation and the mixing-induced component is consistent with that measured in $b \to c\bar{c}s$ decays, $\mathcal{S}_{CP} = 0.672 \pm 0.024$ [20]. These results are also in agreement with the previous Belle results [46] and the recent results from BaBar [45].

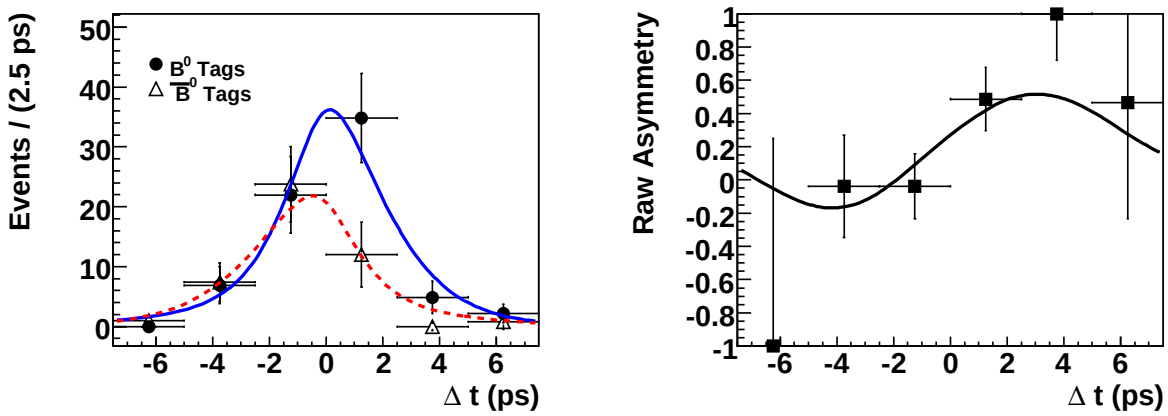


Figure 5.7: CP violation fit result for $B^0 \to K^0\pi^0$ events with Δt information. The left plot shows the background subtracted Δt distribution for B^0 and $\bar{B}^0$ tags where the solid (dashed) curve represents the Δt curve for B^0 ($\bar{B}^0$) in the good tag region $0.5 < r \leq 1.0$. The right plot shows the background subtracted asymmetry defined as $(N^{\rm Sig}_{B^0} - N^{\rm Sig}_{\bar{B}^0})/N^{\rm Sig}_{B^0} + N^{\rm Sig}_{\bar{B}^0})$ in each Δt bin where $N^{\rm Sig}_{B^0}$ ($N^{\rm Sig}_{\bar{B}^0}$) is the B^0 ($\bar{B}^0$) tagged signal yield extracted in that Δt bin.

5.6 Systematics Uncertainties

The sources of systematic errors in the CP measurement are described in the following sections. As a general rule, parameters obtained from data are varied by $\pm 1\sigma$ while parameters obtained from MC are varied by $\pm 2\sigma$. Traditionally, asymmetric systematic errors are symmetrised by taking the larger of the two. Unless otherwise stated, all systematic uncertainties are determined simultaneously with $B^0 \to K^0_L\pi^0$.

5.6.1 Vertex Reconstruction

IP Profile

The vertices of $B_{\rm Rec}$ and $B_{\rm Tag}$ were constructed with an IP constraint smeared in the $x - y$ plane by 21 μm as mentioned in §5.2.1. This amount of smearing is varied by ± 10 μm and the CP violation measurement is repeated. The difference between the fit result with varied smearing and the nominal fit result is taken as a systematic uncertainty.

$B_{\rm Tag}$ Track Selection

The track selection criteria on the tag-side involving σ_z and dr, given in §5.2.2, are sequentially altered by $\pm 10\%$. The CP violation measurement is repeated for each modified criteria and the fit differences are summed in quadrature to estimate the systematic uncertainty in

this category.

Scale Error

The charged track parametrisation errors were corrected by scaling factors as noted in §5.2.3. The effect of these corrections is studied by taking the difference in fit results without scaled errors and the final fit result as a systematic error.

Δt Selection

The Δt requirement given in §5.2.3 is varied by ± 30 ps and the fit difference taken as a systematic error.

Vertex Selection

The selection criteria on the B vertex quality, ξ_{Rec} and ξ_{Tag}, as stated in §5.2.3, are modified in sequence by ± 50 and the fit result differences are summed in quadrature.

Δz Bias

After the resolution function is applied to Δz, the average of the resulting distribution may not be zero. The source of this Δz bias is caused by relative misalignment between the SVD and CDC. This effect was studied in $B^0 \to J/\psi K^0_S$ [62] by applying correction functions to account for observed biases and taking the difference between the fit result with corrections and the default fit result. As this bias is considered to be mode-independent, this systematic uncertainty is taken from the $B^0 \to J/\psi K^0_S$ study.

Misalignment

There is thought to be an unknown intrinsic alignment fault within the SVD. This effect can be estimated by generating MC with and without misalignment and taking the difference in CP measurement fit results as a systematic error. Similar to Δz bias, this error is not assumed to be unique to each decay mode, so the systematic uncertainty is also taken from $B^0 \to J/\psi K^0_S$ [62].

5.6.2 Flavour Tagging

Wrong Tag Fractions

The wrong tag fractions and wrong tag fraction difference in each r-bin, w and Δw are varied by their respective errors given in Appendix B. The systematic uncertainty is taken as the quadratic sum of the fit differences.

Decay Mode Dependence

The wrong tag fractions were obtained from control samples that contain at least one primary track on the reconstructed side. We consider a possible effect of the $B^0 \to K^0_S\pi^0$ hadronic environment on the wrong tag fractions by creating wrong tag fractions from signal MC. We take the difference in fit results to signal MC using the $B^0 \to K^0_S\pi^0$ MC-determined wrong tag fractions and the default MC-determined wrong tag fractions as a systematic error.

5.6.3 Resolution Function

Both the neutral and charged B resolution functions for Δt are used and so the parameters in each resolution function given in Appendix C are altered sequentially by their errors. The systematic uncertainty is the quadratic sum of the fit differences.

5.6.4 Fit Bias

Fit bias may occur in samples with low statistics. To account for this, an ensemble of 1000 pseudo-experiments is created using the final fit result as input. The CP measurement is then performed for each experiment and the distribution of fit residuals is itself fitted with a Gaussian. The determined mean is taken as the systematic error due to fit bias.

5.6.5 Physics Parameters

The world-averages for τ_{B^0} and Δm_d [48], were fixed in the signal PDF for the CP measurement. These parameters are varied by their respective errors and the fit differences are summed in quadrature.

5.6.6 Background

Continuum Δt Shape

The parameters of the $q\bar{q}$ background shape in Δt as stated in §5.4.2 are varied by their errors and the systematic uncertainty determined as the quadratic sum of fit differences.

Effective Lifetime

The effective lifetimes found for charged and neutral B events given in §5.4.3 and §5.4.4 are varied by their respective errors.

CP Contribution

We assumed no CP asymmetry in the background as reflected by the choice of CP parameters in the neutral B background PDF noted in §5.4.4. As a conservative estimate of the systematic error associated with this assumption, the CP parameters of the $B^0\bar{B}^0$ PDF are varied within their physical limits and the fit differences summed in quadrature.

5.6.7 Signal Fraction

The r-bin dependent purities and signal shape parameters of §5.4.6 used to calculate the signal and background probabilities are varied by their errors. Histograms were also used in some instances and the effects of these non-parametric shapes are accounted for by simultaneously varying each bin by its error.

Signal Correction Factors

Recall that we obtained signal shape correction factors from $B^+ \to K^+\pi^0$ in §4.5. As these correction factors were not applied in the CP measurement, we perform another fit using the corrected signal shape taken from the branching fraction measurement given in §4.6 and take the fit difference with and without these correction factors as a systematic uncertainty.

5.6.8 Mis-reconstructed Events

A possible effect from mis-reconstruction is studied by comparing the CP fit results of signal MC with its sub-sample that contains only correctly reconstructed vertices. The fit difference is found to be negligible.

5.6.9 Tag-side Interference

Although the flavour of the tagging side is usually determined by flavour specific events, modes that are flavour non-specific also exist. For example, $\bar{B}^0 \to D^+ \pi^-$ is CKM-favoured, while the CKM-suppressed $B^0 \to D^+ \pi^-$ can also exist which leads to CP violation on the tag side [63]. The correction to the CP violation PDF is estimated with $B^0 \to D^* l\nu$, then an ensemble of pseudo-experiments are generated with and without tag-side interference and the mean fit difference is taken as a systematic error.

5.6.10 Total Systematic Uncertainty

A summary of the systematic uncertainties for the CP violation measurement is given in Table 5.5. The errors for each category are summed in quadrature to determine the total systematic uncertainty.

5.7 Pre-Measurement Validity Tests

In this section, we summarise validity tests that we performed before the final time-dependent CP violation fit.

5.7.1 Signal PDF Test

We search for possible reconstruction bias in Δt using $B^0 \to K^0_S \pi^0$ MC by fitting the generated Δt of reconstructed events with an exponential function,

$$\mathcal{P}_{\rm Sig}(\Delta t) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{2\tau_{B^0}}. \tag{5.18}$$

The lifetimes were found to be $\tau_{B^0} = 1.533 \pm 0.007$ ps for SVD1 and $\tau_{B^0} = 1.538 \pm 0.006$ ps for SVD2 which are in agreement with the MC generated lifetime, $\tau_{B^0} = 1.534$ ps. The fit results are shown in Fig. 5.8. Thus, there is no indication that the reconstruction procedure favours a different lifetime.

To test the resolution function, a lifetime fit is performed to the reconstructed Δt of the same signal MC sample with a lifetime PDF convolved with the Δt resolution function,

$$\mathcal{P}_{\rm Sig}(\Delta t) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{2\tau_{B^0}} \otimes R_{\rm Sig}. \tag{5.19}$$

The extracted lifetimes are $\tau_{B^0} = 1.513 \pm 0.009$ ps for SVD1 and $\tau_{B^0} = 1.533 \pm 0.007$ ps for SVD2 which are also in agreement with the MC generated lifetime. The fit results are shown in Fig. 5.9. At this point, the Δt resolution function appears to adequately describe

Table 5.5: CP violation measurement systematic uncertainties.

Category	Sub-category	$\delta(\mathcal{A}_{CP})$	$\delta(\mathcal{S}_{CP}^{\text{eff}})$
Vertex Reconstruction	Sub-total	0.022	0.013
	IP Profile	0.002	0.004
	B_{Tag} Track Selection	0.001	0.004
	Scale Error	0.003	0.000
	Δt Selection	0.007	0.001
	Vertex Selection	0.000	0.010
	Δz Bias	0.020	0.005
	Misalignment	0.004	0.006
Flavour Tagging	Sub-total	0.005	0.007
	Wrong Tag Fractions	0.005	0.007
	Decay Mode Dependence	0.003	0.001
Resolution Function		0.007	0.063
Fit Bias		0.010	0.020
Physics Parameters		0.001	0.007
Background	Sub-total	0.006	0.015
	$q\bar{q}$ Δt Shape	0.002	0.015
	Effective Lifetime	0.000	0.000
	CP Contribution	0.006	0.001
Signal Fraction	Sub-total	0.023	0.052
	Background Fractions	0.000	0.000
	Histograms	0.006	0.012
	Purity	0.006	0.018
	Signal Correction Factors	0.021	0.047
	K_L Signal Prob	0.002	0.000
Mis-reconstructed Events		0.000	0.000
Tag-side Interference		0.054	0.014
Total		**0.064**	**0.077**

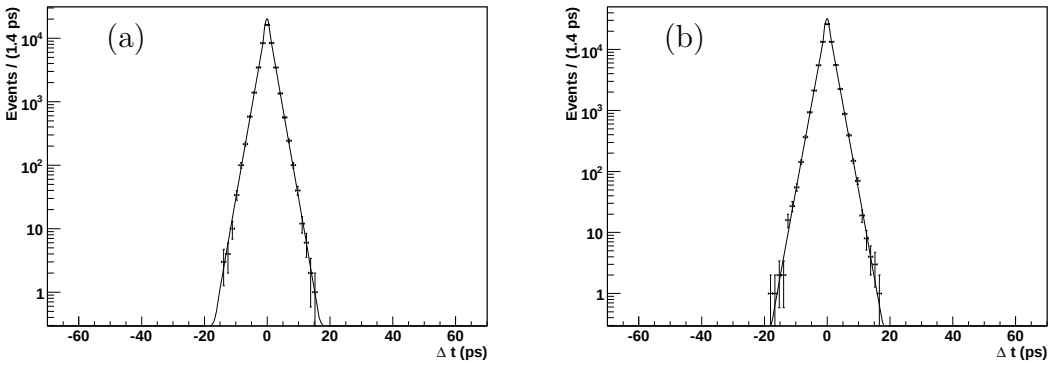


Figure 5.8: Lifetime fit result of generated Δt of reconstructed events. (a) shows the results for SVD1 and (b) shows SVD2.

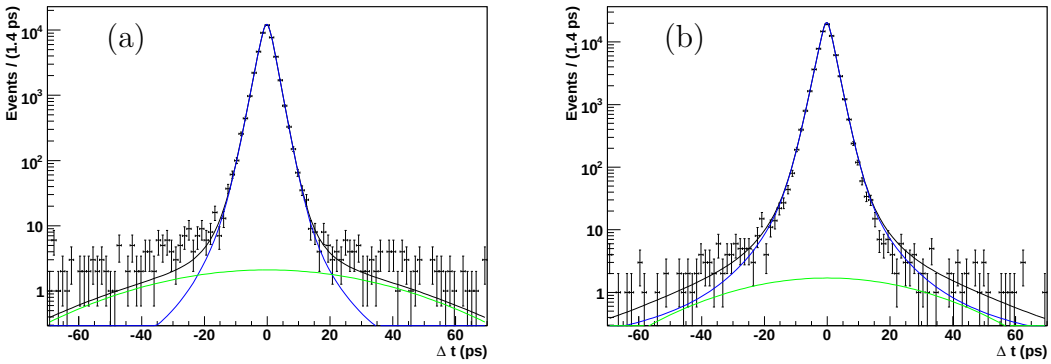


Figure 5.9: Lifetime fit result to reconstructed events. The black curve shows the fit result, the blue curve shows the signal component and the green curve shows the contribution from an additional Gaussian which describes events not modelled by the resolution function. (a) shows the results for SVD1 and (b) shows SVD2.

Δt smearing.

To test the reconstruction procedure and performance of the Δt resolution function with the time-dependent CP violation PDF, Eq. (5.4), we generate signal MC samples, each containing $\sim 5\,000$ reconstructed events using $\mathcal{S}_{CP}^{\text{eff}} = -0.9$ to $\mathcal{S}_{CP}^{\text{eff}} = 0.9$ in intervals of 0.2 as input parameters. A fit to each sample is performed and the $\mathcal{S}_{CP}^{\text{eff}}$ fit residual, $\mathcal{S}_{CP}^{\text{Fit}} - \mathcal{S}_{CP}^{\text{Gen}}$, is plotted as a function of the generated $\mathcal{S}_{CP}^{\text{eff}}$ as can be seen in Fig. 5.10. This distribution is then fit with a first order polynomial and its consistency with a flat line shows that $\mathcal{S}_{CP}^{\text{Fit}}$ agrees with the generated $\mathcal{S}_{CP}^{\text{Gen}}$ over a broad range. Therefore, we do not reject the Δt reconstruction procedure or the description of resolution effects in Δt.

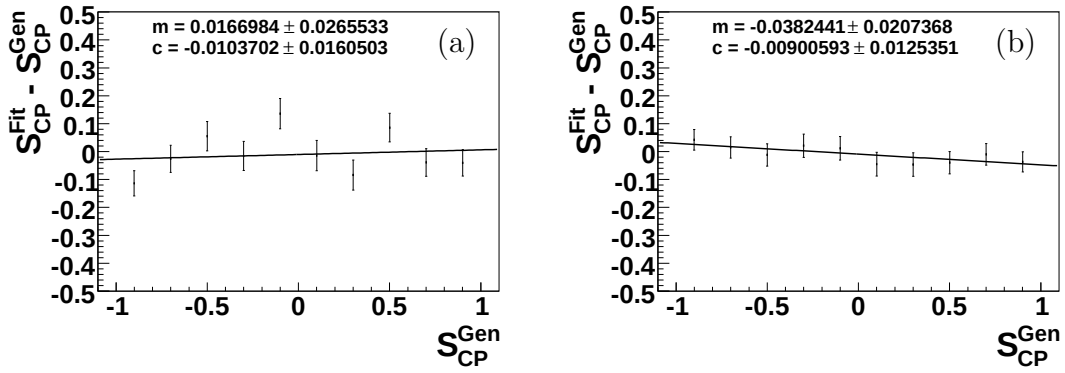


Figure 5.10: GEANT MC linearity test. (a) shows the results for SVD1 and (b) shows SVD2.

5.7.2 Toy MC Test

We generate and fit pseudo-experiments equivalent to the number of events in data reconstructed in the signal region. The input CP parameters, $\mathcal{A}_{CP}$ and $\mathcal{S}_{CP}^{\text{eff}}$ are independently varied within the physical boundaries in intervals of 0.1 and 1 000 pseudo-experiments are generated at each point. The distribution of fit results at each given point is itself fitted with a Gaussian to determine the mean fit result and its error as shown in Figs. D.1 and D.4. The residual, defined as the difference between the mean fit result and the generated value, is plotted as a function of the generated value and is shown in Fig. 5.11. We fit this plot

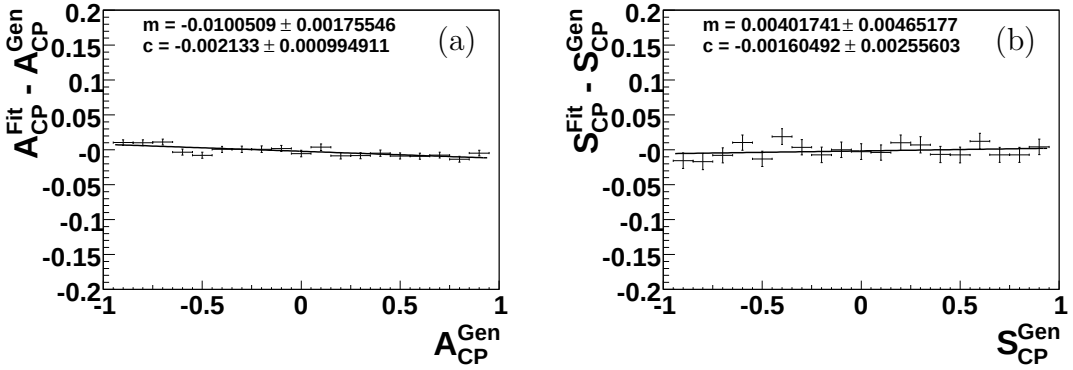


Figure 5.11: Toy MC linearity test. (a) shows the results for $\mathcal{A}_{CP}$ and (b) shows $\mathcal{S}_{CP}^{\text{eff}}$.

with a first order polynomial and its consistency with a flat line in both cases shows there is no significant fit bias for $\mathcal{A}_{CP}$ and $\mathcal{S}_{CP}^{\text{eff}}$.

The expected error for $\mathcal{A}_{CP}$ and $\mathcal{S}_{CP}^{\text{eff}}$ is estimated by taking the arithmetic mean of the distribution of fit errors as shown in Figs. D.2 and D.5. Using the point closest to the Standard Model expectation, the expected errors are,

$$\delta\mathcal{A}_{CP} = 0.14,$$

$\delta\mathcal{S}_{CP}^{\text{eff}} = 0.31$.

The pull distribution is defined as,

$$\text{Pull}(x) \equiv \begin{cases} (x_{\text{Fit}} - x_{\text{Gen}})/\delta_x^- & \text{if } x_{\text{Fit}} - x_{\text{Gen}} > 0 \\ (x_{\text{Fit}} - x_{\text{Gen}})/\delta_x^+ & \text{if } x_{\text{Fit}} - x_{\text{Gen}} < 0, \end{cases} \tag{5.20}$$

where x_{Fit} represents the fitted value of $\mathcal{A}_{CP}$ or $\mathcal{S}_{CP}^{\text{eff}}$, and x_{Gen} is the generated value. The use of the asymmetric MINOS error of the fit, $\delta_x^{\pm}$, depends on the sign of the residual as the MINOS error with the sign opposite to the residual gives the statistical significance of that deviation. Fig. D.2 and Fig. D.5 show that the means and widths of the pull distributions are consistent with zero and unity respectively, which demonstrates that the errors have been determined correctly and that any fit biases are weak.

5.7.3 Lifetime Measurement

A lifetime measurement is performed to the control sample and signal mode data to check the full time-dependent PDF. In these fits, the CP parameters are fixed to zero while the B lifetime is released as a free parameter. We extract the lifetimes, $\tau_{B^+} = 1.81 \pm 0.09$ ps for $B^+ \to K_S^0\pi^+$ and $\tau_{B^0} = 1.46 \pm 0.18$ ps for $B^0 \to K_S^0\pi^0$, which are both in agreement with the current world-averages, $\tau_{B^+} = 1.638 \pm 0.011$ ps and $\tau_{B^0} = 1.530 \pm 0.009$ ps [48]. The fit results are shown in Fig. 5.12.

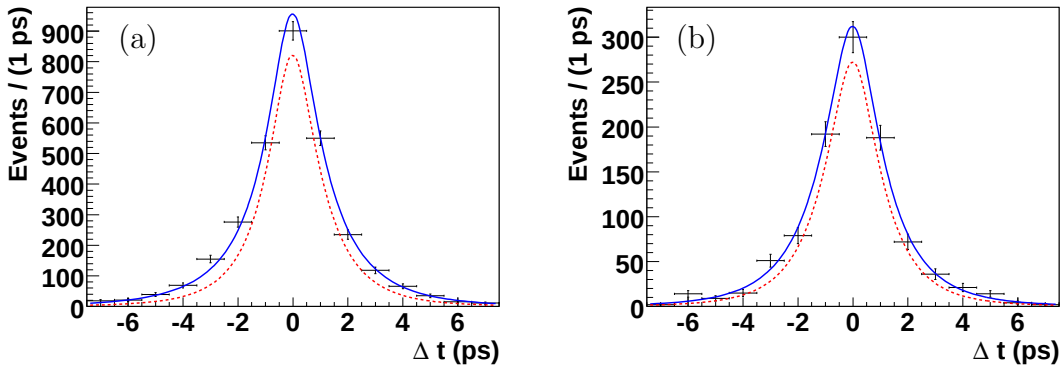


Figure 5.12: Lifetime fit results. The blue curve shows the fit result, and the red curve shows the background contribution. (a) shows the results for $B^+ \to K_S^0\pi^+$ and (b) shows $B^0 \to K_S^0\pi^0$.

5.7.4 Time-dependent CP Violation in $B^+ \to K^0_S\pi^+$

We perform a time-dependent CP violation fit to the control sample and measure the CP parameters,

$$\begin{aligned} \mathcal{A}_{CP} &= +0.007 \pm 0.065 \text{ (stat)}, \\ \mathcal{S}^{\text{eff}}_{CP} &= -0.112 \pm 0.137 \text{ (stat)}, \end{aligned} \tag{5.21}$$

where the fit result is shown in Fig. 5.13. These results are consistent with zero which is expected in the Standard Model for charged B modes.

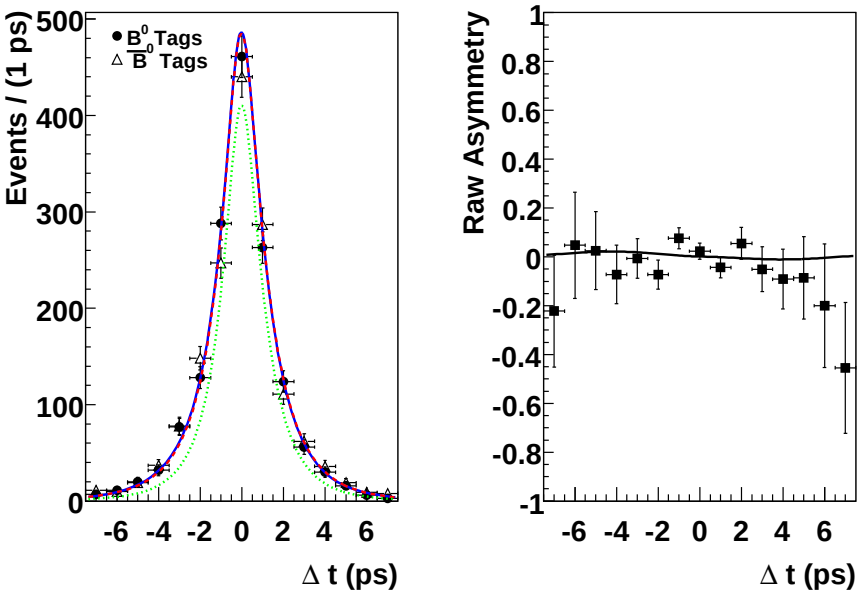


Figure 5.13: CP violation fit result for $B^+ \to K^0_S\pi^+$. The left plot shows the Δt distribution for B^0 and $\bar{B}^0$ tags. The blue (red) curve represents the Δt curve for B^0 ($\bar{B}^0$) tags and the green curve shows the background contribution. The right plot shows the raw asymmetry defined as $(N_{B^0} - N_{\bar{B}^0})/N_{B^0} + N_{\bar{B}^0})$ where N_{B^0} ($N_{\bar{B}^0}$) is the total number of B^0 ($\bar{B}^0$) tags.

5.8 Post-Measurement Validity Tests

A fit to the $B^0 \to K^0_S\pi^0$ sub-sample gives the CP parameters,

$$\begin{aligned} \mathcal{A}_{CP} &= +0.151 \pm 0.131 \text{ (stat)}, \\ \mathcal{S}^{\text{eff}}_{CP} &= +0.668 \pm 0.313 \text{ (stat)}, \end{aligned} \tag{5.22}$$

which is consistent with the combined fit result. Another check is performed to measure $\mathcal{A}_{CP}$ of the entire $B^0 \to K^0\pi^0$ sample using only flavour information. This fit only uses Eq. (5.5) as the signal PDF and gives $\mathcal{A}_{CP} = +0.123 \pm 0.135$, which is also consistent with the final result.

Chapter 6

Discussion

6.1 Branching Fraction

We combine our branching fraction measurement with the latest results from BaBar [43] and CLEO [64] and obtain the new world average,

$$\mathcal{B}(B^0 \to K^0\pi^0) = (9.4 \pm 0.5) \times 10^{-6}. \tag{6.1}$$

6.2 CP Measurement

6.2.1 Shift From Previous Measurement

We use an ensemble test to determine the probability of the central value shift from the results obtained with 535 million $B\bar{B}$ pairs, given in Eq. (1.68). One thousand pseudo-experiments, each containing the number of $B^0 \to K^0\pi^0$ candidates corresponding to 657 million $B\bar{B}$ pairs, are generated using our fit result as input. Each experiment is divided into 535 and the remaining 122 million $B\bar{B}$ pairs and separate fits are performed to these sub-samples for all experiments. The distribution of fit result differences between 535 and the remaining 122 million $B\bar{B}$ pairs is subsequently fit with a Gaussian as shown in Fig. 6.1. From this, the probability of the shift is determined as twice the integral of the tail region bounded by the difference in data between 535 and the remaining 122 million $B\bar{B}$ pairs, as the lower limit. Thus, the shift probability in $\mathcal{A}_{CP}$ is calculated to be 18% and the shift in $\mathcal{S}_{CP}^{\rm eff}$ to be 24% which seems reasonable.

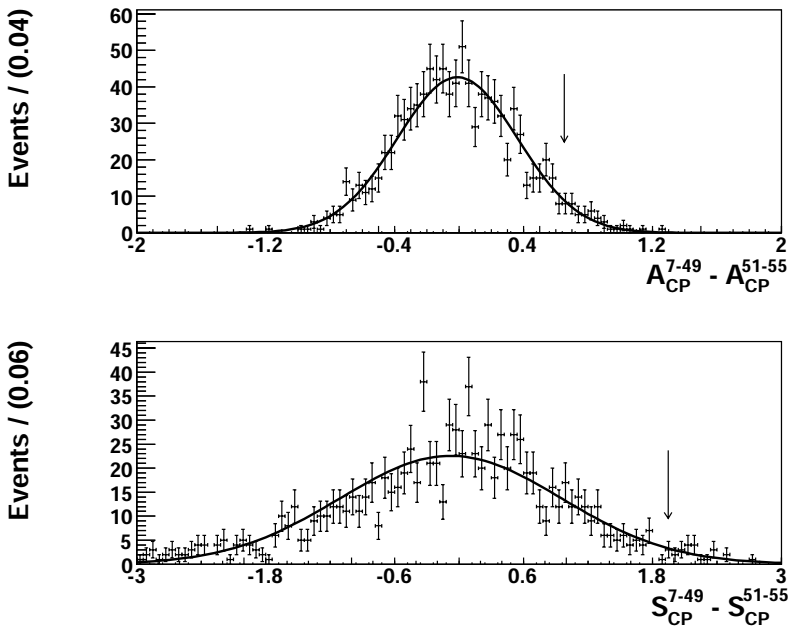


Figure 6.1: Probability of central value shift for $\mathcal{A}_{CP}$ and $\mathcal{S}_{CP}^{\text{eff}}$. The arrow corresponds to the central value difference in data between 535 and the remaining 122 million $B\bar{B}$ pairs.

6.2.2 Sum Rule

We find the measurement of $\mathcal{A}_{CP}$ to deviate from the $K-\pi$ sum rule expectation, $\mathcal{A}_{CP}(K^0\pi^0) = -0.148 \pm 0.044$ [65], by 1.9σ.

6.2.3 Combination With Other Measurements

Our CP violation parameters are combined with the most recent result from BaBar [45] to create the new world average,

$$\begin{aligned} \mathcal{A}_{CP} &= -0.01 \pm 0.10, \\ \mathcal{S}_{CP}^{\text{eff}} &= +0.57 \pm 0.17. \end{aligned} \tag{6.2}$$

The new average for each CP parameter is shown in Fig. 6.2. These comparisons are shown again in Fig. 6.3 in the CP violation plane where the ellipses indicate a possible deviation from zero. The combined results show 3.4σ evidence of CP violation. Figure 6.4 which is an update of the figure shown in §1.6 shows our measurement alongside the other $b \to s$ modes. Although most measurements have central values below that of the $b \to c\bar{c}s$ average, the naive $b \to q\bar{q}s$ average is not significantly below.

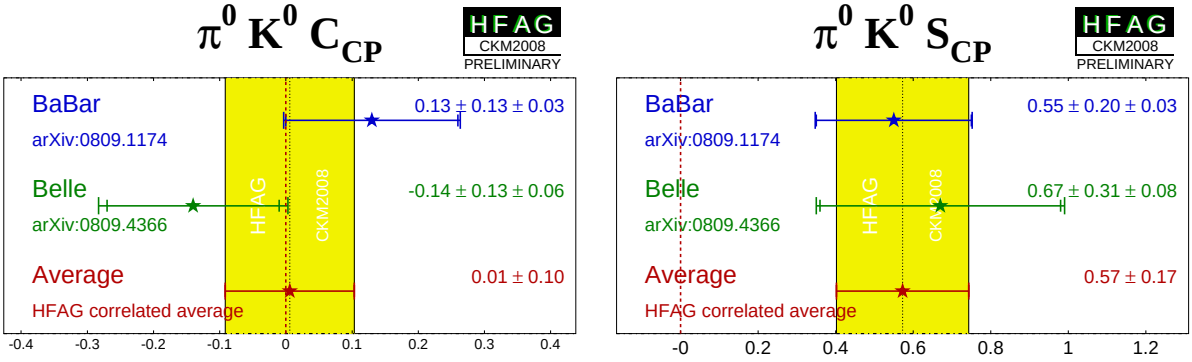


Figure 6.2: The new world averages for direct and mixing-induced CP violation. The left figure shows C_{CP}=-$\mathcal{A}_{CP}$ and the right shows $\mathcal{S}_{CP}^{\rm eff}$.

6.3 Future Prospects

To estimate what the future holds for $B^0 \to K^0\pi^0$, we produce plots of the expected errors of each CP parameter up to 100 ab^{-1} shown in Fig. 6.5. For the statistical errors which are inversely proportional the square root of the luminosity, this is achieved with the expected errors from toy MC given in §5.7.2. For the systematic errors, the errors that cannot be reduced by increased statistics, which include those from vertex reconstruction, resolution function and tag-side interference, are kept as a constant while the remaining contribution can vary with the luminosity in the same way as the statistical error. We can see that for $\mathcal{A}_{CP}$, the systematic error should begin to dominate at around 2 ab^{-1}. Assuming the SM, the error on $\mathcal{A}_{CP}$ should be less than 0.05 to claim evidence for direct CP violation in $B^0 \to K^0\pi^0$ decays. This plot indicates that this is unlikely even with a super B factory unless a method to reduce the systematic error can be found. The outlook on $\mathcal{S}_{CP}^{\rm eff}$ is not so grim. Assuming a SM $\mathcal{S}_{CP}^{\rm eff}$, mixing-induced CP violation could be found as early as 2 ab^{-1}.

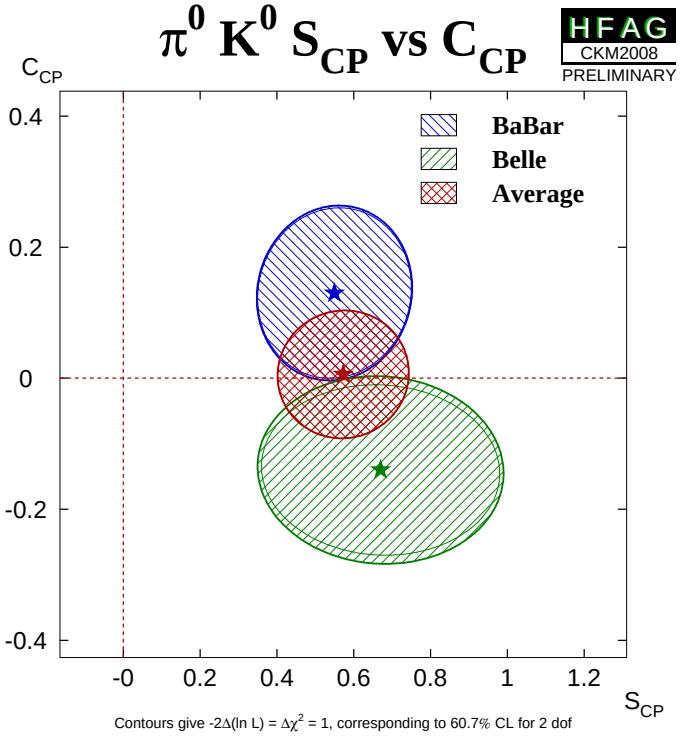


Figure 6.3: Comparison between the latest measurements from BaBar and Belle and their combined average in the CP violation plane.

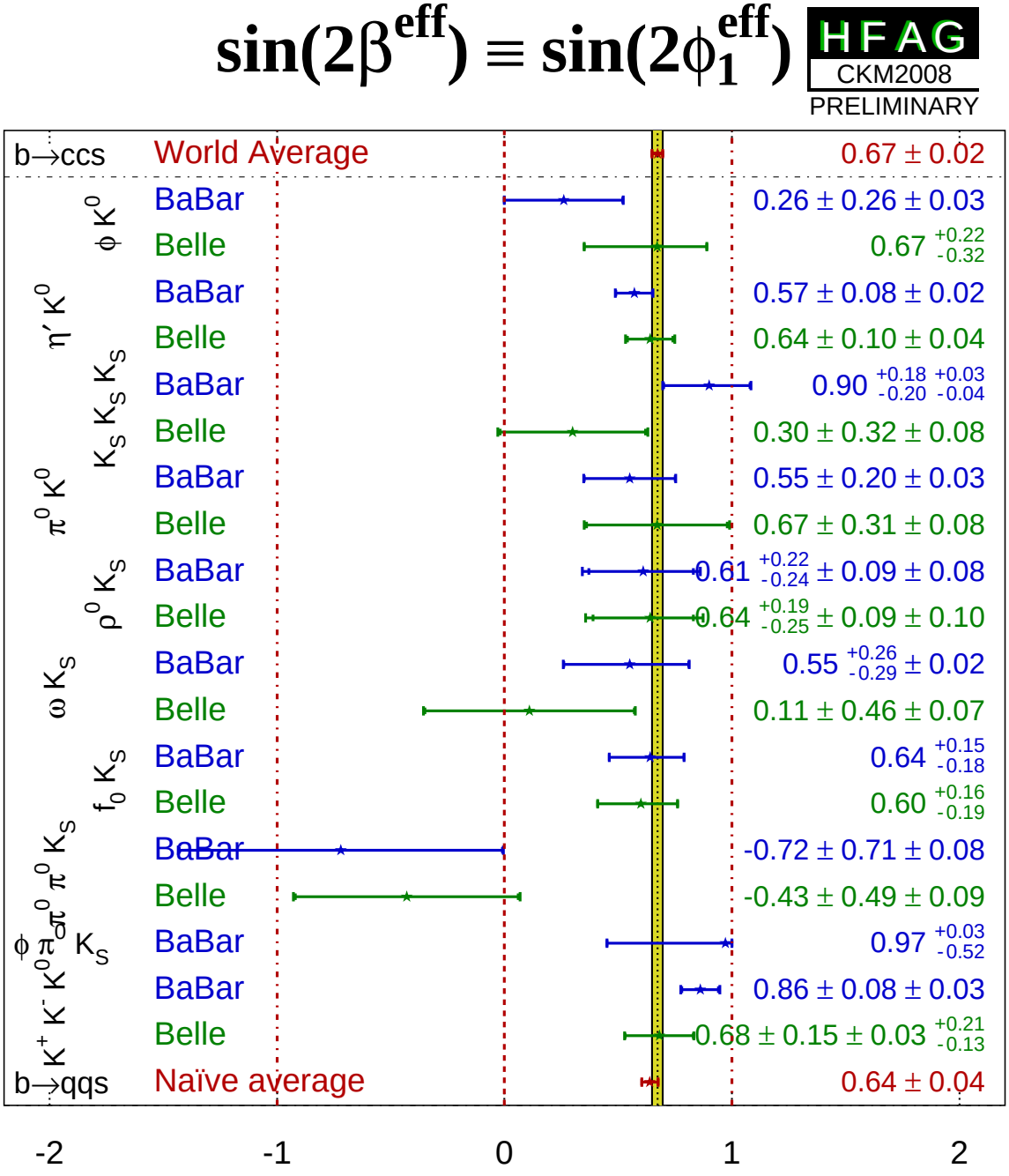


Figure 6.4: Summary of $\mathcal{S}_{CP}^{\rm eff}$ for all $b \to s$ modes.

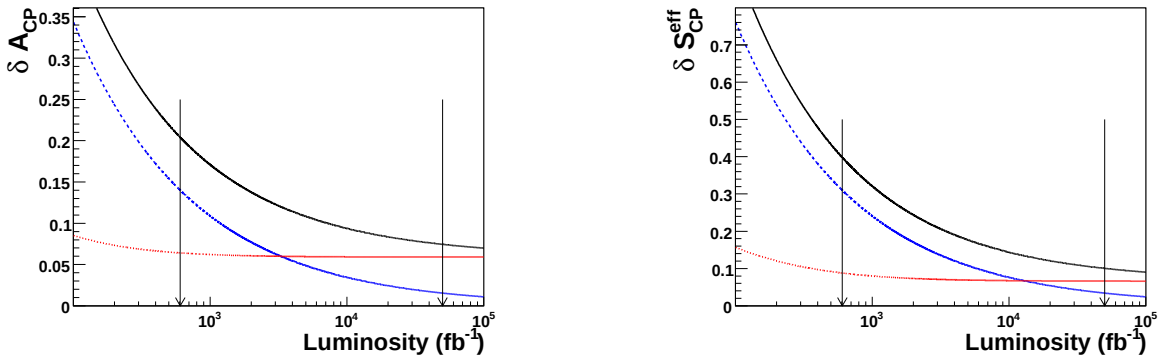


Figure 6.5: The left (right) plot shows the expected errors of $\mathcal{A}_{CP}(\mathcal{S}_{CP}^{\rm eff})$ as a function of the integrated luminosity. The blue (red) curve shows the statistical (systematic) error and the black curve shows the total error. In each plot, the left arrow indicates the current luminosity and the right arrow shows the target luminosity of an upgraded B factory at KEK.

Chapter 7

Conclusion

We have measured the branching fraction and time-dependent CP violation parameters of $B^0 \to K^0\pi^0$ decays. These results are obtained from a data sample that contains 657×10^6 $B\bar{B}$ pairs collected near the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB asymmetric e^+e^- collider.

We obtain the branching fraction,

$$\mathcal{B}(B^0 \to K^0\pi^0) = [8.72\ ^{+0.51}_{-0.50}\ (\text{stat})\ ^{+0.46}_{-0.40}\ (\text{syst})] \times 10^{-6},$$

which is in agreement with the current world average.

We measure the CP parameters,

$$\mathcal{A}_{CP} = +0.14 \pm 0.13\ (\text{stat})\ \pm 0.06\ (\text{syst}),$$
$$\mathcal{S}_{CP}^{\text{eff}} = +0.67 \pm 0.31\ (\text{stat})\ \pm 0.08\ (\text{syst}),$$

where no evidence for direct CP violation is found and the mixing-induced component is consistent with the expectation from the Kobayashi-Maskawa theory of CP violation. We find the measurement of $\mathcal{A}_{CP}$ to deviate from the $K - \pi$ sum rule expectation by 1.9σ.

The analysis technique will need to be improved in order to find evidence of direct CP violation in the future, however, evidence of mixing-induced CP violation in the SM is possible with a super B factory using our method.

Appendix A

Maximum Likelihood Method

The unbinned maximum likelihood method has been frequently employed in this analysis and is the preferred method for data samples with low statistics. This is because the usual χ^2 technique assumes a Gaussian distribution of errors in each bin instead of a Poisson distribution which can lead to a bias in the estimation of parameters and their uncertainties.

Consider a set of N independently measured quantities, x_i, which follows a probability density function (PDF), $P(x;\alpha_\mu)$, where α_μ is a set of μ unknown parameters. The maximum likelihood method is designed to determine the most probable set of α_μ, $\alpha_\mu^{\rm fit}$, for which the PDF product over all data points, $\mathcal{L}(\alpha_\mu)$, is maximised,

$$\mathcal{L}(\alpha_\mu) = \prod_i^N P(x_i;\alpha_\mu). \tag{A.1}$$

However, since $\mathcal{L}(\alpha_\mu)$ is small, it is computationally easier to work with $\log\mathcal{L}(\alpha_\mu)$, which incidentally is also maximised for the same set of $\alpha_\mu^{\rm fit}$. The set of α_μ which maximises $\mathcal{L}(\alpha_\mu)$ is the one for which

$$\frac{\partial\mathcal{L}(\alpha_\mu)}{\partial\alpha_\mu} = 0 \tag{A.2}$$

is satisfied. We construct the quantity, $-2\log\mathcal{L}(\alpha_\mu)$, which behaves like a statistical χ^2 distribution, therefore, the set of α_μ for which the difference between $-2\log\mathcal{L}(\alpha_\mu)$ and $-2\log\mathcal{L}(\alpha_\mu^{\rm fit})$ is unity, corresponds to one standard deviation from $\alpha_\mu^{\rm fit}$ and thus provides the basic error estimate of $\alpha_\mu^{\rm fit}$.

The most probable set of parameters, $\alpha_\mu^{\rm fit}$, is determined numerically with the fitting package, MINUIT [66]. This package also can determine asymmetric errors that take into account non-linearities and correlations between parameters and are known as MINOS errors.

Appendix B

Flavour Tagging Routine

A brief summary of the flavour tagging procedure is provided here, and a more detailed discussion can be found in [67].

B.1 Track-Level Flavour Tagging

After reconstructing B_{Rec}, the remaining final state particles which belong to B_{Tag} are used to determine flavour information. At the simplest level, the flavour, q, can be inferred by searching for particular charged tracks within the set of final states. Some of these processes are shown in Fig. B.1. With the exception of charged tracks assigned to the Λ and kaon categories, these tracks are required to originate from the IP: within 2 cm in the $x-y$ plane and 10 cm in the z direction.

B.1.1 Slow Pion Category

A charged track with a momentum smaller than 0.25 GeV/c in the CMS is assigned to this category. The flavour information from the charge of the slow pion track is used to identify the b-flavour of the tag side. Several variables are included as discriminants: the charge of the track, the momentum in the laboratory frame, the polar angle in the laboratory frame, the angle between the slow pion and the thrust axis of rest of the tag side particles in the CMS, θ_{thr}, and a pion/electron identification likelihood ratio, $R_{\pi/e}$. The angle, θ_{thr}, is used to suppress the background from non-D decays. The ratio, $R_{\pi/e}$, is used to remove background electrons from photon conversion. This category has a low tagging efficiency and poor tagging quality.

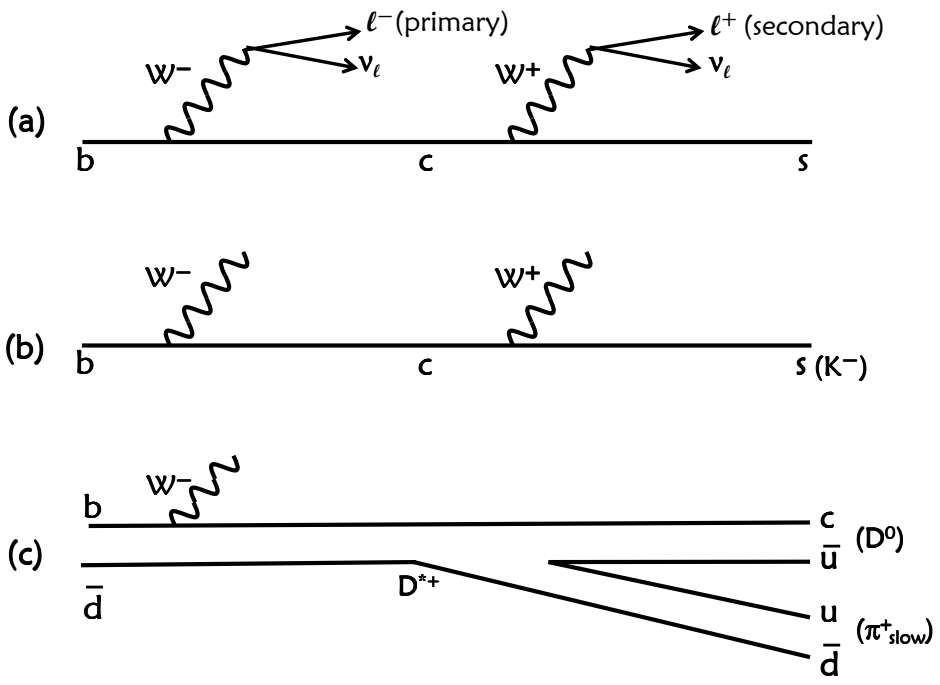


Figure B.1: Some of the decay processes from which the flavour of $B_{\rm Tag}$ can be determined. (a) Lepton category, (b) Kaon category, (c) Slow Pion category.

B.1.2 Lambda Category

A charged track that forms a Λ candidate with another charged track with the opposite charge is assigned to this category. The flavour information from Λ, Λ or $\bar{\Lambda}$, is used to identify the b-flavour of the tag side. One of the tracks should be identified as a proton. The Λ candidate is required to have an invariant mass, $M_{p\pi}$, between 1.1108 GeV/c^2 and 1.1208 GeV/c^2. The angle between the momentum vector of the Λ candidate and the vector formed by the IP and the vertex of the Λ candidate, $\theta_{\rm defl}$, should be less than 30°. The minimum distance between the tracks in z axis, Δz, should be less than 4 cm, and the flight length of the Λ in the $x-y$ plane should be longer than 0.5 cm. The flavour of Λ, $M_{p\pi}$, $\theta_{\rm defl}$ and the presence of K^0_S candidates are included as discriminants.

B.1.3 Kaon Category

A charged track that is not positively identified as a lepton or proton is assigned to this category. The flavour information of the kaon in $b \to c \to s$ decays is mainly used to identify the b-flavour of the tag side. Fast pions from $B \to D^{(*)}\pi^{\pm}$ are also included. The charge of the track, the presence of K^0_S candidates, the track momentum in the CMS, the polar

angle of the track in the laboratory frame and the K/π ID likelihood ratio are included as discriminants. This category has a high tagging efficiency but poor tagging quality.

B.1.4 Lepton Category

A charged track with a momentum larger than 0.4 GeV/c and an electron likelihood larger than 0.8 is included as an electron candidate, while a charged track with a momentum larger than 0.8 GeV/c and a muon likelihood larger than 0.95 is included as a muon candidate. The flavour information from high momentum leptons in $B \to X\ell^{\pm}\nu$ decays and intermediate momentum leptons in $B \to D \to K\ell^{\pm}\nu$ decays are used to identify the b-flavour of the tag side. Several variables are included as discriminants: the lepton momentum in the CMS, $p_{\ell}^{\rm cms}$, the polar angle of the track in the laboratory frame, $\theta_{\rm lab}$, the recoil mass, $M_{\rm recoil}$ and the missing momentum in the CMS, $p_{\rm miss}^{\rm cms}$. The variables, $M_{\rm recoil}$ and $p_{\rm miss}^{\rm cms}$ can discriminate primary leptons from secondary leptons since $M_{\rm recoil}$ can indicate the presence of a D meson and the variable, $p_{\rm miss}^{\rm cms}$, can indicate the existence of neutrinos. This category has a low tagging efficiency but high tagging quality.

B.2 Event-Level Flavour Tagging

Clearly, this flavour tagging method will not always return the correct flavour of $B_{\rm Tag}$, and this is due to incorrect PID and suppressed physics processes. We introduce the MC determined flavour reliability, r, which is defined as

$$q \cdot r \equiv \frac{{\rm N}(B^0) - {\rm N}(\bar{B}^0)}{{\rm N}(B^0) + {\rm N}(\bar{B}^0)}, \tag{B.1}$$

and ranges from $r = 0$ for no flavour discrimination to $r = 1$ for unambiguous flavour assignment. The variables, ${\rm N}(B^0)$ and ${\rm N}(\bar{B}^0)$, are the number of B^0 and $\bar{B}^0$ events that occupy the "cell". The cell is an element in the multi-dimensional flavour tagging table prepared from MC that depends on the flavour tagging method and the discriminants described above.

The event-level flavour tagging combines the results, $(q \cdot r)_x$ where x is the tagging category, for each track to determine a single highest $q \cdot r$ output for each event. For the lepton and slow pion categories, the track with the highest r in each category is chosen as an input. For the kaon and Λ categories, the flavour and dilution factors of each track are combined by calculating the product of the flavour dilution factors in order to account for possible cases where there are multiple s quarks in the final state. By using a three-dimensional look-up table, the correlations between the flavour information for the four categories in track-level are correctly taken into account. A schematic of these procedures is shown in Fig. B.2.

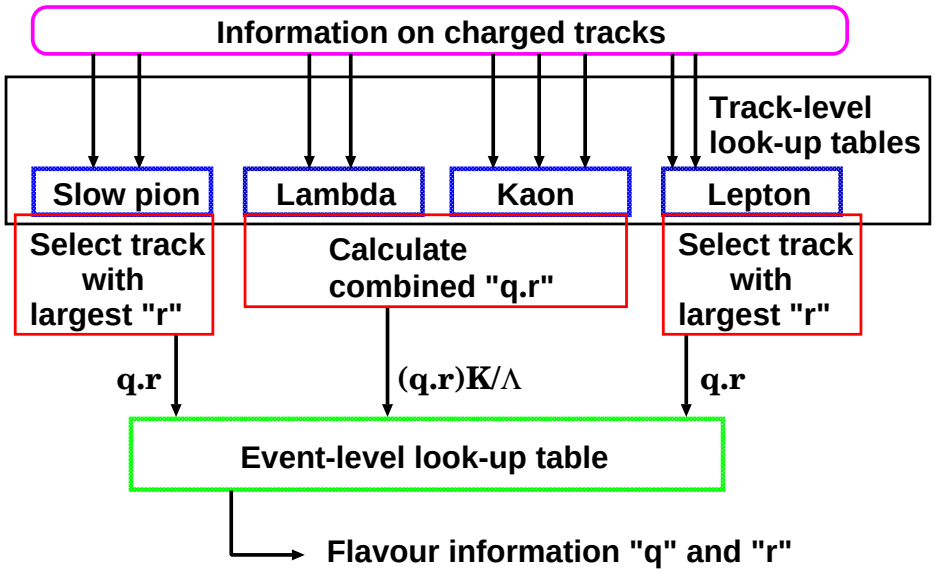


Figure B.2: Flavour tagging algorithm.

B.2.1 Flavour Tagging Performance

The performance of flavour tagging is characterised by the two parameters: ϵ and w. The parameter, ϵ, is the raw tagging efficiency and the wrong tag fraction, w, is the probability that the flavour tag is wrong.

In practise, the observed time-dependence of a CP eigenstate becomes

$$\mathcal{P}_{\rm Sig}^{\rm obs}(\Delta t, q, w) = \epsilon \cdot [(1-w)\mathcal{P}_{\rm Sig}(\Delta t, q) + w(\mathcal{P}_{\rm Sig}(\Delta t, -q))], \tag{B.2}$$

and the observed CP asymmetry is rewritten as

$$\mathcal{A}_{CP}^{\rm obs} = \frac{\mathcal{P}_{\rm Sig}^{\rm obs}(\Delta t, q, w) - \mathcal{P}_{\rm Sig}^{\rm obs}(\Delta t, -q, w)}{\mathcal{P}_{\rm Sig}^{\rm obs}(\Delta t, q, w) + \mathcal{P}_{\rm Sig}^{\rm obs}(\Delta t, -q, w)} = (1-2w)\mathcal{A}_{CP}. \tag{B.3}$$

The observed CP asymmetry is diluted by the factor, $1-2w$.

Since the statistical significance of the measured asymmetries are proportional to $(1-2w)\sqrt{\epsilon}$, the effective number of events are proportional to the factor, $\epsilon(1-2w)^2$. Thus, an effective efficiency, $\epsilon_{\rm eff} = \epsilon(1-2w)^2$, is introduced. The tagging algorithm has been developed to maximise the effective efficiency, $\epsilon_{\rm eff}$. The total effective tagging efficiency for MC is estimated to be $29.72 \pm 0.17\%$.

Note that r is determined from MC-based information only and if the MC information represents data perfectly, then $r = 1-2w$. To maximise the effective tagging efficiency while using w values measured from data, events are subdivided into seven regions based on their r values: $0 < r \leq 0.1$, $0.1 < r \leq 0.25$, $0.25 < r \leq 0.5$, $0.5 < r \leq 0.625$, $0.625 < r \leq 0.75$, $0.75 < r \leq 0.875$ and $0.875 < r \leq 1$. These bins have corresponding labels from 0 to 6. The wrong tag fractions, w, and the difference in w between B^0 and $\bar{B}^0$, Δw, are obtained by

fitting the time-dependent $B^0\bar{B}^0$ mixing oscillation of $B^0 \to D^{*-}l^+\nu_l$, $D^{(*)-}\pi^+$ and $D^{*-}\rho^+$ self-tagging events as shown in Fig. B.3.

The time evolution of $B^0\bar{B}^0$ pairs with opposite flavour (OF) and same flavour (SF) is given by

$$\mathcal{P}_{\rm SF}^{\rm OF} = \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}}[1 \mp q\Delta w \pm (1-2w)\cos(\Delta m_d \Delta t)], \tag{B.4}$$

and the OF-SF asymmetry is given by

$$\mathcal{A}_{\rm Mix} = \frac{\mathcal{P}_{\rm OF} - \mathcal{P}_{\rm SF}}{\mathcal{P}_{\rm OF} + \mathcal{P}_{\rm SF}} = -q\Delta w + (1-2w)\cos(\Delta m_d \Delta t). \tag{B.5}$$

The obtained w and Δw are summarised in Table B.1 for the SVD1 and SVD2 data samples and are used for analysis. The total effective tagging efficiency for data is $29.20 \pm 1.37\%$.

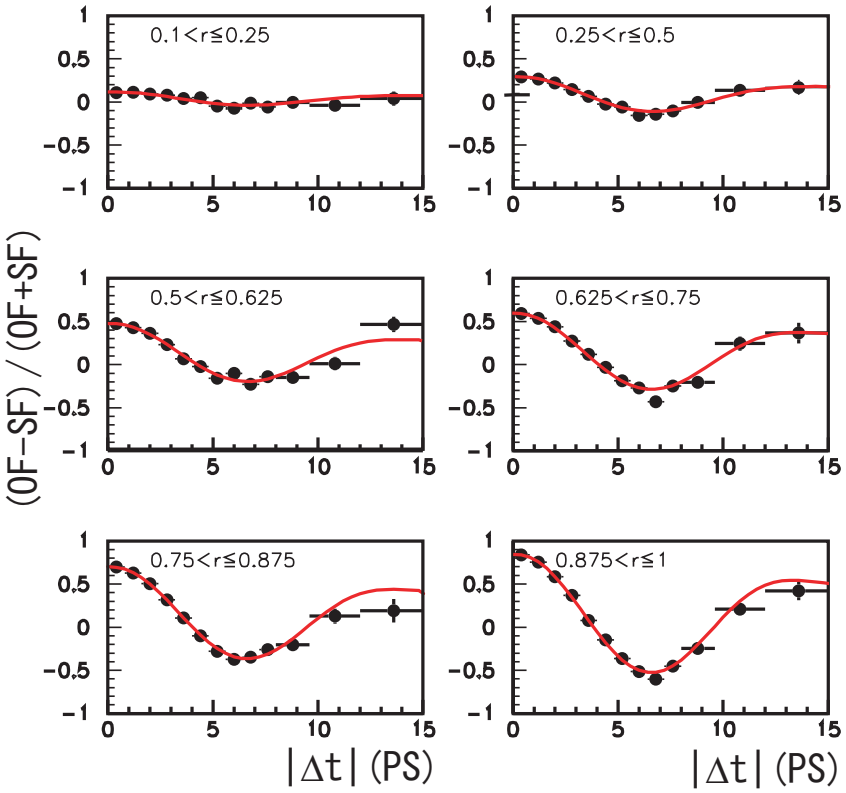


Figure B.3: Time-dependent $B^0\bar{B}^0$ mixing oscillation fit result to the control sample data. Each plot from the top-left to the bottom-right, corresponds to the subsample in each r-bin region from 1 to 6. The amplitudes in the oscillation become larger due to less dilution from incorrect tagging.

Table B.1: Wrong tag fractions and wrong tag fraction differences for each r-bin.

r-bin	w(SVD1)	w(SVD2)	Δw(SVD1)	Δw(SVD2)
$0.0 < r \leq 0.1$	0.5 ± 0.0	0.5 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
$0.1 < r \leq 0.25$	$0.423^{+0.008}_{-0.007}$	0.429 ± 0.005	$0.058^{+0.010}_{-0.009}$	$-0.039^{+0.006}_{-0.007}$
$0.25 < r \leq 0.5$	0.337 ± 0.008	0.327 ± 0.006	0.012 ± 0.010	-0.036 ± 0.006
$0.5 < r \leq 0.625$	$0.235^{+0.010}_{-0.008}$	$0.223^{+0.011}_{-0.006}$	-0.012 ± 0.010	0.018 ± 0.007
$0.625 < r \leq 0.75$	$0.166^{+0.008}_{-0.007}$	$0.161^{+0.010}_{-0.006}$	$-0.011^{+0.010}_{-0.009}$	0.002 ± 0.006
$0.75 < r \leq 0.875$	$0.105^{+0.008}_{-0.007}$	$0.105^{+0.007}_{-0.008}$	0.008 ± 0.009	-0.027 ± 0.006
$0.875 < r \leq 1.0$	$0.026^{+0.006}_{-0.005}$	$0.019^{+0.004}_{-0.005}$	0.003 ± 0.006	-0.001 ± 0.004

Appendix C

Δt Resolution Function

A brief description of the resolution function is provided here, and a more detailed discussion can be found in [68]. As mentioned in §5.2.3, the sources of Δt smearing are detector resolution, $B_{\rm Tag}$ vertex smearing due to the inclusion of non-primary tracks on the tag side and the approximation where the B mesons are treated at rest in the CMS frame.

C.1 Detector Resolution

The detector resolutions, $R_{\rm Rec}$ and $R_{\rm Tag}$, describe the smearing of primary tracks originating from the B mesons. Because the resolution of multi-track vertices is better than single-track vertices, they are treated separately.

The detector resolution function for the $B_{\rm Rec}$ vertex with more than one track is given by a double Gaussian with zero mean,

$$R_{\rm Rec}(\delta z_{\rm Rec}) \equiv (1 - f^{\rm tail}_{\rm Rec})G[\delta z_{\rm Rec}; (S^0_{\rm Rec} + \xi_{\rm Rec} S^1_{\rm Rec})\sigma_{\rm Rec}] + f^{\rm tail}_{\rm Rec} G[\delta z_{\rm Rec}; S^{\rm tail}_{\rm Rec} S^{\rm main}_{\rm Rec}]. \quad \text{(C.1)}$$

Smearing of the multi-track tag side vertex is modelled with a single Gaussian,

$$R_{\rm Tag}(\delta z_{\rm Tag}) \equiv G[\delta z_{\rm Tag}; (S^0_{\rm Tag} + \xi_{\rm Tag} S^1_{\rm Tag})\sigma_{\rm Tag}]. \quad \text{(C.2)}$$

The single-track vertex resolution function for both the CP and tag side is given by the double Gaussian,

$$R_i(\delta z_i) \equiv (1 - f^{\rm tail}_i)G[\delta z_i; S^{\rm main}_i \sigma_i] + f^{\rm tail}_i G[\delta z_i; S^{\rm tail}_i \sigma_i], \qquad i = {\rm Rec, Tag}. \quad \text{(C.3)}$$

In these detector resolution functions, δz_i is defined as being the difference between the reconstructed vertex position and its true position, ξ_i is the vertex quality as defined in Eq. 5.2, and σ_i is vertex z-position error. Note that for the multi-track resolution functions width, $S^{\rm main}_i \equiv (S^0_i + \xi_i S^1_i)\sigma_i$, the vertex error, σ_i, is corrected with a linear polynomial in the vertex quality, ξ_i, because they are correlated. The detector resolution function parameters are determined from a lifetime fit to the data of many control samples such as, $B^0 \to D^* l \nu_l$, $D^{(*)}\pi$ and $D^*\rho$, and are given in Table C.1.

Table C.1: $R_{\rm Det}$ resolution function parameters.

	Parameter	Value(SVD1)	Value(SVD2)
Multi-track	S^0_{CP}	$0.910^{+0.463}_{-0.080}$	$0.656^{+0.453}_{-0.098}$
	S^1_{CP}	0.023 ± 0.004	$0.035^{+0.008}_{-0.006}$
	$f^{\rm tail}_{CP}$	–	$0.101^{+0.040}_{-0.053}$
	$S^{\rm tail}_{CP}$	–	$4.752^{+0.428}_{-1.028}$
	$S^0_{\rm Tag}$	$0.667^{+0.339}_{-0.398}$	$0.790^{+0.494}_{-0.260}$
	$S^1_{\rm Tag}$	$0.010^{+0.001}_{-0.004}$	$0.014^{+0.003}_{-0.006}$
Single-track	$S^{\rm main}_i$	$0.906^{+0.278}_{-0.069}$	$1.022^{+0.445}_{-0.079}$
	$f^{\rm tail}_i$	–	$0.108^{+0.042}_{-0.045}$
	$S^{\rm tail}_i$	–	$3.596^{+4.462}_{-0.803}$

C.2 $B_{\rm Tag}$ Vertex Smearing From Non-Primary Tracks

Due to the vertex reconstruction algorithm on the tag side mentioned in §5.2.2, the $B_{\rm Tag}$ vertex resolution may be smeared by contamination from non-primary tracks as illustrated in Fig. C.1.

The resolution function, $R_{\rm NP}$, is given by

$$R_{\rm NP}(\delta z_{\rm Tag}) \equiv f_\delta\ \delta(\delta z_{\rm Tag}) + (1-f_\delta)[f^+E^+(\delta z_{\rm Tag};\tau^+_{\rm NP}) + (1-f^+)E^-(\delta z_{\rm Tag};\tau^-_{\rm NP})], \quad \text{(C.4)}$$

which consists of two parts. The first component represents the vertices without non-primary track contamination and is given by the Dirac δ-function. The second part describes the lifetimes of secondary vertices, E^+ and E^-, on both sides of the $\delta z_{\rm Tag}$ distribution,

$$E^+(\delta z_{\rm Tag};\tau^+_{\rm NP}) \equiv \begin{cases} \frac{1}{\tau^+_{\rm NP}} e^{-\delta z_{\rm Tag}/\tau^+_{\rm NP}} & \text{if } \delta z_{\rm Tag} > 0 \\ 0 & \text{if } \delta z_{\rm Tag} < 0, \end{cases}$$

$$E^-(\delta z_{\rm Tag};\tau^-_{\rm NP}) \equiv \begin{cases} 0 & \text{if } \delta z_{\rm Tag} > 0 \\ \frac{1}{\tau^-_{\rm NP}} e^{-\delta z_{\rm Tag}/\tau^-_{\rm NP}} & \text{if } \delta z_{\rm Tag} < 0. \end{cases} \quad \text{(C.5)}$$

In this case, $\delta z_{\rm Tag}$ is defined as the difference in the reconstructed vertex with and without non-primary track contamination. The effective decay lengths, $\tau^\pm_{\rm NP}$, depend on the vertex quality, $\xi_{\rm NP}$, because secondary tracks from longer lived particles give a larger distortion of the vertex and its error. So, for a multi-track $B_{\rm Tag}$ vertex, accounting for the correlation between $\sigma_{\rm NP}$ and $\xi_{\rm NP}$, $\tau^\pm_{\rm NP}$ are parametrised as,

$$\begin{aligned} \tau^+_{\rm NP} &\equiv S^3_{\rm Tag}[\tau^+_0 + \tau^+_1(1+S^2_{\rm Tag}\xi)\sigma_{\rm Tag}], \\ \tau^-_{\rm NP} &\equiv S^3_{\rm Tag}[\tau^-_0 + \tau^-_1(1+S^2_{\rm Tag}\xi)\sigma_{\rm Tag}]. \end{aligned} \quad \text{(C.6)}$$

Again, a different parametrisation for $\tau^\pm_{\rm NP}$ is needed for single-track vertices because its vertex quality, ξ, cannot be defined,

$$\begin{aligned} \tau^+_{\rm NP} &\equiv S^3_{\rm Tag}[\tau^+_0 + \tau^+_1\sigma_{\rm Tag}], \\ \tau^-_{\rm NP} &\equiv S^3_{\rm Tag}[\tau^-_0 + \tau^-_1\sigma_{\rm Tag}]. \end{aligned} \quad \text{(C.7)}$$

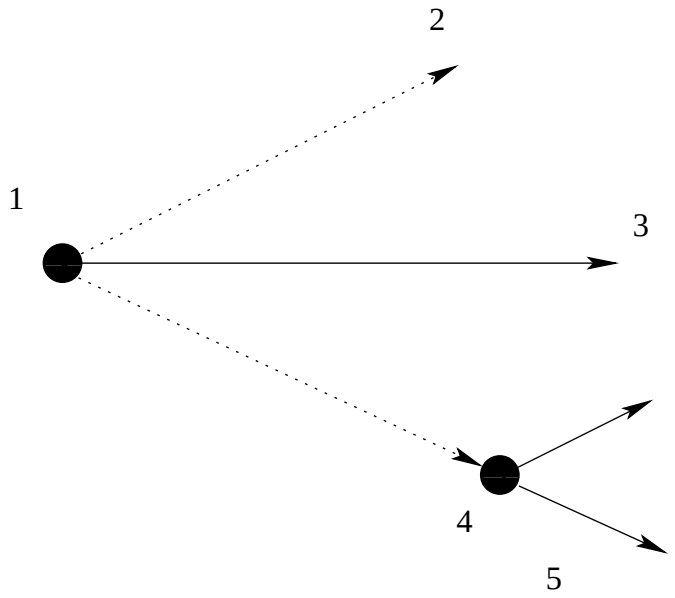


Figure C.1: Schematic showing how non-primary tracks displaced from the true B^0 vertex will cause a smearing of the reconstructed $B_{\rm Tag}$ vertex.

The $R_{\rm NP}$ is determined by comparing two set of MC. One is the usual generic $B\bar{B}$ MC and the other has been specially generated so that all B meson daughters have zero lifetimes. That means all secondary tracks and so forth, will now decay at the B vertex allowing the $\delta z_{\rm Tag}$ distribution to be studied. The resolution function parameters associated with smearing of the $B_{\rm Tag}$ vertex from non-primary tracks are given in Tables C.2 and C.3.

C.3 Kinematic Approximation

The Δt distribution is calculated from Eq. 5.1, which ignores the motion of the B mesons in the CMS frame. The smearing of Δt that arises from this assumption does not need to be parametrised because it can be calculated analytically. One can show that the difference between the reconstructed Δt, Eq. 5.1, and the true Δt, $\Delta t_{\rm True} \equiv t_{\rm Rec} - t_{\rm Tag}$, is given by

$$x \equiv \Delta t_{\rm Rec} - \Delta t_{\rm True} = [(\beta\gamma)_{\rm Rec}/(\beta\gamma)_{\Upsilon(4S)} - 1]t_{\rm Rec} - [(\beta\gamma)_{\rm Tag}/(\beta\gamma)_{\Upsilon(4S)} - 1]t_{\rm Tag}, \qquad \text{(C.8)}$$

where $(\beta\gamma)_{\rm Rec}$ and $(\beta\gamma)_{\rm Tag}$ are the Lorentz boost factors of $B_{\rm Rec}$ and $B_{\rm Tag}$, respectively. Their ratios to the $\Upsilon(4S)$ Lorentz boost are,

$$\begin{aligned} \frac{(\beta\gamma)_{\rm Rec}}{(\beta\gamma)_{\Upsilon(4S)}} &= \frac{E_B^{\rm CMS}}{m_B} + \frac{p_B^{\rm CMS}\cos\theta_B^{\rm CMS}}{\beta_{\Upsilon(4S)} m_B} \equiv a_{\rm K} + c_{\rm K}, \\ \frac{(\beta\gamma)_{\rm Tag}}{(\beta\gamma)_{\Upsilon(4S)}} &= \frac{E_B^{\rm CMS}}{m_B} - \frac{p_B^{\rm CMS}\cos\theta_B^{\rm CMS}}{\beta_{\Upsilon(4S)} m_B} \equiv a_{\rm K} - c_{\rm K}, \end{aligned} \qquad \text{(C.9)}$$

where $E_B^{\rm CMS}$, m_B, $p_B^{\rm CMS}$ and $\cos\theta_B^{\rm CMS}$ are the energy, mass, momentum and polar angle of $B_{\rm Rec}$ in the CMS frame. Since the $t_{\rm Rec}$ and $t_{\rm Tag}$ distributions follow an exponential decay,

Table C.2: $R_{\rm NP}$ resolution function parameters for B^0.

	Parameter	Value(SVD1)	Value(SVD2)
Global	$S^2_{\rm Tag}$	0.020 ± 0.0004	0.034 ± 0.0004
	$S^3_{\rm Tag}$	$0.954^{+0.104}_{-0.053}$	$0.877^{+0.201}_{-0.077}$
Multi-track	f_δ	$0.251^{+0.062}_{-0.067}$	$0.297^{+0.088}_{-0.038}$
	f^+	0.792 ± 0.005	0.770 ± 0.002
	τ_0^+(cm)	0.071 ± 0.004	0.111 ± 0.002
	τ_1^+	0.070 ± 0.005	0.101 ± 0.002
	τ_0^-(cm)	0.062 ± 0.008	0.129 ± 0.004
	τ_1^-	0.011 ± 0.009	0.092 ± 0.004
Single-track	f_δ	$0.598^{+0.041}_{-0.141}$	$0.706^{+0.029}_{-0.082}$
	f^+	0.734 ± 0.010	0.841 ± 0.012
	τ_0^+(cm)	0.641 ± 0.047	1.450 ± 0.030
	τ_1^+	$0.314^{+0.066}_{-0.063}$	0.886 ± 0.033
	τ_0^-(cm)	$0.314^{+0.065}_{-0.060}$	$1.002^{+0.074}_{-0.070}$
	τ_1^-	$0.196^{+0.097}_{-0.085}$	$0.519^{+0.081}_{-0.073}$

Table C.3: $R_{\rm NP}$ resolution function parameters for B^+.

	Parameter	Value(SVD1)	Value(SVD2)
Global	$S^2_{\rm Tag}$	0.020 ± 0.0004	0.034 ± 0.0004
	$S^3_{\rm Tag}$	$0.954^{+0.104}_{-0.053}$	$0.877^{+0.201}_{-0.077}$
Multi-track	f_δ	$0.166^{+0.090}_{-0.077}$	$0.385^{+0.082}_{-0.052}$
	f^+	0.775 ± 0.007	$0.762^{+0.003}_{-0.003}$
	τ_0^+(cm)	0.607 ± 0.009	$0.640^{+0.006}_{-0.006}$
	τ_1^+	0.627 ± 0.010	$0.571^{+0.006}_{-0.006}$
	τ_0^-(cm)	$0.530^{+0.016}_{-0.015}$	$0.589^{+0.009}_{-0.009}$
	τ_1^-	$0.617^{+0.020}_{-0.019}$	$0.561^{+0.010}_{-0.010}$
Single-track	f_δ	$-$	$-$
	f^+	$0.745^{+0.017}_{-0.018}$	$0.834^{+0.026}_{-0.028}$
	τ_0^+(cm)	$1.045^{+0.055}_{-0.028}$	$-$
	τ_1^+	$1.057^{+0.086}_{-0.083}$	$-$
	τ_0^-(cm)	1.199 ± 0.089	$-$
	τ_1^-	$1.307^{+0.154}_{-0.155}$	$-$

Table C.4: $P_{\rm Out}$ resolution function parameters.

	Parameter	Value(SVD1)	Value(SVD2)
Global	$\sigma_{\rm Out}$(ps)	$37.429^{+6.796}_{-9.012}$	$33.623^{+6.801}_{-8.971}$
Multi-track	$f_{\rm Out}$	0.0003 ± 0.0001	0.0002 ± 0.0001
Single-track	$f_{\rm Out}$	$0.042^{+0.007}_{-0.008}$	$0.026^{+0.005}_{-0.004}$

E_i^+, the probability density of simultaneously obtaining x and $\Delta t_{\rm True}$ is given by

$$\begin{aligned}\mathcal{P}(x, \Delta t_{\rm True}) &= \int_0^\infty \int_0^\infty dt_{\rm Rec}\, dt_{\rm Tag}\; E^+(t_{\rm Rec}; \tau_B) E^+(t_{\rm Tag}; \tau_B)\; \delta(\Delta t_{\rm True} - [t_{\rm Rec} - t_{\rm Tag}]) \\ &\qquad \delta(x - \{[a_{\rm K} + c_{\rm K} - 1] t_{\rm Rec} - [a_{\rm K} - c_{\rm K} - 1] t_{\rm Tag}\}),\end{aligned} \tag{C.10}$$

and the probability density of obtaining $\Delta t_{\rm True}$ is

$$\mathcal{P}(\Delta t_{\rm True}) = \int_0^\infty \int_0^\infty dt_{\rm Rec}\, dt_{\rm Tag}\; E^+(t_{\rm Rec}; \tau_B) E^+(t_{\rm Tag}; \tau_B)\; \delta(\Delta t_{\rm True} - [t_{\rm Rec} - t_{\rm Tag}]). \tag{C.11}$$

The resolution function that accounts for the kinematic approximation, $R_{\rm K}$, is defined as the probability density of obtaining x for a given $\Delta t_{\rm True}$. Thus, it is expressed as $R_{\rm K}(x) = \mathcal{P}(x, \Delta t_{\rm True})/\mathcal{P}(\Delta t_{\rm True})$, which evaluates to,

$$R_{\rm K}(x) = \begin{cases} E^+(x - \{(a_{\rm K} - 1)\Delta t_{\rm True} + c_{\rm K}|\Delta t_{\rm True}|\}; |c_{\rm K}|\tau_B) & \text{if } c_{\rm K} > 0 \\ \delta(x - \{a_{\rm K} - 1\}\Delta t_{\rm True}) & \text{if } c_{\rm K} = 0 \\ E^-(x - \{(a_{\rm K} - 1)\Delta t_{\rm True} + c_{\rm K}|\Delta t_{\rm True}|\}; |c_{\rm K}|\tau_B) & \text{if } c_{\rm K} < 0. \end{cases} \tag{C.12}$$

C.4 Outlier

There still exists a long tail in Δt that is not described by the resolution functions above or background functions. We model the tail with a Gaussian with zero mean and a width independent of the event,

$$P_{\rm Out}(\Delta t) \equiv G(\Delta t; \sigma_{\rm Out}). \tag{C.13}$$

The outlier receives a larger contribution from vertices reconstructed from a single track, because they have a worse resolution. Therefore, the outlier fraction, $f_{\rm Out}$, depends on whether a multi-track or single-track vertex was reconstructed.

The outlier parameters are determined from data in the same lifetime fit that extracts the $R_{\rm Det}$ parameters and are given in Table C.4.

Appendix D

$B^0 \to K^0_S \pi^0$ Toy MC Study

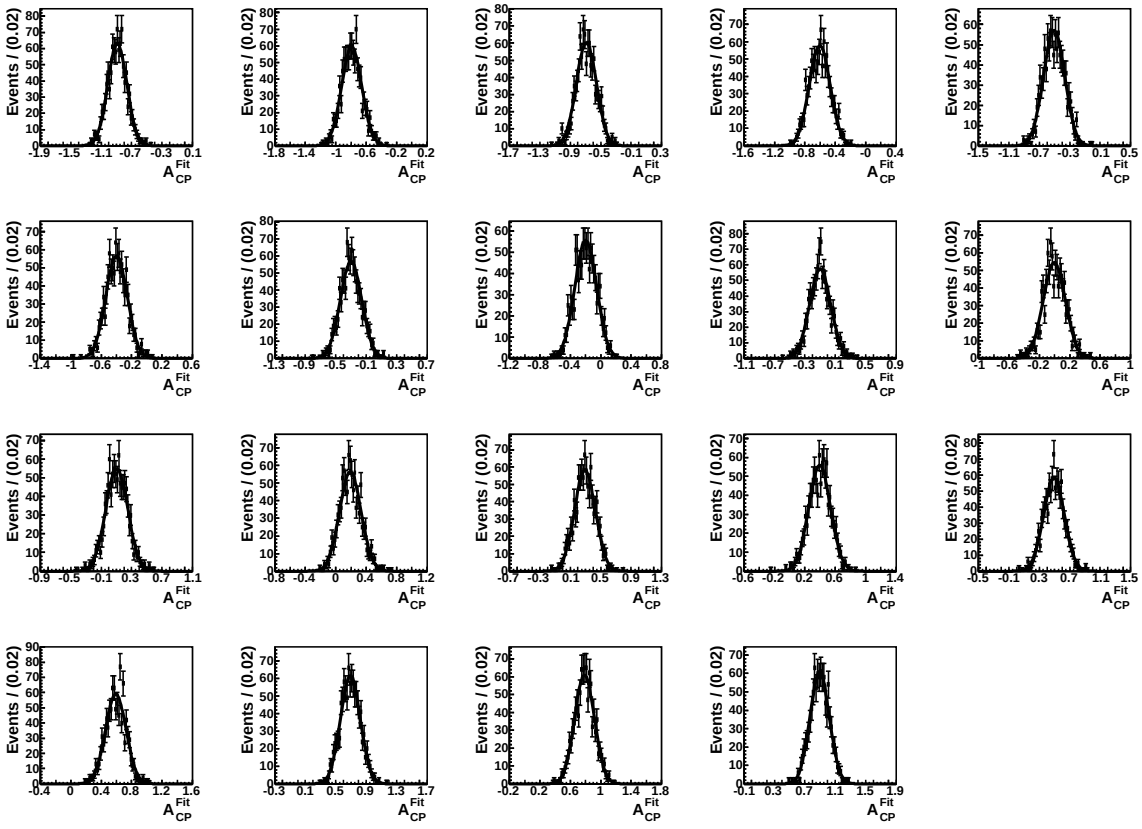


Figure D.1: These plots from left to right, top to bottom range from $\mathcal{A}_{CP}^{\rm Gen} = -0.9$ to $\mathcal{A}_{CP}^{\rm Gen} = 0.9$ and show the distributions of fit results, $\mathcal{A}_{CP}^{\rm Fit}$. The mean and its error from the Gaussian fit to these distributions form the points in the toy MC linearity test (Fig. 5.11).

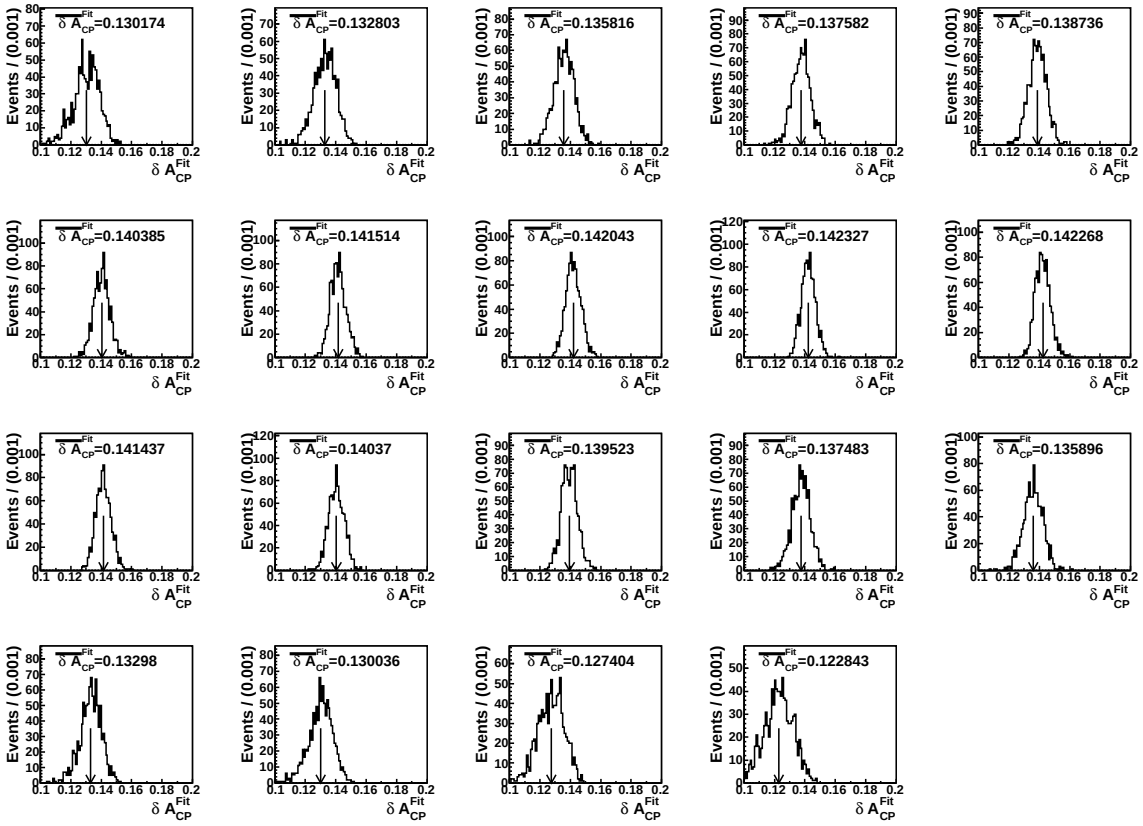


Figure D.2: These plots from left to right, top to bottom range from $\mathcal{A}_{CP}^{\rm Gen} = -0.9$ to $\mathcal{A}_{CP}^{\rm Gen} = 0.9$ and show the distributions of fit errors, $\delta\mathcal{A}_{CP}^{\rm Fit}$. The expected error on $\mathcal{A}_{CP}$ is estimated using the Standard Model input, $\mathcal{A}_{CP}^{\rm Gen} = 0.0$.

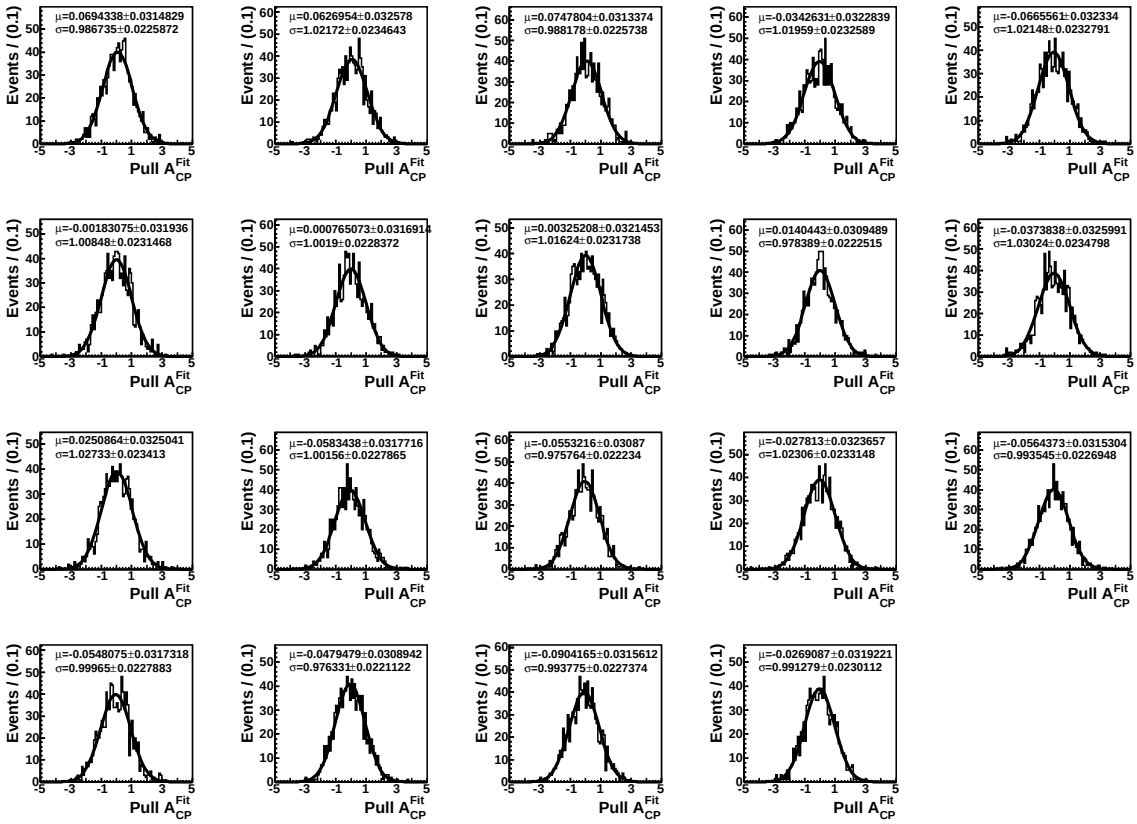


Figure D.3: These plots from left to bottom, top to bottom range from $\mathcal{A}_{CP}^{\rm Gen} = -0.9$ to $\mathcal{A}_{CP}^{\rm Gen} = 0.9$ and show the $\mathcal{A}_{CP}$ pull distributions. A Gaussian fit to these distributions demonstrate consistency with zero shift and unity width.

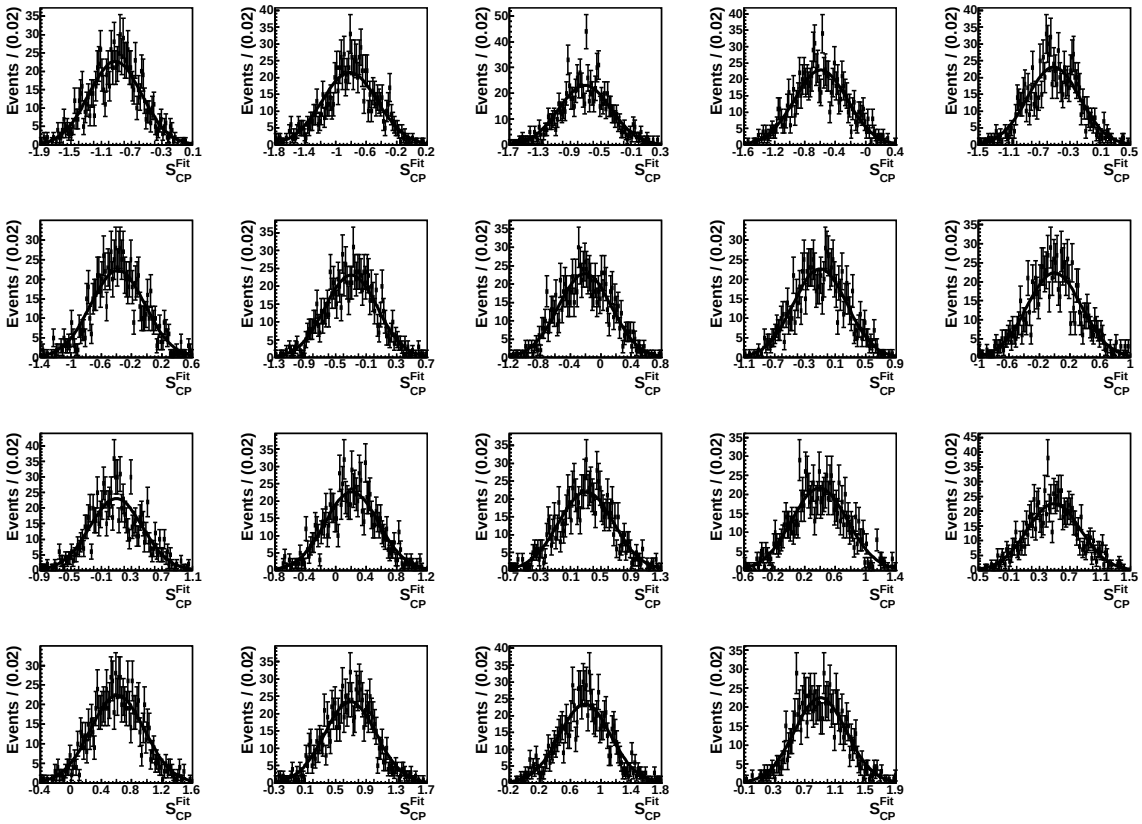


Figure D.4: These plots from left to right, top to bottom range from $\mathcal{S}_{CP}^{\rm Gen} = -0.9$ to $\mathcal{S}_{CP}^{\rm Gen} = 0.9$ and show the distributions of fit results, $\mathcal{S}_{CP}^{\rm Fit}$. The mean and its error from the Gaussian fit to these distributions form the points in the toy MC linearity test (Fig. 5.11).

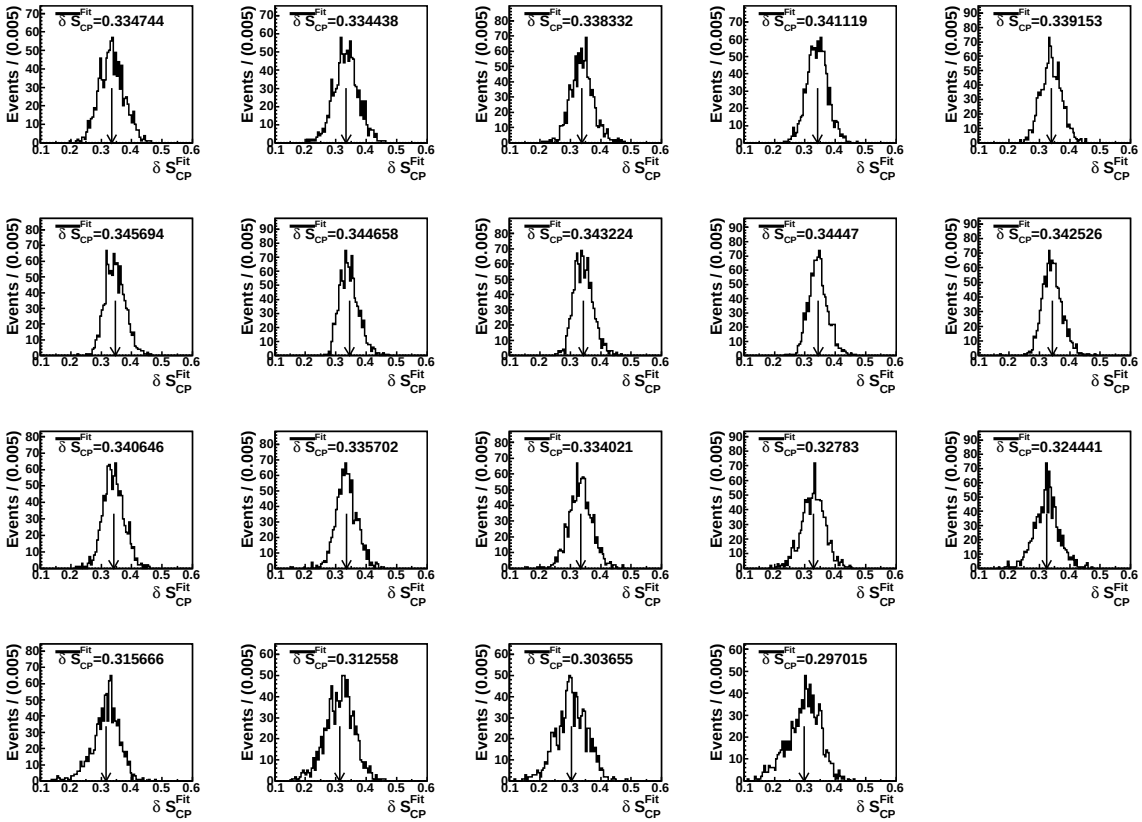


Figure D.5: These plots from left to right, top to bottom range from $\mathcal{S}_{CP}^{\rm Gen} = -0.9$ to $\mathcal{S}_{CP}^{\rm Gen} = 0.9$ and show the distributions of fit errors, $\delta\mathcal{S}_{CP}^{\rm Fit}$. The expected error on $\mathcal{S}_{CP}^{\rm eff}$ is estimated using the near Standard Model input, $\mathcal{S}_{CP}^{\rm Gen} = 0.7$.

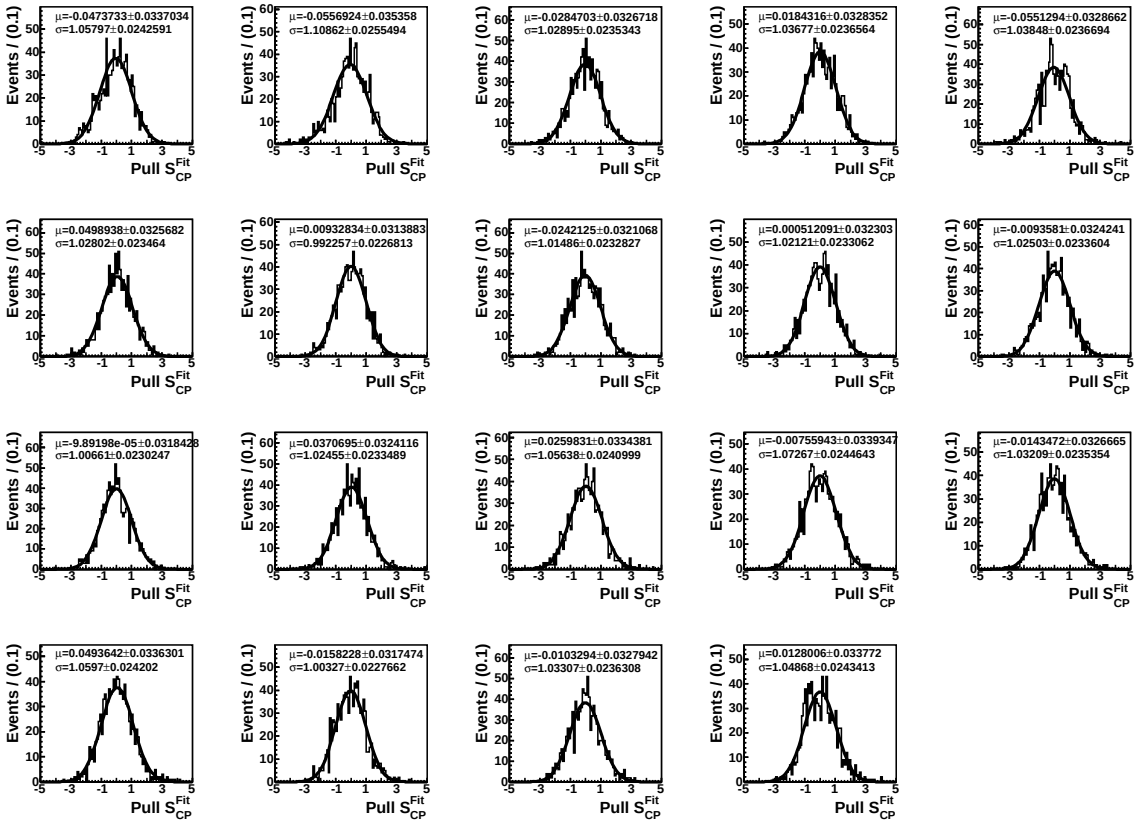


Figure D.6: These plots from left to right, top to bottom range from $\mathcal{S}_{CP}^{\rm Gen} = -0.9$ to $\mathcal{S}_{CP}^{\rm Gen} = 0.9$ and show the $\mathcal{S}_{CP}^{\rm eff}$ pull distributions. A Gaussian fit to these distributions demonstrate consistency with zero shift and unity width.

Appendix E

$B^0 \to K^0_L \pi^0$ Study

E.1 Event Selection

The K^0_L candidates are selected if their momentum in the CMS, $p^*_{K^0_L}$, is greater than 2.2 GeV/c, and $\cos\theta_{\rm miss} > 0.6$, where $\theta_{\rm miss}$ is the angle between the K^0_L and the missing momentum in the laboratory frame. For K^0_L candidates with KLM and ECL information (KLMECL candidates), a likelihood ratio, $\mathcal{L}_{\rm KLMECL}$, is constructed from four variables: the ratio between the energy deposit in 9 CsI crystals and 25 crystals around the shower centroid of the ECL cluster, $E9/E25$, the shower width is the lateral ECL shower spread, the shower invariant mass calculated from all CsI hits in the ECL cluster, and the distance between K^0_L candidate cluster and the nearest charged track. We require $\mathcal{L}_{\rm KLMECL} > 0.6$. We don't use K^0_L candidates identified only by ECL information (ECL candidates) because of the poor signal-to-noise ratio. The π^0 candidates are reconstructed using the same procedure as for $B^0 \to K^0_S \pi^0$ given in §3.6 except for the χ^2 condition.

Since the energy of the K^0_L cannot be measured, we calculate an expected K^0_L direction using the four-momentum of a reconstructed π^0 candidate and the IP, and the direction determined by the the K^0_L cluster in the detector. For calculation of the momentum of the K^0_L, we use energy and momentum conservation laws in B decays,

$$E_{B^0} = E_{K^0_L} + E_{\pi^0}, \tag{E.1}$$
$$\vec{p}_{B^0} = \vec{p}_{K^0_L} + \vec{p}_{\pi^0}, \tag{E.2}$$

and the energy-momentum relations for each particle,

$$E_{B^0} = \sqrt{|\vec{p}_{B^0}|^2 + M^2_{B^0}}, \tag{E.3}$$
$$E_{K^0_L} = \sqrt{|\vec{p}_{K^0_L}|^2 + M^2_{K^0_L}}, \tag{E.4}$$
$$E_{\pi^0} = \sqrt{|\vec{p}_{\pi^0}|^2 + M^2_{\pi^0}}, \tag{E.5}$$

where E, $\vec{p}$ and M are the energy, momentum and invariant mass of the particles, respectively. From Eq. (E.2) and (E.3),

$$\begin{aligned} E_{B^0}^2 &= |\vec{p}_{B^0}|^2 + M_{B^0}^2 \\ &= |\vec{p}_{K_L^0}|^2 + |\vec{p}_{\pi^0}|^2 + 2|\vec{p}_{K_L^0}||\vec{p}_{\pi^0}|\cos\theta_{K-\pi} + M_{B^0}^2, \end{aligned} \tag{E.6}$$

where $\theta_{K-\pi}$ is an angle between the π^0 and K_L^0 flight directions. From Eq. (E.1), (E.4) and (E.5),

$$\begin{aligned} E_{B^0}^2 &= (E_{K_L^0} + E_{\pi^0})^2 \\ &= |\vec{p}_{K_L^0}|^2 + M_{K_L^0}^2 + |\vec{p}_{\pi^0}|^2 + M_{\pi^0}^2 + 2\sqrt{(|\vec{p}_{K_L^0}|^2 + M_{K_L^0}^2)E_{\pi^0}}. \end{aligned} \tag{E.7}$$

From Eq. (E.6) and (E.7),

$$M_{B^0}^2 - M_{K_L^0}^2 - M_{\pi^0}^2 + 2|\vec{p}_{K_L^0}||\vec{p}_{\pi^0}|\cos\theta_{K-\pi} = 2\sqrt{(|\vec{p}_{K_L^0}|^2 + M_{K_L^0}^2)E_{\pi^0}}. \tag{E.8}$$

Squaring of both sides and defining $\Delta M \equiv M_{K_L^0}^2 + M_{\pi^0}^2 - M_{B^0}^2$,

$$\begin{aligned} &\Delta M^2 + 4|\vec{p}_{K_L^0}|^2|\vec{p}_{\pi^0}|^2\cos^2\theta_{K-\pi} - 4\Delta M|\vec{p}_{K_L^0}||\vec{p}_{\pi^0}|\cos\theta_{K-\pi} = 4(|\vec{p}_{K_L^0}|^2 + M_{K_L^0}^2)E_{\pi^0}, \\ &[1 - (|\vec{p}_{\pi^0}|\cos\theta_{K-\pi}/E_{\pi^0})^2]|\vec{p}_{K_L^0}|^2 + \Delta M|\vec{p}_{\pi^0}|\cos\theta_{K-\pi}/E_{\pi^0}^2|\vec{p}_{K_L^0}| \\ &\qquad + M_{K_L^0}^2 - (\Delta M/2E_{\pi^0})^2 = 0. \end{aligned} \tag{E.9}$$

Since all variables except for $|\vec{p}_{K_L^0}|$ are observables, we can obtain magnitude of the K_L^0 momentum by solving Eq. (E.9). The expected K_L^0 direction is calculated from $\vec{p}_{\pi^0}$ assuming the B meson is at rest. The angle between observed and expected K_L^0 direction $\cos\theta_{\rm exp}$ is required to be greater than 0.9. For continuum event suppression, $\mathcal{L}_{s/b} > 0.5$ is required where $\mathcal{R}_{s/b}$ is based on the Fisher discriminating variables used for $B^0 \to K_S^0\pi^0$, the number of hit layers in KLM and the reconstructed to measured energy ratio in the ECL for KLMECL candidates.

We retain events which satisfy $M_{\rm bc} > 5.22$ GeV/c^2. Multiple candidates are predominantly due to several K_L^0 candidates in the event and the contribution from multiple pion candidates is found to be small. In the case of multiple candidates, we take the K_L^0 candidate with the smallest $\theta_{\rm exp}$ in the event.

E.2 Signal Yield Extraction

The signal yield is determined from an unbinned two-dimensional maximum likelihood fit to the $M_{\rm bc} - \mathcal{L}_{s/b}$ distribution in the range 5.22 GeV/$c^2 < M_{\rm bc} <$ 5.29 GeV/c^2 and $\mathcal{L}_{s/b} > 0.5$ in each r-bin.

The signal PDF shape is constructed from a two-dimensional histogram in the $M_{\rm bc} - \mathcal{L}_{s/b}$ plane obtained from MC and its fraction is a free parameter in the fit. The ratio of signal yields among the different r-bins is fixed to that obtained from signal MC.

For the continuum background, we use the ARGUS parameterisation for $M_{\rm bc}$ and a histogram for $\mathcal{R}_{s/b}$ determined from off-resonance data samples.

The $B\bar{B}$ background distribution is also obtained from MC. We found that the $B\overline{B}$ background is predominantly due to rare charmless B decays, while the generic $B\overline{B}$ events, which are dominated by B decays caused by the leading $b \to c$ transitions, bring small background contributions compared to the continuum background. The major contributions are from three-body charmless decays ($K\pi\pi, K^*\pi$) in which one pion has low momentum.

The signal extraction yields 285 ± 52 (stat) $\pm$ 57 (syst) events which is of 3.7σ significance and the fit result is shown in Fig. E.1.

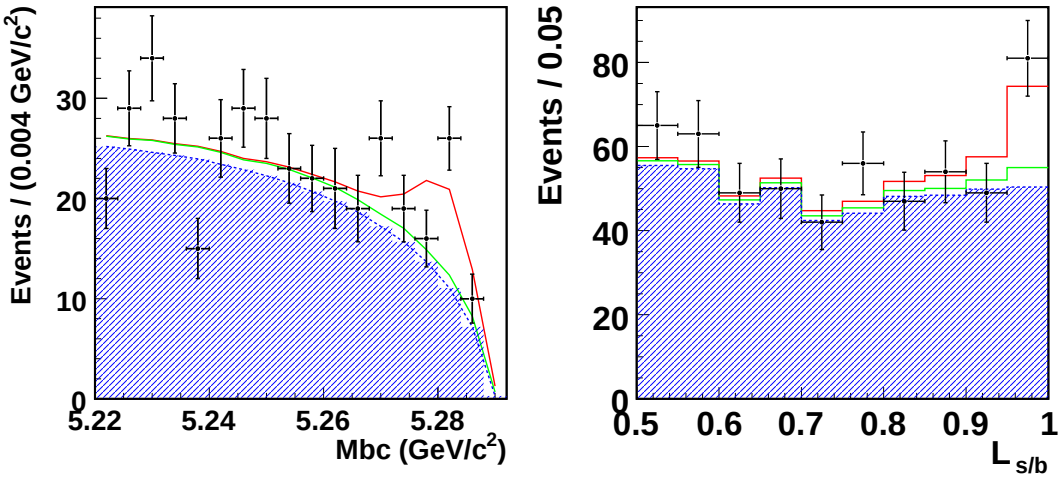


Figure E.1: $M_{\rm bc}$ (left) and $\mathcal{L}_{s/b}$ (right) projections in the best r-bin region, $r > 0.875$. The fit results are indicated by smoothed line or histograms. The colour indicates the category of sample: signal (red), rare B (green) and continuum (blue).

E.3 *CP* Fit Results

Since the decay vertex of $B^0 \to K^0_L\pi^0$ cannot be determined, only direct CP violation can be measured. Using Eq. 5.16, the direct CP of the $B^0 \to K^0_L\pi^0$ sub-sample is found to be,

$$\mathcal{A}_{CP} = -0.013 \pm 0.448 \text{ (stat)}. \tag{E.10}$$

Bibliography

[1] "CERN - The world's largest particle physics laboratory". http://www.cern.ch/.

[2] A.D. Sakharov. JETP Lett. **5**, 24 (1967).

[3] J. H. Christenson *et al.* Phys. Rev. Lett. **13**, 138 (1964).

[4] S.L. Glashow. Nucl. Phys. **22**, 579 (1961).

[5] S. Weinberg. Phys. Rev. Lett. **19**, 1264 (1967).

[6] G. t'Hooft. Nucl. Phys. **33**, 173 (1971); **35**, 167 (1971).

[7] M. Kobayashi and T. Maskawa. Prog. Theor. Phys. **49**, 652 (1973).

[8] A. Carter and A.I. Sanda. Phys. Rev. Lett. **45**, 952 (1980); Phys. Rev. D **23**, 1567 (1981).

[9] I.I. Bigi and A.I. Sanda. Nucl. Phys. **193**, 851 (1981).

[10] B. Aubert *et al.* BABAR Collaboration. Phys. Rev. Lett. **87**, 091801 (2001).

[11] K. Abe *et al.* Belle Collaboration. Phys. Rev. Lett. **87**, 091802 (2001).

[12] G. Segre S.M. Barr and H.A. Weldon. Phys. Rev. D **20**, 2494 (1979).

[13] M. E. Peskin and D. V. Schroeder. An introduction to quantum field theory. Addison-Wesley Advanced Book Program (1995).

[14] L. Wolfenstein. Phys. Rev. Lett. **51**, 1945 (1983).

[15] N. Cabibbo. Phys. Rev. Lett. **8**, 214 (1964).

[16] T. Inami and C.S. Lim. Prog. Theor. Phys. **65**, 297 (1981).

[17] F. Gilman and M. Wise. Phys. Rev. D **27**, 1128 (1983).

[18] M. Jamin A.J. Buras and P.H. Weisz. Nucl. Phys. B **347**, 491 (1990).

[19] J.S. Hagelin. Nucl. Phys. B **193**, 123 (1981).

[20] E. Barberio *et al.* (HFAG). arXiv:0808.1297 (2008).

[21] J. Charles *et al.* (CKMfitter Group). Eur. Phys. J. C **41**, 1-131 (2005).

[22] B. Aubert *et al.* (BaBar Collab.). Phys. Rev. D **71**, 032005 (2005).

[23] B. Aubert *et al.* (BaBar Collab.). Phys. Rev. D **74**, 091101 (2006).

[24] B. Aubert *et al.* (BaBar Collab.). Phys. Rev. Lett. **99**, 231802 (2007).

[25] R. Itoh *et al.* (Belle Collab.). Phys. Rev. Lett. **95**, 091601 (2005).

[26] P. Krokovny *et al.* (Belle Collab.). Phys. Rev. Lett. **97**, 081801 (2006).

[27] J. Dalseno *et al.* (Belle Collab.). Phys. Rev. D **76**, 072004 (2007).

[28] Y. Nir Y. Grossman, Z. Ligeti and H. Quinn. Phys. Rev. D **68**, 015004 (2003).

[29] Y. Grossman M. Gronau and J. L. Rosner. Phys. Lett. B **579**, 331 (2004).

[30] J. L. Rosner M. Gronau and J. Zupan. Phys. Lett. B **596**, 107 (2004).

[31] C.-K. Chua H.-Y. Cheng and A. Soni. Phys. Rev. D **72**, 014006 (2005).

[32] M. Gronau and J. L. Rosner. Phys. Rev. D **71**, 074019 (2005).

[33] M. Beneke. Phys. Lett. B **620**, 143 (2005).

[34] Y. Nir G. Engelhard and G. Raz. Phys. Rev. D **72**, 075013 (2005).

[35] C.-K. Chua H.-Y. Cheng and A. Soni. Phys. Rev. D **72**, 094003 (2005).

[36] A. R. Williamson and J. Zupan. Phys. Rev. D **74**, 014003 (2006).

[37] S.-W. Lin *et al.* (The Belle collaboration). Nature **452**, 332 (2008).

[38] B. Aubert *et al.* (The BaBar Collaboration). arXiv:0807.4226 (2008).

[39] M. Gronau. Phys. Lett. **B627**, 82 (2005).

[40] M. Gronau and J. L. Rosner. arXiv:0807.3080 (2008).

[41] R. Fleischer *et al.* arXiv:0806.2900 (2008).

[42] J. Zupan. arXiv:0707.1323 (2007).

[43] B. Aubert *et al.* (The BaBar Collaboration). arXiv:0807.4226 (2008).

[44] S.-W. Lin *et al.* (The Belle Collaboration). Phys. Rev. Lett. **99**, 121601 (2007).

[45] B. Aubert *et al.* (The BaBar Collaboration). arXiv:0809.1174 (2008).

[46] Y. Chao *et al.* (The Belle Collaboration). Phys. Rev. D **76**, 091103(R) (2007).

[47] National Laboratory for High Energy Physics. "KEKB B-factory Design Report". Phys. Rev. D **43**, 2193 (1993).

[48] S. Eidelman *et al.* (Particle Data Group). Phys. Lett. B **592**, 1 (2004).

[49] A. Abashian *et al.* Belle Collaboration. Nucl. Instr. and Meth. A **479**, 117 (2002).

[50] V. Chabaud *et al.* Nucl. Instr. and Meth. A **368**, 314 (1996).

[51] B. Casey. "HadronB". Belle Note #390.

[52] J. Tanaka. "Kinematic Fitting". Belle Note #194.

[53] F. Fang. "Study of $K_S \to \pi^+\pi^-$ Selection". Belle Note #323.

[54] K.-F. Chen *et al.* Belle note 329, unpublished.

[55] R.A. Fisher. Annals of Eugenics **7**, 179 (1936).

[56] "The Number of B events in the Belle Hadronic Skims". http://belle.kek.jp/secured/nbb/nbb.html.

[57] "The EvtGen package home page". http://www.slac.stanford.edu/~lange/EvtGen/.

[58] R. Brune *et al.* "GEANT 3.21". CERN DD/EE/84-1 (1984).

[59] P. Chang S.W. Lin and H. C. Huang. "Update of π^0 Systematics Using Inclusive η (78/fb)". Belle Note #645.

[60] H. Albrecht *et al.* ARGUS Collaboration. Z. Phys. C **48**, 543 (1990).

[61] M. Fujikawa and Y. Yusa *et al.* (Belle Collab.). arXiv:0809.4366 [hep-ex].

[62] K.F. Chen *et al.* "Time-Dependent CP Asymmetries in $B^0 \to \eta'\, K^0$, $\phi\, K^0$, $K^0_S\, K^0_S\, K^0_S$ and $B^0 \to J/\psi\, K^0$ decays with 532 Million $B\overline{B}$ Pairs". Belle Note #929.

[63] R. N. Cahn O. Long, M. Baak and D. Kirkby. Phys. Rev. D **68**, 034010 (2003).

[64] A. Bornheim *et al.* (The CLEO Collaboration). Phys. Rev. D **68**, 052002 (2003).

[65] M. Gronau and J. L. Rosner. arXiv:0807.3080 (2008).

[66] "MINUIT, Function Minimization and Error Analysis". http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/minmain.html.

[67] H. Kakuno *et al.* Nucl. Instr. and Meth. A **533**, 516 (2004).

[68] H. Tajima *et al.* Nucl. Instr. and Meth. A **533**, 370 (2004).