Investigation of B^+ mesons decay to $K^+K^-\pi^+$ at the Belle experiment

Chia-Ling Hsu

orcid.org/0000-0002-1641-430X

Submitted in total fulfilment of the requirements
of the degree of Doctor of Philosophy

June 2017

School of Physics
The University of Melbourne

Abstract

Charmless decays of B mesons to three charged hadrons are suppressed in the Standard Model, and thus provide an opportunity to search for physics beyond the Standard Model. An unexpected excess and a large CP asymmetry in the low invariant mass spectrum of the K^+K^- system for the decay $B^+ \to K^+K^-\pi^+$ were observed by BaBar and LHCb in recent years.

We present the measurements of branching fraction and direct CP asymmetry of the charmless decay $B^+ \to K^+K^-\pi^+$. This analysis is performed on a data sample of $772 \times 10^6 B\overline{B}$ pairs produced at the $\Upsilon(4S)$ resonance by the KEKB asymmetric-energy e^+e^- collider and collected by the Belle detector.

We perform a blind analysis, examining signal reconstruction and background suppression with Monte Carlo simulated samples, and extract signal yield and direct CP asymmetry with a 2D extended maximum likelihood fit to the data. The measured branching fraction and direct CP asymmetry are

$$\mathcal{B}(B^+ \to K^+K^-\pi^+) = (5.38 \pm 0.40 \pm 0.35) \times 10^{-6}$$

and

$$\mathcal{A}_{CP} = -0.170 \pm 0.073 \pm 0.017,$$

respectively, where the first uncertainties are statistical and the second are systematic. These results are in agreement with the current world average.

We extract the branching fraction and direct *CP* asymmetry as a function of the K^+K^- invariant mass. The K^+K^- invariant mass distribution of the signal candidates shows an excess in the region below 1.5 GeV/c^2, which is consistent with the previous studies from BaBar and LHCb. Strong evidence of a large direct *CP* asymmetry of $-0.90 \pm 0.17 \pm 0.03$ with 4.8σ significance is found in the K^+K^- low-invariant-mass region.

Declaration

This is to certify that,

1. the thesis comprises only my original work towards the PhD except where indicated.
2. Due acknowledgement has been made in the text to all other material used.
3. The thesis is fewer than 100,000 words in length, exclusive of tables, bibliographies and appendices.

Chia-Ling Hsu

Acknowledgements

It is impossible to list all people I would like to thank here. To all the people that I do not list by name here, please know, you will be in my heart.

First and foremost, I would like to thank my supervisor, Martin Sevior. He is always patient, optimistic and active guiding me whenever I need. Without his support and encouragement I would not be able to finish the analysis or fulfill the graduation requirements. Also I would like to express my sincere thanks to Phillip Urquijo for his suggestions and patient advising which helped me a lot through my research.

This data analysis work is truly a collaborative effort of hundreds of Belle and KEKB members. I must extend thanks to all the people in the Belle collaboration and the KEKB accelerator team. I appreciate the excellent insights and suggestions from Belle experts who gave me useful suggestions on my analysis and writing of the journal article, special thanks to K. Miyabayashi, Martin Ritter, M. Takahashi, Y. Sakai, Gagan, Karim, for their detailed reviewing this analysis. Also thanks to Yun-Tsung Lai, Hulya, Xiaolong Wang, and all the friends that spent some time with me during my stay at KEK.

I also express my gratitude to the post-doc researchers, Geng-Yuan, Yi, Chunhau, and David, who gave advice and discussed with me when I met obstacles during my study. I am grateful to all my fellow students in CoEPP that spent some time with me in the office for a few months or years. I probably spent most of my time in the office, and I leant a lot from all of you guys. Especially thanks to Laurence, Eiasha, Marco, Alexander, it's always fun to talk with you, and those discussions after work (mostly during work) have opened my mind.

I must thank Anton and Tristan, who helped my read through the drafts of my belle note, paper, and thesis. You guys continue to help me a lot on

correcting English, and teach me how to use “proper English”. I will buy you many drinks!

I am grateful to all my friends in Melbourne for their company. Special thanks to Kei, Eilonwy, Kuan-Yin, Leya, and all the members at the Nonbei club (aka. Drinker club). Because of you, my life in Melbourne has been colorful. I am lucky to have all of you in my life!

Last but not least, I would like to extend a sincere thanks to my family for the love, support, and constant encouragement I have received over the years.

Preface

This thesis summarizes my original work on the analysis of $B^+ \to K^+K^-\pi^+$ decay, which I conducted as a member of the Belle collaboration under the supervision of Assoc. Prof. Martin Sevior.

Chapter 1 provides an original overview of the Standard Model and *CP* violation as it pertains to this analysis.

Chapter 2 is an original overview of the Belle experiment and the KEKB accelerator whose data was used in this analysis.

Chapter 3, 4 and 5 presents the $B^+ \to K^+K^-\pi^+$ measurements which was performed by this author which formed the basis of this thesis and the paper "Measurement of branching fraction and direct *CP* asymmetry in charmless $B^+ \to K^+K^-\pi^+$ decays at Belle" has been published on the journal Physical Review D(RC) [1].

The appendix presents the work which was done by the author to extract the signal distribution in M_{KK}. However as described in A.2, this method was abandoned as unusable in favor of the method described in Sec. 3.3.2.

Otherwise all work was performed by the author unless stated.

Contents

List of figures

List of tables

Chapter 1.

Introduction

1.1. Particle Physics

The innate curiosity of humans leads us to ask "What's the universe made of?". In the fifth century BC, the Greek philosopher Leucippus and his student Democritus proposed that the world was made of atoms, from the Greek word "atomos" which means uncuttable.

In the early 19th century, John Dalton presented Atomic Theory, in which he claimed that all matter is composed of atoms with definite weights. Although at that time, the atoms could not be observed directly because of the technological limitations, Atomic Theory predicted some concrete characteristics based on experiments. A few decades later, Dmitri Mendeleev presented a periodic table of elements on which the elements were arranged according to their chemical properties. The regularity in the periodic table indicated that the elements might be made of smaller particles.

Until the discovery of the electron in 1897 by J. J. Thomson, atoms were regarded as the fundamental particles. A few decades later, Ernest Rutherford and Sir James Chadwick discovered the proton and neutron. These three types of particles construct an atom, it seemed the question of the foundation of matter was answered.

Yet, a question remained about the nucleus: what kind of force is binding protons and neutrons together in a nucleus? Hideki Yukawa proposed the theory of mesons and predicted the existence of mesons in 1935. In the theory

of mesons, mesons are interchanged by nucleons to provide the needed force. The pion meson (π) was discovered about 10 years later. The prediction and existence of a subatomic particle led to a rapid development of particle physics, and more and more kinds of particles were found in the following decades.

Until the 1940's, particle physicists could only do experiments via high energy cosmic rays, which are uncontrollable and infrequent. While energetic cosmic rays might scatter particles in the atmosphere, such interactions could produce subatomic particles. To provide a controllable interaction source, scientists built particle accelerators to propel charged particles (electrons, positrons or protons) to high energy, which were then collided with a target or another energetic particle beam.

1.2. The Standard Model

The Standard Model (SM) of particle physics is the theory describing the characteristics and interactions of fundamental particles. It's one of most successful theories and has passed numerous experimental tests. In the SM, the elementary particles are classified in four categories: 6 quarks, 6 leptons, 4 vector gauge bosons, and the scalar Higgs boson. Based on the SM, all matter consists of quarks and leptons. The quarks carry fractional unit charges - $+2/3$ or $-1/3$, and leptons carry integral charge. All of these are spin-1/2 particles (fermions), and can be classified in three generations. The properties of fermions are shown in Table 1.1. The gauge bosons are exchanged between particles, mediating the interactions: photons for the electromagnetic force, $W^{\pm}$ and Z^0 bosons for the weak interaction, 8 gluons for the strong interaction, and the scalar Higgs particle which provides mass to the fundamental fermions as well as the W and Z. Still, the SM is not a complete theory of fundamental particle interactions. Not only does it not incorporate the physics of general relativity, such as gravitation and dark energy, but also some properties, such as the masses of the fermions and bosons, are not predicted and have to be measured from experiments.

Table 1.1.: Organization of Leptons and Quarks. [2]

		Name	Symbol	Charge(e)	Mass(MeV/c^2)
First generation	Quark	Up	u	$+\frac{2}{3}$	1.5 – 3.3
		Down	d	$-\frac{1}{3}$	3.5 – 6.0
	Lepton	Electron	e	-1	0.511
		Electron neutrino	ν_e	0	< 0.0000022
Second generation	Quark	Charm	c	$+\frac{2}{3}$	1160 – 1340
		Strange	s	$-\frac{1}{3}$	70 – 130
	Lepton	Muon	μ	-1	105.7
		Muon neutrino	ν_μ	0	< 0.17
Third generation	Quark	Top	t	$+\frac{2}{3}$	169100 – 173300
		Bottom	b	$-\frac{1}{3}$	4130 – 4370
	Lepton	Tau	τ	-1	1777
		Tau neutrino	ν_τ	0	< 15.5

Table 1.2.: Properties of gauge and Higgs bosons. [2]

Mediator	Charge(e)	Mass(GeV/c^2)	Interaction
Gluon(g)	0	0	strong
Photon(γ)	0	0	electromagnetic
$W^\pm$	± 1	80.4	(charged) weak
Z^0	0	91.2	(neutral) weak
H^0	0	125.1	

1.2.1. *CP* asymmetry

The origin of the universe is generally believed to start from the Big Bang. According to Einstein's mass-energy equivalence formula ($E = mc^2$) and pair production ($\gamma \rightarrow e^+e^-, q\bar{q}, ...$), matter and anti-matter should be equally produced from energy. In the other words, there should be equal amount of matter and anti-matter in the world. In reality, the universe is chiefly made of matter rather than consisting of equal parts of matter and anti-matter.

In 1967, Andrei Sakharov proposed three necessary conditions for producing matter and anti-matter at a different rate [3]. The three necessary Sakharov conditions are:

- Baryon number violation.
- *CP*-symmetry violation.
- Interaction out of thermal equilibrium.

The *C* stands for the charge conjugate operator, which flips all quantum charges. For example, the B^+ meson's charge conjugate is the B^- meson. From here on the inclusion of the charge-conjugate decay is implied, unless explicitly stated otherwise. The *P* stands for the parity operator, which creates the mirror image of the system in spatial dimensions. Until 1956, parity conservation was believed to be one of the fundamental conservation laws. T.-D. Lee and C.-N. Yang revealed that while parity conservation had been proved in decays mediated by the strong or electromagnetic interactions, it was unverified in the weak interaction [4]. Soon, parity violation was discovered in the β decay of Co-60 nuclei experiment conducted by C.-S. Wu in 1957 [5]. The *CP* symmetry was also believed to be a conserved quantity until the violation of *CP* symmetry was first observed in the decays of the neutral kaon system by Cronin and Fitch in 1964 [6]. However the violation of *CP* symmetry is small in the kaon system, at the order of 10^{-3} only, and is insufficient to explain the asymmetry of matter and anti-matter in the universe.

The violation of *CP* symmetry is incorporated in the Standard Model by including the complex phase in the Cabibbo-Kobayashi-Maskawa matrix which describes the quark mixing.

1.2.2. The Cabibbo-Kobayashi-Maskawa quark-mixing Matrix

In order to explain the observed suppression of strangeness changing ($\Delta S = 1$) transitions, N. Cabibbo introduced the idea of a mixing angle, θ_c [7]. This applied different couplings to $\Delta S = 1$ and $\Delta S = 0$ weak currents by having the superposition of the two different currents be the overall current in the interactions. Or in quark form, the down-type quarks form a weak eigenstate, which is a superposition from their mass eigenstates via a rotation of the form

$$d' = d\cos\theta_c + s\sin\theta_c. \tag{1.1}$$

The GIM mechanism proposed by Glashow, Iliopoulos, and Maiani in 1970 extended this to a 2×2 unitary rotation matrix, in order to provide a mechanism for flavour changing neutral currents (FCNC) being heavily suppressed in loop diagrams [8].

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}. \tag{1.2}$$

This predicted the existence of the 4th (charm) quark and placed a unitary quark mixing matrix into the SM.

In order to explain the *CP* violation in the weak interaction, M. Kobayashi and T. Maskawa realized that a complex mixing matrix was necessary for *CP* violation, which is required to be at least 3×3 to have the necessary degrees of freedom. This predicts the existence of the third generation of quarks [9]. They generalized the Cabibbo matrix into the Cabibbo-Kobayashi-Maskawa (CKM) matrix by introducing a complex term into the mixing matrix. The CKM matrix is often expressed in terms of a 3×3 unitary matrix V operating on the charge $-e/3$ quark mass eigenstates (d, s, and b):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \mathbf{V_{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1.3}$$

where magnitudes of the elements are the coupling strength between quarks and the W-boson. The values of elements are determined by the measured decay widths of weak interactions of the relevant quarks.

There are several parameterizations of the CKM matrix. Kobayashi and Maskawa chose a parametrization involving the four angles θ_1, θ_2, θ_3, and δ:

$$\mathbf{V_{CKM}} = \begin{pmatrix} c_1 & -s_1c_3 & -s_1s_3 \\ s_1c_2 & c_1c_2c_3 - s_2s_3e^{i\delta} & c_1c_2c_3 + s_2c_3e^{i\delta} \\ s_1s_2 & c_1s_2c_3 + c_2s_3e^{i\delta} & c_1s_2s_3 - c_2c_3e^{i\delta} \end{pmatrix}, \tag{1.4}$$

where c_i and s_i are the cosine and sine value of θ_i for $i = 1, 2, 3$. When $\theta_2 = \theta_3 = 0$, it reduces to the Cabibbo mixing with $\theta_1 = \theta_c$.

The standard parametrization of V that uses angles θ_{12}, θ_{23}, θ_{13}, and δ_{13}:

$$\mathbf{V_{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{13} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{13} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix}, \tag{1.5}$$

where c_{ij} and s_{ij} are the cosine and sine values of θ_{ij} for the "generation" labels $i, j = 1, 2, 3$. With this expression, the rotation angles are defined and labeled in a way which relates the mixing of two specific generations. If one of these angles vanishes, so does the mixing between those two generations.

A popular parameterization proposed by Wolfenstein emphasizes the hierarchy in the size of the angle, $s_{12} \gg s_{23} \gg s_{13}$, written in the form of a Taylor expansion in λ [10]:

$$\mathbf{V_{CKM}} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4), \tag{1.6}$$

where λ is the sine value of the Cabibbo angle, and A, ρ and η are real numbers that are of order unity. This approximation is widely utilized for B physics, and implies that $b \to c$ transitions are more probable than $b \to u$.

These four quantities in the Wolfenstein parameterization can be determined from experimental measurement results [2]:

$$\begin{aligned} \lambda &= 0.22506 \pm 0.00050, \quad A = 0.811 \pm 0.026, \\ \bar{\rho} &= 0.124^{+0.019}_{-0.018}, \qquad\quad \bar{\eta} = 0.356 \pm 0.011. \end{aligned} \tag{1.7}$$

The magnitudes of nine CKM elements are

$$\mathbf{V_{CKM}} = \begin{pmatrix} 0.97434^{+0.00011}_{-0.00012} & 0.22506 \pm 0.00050 & 0.00357 \pm 0.00015 \\ 0.22492 \pm 0.00050 & 0.97351 \pm 0.00013 & 0.0411 \pm 0.0013 \\ 0.00875^{+0.00032}_{-0.00032} & 0.0403 \pm 0.0013 & 0.99915 \pm 0.00005 \end{pmatrix}. \tag{1.8}$$

As this matrix is unitary, there are 6 sets of equations relating the components, one is famous in B physics:

$$V_{ud}V_{ub}^{*} + V_{cd}V_{cb}^{*} + V_{td}V_{tb}^{*} = 0. \tag{1.9}$$

This equation can be represented geometrically as a unitarity triangle in the complex plane with corresponding angles as shown in Fig 1.1. In the standard parametrization, V_{cb} is real and so is V_{cd}, which provides a very good approximation. In this relation, all sides are approximately the same size since each is proportional to λ^3.

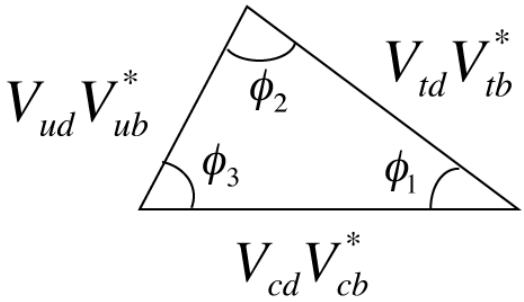


Figure 1.1.: The unitarity triangle in the complex plane.

Table 1.3.: Properties of B mesons [2].

Type	Constituents	$I(J^P)$	Mass(MeV/c^2)	Lifetime(ps)
B^+	$u\overline{b}$	$\frac{1}{2}(0^-)$	5279.17 ± 0.29	1.641 ± 0.008
B^0	$d\overline{b}$	$\frac{1}{2}(0^-)$	5279.50 ± 0.30	1.519 ± 0.007
B_s^0	$s\overline{b}$	$0(0^-)$	5366.3 ± 0.6	1.425 ± 0.041
B_c^+	$c\overline{b}$	$0(0^-)$	6277 ± 6	0.453 ± 0.041

1.3. $\Upsilon(4S)$ and B meson

In 1977 a new resonance named Υ was discovered by E288 at Fermilab [11], which turned out to be a flavorless meson formed by a third generation quark and its anti-quark. In the following years, three more resonances, $\Upsilon(2S)$, $\Upsilon(3S)$, and $\Upsilon(4S)$, were found. Of these resonances, $\Upsilon(4S)$ is the most important for B physics as its mass is above the threshold of a $B\overline{B}$ pair. Its dominant decay mode is to B-meson pairs. The branching fraction of $\Upsilon(4S) \to B^+B^-$ is almost the same as $\Upsilon(4S) \to B^0\overline{B^0}$.

In the Standard Model, B mesons are the bound states of a $\overline{b}$ quark and either a u (B^+), d (B^0), s (B_s^0) or c (B_c^+) quark. The properties of B mesons are shown in Table 1.3. B mesons mainly decay via $b \to c$ transitions. The decay modes of B mesons which do not occur through $b \to c$ transitions are called charmless B decays. The charmless B decays are suppressed by the small magnitude of the element V_{ub} in CKM matrix and are good probes for CP asymmetry (see 1.5) and new physics. However, the branching fractions are very small in charmless B decays, usually $O(10^{-5})$, requiring a very large sample of $B\overline{B}$ pairs to study them. To measure charmless decays and CP violation in B mesons, KEK (Tsukuba, Japan) and SLAC (California, USA) built energy-asymmetry e^+e^- colliders, which were designed to produce a large number of B mesons.

1.4. CP violation in B decays

The internal angles of the CKM triangle can be determined by observing asymmetries in B decays. To see how these asymmetries can arise, we consider

a generic neutral meson. Following the formalism of Bigi and Sanda [12], we develop equations 1.10 to equations 1.16. We start by decomposing the flavor eigenstates of a neutral meson in terms of its mass eigenstates:

$$|P^0\rangle = \frac{1}{2p}(|P_H\rangle + |P_L\rangle) \tag{1.10}$$

$$|\overline{P}^0\rangle = \frac{1}{2q}(|P_H\rangle - |P_L\rangle), \tag{1.11}$$

where P_H and P_L are the heavy and light mass eigenstates, respectively, and p and q are complex parameters with a normalization of $|p|^2 + |q|^2 = 1$. In the B-system, the heavy mass eigenstate $|P_H\rangle$ is CP odd, which means $CP|P^0\rangle = -|\overline{P}^0\rangle$.

The inclusion of time evolution leads to oscillation between P^0 and $\overline{P}^0$ due to their being a linear combination of the mass eigenstates. The oscillation rate is defined by box diagrams involving two W bosons as shown in Fig. 1.2. The time evolution of the flavor eigenstates starting out as P^0 and $\overline{P}^0$ are given by:

$$\begin{aligned} |P^0(t)\rangle &= f_+(t)|P^0\rangle + \frac{q}{p}f_-(t)|\overline{P}^0\rangle \\ |\overline{P}^0(t)\rangle &= f_+(t)|\overline{P}^0\rangle + \frac{p}{q}f_-(t)|P^0\rangle, \end{aligned} \tag{1.12}$$

where $m_{H(L)}$ and $\Gamma_{H(L)}$ are the mass and decay width of heavy (light) eigenstate, and the $f_\pm(t)$ are the time evolution equations of the even and odd CP-states of the B mesons.

$$f_\pm(t) = \frac{1}{2}e^{-iMt}e^{-\frac{1}{2}\Gamma t}(1 \pm e^{-i\Delta m t}e^{-\frac{1}{2}\Delta\Gamma t}), \tag{1.13}$$

where $\Gamma \equiv (\Gamma_L + \Gamma_H)/2$, $M \equiv (m_L + m_H)/2$, $\Delta m \equiv m_H - m_L$, and $\Delta\Gamma \equiv \Gamma_H - \Gamma_L$, where $\Delta\Gamma$ can have either sign.

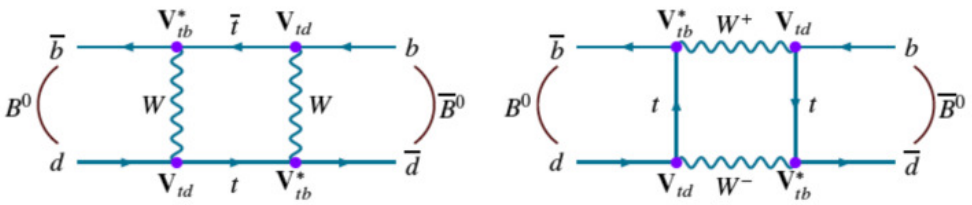


Figure 1.2.: Feynman diagrams for $B^0 - \overline{B}^0$ oscillation.

The amplitude for the decay P^0 and $\overline{P}^0$ to a final state f are denoted by $A(f)$ and $\overline{A}(f)$:

$$\begin{aligned} A(f) &= \langle f|H|P^0\rangle \\ \overline{A}(f) &= \langle f|H|\overline{P}^0\rangle, \end{aligned} \tag{1.14}$$

here H stands for a weak Hamiltonian.

The decay rates can be expressed as follows:

$$\begin{aligned} \Gamma(P^0(t) \to f) \propto\ & e^{-\Gamma t}|A_f|^2 \\ & (\cosh\frac{\Delta\Gamma t}{2} + D_f \sinh\frac{\Delta\Gamma t}{2} + C_f\cos\Delta m t - S_f \sin\Delta m t) \\ \Gamma(\overline{P}^0(t) \to f) \propto\ & e^{-\Gamma t}|A_f|^2|\frac{p}{q}|^2 \\ & (\cosh\frac{\Delta\Gamma t}{2} + D_f \sinh\frac{\Delta\Gamma t}{2} - C_f\cos\Delta m t + S_f \sin\Delta m t), \end{aligned} \tag{1.15}$$

where

$$D_f \equiv -\frac{2\Re(\lambda_f)}{1+|\lambda_f|^2}, \qquad C_f \equiv -\frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}, \qquad S_f \equiv -\frac{2\Im(\lambda_f)}{1+|\lambda_f|^2}, \tag{1.16}$$

where the parameter λ_f is defined as $\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f}$. The sinh- and sin- terms are associated with the interference between the decays with and without oscillation.

CP violation can be classified into three categories:

- **Direct *CP* violation**
  This type of *CP* violation occurs in the presence of a different decay rate for a particle and its *CP* conjugate, i.e. $|A(f)| \neq |\overline{A}(\overline{f})|$. Assuming only two diagrams contributing to the final state, the total amplitude of particle and its anti-particle can be written as:

$$A(P \to f) = |D_1|e^{i\phi_1}e^{i\delta_1} + |D_2|e^{i\phi_2}e^{i\delta_2} \tag{1.17}$$

$$\overline{A}(\overline{P} \to \overline{f}) = |D_1|e^{-i\phi_1}e^{i\delta_1} + |D_2|e^{-i\phi_2}e^{i\delta_2}, \tag{1.18}$$

where D is the decay amplitude, ϕ is the weak (CP-violating) phase, and δ is the strong (CP-conserving) phase. The asymmetry of the decay rate is:

$$|A(P \to f)|^2 = |D_1|^2 + |D_2|^2 + 2|D_1||D_2|\cos(\Delta\delta + \Delta\phi) \quad (1.19)$$

$$|\overline{A}(\overline{P} \to \overline{f})|^2 = |D_1|^2 + |D_2|^2 + 2|D_1||D_2|\cos(\Delta\delta - \Delta\phi) \quad (1.20)$$

$$\begin{aligned}\mathcal{A}_{CP} =& \frac{|\overline{A}(\overline{P} \to \overline{f})|^2 - |A(P \to f)|^2}{|\overline{A}(\overline{P} \to \overline{f})|^2 + |A(P \to f)|^2} \\ =& \frac{2|D_1||D_2|\sin\Delta\delta\sin\Delta\phi}{|D_1|^2 + |D_2|^2 + 2|D_1||D_2|\cos\Delta\delta\cos\Delta\phi}.\end{aligned} \quad (1.21)$$

To have direct CP violation, there must be non-vanishing values for both $\Delta\phi$ and $\Delta\delta$, the weak and strong phases. Furthermore, to obtain large direct CP-violation, $\Delta\delta$ and $\Delta\phi$ must both be near $\pi/2$. In addition, D_1 and D_2 should have similar amplitudes. A necessary condition for direct CP violation, non-zero $\Delta\delta$, can arise from final-state interactions [13]. Final-state interactions in B decays are any interactions between the final state particles after the weak decay. Note that this is the only possible source for CP violation in charged B meson decays within the SM.

- ***CP* violation in mixing**
  Flavor-specific decays are those that can come from either P^0 or $\overline{P}^0$, but not both:

$$P^0 \to f \not\leftarrow \overline{P}^0 \quad (1.22)$$

or

$$P^0 \not\to f \leftarrow \overline{P}^0. \quad (1.23)$$

This occurs when the mass eigenstate of a neutral particle cannot be chosen as its CP eigenstate. The manifestation of this form is independent of the final state, and is related to the aforementioned difference in the oscillation rates between the neutral meson and its antiparticle, i.e.:

$$\Gamma(P^0 \to \overline{P}^0) \neq \Gamma(\overline{P}^0 \to P^0) \quad (1.24)$$

From Eq. 1.12, *CP* violation in mixing occurs when

$$|\frac{q}{p}| \neq 1. \tag{1.25}$$

Experimentally, this is studied in semileptonic neutral B meson decays, as one can determine the flavor from the charge of lepton, e.g. $b \to q\ell^-\overline{\nu}$ and $\overline{b} \to \overline{q}\ell^+\nu$.

- ***CP* violation in interference between decay and mixing**
  Flavor-nonspecific final states are those that are fed by both P^0 and $\overline{P}^0$ decays, however not necessarily with the same rate:

$$P^0 \to f \leftarrow \overline{P}^0. \tag{1.26}$$

This occurs in the common final states to B^0 and $\overline{B}^0$. *CP* violation occurs when the following condition is satisfied:

$$\Gamma(P^0(\to \overline{P}^0) \to f)(t) \neq \Gamma(\overline{P}^0(\to P^0) \to f)(t). \tag{1.27}$$

This type of *CP* violation is from the interference between decays and mixing. Consider the case that $|q/p| = 1$, the following expression is obtained using Eqs. 1.15:

$$\mathcal{A}_{CP}(t) = \frac{\Gamma(P^0 \to f) - \Gamma(\overline{P}^0 \to f)}{\Gamma(P^0 \to f) + \Gamma(\overline{P}^0 \to f)} \tag{1.28}$$

$$= \frac{2C_f \cos\Delta mt - 2S_f \sin\Delta mt}{2\cosh\frac{\Delta\Gamma t}{2} + 2D_f \sinh\frac{\Delta\Gamma t}{2}}. \tag{1.29}$$

We further suppose that $|A_f| = |\overline{A}_f|$ (i.e. $|\lambda_f| = 1$), then:

$$\mathcal{A}_{CP}(t) = \frac{-\Im\lambda_f \sin\Delta mt}{\cosh\frac{\Delta\Gamma t}{2} + \Re\lambda_f \sinh\frac{\Delta\Gamma t}{2}}. \tag{1.30}$$

CP violation still occurs in the case that the imaginary part of λ_f is non-vanishing:

$$\Im\lambda_f = \Im(\frac{q}{p}\frac{A_f}{\overline{A_f}}) \neq 0 \tag{1.31}$$

We conclude that this type of *CP* violation can occur even in the absence of *CP* violation in both mixing and decay.

1.5. Charmless *B* decays

Charmless hadronic B meson decays offer a relevant environment for studying *CP* violation since the final state is reached by the $b \to u$ tree level transition and $b \to (s,\ d)$ loop level penguin transition. Figure 1.3 shows a penguin diagram of $b \to s$ transition. Although the penguin process is loop-suppressed, the tree level process is suppressed by V_{ub} in the CKM matrix, so both processes can have similar amplitudes. If the amplitudes have a phase difference, this leads to direct *CP* violation, which has been observed at Belle in several B decay channels [14]. As these decay amplitudes are suppressed in the SM, they are sensitive to potential branching fraction enhancements from loop diagrams containing beyond-SM particles.

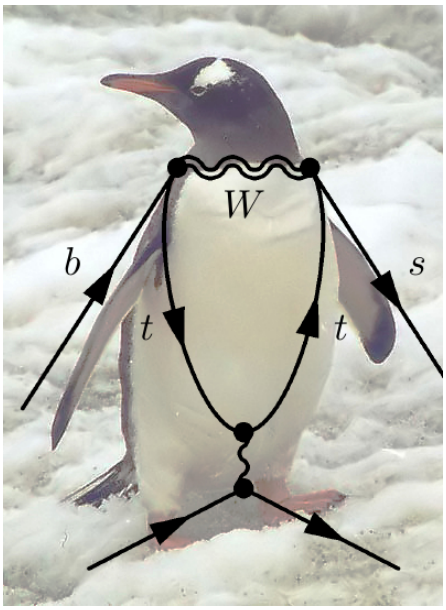


Figure 1.3.: Example of a penguin diagram [15].

1.6. Motivation of charmless $B^+ \to K^+K^-\pi^+$ decays

Charmless B meson decays to hadronic three-body final states offers the possibility to search for CP violation localized in phase space since they are dominated by two-body intermediate states [16, 17]. The LHCb experiment observed sizable asymmetries in localized regions of the phase spaces of $B^\pm \to K^+K^-K^\pm$, $B^\pm \to K^\pm\pi^+\pi^-$, $B^\pm \to K^+K^-\pi^\pm$, and $B^\pm \to \pi^+\pi^-\pi^\pm$ [18–20]. In these studies, one of the most interesting is the measurement of $B^\pm \to K^+K^-\pi^\pm$. Figure 1.4 shows some SM Feynman diagrams that contribute to the $B^+ \to K^+K^-\pi^+$ decay. The dominant process is the Cabibbo-suppressed $b \to u$ tree transition in Fig 1.4(a); the $b \to d$ penguin diagram in Fig. 1.4(d) leading to $B^+ \to \phi\pi^+$ with $\phi \to K^+K^-$ is heavily suppressed due to both the CKM matrix elements and the Okubo-Zweig-Iizuka [21–24] rule. The OZI suppression is due to the colorless $s\bar{s}$ pair forming the ϕ. The branching fraction of $B^+ \to \phi\pi^+$ is predicted to be $(5 - 10) \times 10^{-9}$ [25]. However, it can be enhanced to the order of 10^{-7} if $\omega - \phi$ mixing is considered [26, 27]. If the dynamics of $B^+ \to K^+K^-\pi^+$ decays were sufficiently well-understood, these results could be used as an input to constrain the values of ϕ_3, and in analyses based on flavor SU(3) that can limit the deviation of $\sin(2\beta^{\text{eff}})$ measured in $b \to s$ penguin modes to the reference value obtained in $b \to c\bar{c}s$ transitions (e.g.

$B^+ \to J/\psi K_S^0$) [28, 29]. In previous studies, the three-body inclusive branching fraction $\mathcal{B}(B^+ \to K^+K^-\pi^+) = (5.0 \pm 0.5 \pm 0.5) \times 10^{-6}$, was observed by the BaBar collaboration [30], however no evidence was found for the two-body intermediate modes. The current experimental upper limits for these two-body modes are $\mathcal{B}(B^+ \to \phi\pi^+) < 1.5 \times 10^{-7}$ [31], $\mathcal{B}(B^+ \to K^+\overline{K}^{*0}(892)) < 1.1 \times 10^{-6}$ and $\mathcal{B}(B^+ \to K^+\overline{K}_0^{*0}(1430)) < 2.2 \times 10^{-6}$ [32] all at the 90% confidence level. Even with the lack of K^+K^- resonant states, an unidentified peak was seen near 1.5 GeV/c^2 in the K^+K^- invariant mass spectrum by the BaBar Collaboration. The LHCb studies found a non-zero inclusive *CP* asymmetry of $-0.123 \pm 0.017 \pm 0.012 \pm 0.007$ [20] and a large asymmetry in the measured yields in the same mass region as BaBar. These results suggest that final-state interactions may contribute to *CP* violation [33, 34]. This study attempts to quantify the *CP* asymmetry and branching fraction as a function of the K^+K^- invariant mass.

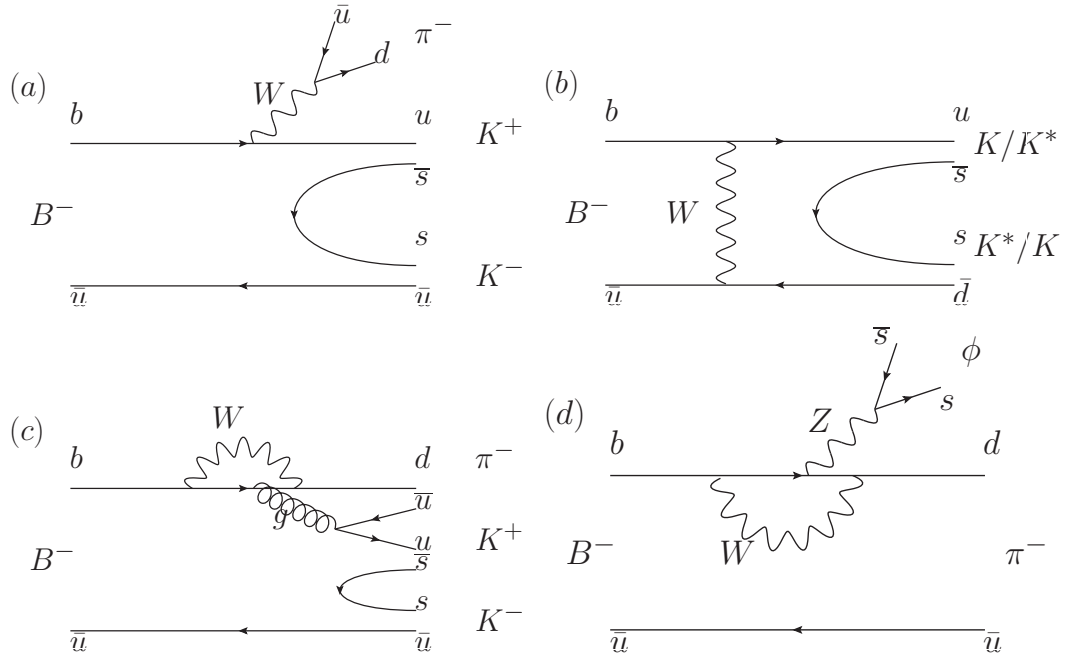


Figure 1.4.: $B^+ \to K^+K^-\pi^+$ Feynman diagrams (all Cabibbo-suppressed). (a) Tree diagram, (b) *W*-exchange diagram leading to KK^* states, (c) strong-penguin diagram, and (d) electroweak penguin leading to the $\phi\pi$ state.

Chapter 2.

B Factory at KEK

KEK is a research organization for high-energy physics located in Tsukuba, Japan. The Belle experiment is one of the KEK projects, the main goal of which is to study *CP* violation. There are two major facilities in the Belle experiment, the KEKB accelerator (Fig. 2.1) and the Belle detector (Fig. 2.3).
The KEKB accelerator is designed to produce B meson pairs efficiently, hence the Belle experiment is known as a B-factory.

2.1. KEKB accelerator

The KEKB accelerator [35, 36] consists of two storage rings: a low-energy ring (LER) for positrons with 3.5 GeV energy and a high-energy ring (HER) for electrons with 8.0 GeV energy. The asymmetric energy beams collide at the interaction point (IP) with a finite crossing angle of ±11 mrad. This crossing angle eliminates the need for the separation-bend magnets, reducing the background due to synchrotron light, and also reduces the number of parasitic collisions near the IP [36].

To avoid the finite crossing angle leading to uncontrollable synchrotron-betatron resonances, the "crab cavities" [37] were installed in 2007 [36]. They were placed near the IP to rotate the beam bunches such that they are oriented head-on at the IP. The beam rotation by "crab cavities" is shown in Fig. 2.2.

The center-of-mass (CM) energy of the collisions is 10.58 GeV. This populates the $\Upsilon(4S)$ resonance which is just above threshold for $B\overline{B}$ production. At this

energy, the cross section for $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ is 1.05 nb, and the continuum process $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$) has a cross section of 3.7 nb. The flight length of a B meson in the CM frame of $\Upsilon(4S)$ is 2 μm. This distance is not enough to resolve the difference between $B\overline{B}$ pair decay lengths, which is required for time-dependent CP analyses. However, the asymmetric beam energies provide a Lorentz boost with a factor $\beta\gamma \sim 0.425$ for B mesons. Because of this, the average distance between decay vertices of the two B mesons in the beam direction in the laboratory frame(the z axis) is about 200 μm. The resolution of the silicon vertex detector is sufficiently good to measure the separation between the decay points of the two B-mesons. KEKB was designed to operate with a peak luminosity of 1×10^{34} cm^{2}s^{-1} and achieved 2.1×10^{34} cm^{2}s^{-1}, corresponding to $\sim 2 \times 10^8$ $B\overline{B}$ pairs per year. The configuration plot of the KEKB accelerator is shown in Fig. 2.1, and the parameters for the KEKB accelerator are summarized in Table 2.1.

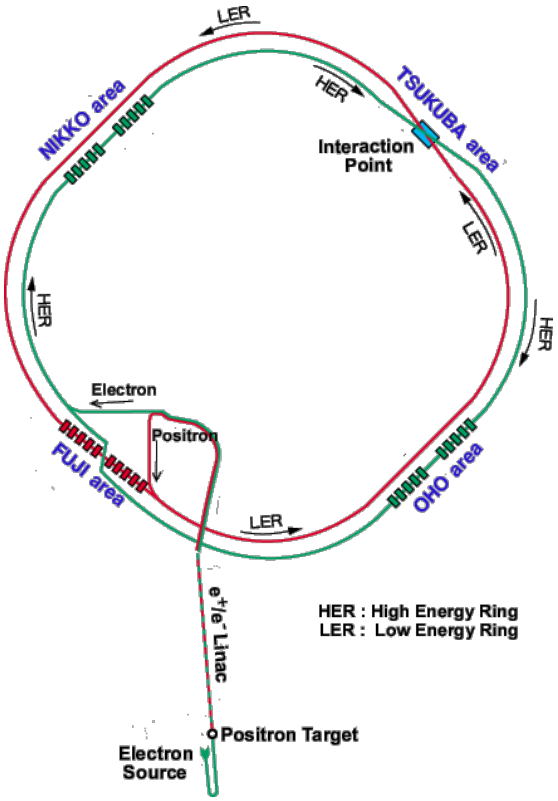


Figure 2.1.: The configuration of the KEKB accelerator [38].

Table 2.1.: Design parameters of the KEKB accelerator [35].

Ring		LER	HER	Unit
Energy	E	3.5	8.0	GeV
Circumference	C	3016.26		m
Luminosity	L	1×10^{34}		$cm^{-2}s^{-1}$
Crossing angle	θ_x	± 11		mrad
Tune shifts	ξ_x/ξ_y	0.039/0.052		
Beta function at IP	β_x^*/β_y^*	0.33/0.01		m
Beam current	I	2.6	1.1	A
Natural bunch length	σ_z	0.4		cm
Energy spread	σ_ε	7.1×10^{-4}	6.7×10^{-4}	
Bunch spacing	s_b	0.59		m
Particles/bunch	N	3.3×10^{10}	1.4×10^{10}	
Emittance	$\varepsilon_x/\varepsilon_y$	$1.8 \times 10^{-8}/3.6 \times 10^{-10}$		
Synchrotron	ν_s	$0.01 \sim 0.02$		
Betatron tune	ν_x/ν_y	45.52/45.08	47.52/43.08	
Momentum compaction factor	α_p	$1 \times 10^{-4} \sim 2 \times 10^{-4}$		
Energy loss/turn	U_0	$0.81^{\dagger}/1.5^{\dagger\dagger}$	3.5	MeV
RF voltage	V_c	$5 \sim 10$	$10 \sim 20$	MV
RF frequency	f_{RF}	508.887		MHz
Harmonic number	h	5120		
Longitudinal damping time	τ_ε	$43^{\dagger}/23^{\dagger\dagger}$	23	ms
Total beam power	P_b	$2.7^{\dagger}/4.5^{\dagger\dagger}$	4.0	MW
Radiation power	P_{SR}	$2.1^{\dagger}/4.0^{\dagger\dagger}$	3.8	MW
HOM power	P_{HOM}	0.57	0.15	MW
Bending radius	ρ	16.3	104.5	m
Length of bending magnet	l_B	0.915	5.86	m

†: without wigglers, ††: with wigglers

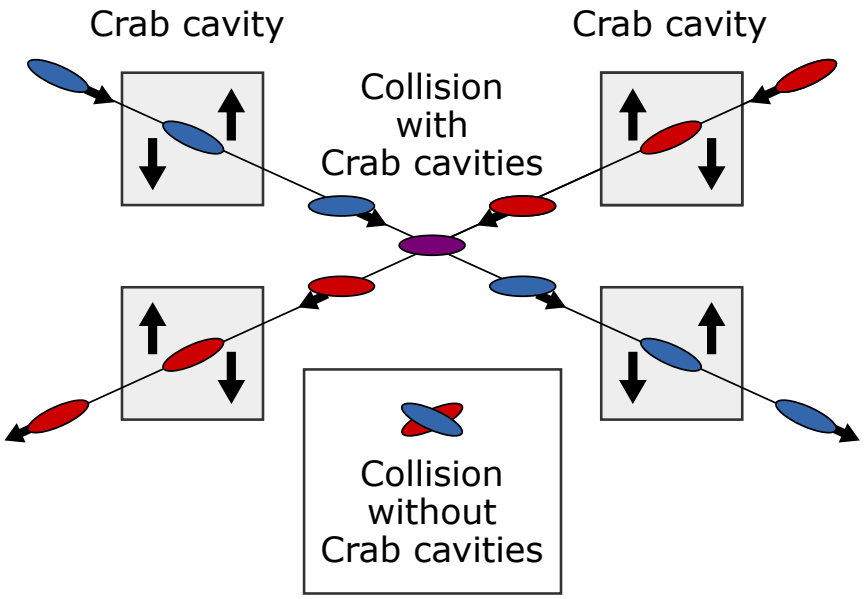


Figure 2.2.: Illustration of beam bunch rotation by crab cavities [39].

2.2. Belle Detector

The Belle detector consists of a silicon vertex detector (SVD), an extreme forward calorimeter (EFC), a 50-layer central drift chamber (CDC), time-of-flight scintillation counters (TOF), an array of aerogel threshold Cherenkov counters (ACC), and a CsI(Tl) electromagnetic calorimeter located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. Outside the coil, the K_L^0 and muon detector (KLM), composed of layers of iron for magnetic flux return interspersed with position sensitive resistive plate chambers, detects K_L^0 mesons and identifies muons. The performance parameters and the configuration of the Belle detector are shown in Table 2.2 and Fig 2.3. The components of the sub-detectors are briefly described in the following sections. More information of the Belle detector is detailed in Ref. [40].

2.2.1. The Beam Pipe

A main requirement of the Belle detector is precise measurement of B meson decays. In order to achieve this goal, the SVD is placed close to the IP. To facilitate this, the beam pipe has an inner radius of 2.0 cm, and an outer wall

Table 2.2.: Performance parameters for the Belle detector. There were two configurations of inner detectors used to collect two data sets, DS-I and DS-II, corresponding to a 3-layer SVD1 and a 4-layer SVD2 with a smaller beam pipe respectively [40].

Detector	Type	Configuration	Readout	Performance
Beam pipe DS-I	Beryllium double wall	Cylindrical, r = 20mm, 0.5/2.5/0.5(mm) = Be/He/Be w/ He gas cooled		
Beam pipe DS-II	Beryllium double wall	Cylindrical, r = 15mm, 0.5/2.5/0.5(mm) = Be/PF200/Be		
EFC	BGO	Photodiode readout Segmentation : 32 in ϕ; 5 in θ	160×2	Rms energy resolution: 7.3% at 8 GeV 5.8% at 2.5 GeV
SVD1	Double-sided Si strip	3-layers: 8/10/14 ladders Strip pitch: 25(p)/50(n)μm	ϕ: 40.96k z: 40.96k	$\sigma(z_{CP}) \sim 78.0\mu$m for $B \to \phi K^0_s$
SVD2	Double-sided Si strip	4-layers: 6/12/18/18 ladders Strip pitch: 75(p)/50(n)μm (layer1-3) 73(p)/65(n)μm (layer4)	ϕ: 55.29k z: 55.296k	$\sigma(z_{CP}) \sim 78.9\mu$m for $B \to \phi K^0_s$
CDC	Small cell drift chamber	Anode: 50 layers Cathode: 3 layers r = 8.3 - 86.3 cm $-77 \le z \le 160$ cm	Anode: 8.4k Cathod: 1.8k	$\sigma_{r\phi} = 130\mu$m $\sigma_z = 200 \sim 1400\mu$m $\sigma_{Pt}/Pt = 0.3\% \sqrt{p_t^2 + 1}$ $\sigma_{dE/dx} = 0.6\%$
ACC	Silica aerogel	960 barrel/228 end-cap FM-PMT readout		$N_{p.e.} \ge 6$ K/π seperation: $1.2 < p < 3.5$GeV/c
TOF	Scintillator	128 ϕ segmentation r = 120 cm, 3-cm long	128×2	$\sigma_t = 100$ ps K/π seperation:
TSC		64 ϕ segmentation	64	up to 1.2 GeV/c
ECL	CsI (Towered-structure)	Barrel: r = 125 - 162 cm End-cap: z = -102 cm and +196cm	6624 1152(F) 960(B)	$\sigma_E/E = 1.3\%/\sqrt{E}$ $\sigma_{pos} = 0.5$ cm$/\sqrt{E}$ (E in GeV)
KLM	Resistive plate counters	14 layers (5 cm Fe + 4cm gap) 2 RPCs in each gap	θ: 16k ϕ: 16k	$\Delta\phi = \Delta\theta = 30$mr for K_L $\sim 1\%$ hadron fake
Magnet	Supercon.	Inner radius = 170 cm		B=1.5T

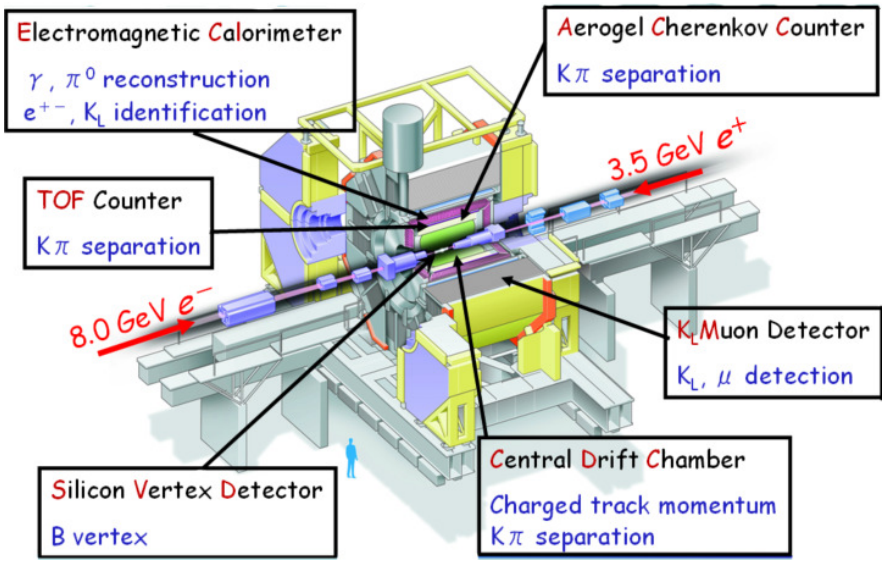


Figure 2.3.: Overall view of Belle detector [38].

2.3 cm in radius. The central part (-4.6 cm $\leq z \leq 10.1$ cm) of the beam pipe consists of double beryllium cylinders of 0.5 mm thickness. The gap between these two beryllium walls provides a channel for helium gas, which is used to cool the pipe.. Figure 2.4 shows the cross-section of the beryllium beam pipe. The maximum temperature increase for the inner beryllium is estimated to be $25K$ assuming a uniformly distributed 100W heat load on the inner wall and a 2 g/s helium gas flow. The beryllium central section is brazed to aluminium pipes which allows the synchrotron x-rays generated in the final-focus quadrupole magnets and other magnets to pass through without hitting the inner beryllium wall. A mask of gold is used to absorb back-scattered photons which have a critical energy less than 2 KeV. Particle backgrounds are critical at KEKB. The rate of particles from both beams hitting the beam pipe is calculated to be around 130 kHz in a 10^{-9} Torr vacuum. Movable masks are installed to reduce the radiation levels at injection. The parameters of these are adjusted during beam tuning, which normally happens during machine development time.

2.2.2. Silicon Vertex Detector (SVD)

The Silicon Vertex Detector (SVD) is employed for the measurement of the decay vertices of B mesons and contributes to charged particle tracking and

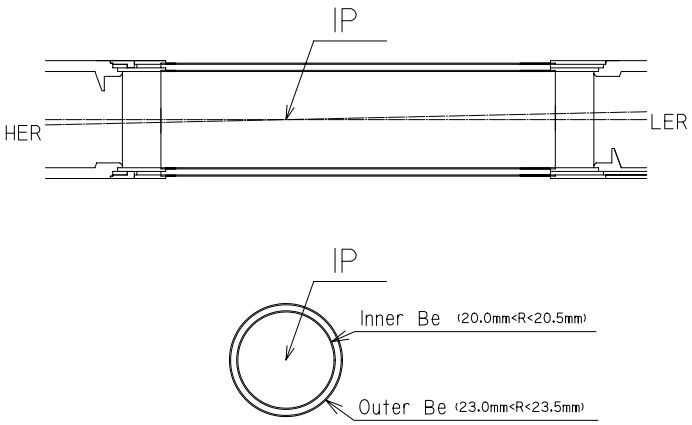


Figure 2.4.: Graphical illustration of the beam pipe [40].

the determination of their momenta. The information of the decay vertices is essential for the measurement of time-dependent *CP*-violation.

The first version of the SVD (SVD1) consists of three layers of double-sided silicon strip detector (DSSD), comprising 8, 10, 14 ladders in the inner, middle and outer layers. The radii of these layers are 30 mm, 45.5 mm and 60.5 mm. It covers $23^\circ < \theta < 139^\circ$, corresponding to 86% solid angle coverage. The DSSDs were originally developed for the DELPHI micro-vertex detector. A DSSD is in essence a $p - n$ junction, which is operated under a reverse bias of 75 V to form a full depletion zone. A charged particle passing through a DSSD creates electron-hole pairs, which drift to the biased side of the DSSD where the charge is deposit and read out.

In each DSSD, there are 1280 striped sensors and 640 readout pads on both sides. The z-strip (ϕ-strip) pitch is 42 (25) μm, and readout z-strip (ϕ-strip) pitch is 84 (50) μm, respectively. The size of the active region in each DSSD is 53.5×32.0 mm^2 on the z-side and 54.5×32.0 mm^2 on the ϕ-side. Figure 2.5 shows the designed configuration of the SVD.

In summer 2003, a new vertex detector replaced SVD1, called SVD2 [41]. SVD2 consists of four detector layers, comprising 6, 12, 18 and 18 ladders in the first, second, third and forth layer, respectively. The SVD2 also has a larger angular acceptance than SVD1, $17^\circ < \theta < 150^\circ$, which is the same as the CDC. The beam pipe is replaced by a smaller(1.5 cm in radius) one. The radii of the four layers are 20, 43.5, 70 and 80 mm, respectively. This design improves the

vertex resolution. The side-view of SVD1 and SVD2 are shown and compared in Fig 2.6.

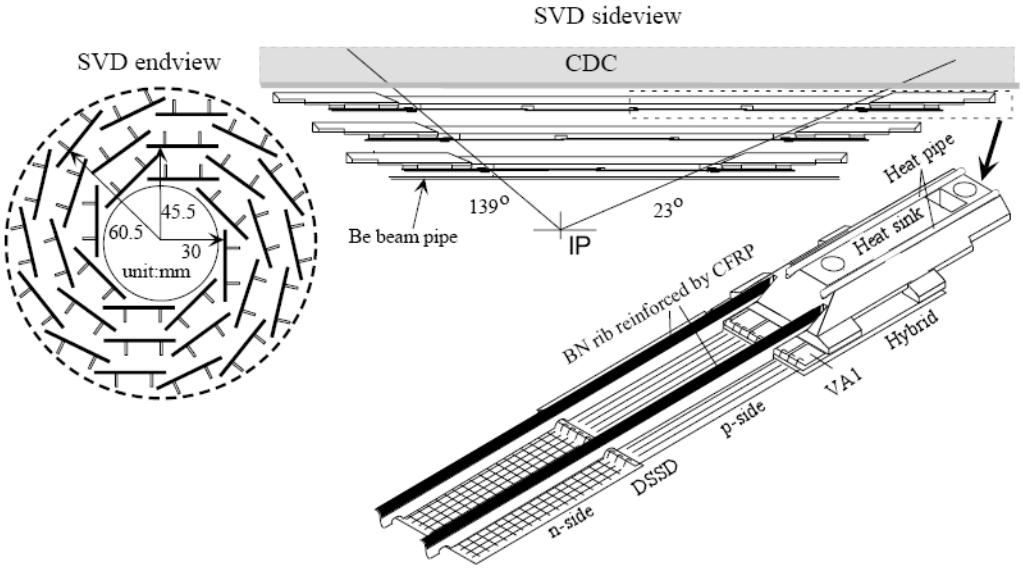


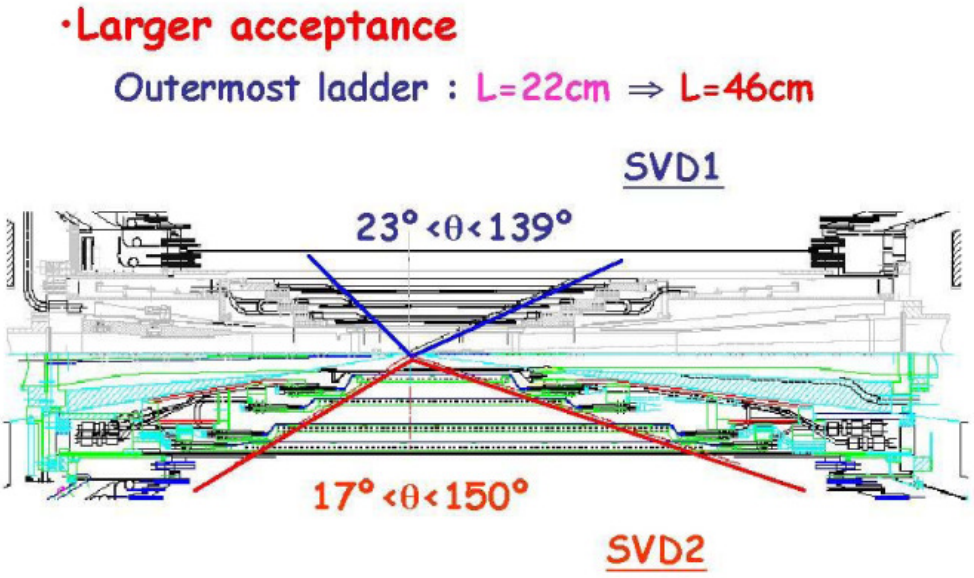
Figure 2.5.: Configuration of the SVD [40].

2.2.3. Extreme Forward Calorimeter (EFC)

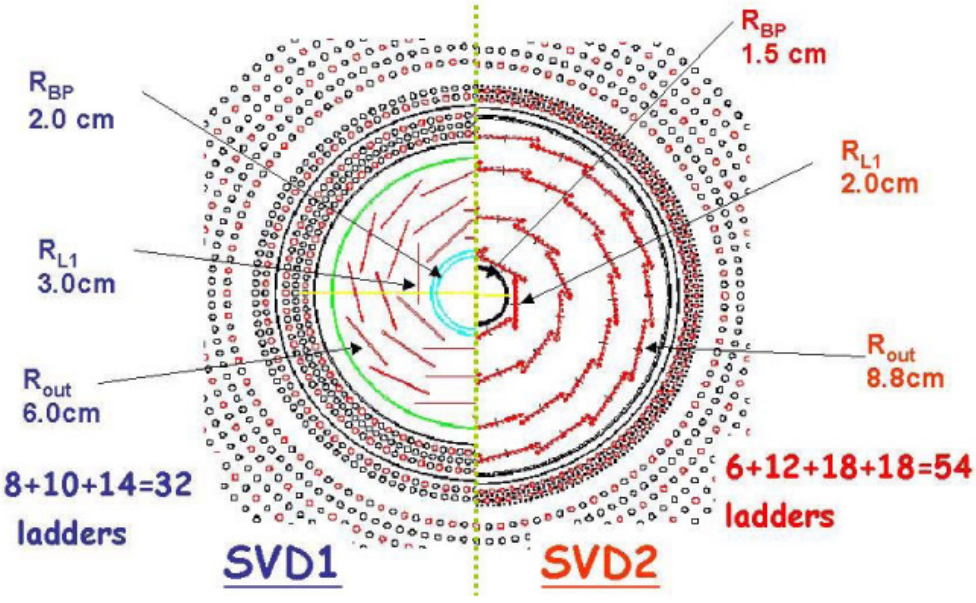
The extreme forward calorimeter (EFC) is designed to improve luminosity monitoring and further extend the polar angle covered by the ECL ($17° < \theta < 150°$). The EFC covers the polar angle range $6.4° < \theta < 11.5°$ in the forward direction and $163.3° < \theta < 171.2°$ in the backward direction. The extended coverage can improve the sensitivity to two-photon physics and $B \to \tau\nu$ decays. Due to the high exposure to radiation, the material of the EFC is radiation-hard BGO (Bismuth Germanate, $Bi_4Ge_3O_{12}$). Figure 2.7 shows an isometric view of the BGO crystal.

2.2.4. Central Drift Chamber (CDC)

The Central Drift Chamber (CDC) is used to reconstruct charged tracks passing through the region of its coverage, $17° < \theta < 150°$, and provides measurements of the momentum and energy deposition(dE/dx) of charged tracks. The curvature of charged particles in the transverse plane is used to



(a) Side view comparison of SVD1 and SVD2.



(b) End view comparison of SVD1 and SVD2.

Figure 2.6.: The graphical illustration of sub-detector SVD1 and SVD2 [38].

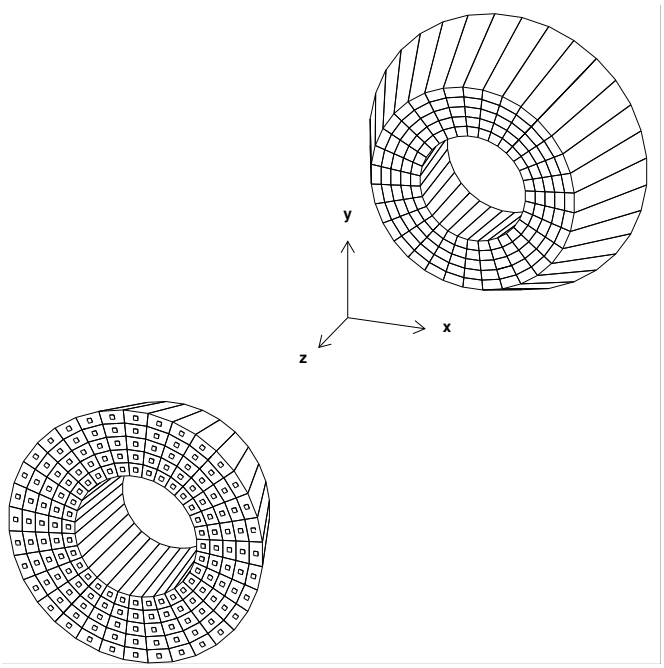


Figure 2.7.: Isometric view of the forward and backward EFC detectors. The BGO crystals and the location of photodiodes are illustrated [40].

determine the transverse momentum (p_T), and combining this with the helical track information, the z component of the momentum is determined as well. This sub-detector is designed to fulfil the requirement of momentum resolution of $\sigma_{p_T}/p_T \sim 0.005\sqrt{1+p_T^2}$ (p_T in GeV/c) for all charged particles with $p_T > 100$ MeV/c.

The CDC consists of 50 cylindrical layers and 3 cathode strip layers, for a total of 8400 drift cells. The structure of the CDC and a cell are shown in Fig. 2.8 and Fig. 2.9. Each cell consists of one anode sensor wire surrounded by six cathode field wires in a low-Z gas. Charged particles passing through the CDC ionize the gas molecules along their path, the resulting electrons drift towards the sensor wires at a velocity of approximately 4 cm/μs, which weakly depends on the strength of the electric field. As the electrons come close to the anode wire, they gain enough energy to ionize gas molecules. This process is repeated by the liberated electrons and an ionization avalanche with total charge proportional to the gas gain and the amount of primary ionizations is created. The resulting pulse produces a detectable image charge on the sensor wire.

Since the majority of the B decay daughters have momenta lower than 1 GeV/c, the minimization of multiple Coulomb scatterings is important for improving the momentum resolution. To enable this, we employ a low-Z gas of a 50% helium and 50% ethane mixture in the CDC. Moreover, the large portion of ethane provides good energy loss resolution (dE/dx) resolution, which is useful in the particle identification, especially for the separation of kaons (K) and pions (π) in the momentum region below 0.5 GeV/c.

The dE/dx information is described by the Bethe formula:

$$-dE/dx = \frac{4\pi N_0 z^2 e^5}{m_e v^2} \frac{Z}{A} [\ln(\frac{2m_e v^2}{I(1-\beta^2)} - \beta^2)], \tag{2.1}$$

where m and e are the mass and the charge of the electron, N_0 is Avogadro's number, z is the charge of the particle, v is the velocity of the particle, Z and A are the atomic number and the mass number of the gas, and I is the effective ionization potential, which equals (10 eV)·Z.

Figure 2.10 shows a plot of dE/dx and particle momentum, together with the expected truncated mean.

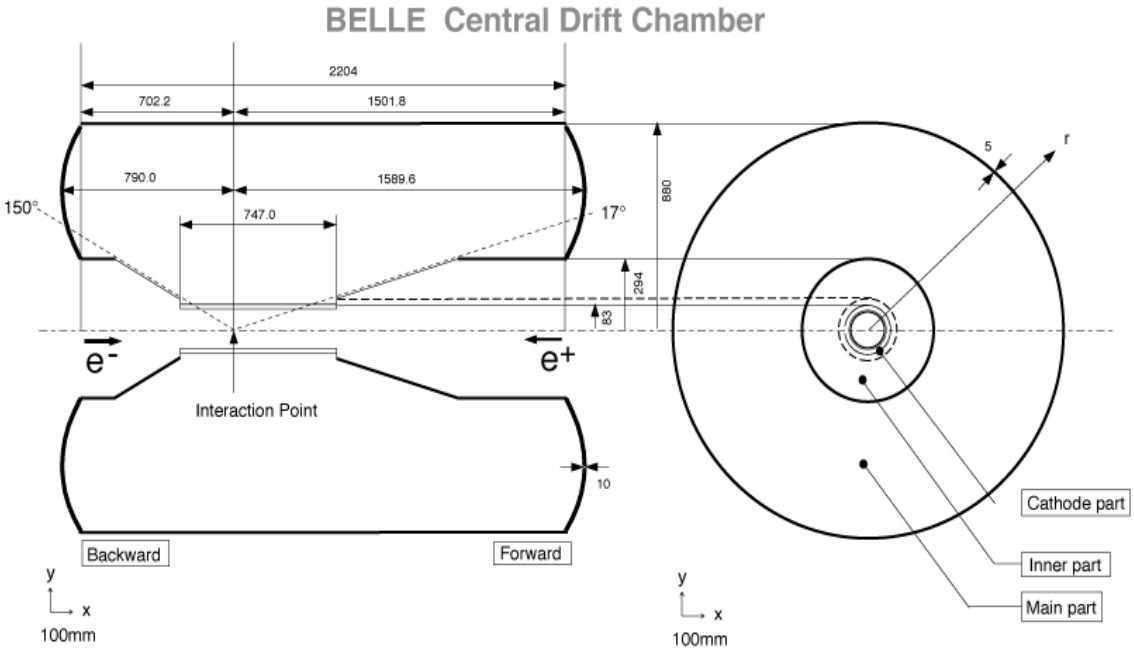


Figure 2.8.: Overview of the CDC structure [40]. The unit is mm.

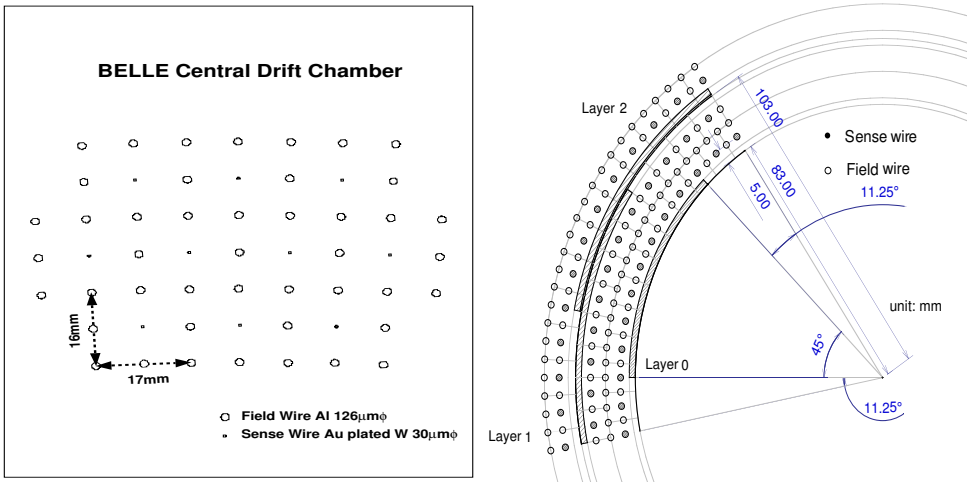


Figure 2.9.: CDC cell structure [40].

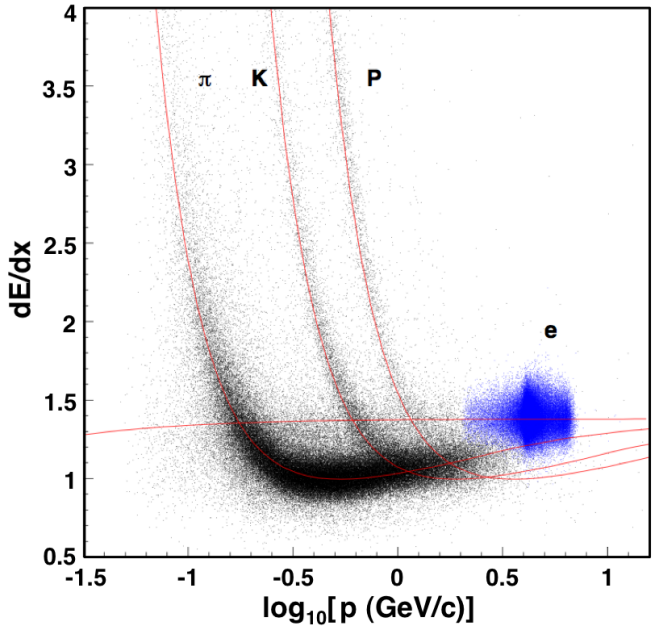


Figure 2.10.: Plot of dE/dx and particle momentum, together with the expected truncated mean [40].

2.2.5. Aerogel Čerenkov Counter (ACC)

The Aerogel Čerenkov Counter (ACC) is used to distinguish $\pi^{\pm}$ and $K^{\pm}$ in the high momentum region (1.2 GeV/c ~ 4.0 GeV/c). Čerenkov light is emitted when charged particles exceed the speed of light in a transparent medium. Consequently, the phenomenon will take place if the velocity of a particle, β, satisfies

$$n > 1/\beta = \sqrt{1 + (m/p)^2} \tag{2.2}$$

where n is the refractive index of the medium, m is the mass of the particle and p is the momentum of the particle.

The ACC is divided into two parts, barrel and end-cap. The barrel part consists of 960 silica aerogel counter modules segmented into 60 cells in the ϕ direction, covering the polar angle of 33.3° to 127.9°, and the end-cap part comprises 288 modules arranged in 5 concentric layers, covering the polar angle of 13.6° to 33.4°. The side-view of the full ACC system is shown in Fig. 2.11. All aerogel counters are arranged in a semi-tower geometry, pointing to the IP. We employ finemesh photomultiplier tubes (FM-PMTs) to detect Čerenkov light since they can operate in a strong magnetic field. The refractive indices are chosen to be between 1.01 to 1.03 depending on their polar angle region. Figure 2.12 shows two types of unit modules.

2.2.6. Time of Flight (TOF)

The time-of-flight (TOF) detector system shown in Fig. 2.13 provides a powerful identification method for lower momentum particles. For a 1.2 m flight path, a system of counters with 100 ps time resolution is effective for particle momenta below about 1.2 GeV/c, which encompasses 90% of the particles produced in $B\overline{B}$ decays. The system functions with the concept of TOF, that if the time T, the path length L and the momentum p are measured, the mass m of a particle can be obtained from the following equation:

$$T = \frac{L}{c}\sqrt{1 + \frac{m}{p}c^2}. \tag{2.3}$$

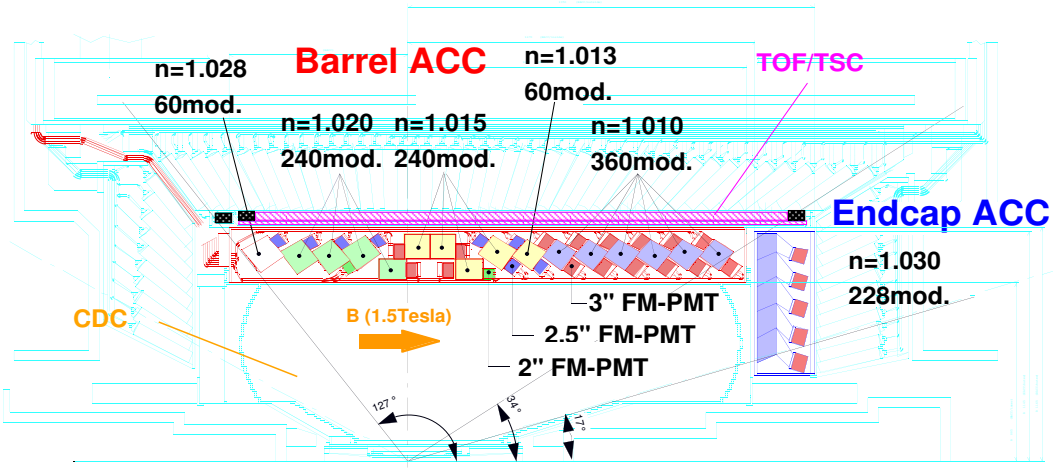


Figure 2.11.: Sideview of ACC system, together with other nearby detectors, the index of refraction (n) is also given for each ACC module [40].

When $E >> mc^2$, the relation of time difference for two particles of different masses is:

$$\Delta T = T_1 - T_2 \approx \frac{Lc}{2p^2}(M_1^2 - M_2^2). \tag{2.4}$$

where $M_{1(2)}$ is the mass of the first (second) particle. Consequently measurement of ΔT provides additional K/π separation for low momentum particles.

2.2.7. Electromagnetic Calorimeter (ECL)

The main purpose of the Electromagnetic Calorimeter (ECL) is to detect the photons from B meson decays with high efficiency and good energy resolution. In addition to photon detection, the ECL is also used to identify electrons and to detect K_L^0. The energy of electrons and photons are determined in the crystal calorimeter which emits scintillation light in proportion to the energy loss of charged particles traversing the detector. Most of the photons are end-products of decay cascades and have relatively low energy, thus, good performance with

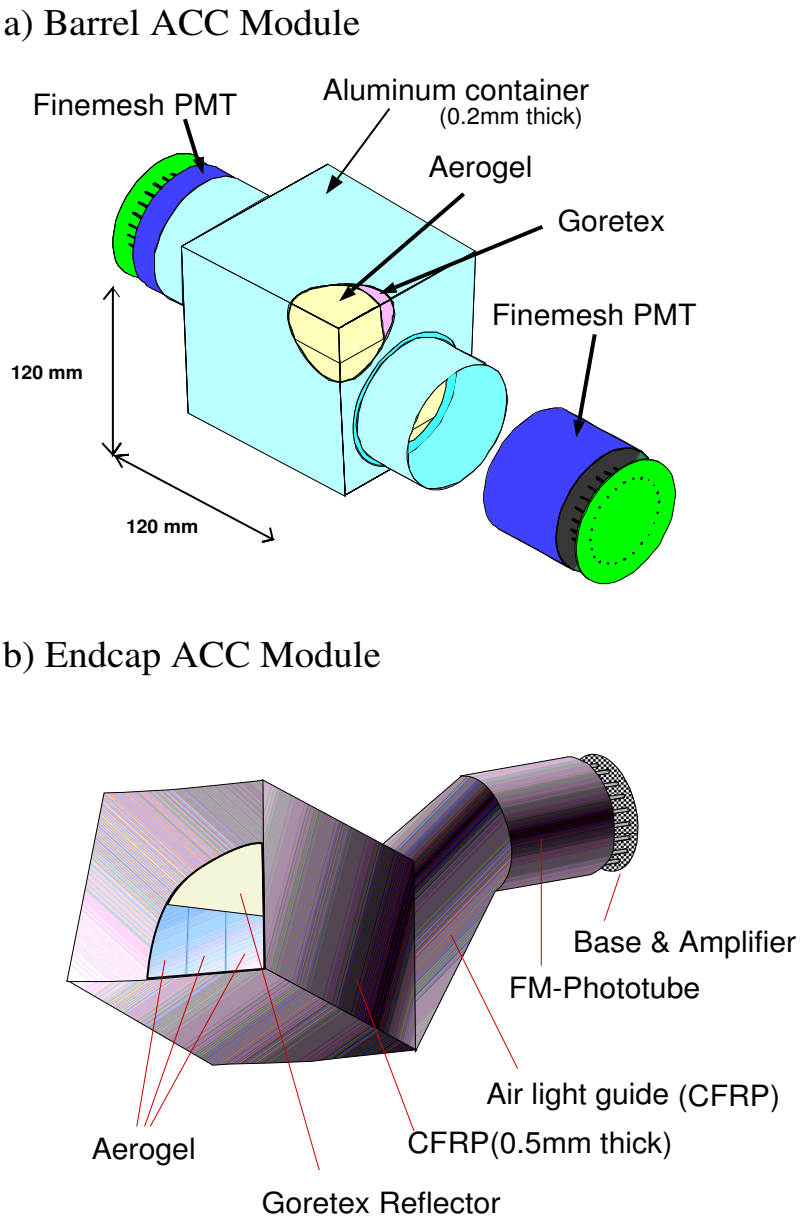


Figure 2.12.: Schematic drawing of a typical ACC counter module: (a) barrel and (b) end-cap ACC [40].

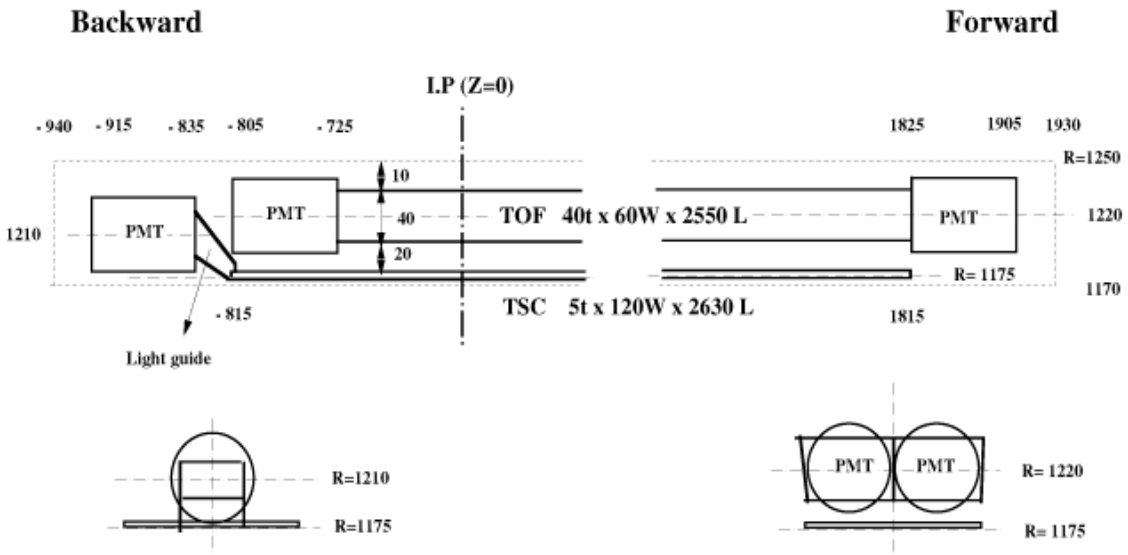


Figure 2.13.: An illustration of a TOF/TSC module [40]. The units of the figure are in mm.

photons below 500 MeV is necessary. Besides these cascade photons, important two-body decay modes such as $B \to K^*\gamma$ and $B^0 \to \pi^0\pi^0$ require a good energy resolution up to 4 GeV. Electron identification in Belle depends primarily on a comparison of the momentum of charged tracks and the energy recorded in the ECL. Good electromagnetic energy resolution results in better hadron rejection. High momentum π^0 detection requires good resolution of positions for two nearby photons and a precise determination of their opening angle. All of these require a fine-grained segmentation in the calorimeter.

The ECL consists of the barrel section of 3.0 m in length with inner radius 1.25 m and annular end-caps at $z = 2.0$ m and $z = -1.0$ m from the interaction point (IP). Each ECL cell consists of a 30 cm long tower-shaped CsI(Tl) crystal and 2 silicon photodiode read outs, installed in a magnetic field of 1.5 T inside a super-conducting solenoid magnet. The energy of electrons and photons are determined in the crystal calorimeter by the electromagnetic shower process. The electromagnetic showers result from pair-production and bremsstrahlung cascades and produce an exponential increase in the number of electrons and positrons and an exponential decrease in their energy until it is so low that ionization energy loss dominates. The produced electrons and positrons excite electron bands in the crystal lattice that correspond to the visible energy, resulting

in scintillation light read out by a photodiode. The crystals have different dimensions that depend on their polar angle position. However the length of all crystals is 30 cm which corresponds to about 16.2 radiation lengths (X_0) [42] for elections and 0.8 nuclear interaction lengths of K_L^0's. Consequently, most of the incident electron and photon energy will be fully deposited in the ECL detector. Neutral hadrons also have a significant detection probability in the ECL. Figure 2.14 shows the configuration of the ECL.

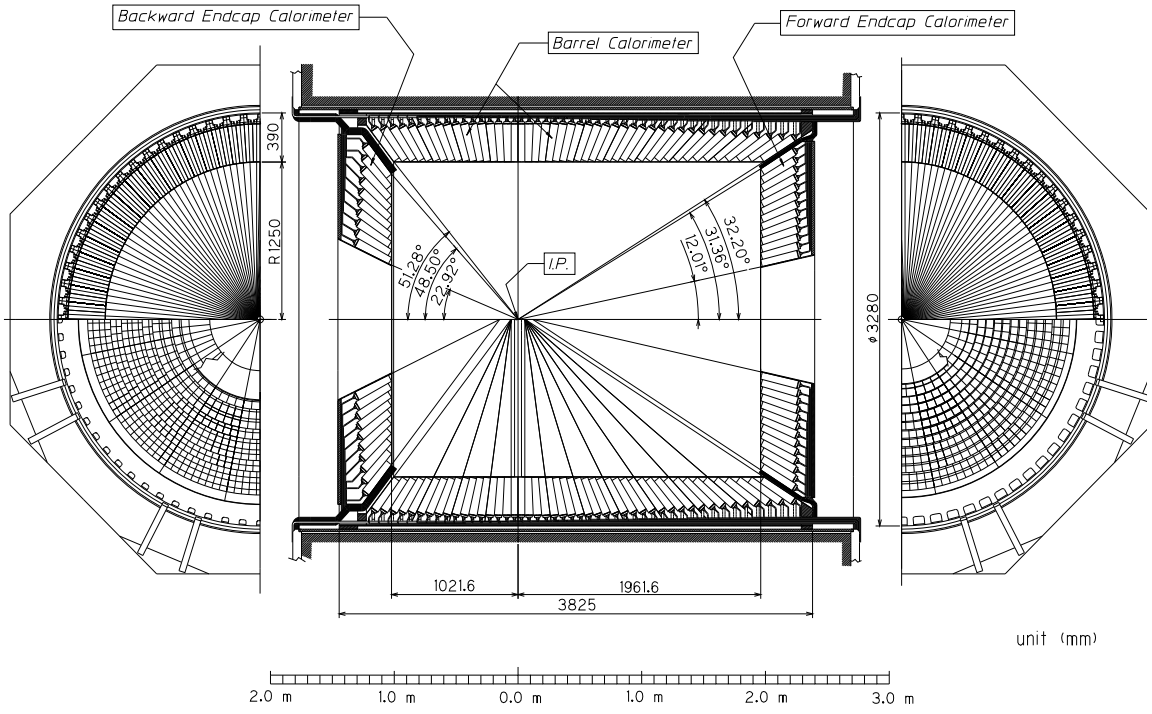


Figure 2.14.: The overall configuration of the ECL [40]. The unit is mm.

2.2.8. K_L and Muon Detector (KLM)

The main purpose of the KLM is to detect neutral kaons (K_L^0) and muons ($\mu^\pm$) with momenta greater than 600 MeV/c. The KLM detector consists of 15 (14) layers of glass-electrode-resistive plate counters (RPCs) and 14 (14) layers of 4.7 cm-thick iron plates arranged alternately in the barrel (end-cap) region. It covers the region from 17° to 155°. An illustration of the KLM is shown in Fig. 2.15. The RPCs consist of two parallel resistive plates with a gas-filled gap. An ionized

particle passing through the gap induces a streamer discharge in the gas, and results in a local discharge of the plates. The discharge generates a signal, and the location and time are recorded. Figure 2.16 shows the arrangement for RPCs. The location of the shower provides the direction of K_L^0 if there is no associated track in the CDC, however the KLM does not have the capability of K_L^0 energy measurement. Since muons only experience ionization energy loss, they are distinguished from strongly interacting charged particles like pions and kaons which are likely to undergo nuclear interactions in the many layers of iron in the KLM. Consequently, muons are identified via their track length through the KLM. Particles that have a corresponding charged track in the CDC and a large penetration depth are muon candidates. For muon detection, an efficiency of greater than 90% and a fake rate of less than 5% for charged tracks above 1.5 GeV/c is achieved.

2.2.9. Trigger and Data Acquisition (DAQ)

The trigger system's purpose is to decide when the information of the various subsystems of the Belle detector should be read out. Once a collision satisfies the trigger criteria, the data from all the subsystems is read out and stored for further analysis. The decision to read out is based on criteria carefully chosen to remove background events while retaining events of interest. Once the decision has been made to read out, the data acquisition (DAQ) system steps in to transfer the raw data from subsystems to the data storage system.

In the Belle experiment, people are mainly interested in hadronic events, which include $e^+e^- \to \Upsilon(4S)$, $e^+e^- \to c\bar{c}$, and $e^+e^- \to \tau\bar{\tau}$ processes. Events of interest are also included from two photon ($e^+e^- \to \gamma\gamma \to hh$), Bhabha ($e^+e^- \to e^+e^-$), and μ-pair ($e^+e^- \to \mu\bar{\mu}$) events, which can be used for luminosity measurements and for detector calibration. The main sources of background events at the Belle experiment were through beam gas, which is the interactions between the KEKB beams with residual gas molecules in the beam pipe, and synchrotron radiation from the beams. At a luminosity of 10^{34} cm^{-2}s^{-1}, the total event rate of these physics processes lies around 100 Hz. The trigger rate due to beam background is calculated to be around 100 Hz. Therefore, the trigger system is designed to

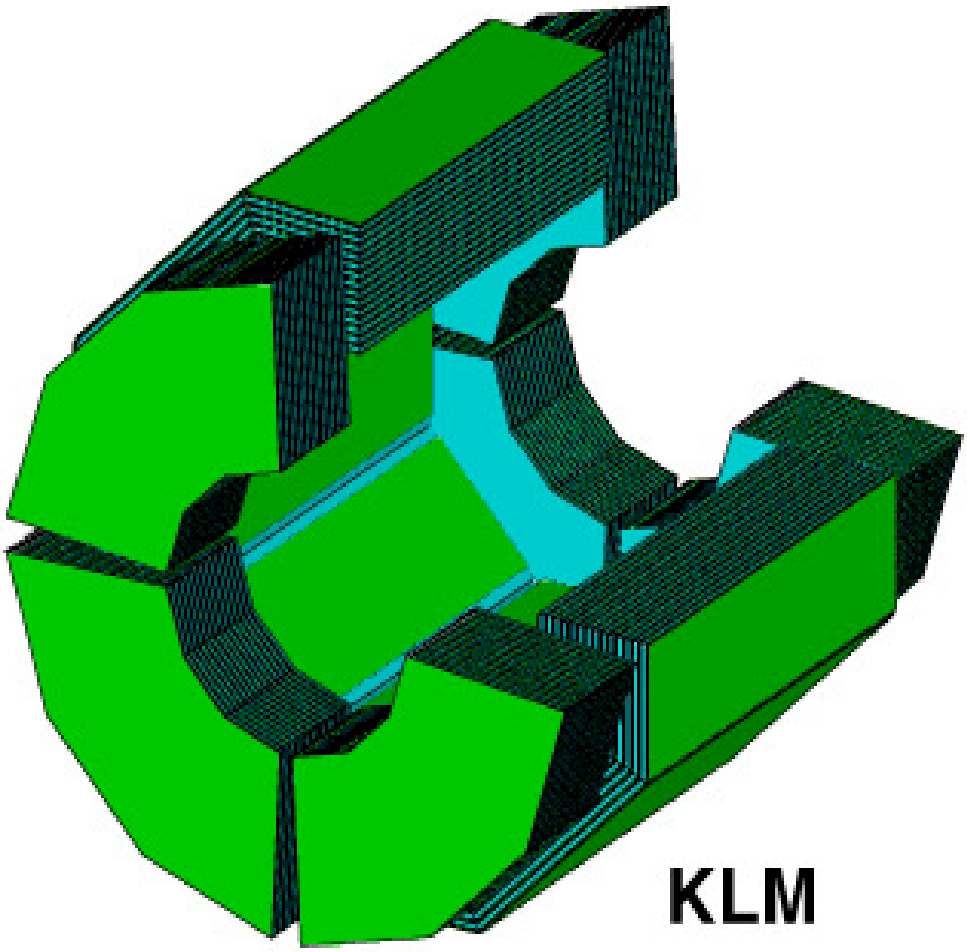
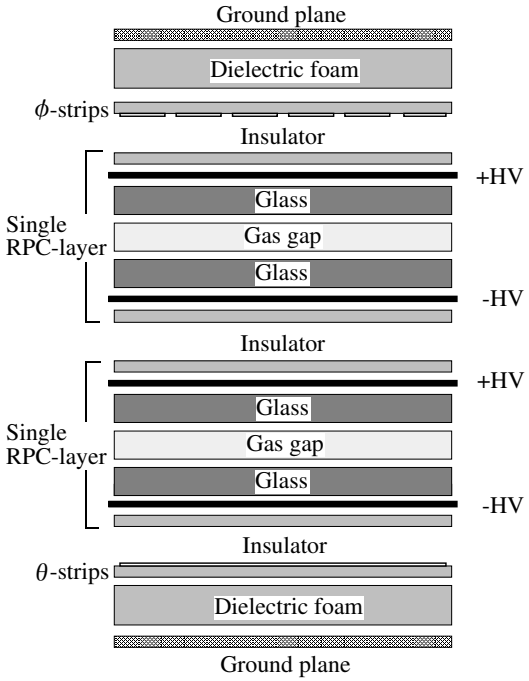
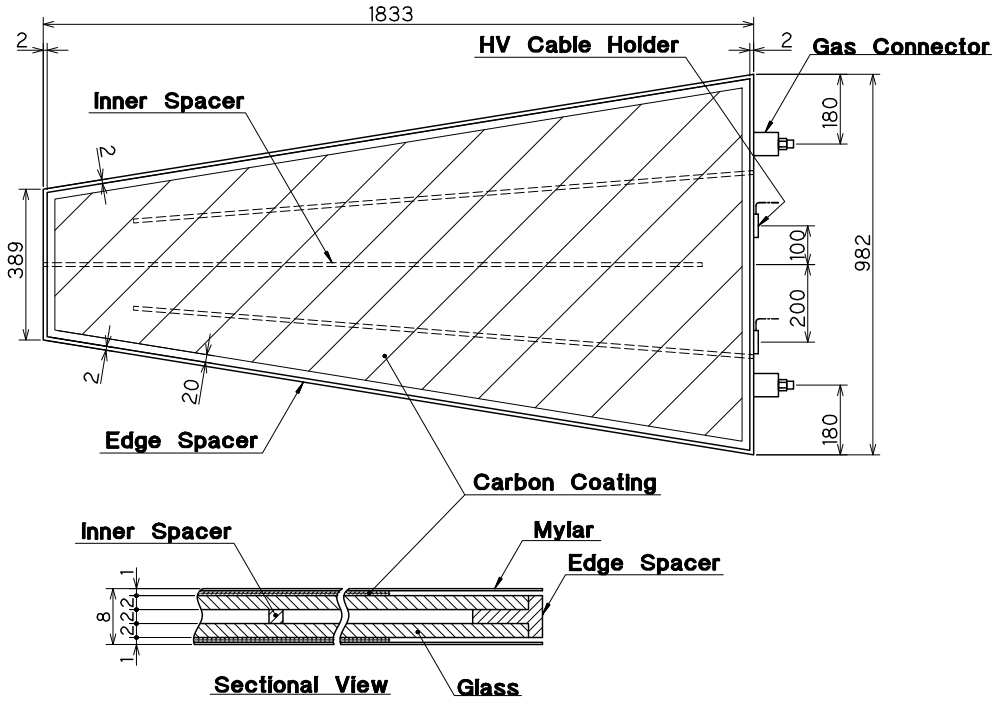


Figure 2.15.: An illustration of the KLM sub-detector [40].



(a) Barrel RPC module



(b) End-cap RPC module

Figure 2.16.: Schematic diagrams of the internal spacer arrangement for barrel and end-cap RPCs [40].

keep the total trigger rate at about 200 Hz. The Belle trigger system consisted of the level-1 (L1) hardware trigger and the level-3 (L3) software trigger.

The overview of the signal flow of the L1 trigger system is shown in Fig. 2.17. The trigger systems of the CDC and the SVD provided charged track signals, that of the TOF provided timing information of an event, and the ECL trigger system provided trigger information for both neutral and charged particles based on total energy deposit and cluster counting of crystal hits. The KLM trigger system provides additional information on muons and the EFC trigger system was used for tagging two photon and Bhabha events. The trigger signals from all sub-detectors were sent to the Global Decision Logic (GDL) [43] to make the global trigger decision. All the sub-trigger signals must arrive at the GDL within 1.85 μs after the event occurred, and the L1 final trigger must be issued 2.2 μs after the event crossing. The timing of the trigger signal must be accurate, because the trigger signal determined the readout timing. The timing of the final trigger was primarily determined by the TOF trigger. The final jitter of the trigger signal is less than 10 ns. The ECL trigger signals was also used as timing signals for events in which the TOF trigger was not available. The efficiency of the L1 trigger was ~100%.

The L3 trigger further reduced the number of background events. It used a fast track fitting algorithm to find tracks coming from the interaction point. Background events that did not produce such tracks were rejected. The L3 trigger had an efficiency of about 99% and an event rate reduction of 40%~50% [44].

The DAQ system was designed to be tolerant of a trigger rate up to 500 Hz with a dead time fraction of less than 10%. The global scheme of the DAQ system is shown in Fig 2.18. The DAQ system is segmented into 7 subsystems running in parallel, each handling the data from a sub-detector. When the GDL issued a trigger signal, the data from each sub-system was combined into a single event record by an event builder. Its output was transferred to an online computer farm, where the raw event data was converted into the offline data format, and then processed through the L3 trigger. The final data was then sent to a mass storage system located at the KEK computer center via optical fibers.

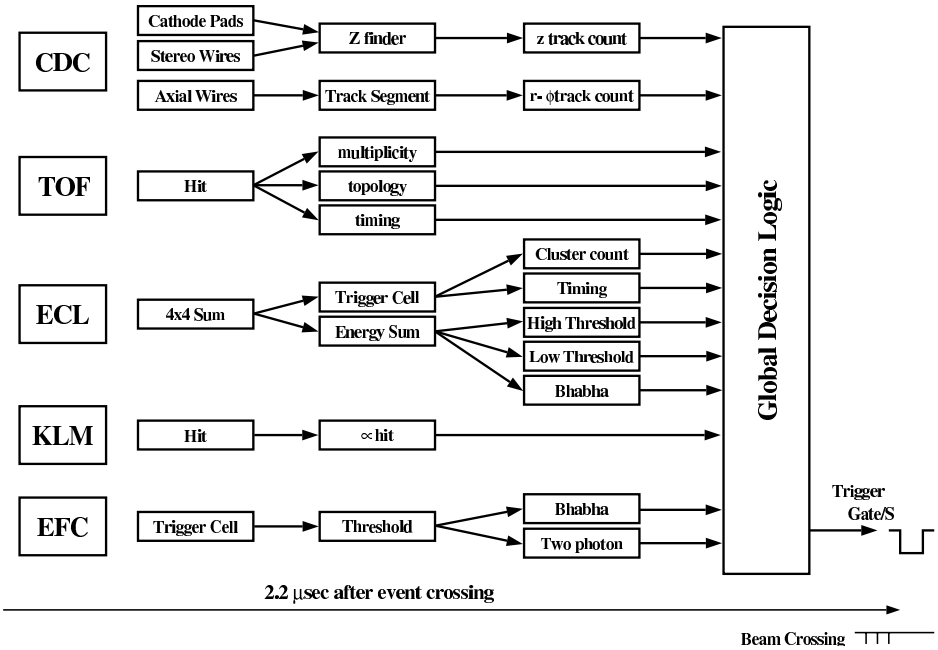


Figure 2.17.: Level-1 trigger system for the Belle detector [40].

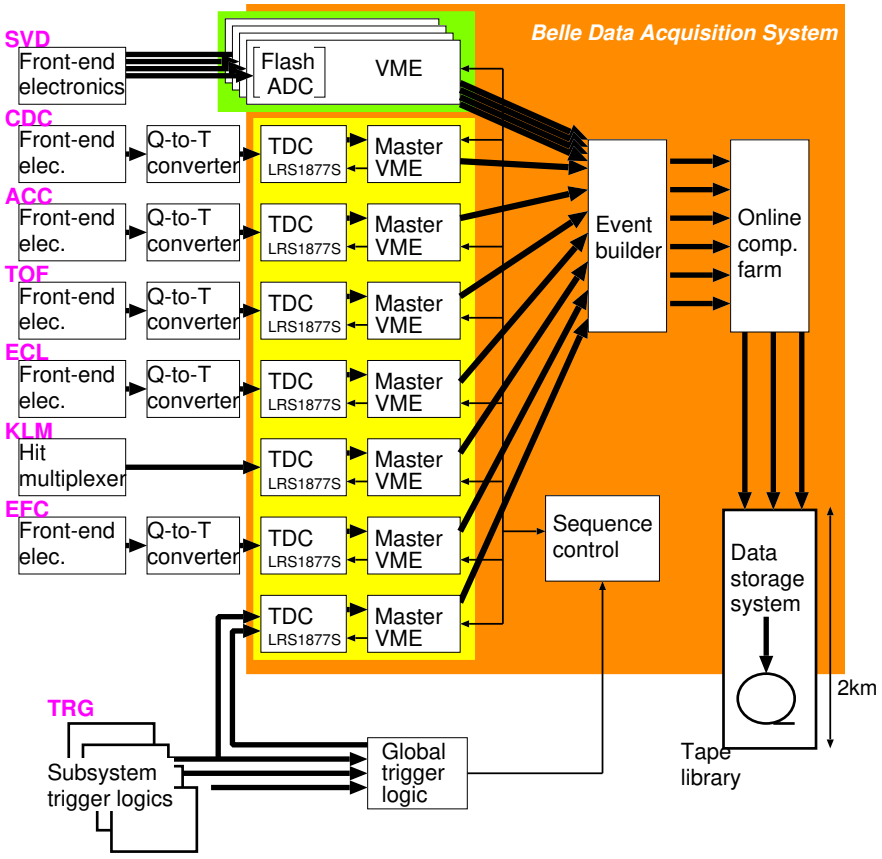


Figure 2.18.: Overview of the DAQ system [40].

2.3. Belle Analysis Framework

Belle Analysis Framework (BASF) is a software framework to handle event processing in the Belle collabration. Using this framework, users only have to write their own analysis code without taking care of the interface with external software, handling of data files, etc. An user's analysis code is written as a "module" which can be plugged into BASF. An analysis program is built by plugging user analysis modules into BASF along with the modules required for data input, output and data processing. Modules can be written in Fortran, C or C++ although BASF itself is written in C++. BASF is used for the DAQ and data processing described above, as well as Monte Carlo generation and event reconstruction described in the next chapter. The global architecture of BASF is shown in Fig. 2.19.

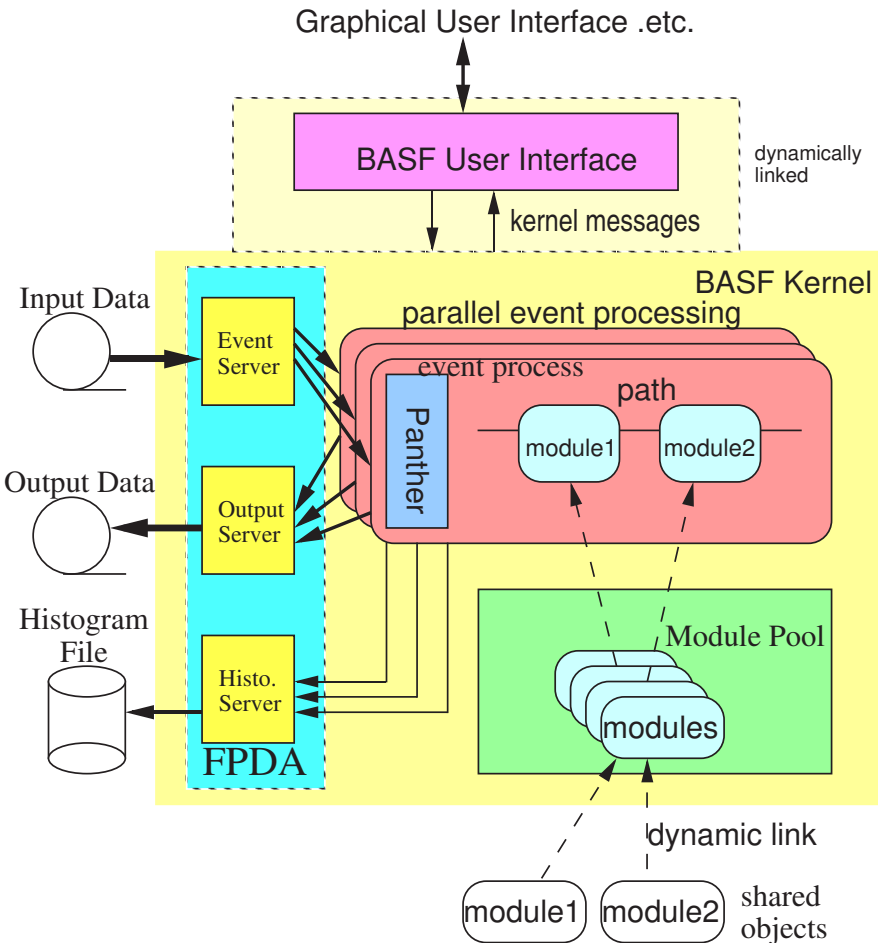


Figure 2.19.: Schematic of the BASF architecture.

2.3.1. Particle Identification (PID)

Particle identification (PID) for K/π is crucial for the analysis of $B^+ \to K^+K^-\pi^+$. The particle identification algorithm used in Belle is based on several measurements from the sub-detectors. A likelihood method is introduced to calculate a combined probability for charged tracks. Probability density functions (PDFs) are introduced for three discriminants, dE/dx from the CDC, timing information from the TOF, and the Čerenkov photons from the ACC. The kaon/pion likelihood ratio is:

$$\mathcal{P}(i) = \mathcal{P}_{dE/dx}(i) \times \mathcal{P}_{\mathrm{TOF}}(i) \times \mathcal{P}_{\mathrm{ACC}}(i), \tag{2.5}$$

where i is one of the particle hypotheses (e.g. K or π).

The likelihood for dE/dx is parametrized as a Gaussian distribution:

$$\mathcal{P}_{dE/dx}(i) \propto \frac{\exp(-\chi^2/2)}{\sigma_{dE/dx}} \tag{2.6}$$

$$\chi^2 = \left[\frac{(dE/dx)_{\mathrm{measured}} - (dE/dx)_{(i)}}{\sigma_{dE/dx}}\right]^2, \tag{2.7}$$

where $(dE/dx)_{\mathrm{measured}}$ is the dE/dx from the CDC (Fig. 2.10), and $(dE/dx)_{(i)}$ is the expected dE/dx for particle hypotheses i (K or π). The expected resolution of dE/dx is denoted by $\sigma_{dE/dx}$.

The likelihood for TOF hits is given by

$$\begin{cases} \mathcal{P}_{\mathrm{TOF}}(j) & \propto \frac{\exp(-\chi^2/2)}{\prod_j \sigma^j} \\ \chi^2 & = \sum_j (t^j_{\mathrm{measured}} - t^j_{(i)})^T (E_j)^{-1} (t^j_{\mathrm{measured}} - t^j_{(i)}), \end{cases} \tag{2.8}$$

where j denotes the j-th TOF hit, and t_{measured} is the vector with two timing signals from the PMT at the two ends of the TOF counter. The expected values for particle species i are stored in $t_{(i)}$, while E is a 2×2 error matrix and σ^j is the expected timing resolution.

The ACC is basically an on-off device. The number of observed photons (N_{pe}) by the FM-PMT depends on the associated particle type and track momentum. It is expected to have a peak at $N_{\mathrm{pe}} = 0$ and a tail corresponding to non-zero

photon numbers for kaons, while a finite number of $N_{\rm pe}$ is expected for pions. The tail is due to possible δ-ray Čerenkov, or light emission in the Goretex reflector in the ACC modules. This complex situation does not allow a simple function to describe the distributions of $N_{\rm pe}$. The likelihood for ACC ($\mathcal{P}_{\rm ACC}$) is parametrized in a simpler way:

$$\mathcal{P}_{\rm ACC}(i) = \begin{cases} \epsilon & (N_{\rm pe} \geq N^{(i)}_{\rm threshold}) \, ; \\ 1-\epsilon & (N_{\rm pe} \leq N^{(i)}_{\rm threshold}) \, , \end{cases} \tag{2.9}$$

where $N^{(i)}_{\rm threshold}$ is the threshold value of $N_{\rm pe}$ for particle hypotheses i, and ϵ is the expected efficiency. The $N^{(i)}_{\rm threshold}$ and ϵ are evaluated by look-up tables with a dependence on the momentum of the associated charged track for each ACC module. The look-up tables are obtained from Monte Carlo simulation, while the threshold values are chosen so as to maximize the separation power between the tracks from two-sides of the threshold.

The likelihood ratio ($\mathcal{R}_{K/\pi}$) is introduced as:

$$\mathcal{R}_{K/\pi} = \frac{\mathcal{P}_K}{\mathcal{P}_K + \mathcal{P}_\pi}. \tag{2.10}$$

The performance of kaon identification is determined by measuring the decay $D^{*+} \to D^0\pi^+(D^0 \to K^-\pi^+)$ [40, 45]. Fig. 2.20 shows the typical $\mathcal{R}_{K/\pi}$ distributions for high momentum charged tracks. It is important to note that the efficiencies and fake rates are always functions of track momentum and the polar angle of the particle. The efficiencies and fake rates are studied in 384 momentum and $\cos\theta$ bins in the lab frame. For example, particles exceeding a threshold, $\mathcal{R}_{K/\pi} > 0.6$, are tagged as kaons and those particles not exceeding this are tagged as pions. At this threshold, the average kaon efficiency for data is approximately 88%, while the average pion background contamination is approximately 8.5%; similarly, the averaged pion efficiency is around 89% and the kaon fake rate is about 8.8% [45]. Figure 2.21(a) shows the two dimensional scatter of $\mathcal{R}_{K/\pi}$ versus momentum of the tracks. The kaon efficiencies and pion fake rates as a function of momentum based on requiring $\mathcal{R}_{K/\pi} > 0.6$ are shown in Fig. 2.21(b).

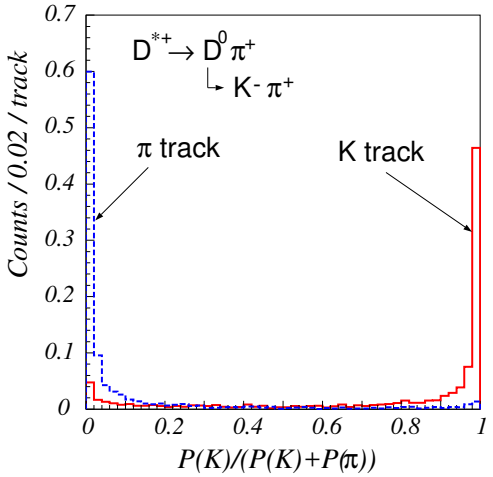
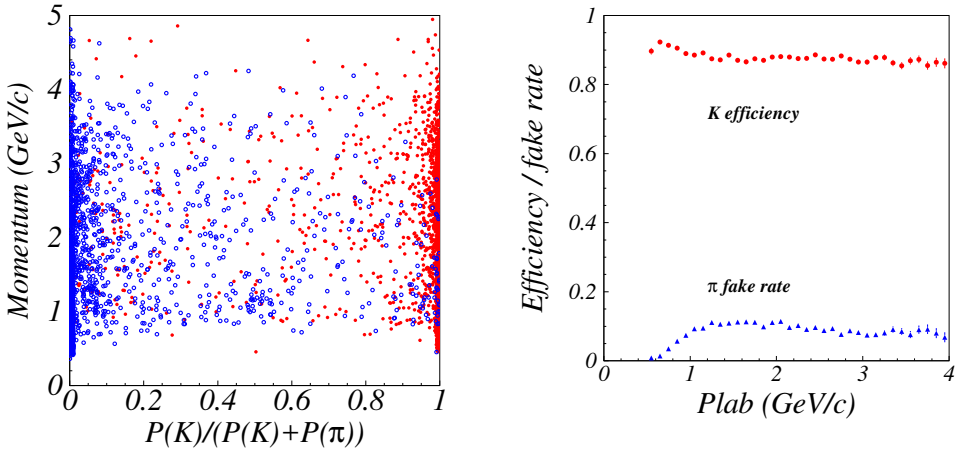


Figure 2.20.: Typical $\mathcal{R}_{K/\pi}$ distributions for high momentum charged tracks for a D^{*+} tagged data sample. The solid histogram is for charged kaon tracks and the dashed histogram shows the distribution for pions [46].



(a) (b)

Figure 2.21.: Performance of kaon identification from D^{*+} tagged $D^0 \to K^-\pi^+$ sample. Figure (a) shows the likelihood ratio $\mathcal{R}_{K/\pi} = P_K/(P_K + P_\pi)$ versus the momentum of tracks, while the filled (open) circles denote kaon (pion) tracks; figure (b) shows the kaon efficiencies and pion fake rates as a function of track momentum based on requiring $\mathcal{R}_{K/\pi} > 0.6$.

Chapter 3.

Analysis of $B^+ \to K^+K^-\pi^+$

The main goal in this study is to measure the branching fraction and direct *CP* asymmetry of $B^+ \to K^+K^-\pi^+$ decays. This chapter describes the procedure to reconstruct signal events, suppress backgrounds, extract signal yield and *CP* asymmetry, and validate the procedure using a control sample analysis.

3.1. Event Selection

Analyses at Belle must be conducted blind, which means all the procedures need to be studied and validated with Monte Carlo (MC) simulation samples before analysing the final data. In order to study the possible background sources, we use the official background MC samples provided by Belle. For possible discrepancies between data and MC sample, we use a control sample analysis (see 3.4). Moreover, an off-resonance data sample is used to validate treatment of continuum background, which is described in later sections.

3.1.1. MC Data Sample

This analysis is based on a data sample that contains $(771.58 \pm 10.57) \times 10^6$ $B\overline{B}$ pairs collected by the Belle detector at the KEKB collider with asymmetric energy e^+e^- (3.5 GeV and 8 GeV, respectively) at the $\Upsilon(4S)$ resonance. Belle data collection is grouped into "experiments" based on different detector conditions.

We study signal and backgrounds using a blind study with the MC simulation method. We generate 771,000 B^+B^- pairs, with one of the B mesons decaying to $K^+K^-\pi^\pm$ events, via the EvtGen [47] package, and the detector simulation is done with the GEANT [48] package for the signal decay study. The MC samples which are simulated with the GEANT package, including the background MC samples described below, are also called GSIM samples. The signal MC events are distributed in several groups from Exp.7 to Exp.65 according to the luminosity of each experiment in the data sample and simulated with the corresponding detector conditions. The decay amplitude is assumed to be phase space in signal MC, which means no specific angular distribution.

For the background study, we use the official background MC samples as following: Generic B ($b \to c$ transition) MC sample equivalent to 5 times the Exp.7 - Exp.65 data set, continuum ($e^+e^- \to q\bar{q}, q = u, d, s, c$) MC sample equivalent to 3 times the data set, and Rare B ($b \to u, d, s$ transition) MC sample equivalent to 50 times the data set. In the Generic B MC sample, both B mesons decay through a $b \to c$ transition. In this Rare B sample, one of $B^\pm$ or $B^0(\overline{B}^0)$ decays through a charmless decay process with a known or estimated branching fraction; while the other B decays generically through $b \to c$ transition.

3.1.2. Control and Off-Resonance Data Samples

To validate the analysis procedure, we perform a control sample study (see 3.4), which involves applying the analysis procedure to a data set of a known decay with large statistics. An MC sample of $1,542,000$ $B\overline{B}$ pairs is generated for the control sample study. For the background samples, we also use the official MC samples: continuum MC equivalent to 3 times the data set, Generic B MC equivalent to the data set, and Rare B MC equivalent to 50 times the data set.

To study possible discrepancies between data and MC in continuum events, a data sample of 89.4 fb^{-1} collected 60 MeV below the $\Upsilon(4S)$ resonance, off-resonance data, is used for the background study as well.

3.1.3. Kaon/Pion Selection

We select charged kaons and pions to reconstruct B mesons based on the likelihood ratio ($\mathcal{R}_{K/\pi}$) from the particle identification algorithm (please see Sec. 2.3.1 for details). Good kaon/pion selection is important to reduce the background from misidentified tracks as they form peaking backgrounds which will be discussed later (See 3.2.2). Both kaons and pions are required to satisfy track quality criteria based on the impact parameter relative to the interaction point (IP). The deviations from the IP position are required to be within ±0.2 cm in the transverse plane (dr), and within ±5 cm in the z axis (dz). A track is identified as kaon if $\mathcal{R}_{K/\pi} > 0.6$ or as pion if it is less than 0.4.

3.1.4. B Reconstruction

The B candidates are reconstructed from a pair of charged kaon mesons with opposite charges, and a charged pion meson. Two kinetic variables are used to identify the reconstructed B meson candidates: the beam constrained mass

$$M_{\rm bc} \equiv \sqrt{E^2_{\rm beam} - p^2_B}, \tag{3.1}$$

and the energy difference

$$\Delta E \equiv E_B - E_{\rm beam}. \tag{3.2}$$

$E_{\rm beam}$ is the beam energy in the center-of-mass (CM) frame, and p_B and E_B are the momentum and energy of the reconstructed B meson in the CM frame. The candidate region is defined by $-0.3 < \Delta E < 0.3$ GeV and $5.24 < M_{\rm bc} < 5.29$ GeV/c^2. The signal box region is defined by $-0.05 < \Delta E < 0.05$ GeV and $5.27 < M_{\rm bc} < 5.29$ GeV/c^2.

3.1.5. B Vertex Fit and Best Candidate Selection

Only one signal B is expected per event, as the expected number of events having both B mesons decay to our signal is only 0.02 in the data set, however multiple B candidates may be reconstructed due to misreconstruction. Therefore,

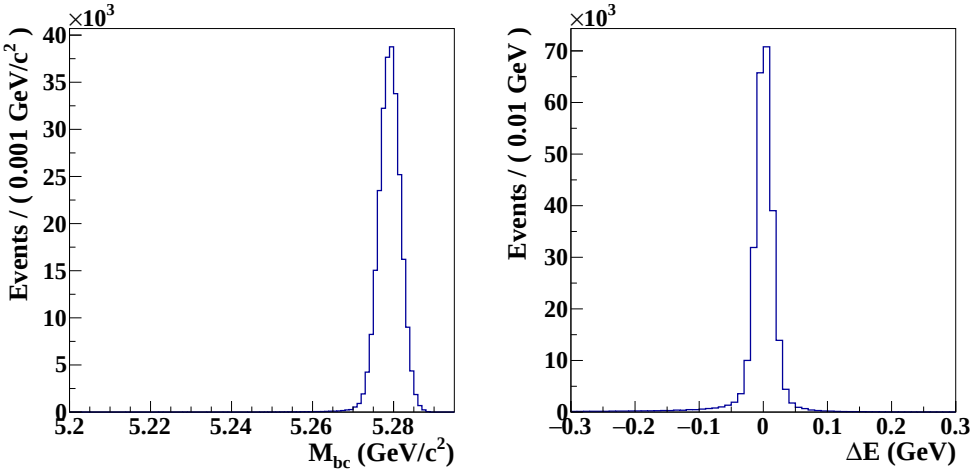
we need a criterion to choose the best candidate. We use the charged daughter particles to determine the B vertex and the quality of the fit is determined via the returned χ^2 value from using the Belle vertexing code, ExKFitter. If there are more than one B meson candidate, the one with the smallest χ^2 is chosen. The signal MC sample shows that multiple candidates occur in 25.8% of events. The average multiplicity is 1.344. With our selection, the selected B candidate corresponds to a true B candidate 92% of the time.

3.1.6. Efficiency measured on MC

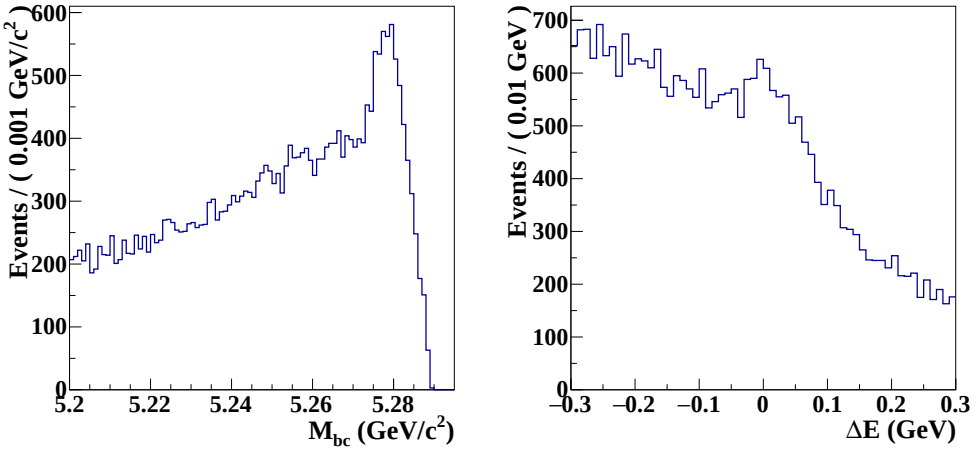
After all selection criteria and the best candidate selection is applied, the efficiency of the signal events is 40.16%. However, even in the pure signal MC, there are some events where the best B candidate is reconstructed with final state particles from the other B, the so called self-cross feed (SCF). Please note that in the following sections the term "true events" refers to correctly reconstructed signal events.
The $M_{\rm bc}$ and ΔE distributions for true events and SCF events are shown in Fig. 3.1. The efficiency of true events is 36.37 ± 0.05%, and the ratio of SCF events is 9.47 ± 0.05%. Consequently, the expected signal events in data is

$$N_{expect} = \epsilon \times N_{B\bar{B}} \times \mathcal{B}(B^\pm \to K^+K^-\pi^\pm) = 1402.1, \tag{3.3}$$

with $\mathcal{B}(B^\pm \to K^+K^-\pi^\pm)$ taken from the PDG as $(5.0 \pm 0.5) \times 10^{-6}$ [49].



(a) True Signal



(b) Self-cross feed

Figure 3.1.: The ΔE and M_{bc} distribution of true signal events (a) and SCF (b) events after the best candidate selection.

3.2. Background Studies

3.2.1. Continuum Suppression

The dominant background in this analysis is from the continuum process $e^+e^- \to q\bar{q}(q = u, d, s, c)$. The event topology difference between continuum and $B\overline{B}$ can be exploited; the jet-like distribution of continuum events makes the event shape variables useful for separating them from the $B\bar{B}$ events which are spherically distributed as shown in Fig. 3.2. Therefore, we introduce the KSFW method and then combine the output of the KSFW method with four other

discriminating variables using the NeuroBayes software package [50] which employs an artificial neural network algorithm. The ΔE and $M_{\rm bc}$ distributions of the continuum background are shown in Fig. 3.3.

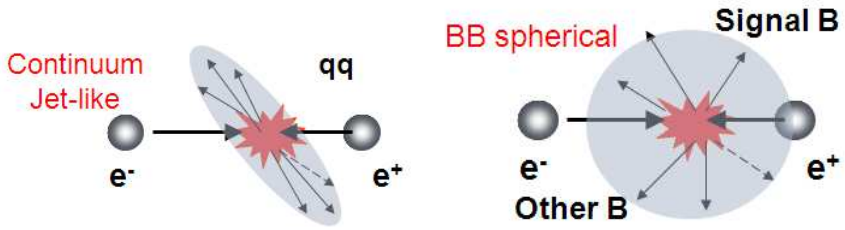


Figure 3.2.: Diagram of jet-like continuum events and spherical-like $B\overline{B}$ events.

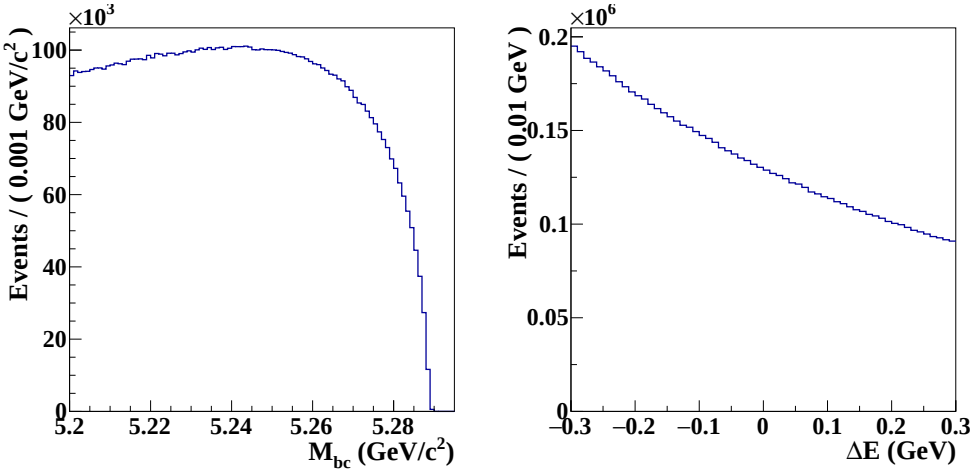


Figure 3.3.: The ΔE and $M_{\rm bc}$ distributions of continuum MC after the best candidate selection.

Kakuno Super Fox-Wolfram moments (KSFW)

The lth-order Fox-Wolfram moment R_l is defined as [51]:

$$H_l = \sum_{i,j} \frac{|\vec{p}_i||\vec{p}_j|}{s} P_l(\cos\theta_{ij}), \tag{3.4}$$

$$R_l = \frac{H_l}{H_0}, \tag{3.5}$$

where $\vec{p_i}$ and $\vec{p_j}$ are the momenta of the ith and jth daughter particles in the event, s is the square of the total energy of the event, P_l denotes the Legendre polynomial of order l, and $\theta_{i,j}$ is the angle between the momenta of the particles. Usually, the moments are normalized to H_0. The most commonly used variable in the set is the second order of the Fox-Wolfram moment, R_2, for which has the most distinguishing power to separate B events and jet-like events.

As an improvement, super Fox-Wolfram (SFW) moments are defined by decomposing the Fox-Wolfram moments into three parts [52]:

$$R_l = R_l^{SO} + R_l^{SS} + R_l^{OO} \tag{3.6}$$

$$\begin{aligned} R_l^m &= \frac{H_l^m}{H_0^m} \ (m = SS, SO, OO) \\ &= \frac{\sum_{i,j} |\vec{p_i}||\vec{p_j}| P_l(\cos\theta_{ij})}{\sum_{i,j} |\vec{p_i}||\vec{p_j}|}, \end{aligned} \tag{3.7}$$

where S denotes the daughter particles from the reconstructed B, and O denotes the remaining particles. The difference for the SFW moments is that we consider all reconstructed particles, including signal B candidates and remaining particles in the double sum. For R^{SS}, the double sum is constrained to the sum over tracks from the B candidate; for R^{SO}, one index runs over signal tracks whiles the other index runs over particles from the other side of the event. Since the R_l^{SS} term is obtained from the momenta of the reconstructed B daughter particles, it is correlated to $M_{\rm bc}$ and ΔE. Since the final signal extraction will use these two kinematic variables, the R^{SS} term is not used in the background suppression. For the odd l, R^{SO} is zero if the signal tracks are back to back, this causes a strong correlation with the beam constrained mass. It is thus removed. The SFW variable is the Fisher discriminant [53] projection of the first to the fourth moments of these two series:

$$SFW = \sum_{l=2,4} a_l R_l^{SO} + \sum_{l=1\sim4} b_l R_l^{OO} \tag{3.8}$$

where a_l and b_l are Fisher coefficients. The definition of Fisher discriminant and coefficients will be discussed in later section.

H. Kakuno developed the KSFW algorithm by introducing the missing momentum as a discriminant, and combining it with the original SFW moments.

The variables of KSFW are defined as

$$KSFW = \sum_{l=0}^{4} R_l^{SO} + \sum_{l=0}^{4} R_l^{OO} + \gamma \sum_{n=0}^{N_t} |(p_t)_n|. \tag{3.9}$$

- R_l^{SO}
  The missing momentum is viewed as one additional particle and then used to calculate the Fox-Wolfram moments of the event. The moments are separated into "c: charged", "n: neutral", and "m: missing" categories:

$$R_l^{SO} = \frac{(a_c)_l H_{cl}^{SO} + (a_n)_l H_{nl}^{SO} + (a_m)_l H_{ml}^{SO}}{E_{\text{beam}} - \Delta E}, \tag{3.10}$$

  where $E_{\text{beam}} - \Delta E = H_0^{SO}|_{mm^2}$. Here mm^2, the missing mass square, is defined as:

$$mm^2 = (E_{\Upsilon(4S)} - \sum_{n=1}^{N_t} E_n)^2 - (\sum_{n=1}^{N_t} \vec{p_n})^2, \tag{3.11}$$

  where $E_{\Upsilon(4S)}$ is the energy of $\Upsilon(4S)$ in the rest frame and E_n and $\vec{p_n}$ present the energy and momentum of n-th particle in a event.

  For $l = 1$ and 3,

$$H_{cl}^{SO} = \sum_i \sum_{jX} b_l^{SO} Q_i Q_{jX} |p_{jX}| P_l \cos\theta_{ijX} \tag{3.12}$$

$$H_{nl}^{SO} = H_{ml}^{SO} = 0, \tag{3.13}$$

  here the index i iterates over the daughter particles in the B candidates and the index jX iterates over the remaining particles in the category X ($X = c, n, m$). Q_i and Q_{jX} are the charge of the particle i and jX.

  For $l = 0, 2, 4$,

$$H_{Xl}^{SO} = \sum_i \sum_{jX} b_l^{SO} |p_{jX}| P_l \cos\theta_{ijX}|. \tag{3.14}$$

There are two parameters of $(a_X)_l \cdot b_l^{SO}$ for $l = 1, 3$ while nine parameters for $l = 0, 2, 4$.

- R_l^{OO}
  For $l = 1$ and 3,

$$R_l^{OO} = \frac{\sum_j \sum_k b_l^{OO} Q_j Q_k |p_j||p_k| P_l \cos\theta_{jk}}{E_{\text{beam}} - \Delta E}, \tag{3.15}$$

  where the index j and k iterates over the remaining particles and Q_j is the charge of the particle j.

  For $l = 0, 2, 4$,

$$R_l^{OO} = \frac{\sum_j \sum_k b_l^{OO} |p_j||p_k| P_l \cos\theta_{jk}}{E_{\text{beam}} - \Delta E}. \tag{3.16}$$

  There are five parameters in total in this term.

- $\sum_{n=0}^{N_t} |(p_t)_n|$
  $\sum_{n=0}^{N_t} |(p_t)_n|$ is the scalar sum of the transverse momenta p_t of all final state particles in both signal B candidate and the remaining particles. N_t is the number of all particles. There is one parameter, γ, in this term.

In total there are 17 parameters in KSFW that will be determined in the training procedures. Because the missing mass is related to the value of KSFW, we split the events into 7 regions of missing mass as shown in Table 3.1.

Table 3.1.: The regions of missing mass of KSFW.

Region	1	2	3	4
mm^2 (GeV/c^2)	< -0.5	$-0.5 \sim 0.3$	$0.3 \sim 1.0$	$1.0 \sim 2.0$
Region	5	6	7	
mm^2 (GeV/c^2)	$2.0 \sim 3.5$	$3.5 \sim 6.0$	> 6.0	

Event Shape Variables

In addition to the KSFW moments, we use other discriminating variables to enhance the separation strength. The additional variables used are:

- B flight direction($\cos\theta_B$)
  θ_B is the angle between the momentum of the B candidate and the e^+e^- vector in the $\Upsilon(4S)$ frame. In the decay $\Upsilon(4S) \to B\overline{B}$, a vector particle decays to two scalar particles, thus the angular distribution would obey the $1 - \cos^2\theta$ formula based on quantum physics, while the distribution in the continuum background is flat.

- Δz
  Δz is the distance in the z axis between the candidate B vertex and the vertex position formed by the charged tracks that are not associated with the candidate B meson. Because of the lifetime of the B-meson, the width of Δz is wider in the signal events than in the continuum events.

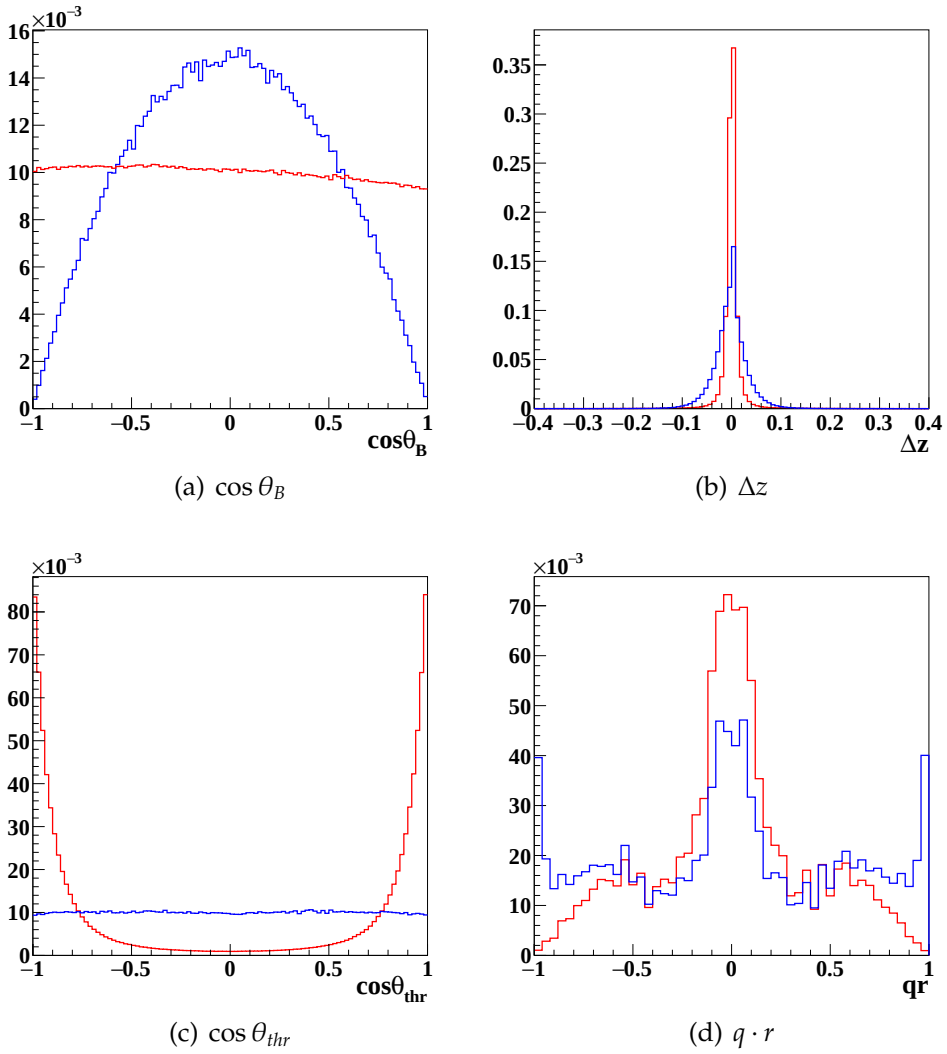
- Thrust angle ($\cos\theta_{thr}$)
  The thrust axis ($\vec{n}$) is determined by maximizing $T(\vec{n})$:

$$T(\overline{n}) = \frac{\sum_i |\vec{p}_i \cdot \vec{n}|}{\sum_i |\vec{p}_i|}, \tag{3.17}$$

  where $\vec{p}_i$ is the momentum of the ith particle from the B candidate or the remaining one in an event. The definition of thrust angle θ_{thr} is the angle of the thrust axis of B candidate with respect to the thrust axis of remaining particles. The distribution of $\cos\theta_{thr}$ will be randomly distributed from -1 to $+1$ if we get a real B event or it will be peaking at ± 1 for continuum events due to the jet-like event shape.

- Tagging variable $q \cdot r$
  $q \cdot r$, which is from a B flavour-tagging algorithm [54], provides a tagging quality and the preferred flavour of the other B meson according to information from the particles not associated with the B candidate. The preferred flavour q equals $+1$ for B^+ or B^0 and -1 for B^- or $\overline{B}^0$. The quality factor "r" ranges from 0 (for no flavour information) to 1 (for unambiguous flavour assignment). If $q \cdot r$ approaches $+1$ (-1), the tagged B is similar to B^+/B^0 ($B^-/\overline{B}^0$). Signal events are likely to peak around ± 1.

The distributions of above variables are shown in Fig. 3.4



(a) $\cos\theta_B$ (b) Δz

(c) $\cos\theta_{thr}$ (d) $q \cdot r$

Figure 3.4.: The shape variables used in continuum suppression from signal MC (blue histogram) and continuum MC (red histogram). The top left and right are $\cos\theta_B$ and Δz, and the bottom two are $\cos\theta_{thr}$ and $q \cdot r$.

Multi-variate Analysis Methods

We introduce two multi-variate analysis methods for the background suppression in this analysis. First we choose 17 KSFW moments as the input variables for a Fisher discriminant. The Fisher discriminator, $\mathcal{F}$, is a method to combine n-dimensional variables into one dimension by a linear weighted sum, where the coefficient for each variable is optimized to separate the signal and

background.

$$\mathcal{F} = \vec{W} \cdot \vec{V}, \tag{3.18}$$

where $\vec{W}$ is a row vector consisting of the input variables. The Fisher discriminant outputs are used to calculate the likelihood ratios $\mathcal{LR} = \mathcal{L}_S/(\mathcal{L}_S + \mathcal{L}_B)$, where "$S$" stands for signals and "B" stands for the backgrounds. The distributions of the KSFW discriminants for all of the mm^2 bins are shown in Fig. 3.5, and the normalized $\mathcal{LR}$ distribution of signal MC and continuum MC is shown in Fig. 3.6.

In addition to $\mathcal{LR}$, four variables mentioned above ($\cos\theta_B$, Δz, $\cos\theta_{thr}$, and $q \cdot r$) are combined in an artificial neural network provided by the NeuroBayes software package. To optimize the suppression power in the signal region, only the events in the signal box are used for both KSFW and NeuroBayes methods. The NeuroBayes output (NN) requirement is determined by maximizing the figure-of-merit (FOM), which is equal to the statistical significance:

$$FOM = \frac{S}{\sqrt{S+B}}, \tag{3.19}$$

where S denotes the number of signal events estimated by assuming the branching fraction to be 5×10^{-6}, and B denotes the sum of number of background events, of which the continuum background events are obtained from the off-resonance data sample (please see detailed description in below). The NN distribution obtained from the signal and continuum events and the distribution of FOM are shown in Fig. 3.7. The highest value of the FOM is at $NN > 0.88$. Therefore we apply a selection criterion of $NN > 0.88$, rejecting 98.71% continuum events while keeping 47.99% signal events.

Consistency Check with Off-resonance Data

In order to study the discrepancy between continuum MC simulation and real data, the off-resonance data is used to validate the PDF shapes obtained from continuum MC (detail in 3.3.2), and also to acquire the expected number of continuum events. Due to the different beam energy used in the off-resonance

data, the beam energy used in the Eq. 3.1 is fixed at 5.289 GeV in order to compare the distribution of $M_{\rm bc}$. As shown in Fig 3.21(b), no significant difference is seen in the check. After accounting for the different luminosity of the off-resonance data sample, we found the number of events obtained is 1.7 times the expected number from simulated MC. This indicates that the yield of $e^+e^- \to q\bar{q}$ is different to the predictions from continuum MC simulation. Therefore, we used the rescaled off-resonance rates for the continuum component in the FOM calculation and the ensemble tests.

3.2.2. $B\overline{B}$ background study

There are two kinds of $B\overline{B}$ backgrounds, Generic B, which denotes the $b \to c$ transition; and Rare B, which denotes $b \to u, d, s$ transitions. The MC samples used in this study is described in Sec. 3.1.1

Generic B background

First we look at the $M_{\rm bc}$ and ΔE distributions of Generic B events, which are shown in Fig. 3.8. The ΔE distribution is found to be strongly peaking in the signal region. To investigate the source of the peaks, we check the invariant mass distributions of K^+K^- and $K^\pm\pi^\mp$ systems. The invariant mass distribution of the $K^\pm\pi^\mp$ system shown in Fig 3.9 has a small bump at the nominal D^0 mass. The K^+K^- invariant mass shown in Fig. 3.10 has two peaks that correspond to the contribution from charm decays: the one peaking at the nominal D^0 mass comes from the decay $D^0 \to K^+K^-$ and the other one with the peak slightly shifted from the D^0 mass comes from $D^0 \to K^-\pi^+$ when a π is misidentified as a K meson. The contribution of $D^0 \to K^-\pi^+$ in the K^+K^- system is confirmed by changing the assigned mass of the K to that of a π. To remove the charm backgrounds, we apply a D^0 mass selection criterion to veto the events within $[1.85, 1.88]\text{GeV}/c^2$, which corresponds to 3.75σ around the mean. The mass window is decided by fitting the KK and $K\pi$ invariant masses with a Gaussian function shown in Fig. 3.11. Figure 3.12 shows the 2D distribution of the $K\pi$ invariant masses of the misidentified events. After applying the D^0 mass veto in M_{KK}, a bump appears around 3.4 GeV/c^2 in the K^+K^- system in the right

plot in Fig. 3.10, which comes from $\chi_{c0} \to K^+K^-$. A mass window criterion of [3.375, 3.475] GeV/c^2 is also introduced for the χ_{c0} decay. Further charmonium backgrounds are studied by releasing the lepton identification requirements. As shown in Fig. 3.13, there is no apparent peaking in the K^+K^- system, however, a peak is found at the nominal J/ψ mass in the $K^\pm\pi^\mp$ system. Therefore, a mass window selection [3.06, 3.14] GeV/c^2, which corresponds to 4σ as shown in Fig 3.14, is included in the J/ψ veto. The veto windows are summarized in Table 3.2. Figure 3.15 and Figure 3.16 show the Dalitz plot and the $M_{\rm bc}$ and ΔE distributions after the charm veto is applied, respectively.

Table 3.2.: The mass windows of the charm veto.

Mass windows	Variables
$1.85 < M_{D^0} < 1.88$	$M_{KK}, M_{K\pi}, M_{K^+\pi^-}, M_{\pi^+K^-}$
$3.06 < M_{J/\psi} < 3.14$	$M_{ee}, M_{\mu\mu}$
$3.375 < M_{\chi_{c0}} < 3.475$	M_{KK}

Rare B background

In the Rare B MC, which are $B\overline{B}$ events with charmless processes, there are several quasi-two-body decays found to contribute in our signal channel, listed in Table 3.3. After removing these decays from the Rare B MC samples, two peaks are found in the ΔE distribution shown in Fig. 3.17, and these come from $B^+ \to K^+K^-K^+$ (left) and $B^+ \to K^+\pi^+\pi^-$ (right), with a kaon faking to a pion in the former case and a pion faking to a kaon in the latter case. To confirm the ratios of the two peaks and the rest of combinatorial events, we examine the amount of peaking components with a 1-D fit method on the ΔE distribution. Then we compare the resultant yields between the Rare B MC and the generated MC samples for these decays. We generate 771000 events for both channels. The ΔE distributions for $B^+ \to K^+K^-K^+$ and $B^+ \to K^+\pi^+\pi^-$ are described by a double Gaussian and a second order Chebyshev polynomial; for the remaining events in the Rare B MC, a smoothed histogram is used. Since the branching fractions of the peaking channels in the Rare B MC sample are not the same as the PDG variables, we need to rescale the number of events. In the Rare B sample, the branching fractions are 3.24×10^{-5} for $B^+ \to K^+K^-K^+$ and 5.9×10^{-5}

for $B^+ \to K^+\pi^+\pi^-$, while the PDG values are 3.40×10^{-5} for $B^+ \to K^+K^-K^+$ and 5.1×10^{-5} for $B^+ \to K^+\pi^+\pi^-$ [49]. We extract the yields for the peaking channels and rescale them with the corresponding branching fraction in the Rare B MC sample. The resultant yields of peaking backgrounds are consistent with the expected number of events in the Rare B sample. The fitting results are shown in Fig. 3.18 and Table 3.4. From now on, Rare B refers to Rare B samples without the peaking components ($B^+ \to K^+K^-K^+$ and $B^+ \to K^+\pi^+\pi^-$), unless otherwise stated.

Table 3.3.: List of possible signal channels in the Rare B MC sample.

No.	Decay channel
1.	$B^+ \to a_0^0\pi^+$
2.	$B^+ \to a_2^0\pi^+$
3.	$B^+ \to f_0\pi^+$
4.	$B^+ \to f_0^{'}\pi^+$
5.	$B^+ \to f_1\pi^+$
6.	$B^+ \to f_1^{'}\pi^+$
7.	$B^+ \to K_2^{*0}K^+$

Table 3.4.: The fitted yields of $B^+ \to K^+K^-K^+$, and $B^+ \to K^+\pi^+\pi^-$ modes in the Rare B MC and the generated MC. The yields are calculated with the branching fraction in the Rare B sample.

MC Sample	Yield
$B^+ \to K^+K^-K^+$ (generated MC)	1450.7 ± 8.9
$B^+ \to K^+K^-K^+$ (Rare B MC)	1452.2 ± 15.2
$B^+ \to K^+\pi^+\pi^-$ (generated MC)	857.5 ± 12.1
$B^+ \to K^+\pi^+\pi^-$ (Rare B MC)	821.3 ± 11.8

3.2.3. Estimated signal events in data

After applying all the background suppression selection, the signal efficiency on MC becomes 15.92 ± 0.04%, and the ratio of SCF events is 3.64 ± 0.04%. The

expected signal is 613.9 ± 85.9 events in the entire fitting region. Please note that the expected branching fraction of signal is 5×10^{-6} [2].

Table 3.5.: Expected number of background events after all selection criteria.

	Fitting Region	Signal Box
Continuum	27886.0 ± 118.1	1218.0 ± 24.7
Off-res. data (rescaled)	48210.0 ± 601.3	1950.0 ± 120.9
Generic B	20073.8 ± 57.8	969.17 ± 12.71
Rare B	3166.6 ± 8.0	151.4 ± 1.7
$B^+ \to K^+K^-K^+$	972.3 ± 40.4	215.5 ± 9.3
$B^+ \to K^+\pi^+\pi^-$	580.4 ± 33.4	86.5 ± 5.3

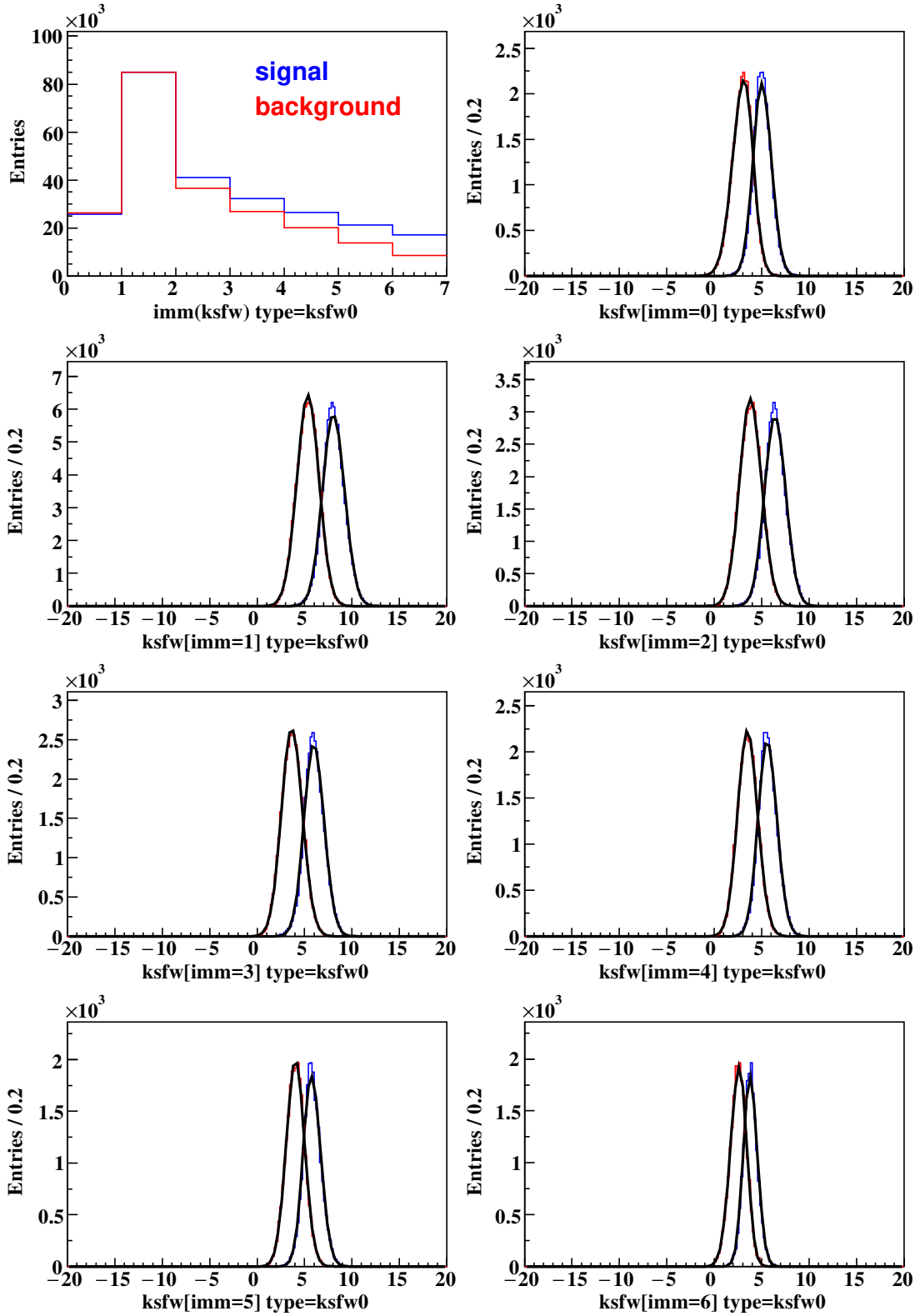


Figure 3.5.: First plot is the distribution of events in the mm^2 bins. The other plots show the distribution of the Fisher discriminant in each mm^2 bin. The blue line indicates the signal events, the red line indicates the continuum events, and the black line is the fitted distribution.

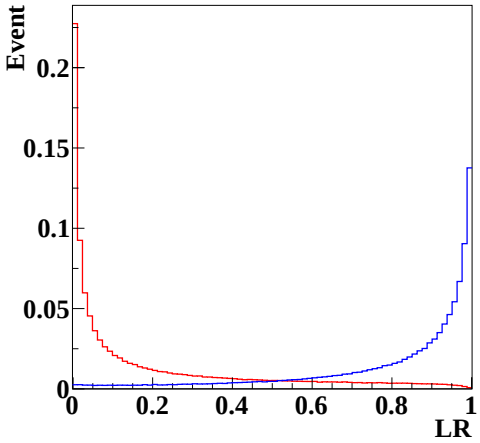
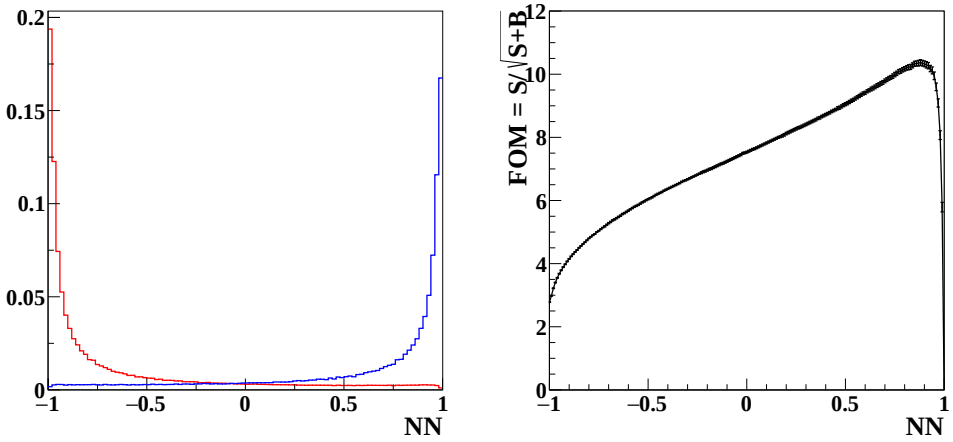
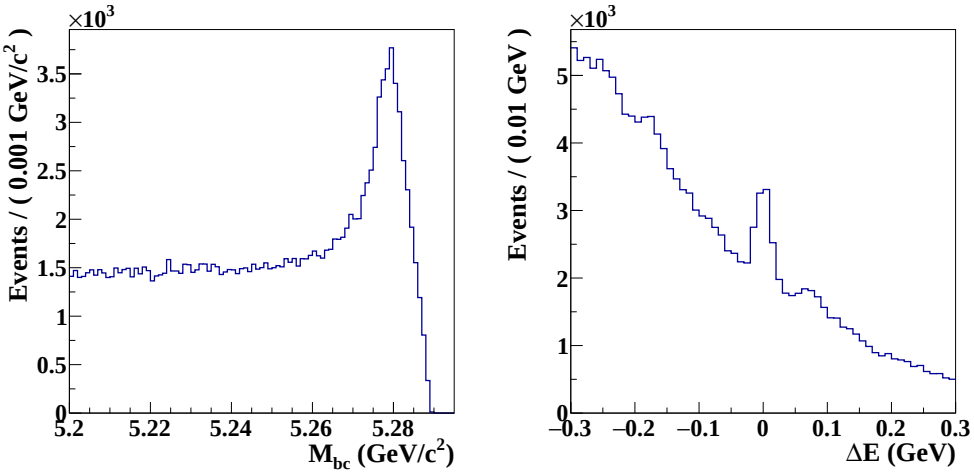


Figure 3.6.: The normalized $\mathcal{LR}$ distribution of the KSFW method for background suppression of the signal MC (blue line) and continuum MC (red line).

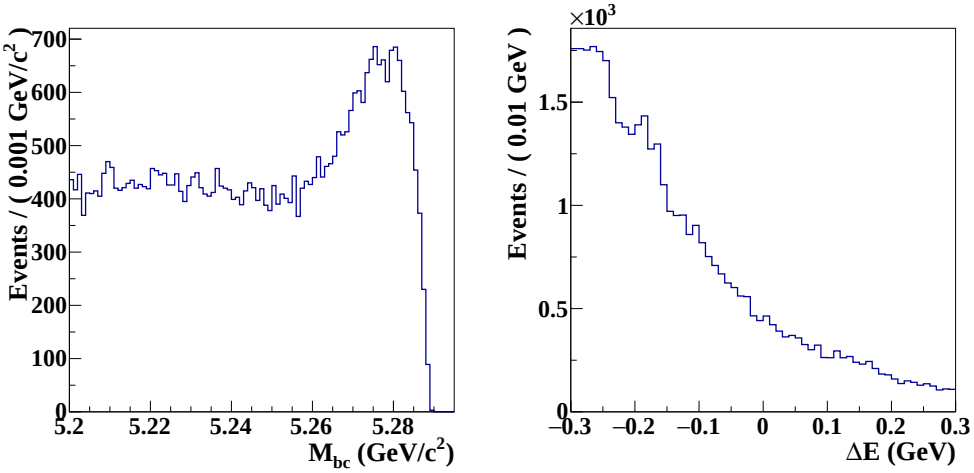


(a) Neurobayes Output (b) FOM v.s. *NN*

Figure 3.7.: Left plot shows the *NN* distribution for signal events (blue line) and continuum events (red line). The right plot shows the FOM distribution with different *NN* selection criteria.



(a) Generic $B^{\pm}$



(b) Generic B^0

Figure 3.8.: The $M_{\rm bc}$ and ΔE distributions of Generic B MC sample. The top plots are for the $B^{\pm}$ events, and the bottom plot are for the B^0 events.

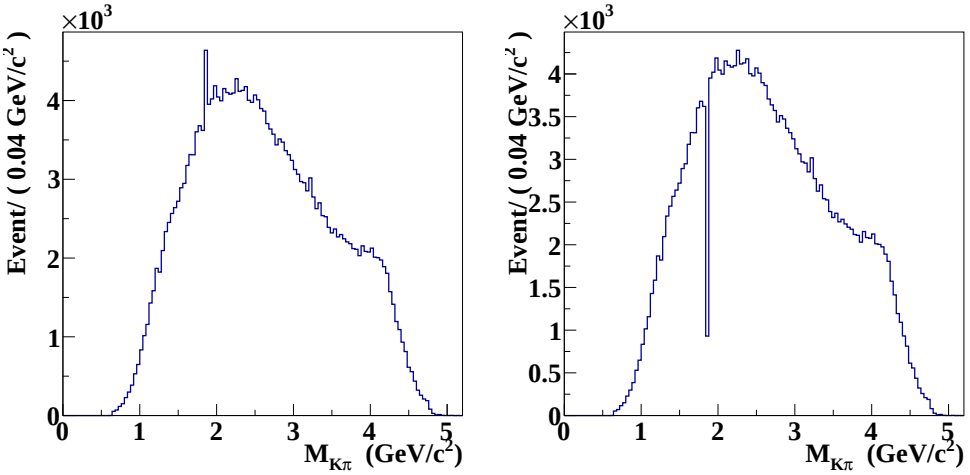


Figure 3.9.: The $K^{\pm}\pi^{\mp}$ invariant mass distributions from Generic B MC sample before (left) and after (right) D^0 mass veto.

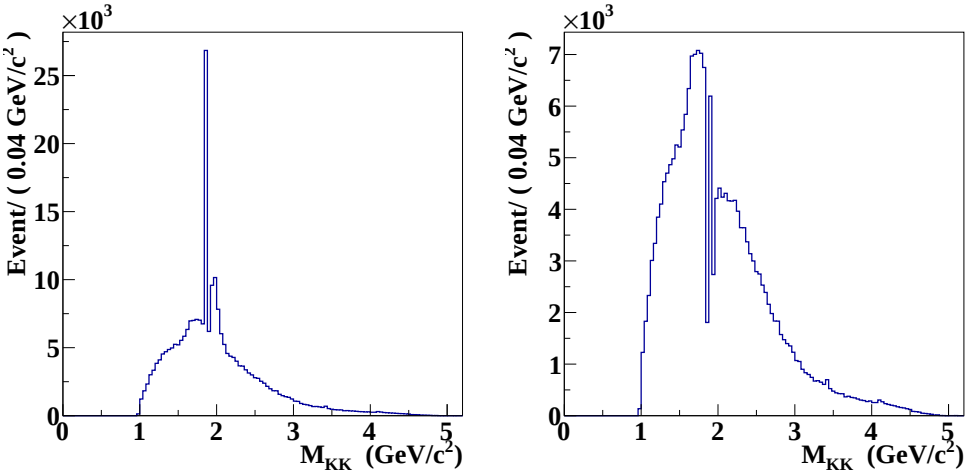


Figure 3.10.: The K^+K^- invariant mass distributions from Generic B MC sample before (left) and after (right) D^0 mass veto.

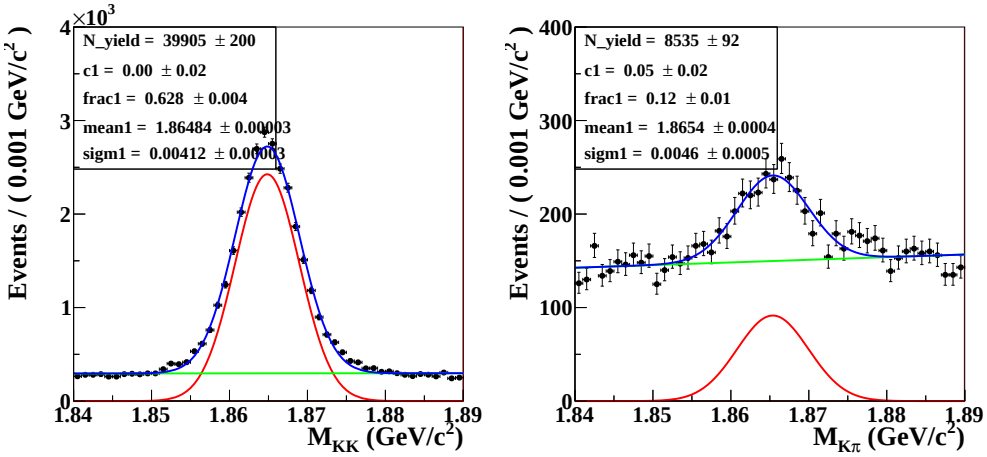


Figure 3.11.: Fitted mass distributions of the KK (left) and $K\pi$ (right) systems in Generic B MC sample.

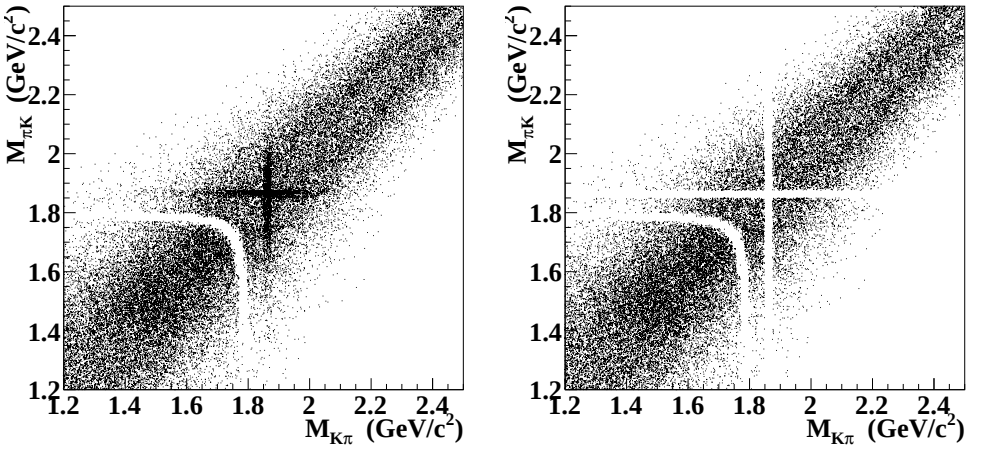


Figure 3.12.: The $M_{K^+\pi^-}$ and $M_{\pi^+K^-}$ distributions in the K^+K^- system before (left) and after (right) the charm veto in Generic B MC sample.

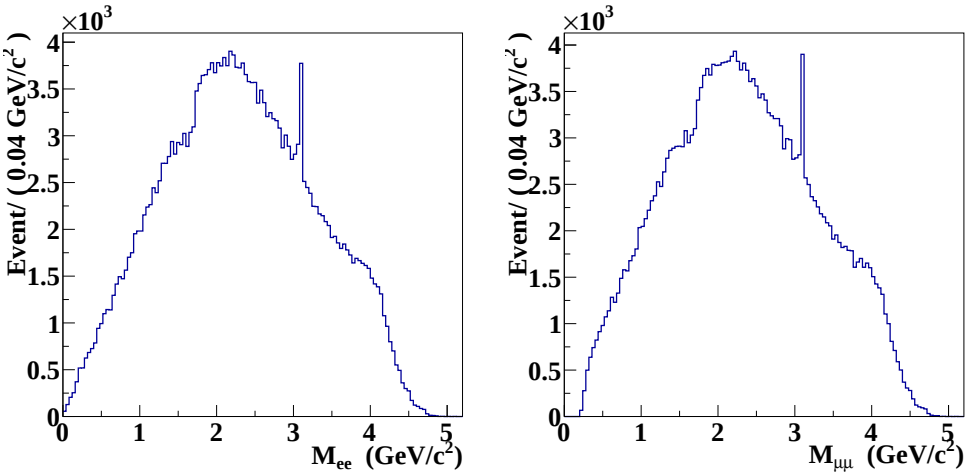


Figure 3.13.: The M_{ee} and $M_{\mu\mu}$ distributions in the $K^{\pm}\pi^{\mp}$ system in Generic B MC sample.

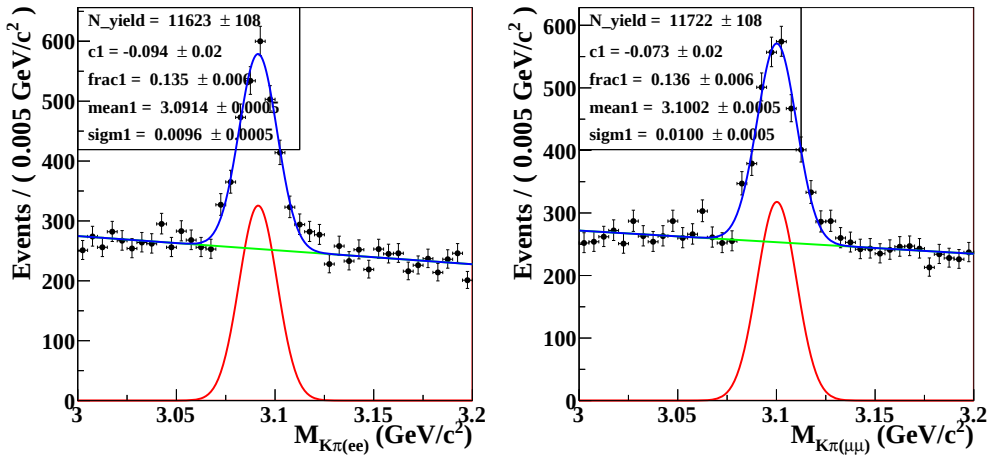


Figure 3.14.: Fitted mass distributions of the $K\pi$ system in Generic B MC sample. The left plot is assuming that K and π are electrons, and the right plot is assuming that they are muons.

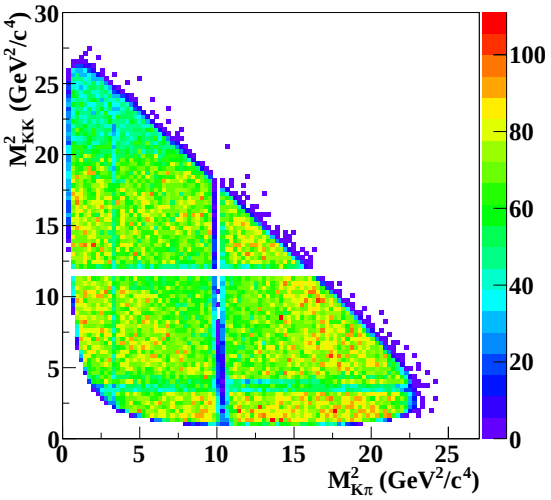
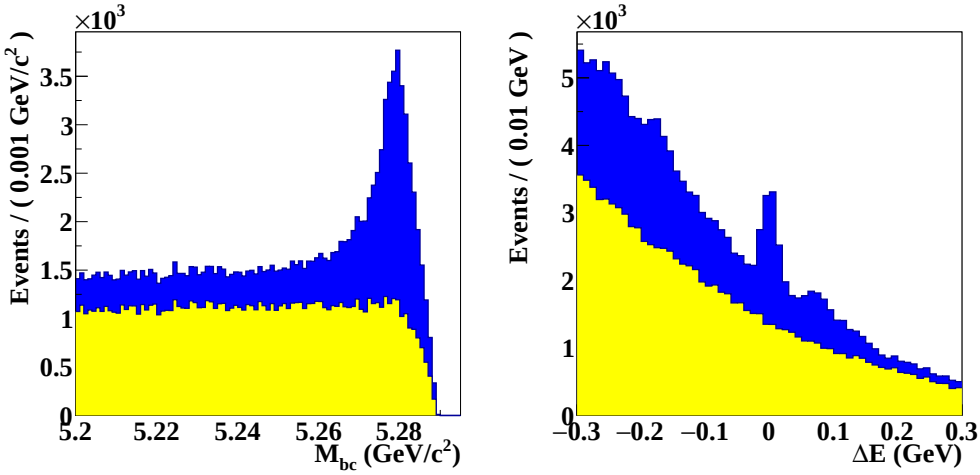
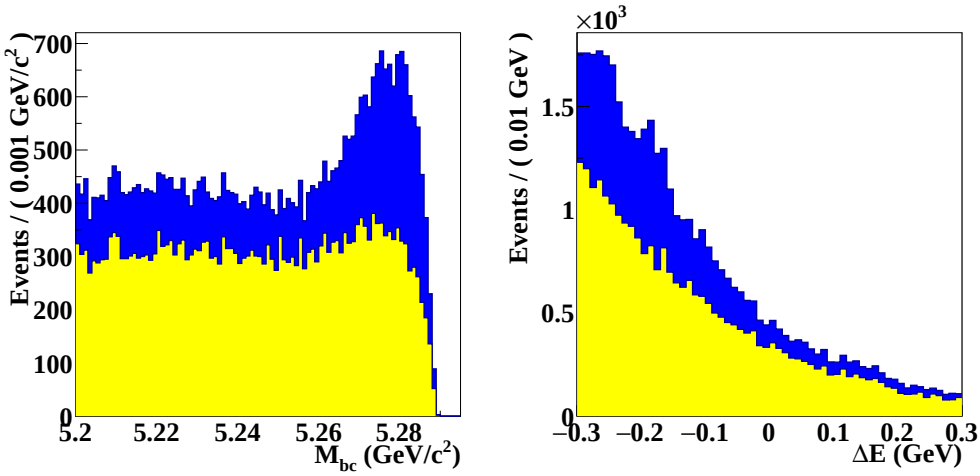


Figure 3.15.: The Dalitz plot [55] in the signal MC after the charm veto was applied.

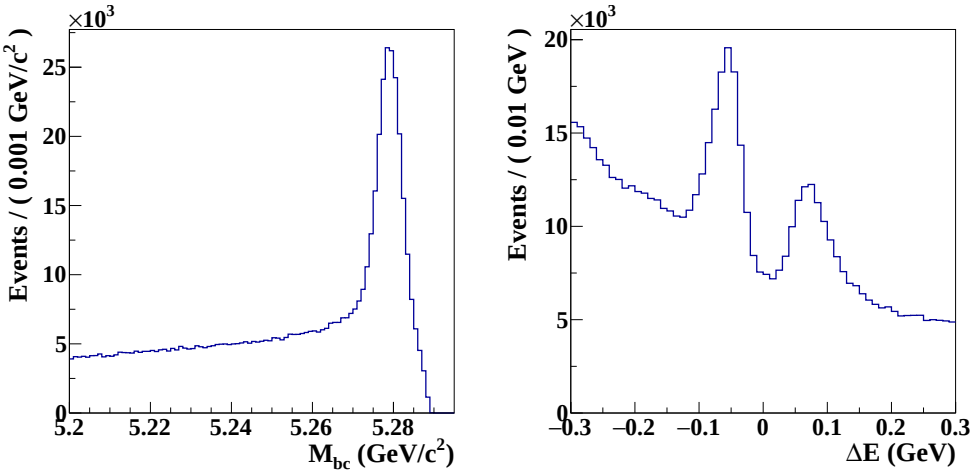


(a) Generic $B^{\pm}$

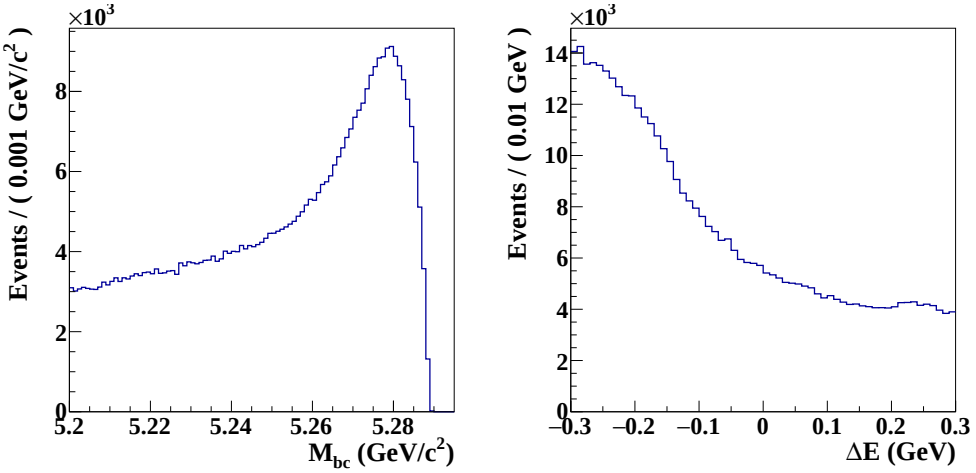


(b) Generic B^0

Figure 3.16.: The ΔE and M_{bc} distributions in the Generic B MC sample before and after (blue and yellow respectively) applying the charm veto. The top plots are for $B^{\pm}$ events, and the bottom plot are for B^0 events.

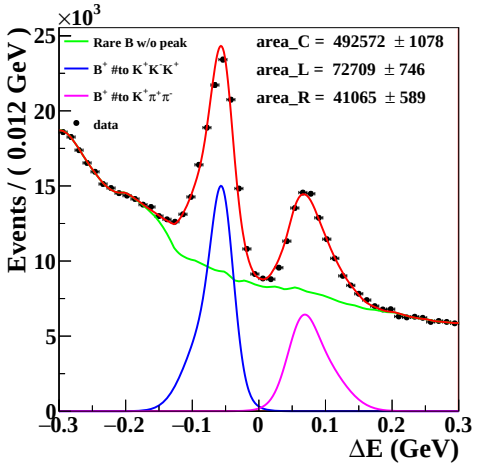


(a) Rare $B^{\pm}$

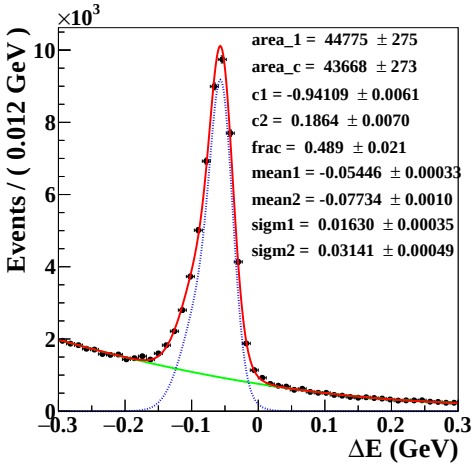


(b) Rare B^0

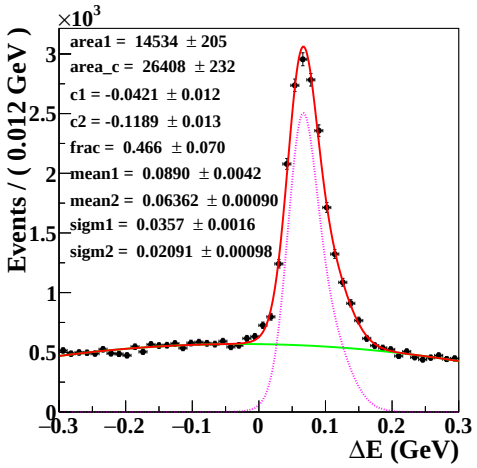
Figure 3.17.: The ΔE and M_{bc} distributions of the Rare B MC. The top plots are for the charged B events and the bottom two are for the neutral B events. These distributions are from a sample 50 times larger than the expected events.



(a) Rare B MC



(b) $B^+ \to K^+K^-K^+$ MC



(c) $B^+ \to K^+\pi^+\pi^-$ MC

Figure 3.18.: Fitted ΔE distributions of the Rare B MC (a), $B^+ \to K^+K^-K^+$ (b), and $B^+ \to K^+\pi^+\pi^-$ (c).

3.3. Signal Extraction

This section describes the fitting strategy that is used to extract the signal yield and *CP* asymmetry, and the validation of the fitting method with ensemble tests. First we perform a fit with the inclusive $M_{\rm KK}$ sample to obtain the number of signal events. Then, to investigate the enhancement on branching fraction and *CP* asymmetry in the low $M_{\rm KK}$ region, we determine the signal yield and *CP* asymmetry in bins of $M_{\rm KK}$. The inclusive branching fraction and direct *CP* asymmetry are obtained from the average of the bin-by-bin results.

3.3.1. Fitting Strategy

To determine the signal yield and evaluate the *CP* asymmetry ($\mathcal{A}_{CP}$), we perform an extended unbinned maximum likelihood fit to the $M_{\rm bc}$ and ΔE distributions with the RooFit package [56]. The *CP* asymmetry is defined to be:

$$\mathcal{A}_{CP} = \frac{N_{B^-} - N_{B^+}}{N_{B^+} + N_{B^-}}, \tag{3.20}$$

where $N_{B^\pm}$ is the number of signal events in $B^\pm$ decays.

The likelihood function and probability density function (PDF) are defined as

$$\mathcal{L} = \frac{e^{-\sum_j n_j}}{N!} \prod_i^N (\sum_j n_j \mathcal{P}_j^i(M_{\rm bc}, \Delta E)) \tag{3.21}$$

and

$$\mathcal{P}_j^i(M_{\rm bc}, \Delta E) = \frac{1}{2}(1 - q_i \cdot \mathcal{A}_{CP}) \times \mathcal{P}_j(M_{\rm bc}, \Delta E), \tag{3.22}$$

where i is the event identifier; n_j is the yield for category j, which corresponds to either signal, Generic B, Rare B, $B^+ \to K^+K^-K^+$, $B^+ \to K^+\pi^+\pi^-$, or continuum background; q_i is the (charge) flavour of the event ($q = \pm 1$ for $B^\pm$) and $\mathcal{P}_j$ denotes the PDF for the j-th category. The PDF construction and tests of the fitter are presented in the following section.

The obtained signal yield is converted to a branching fraction ($\mathcal{B}$) using:

$$\mathcal{B} = \frac{N_{\rm sig}}{N_{B\overline{B}} \times \epsilon_{MC} \times C_{PID}}, \tag{3.23}$$

where $N_{\rm sig}$ is the number of signal obtained from the fit, $N_{B\overline{B}}$ is the number of $B\overline{B}$ pairs in the data set, ϵ_{MC} is the signal efficiency, and C_{PID} is the correction fraction for the difference between the MC and data efficiencies while applying the particle identification (PID) selection, which will be discussed later (See 4.1.2).

To obtain the distributions of differential branching fractions and $\mathcal{A}_{CP}$ in $M_{\rm KK}$, we divide the data sample into five bins according to $M_{\rm KK}$, and apply the same method as in the overall fit in each bin. The range of each $M_{\rm KK}$ bin is shown in Table 3.6.

Table 3.6.: Indice and bin sizes in $M_{\rm KK}$ for the bin-by-bin fit. (Unit: GeV/c^2)

Index	$M_{\rm KK}$ range	Bin size
0	$0.8 < M_{\rm KK} < 1.1$	0.3
1	$1.1 < M_{\rm KK} < 1.5$	0.4
2	$2.5 < M_{\rm KK} < 3.5$	1.0
3	$2.5 < M_{\rm KK} < 3.5$	1.0
4	$3.5 < M_{\rm KK} < 5.3$	1.8

3.3.2. Probability Density Functions

Overall $M_{\rm KK}$ Fit

To model the PDF for true signal events, we use a double Gaussian to represent the $M_{\rm bc}$ distribution, and a triple Gaussian for the ΔE distribution. All the non-yield parameters of the signal PDF are fixed in the fit, including the fractions between the Gaussian functions. The parameters of these PDFs will be modified to account for the potential discrepancies between data and MC by using the control samples. For the self-cross feed background, because of the correlation between the two dimensions of $M_{\rm bc}$ and ΔE shown in Fig. 3.19, we use a smoothed 2D histogram generated using RooFit for $M_{\rm bc}$ and ΔE. The

ratio between signal and self-cross feed will be fixed when fitting. Figure 3.20 shows the fitting results for true signal events and self-cross feed background. Also shown in figures 3.20-3.24 are the residual-pulls, r_i, between the projected PDF and data.

$$r_i = \frac{PDF_i - x_i}{\sigma_i} \tag{3.24}$$

where PDF_i is the fitted PDF value for data point i, x_i is the value of data point i, and σ_i is the standard deviation of the data point i.

For continuum background PDF modeling, we choose the parameterization first used by the ARGUS collaboration [57], $f(M_{\text{bc}}) \propto M_{\text{bc}}\sqrt{1-x^2}e^{\xi(1-x^2)}$, for the M_{bc} distribution, where x is given by $M_{\text{bc}}/E_{\text{beam}}$ and ξ is a fit parameter. For the ΔE distribution, we use a second order Chebyshev polynomial whose coefficients are fixed when fitting. The fit results with continuum background are shown in Fig. 3.21.

Because of the correlation between the two dimensions of M_{bc} and ΔE in other backgrounds as shown in Fig. 3.19, they are modeled separately by a smoothed 2D histogram generated by RooFit. The correlation factor, returned by ROOT [58], is presented in Table 3.7 The fit results of Generic B, Rare B, $B^+ \to K^+K^-K^+$, and $B^+ \to K^+\pi^+\pi^-$ backgrounds are shown in Fig. 3.22, Fig. 3.23, and Fig. 3.24.

Bin-by-bin Fit

Since the $B^+ \to K^+K^-K^+$ and $B^+ \to K^+\pi^+\pi^-$ backgrounds have resonant states in the M_{KK} spectrum and the branching fractions are not reliable in the Rare B decay table, we float the yields for these two components in the bin-by-bin fit. The expected yields of these two components are estimated from the Rare B MC sample when we test the bin-by-bin fitting method. Figures 3.25-3.27 show the ΔE and M_{bc} distributions for signal, SCF background, and continuum background in different M_{KK} bins. As the PDFs for signal and continuum events for the whole region match the distributions in each bin well, we use the PDFs in the overall fit for the bin-by-bin fit. The ΔE and M_{bc} distributions for all the $B\overline{B}$ backgrounds are highly correlated with M_{KK}, so different histogram PDFs are

generated in different M_{KK} bins. Because the self-cross feed events are around 3% of signal events, there are not enough statistics to build a PDF in each M_{KK} bin. Thus we use the same self-cross feed PDF in the overall fit as described above. A summary of PDF modeling is shown in Table 3.8.

Table 3.7.: Correlation between M_{bc} and ΔE for different samples.

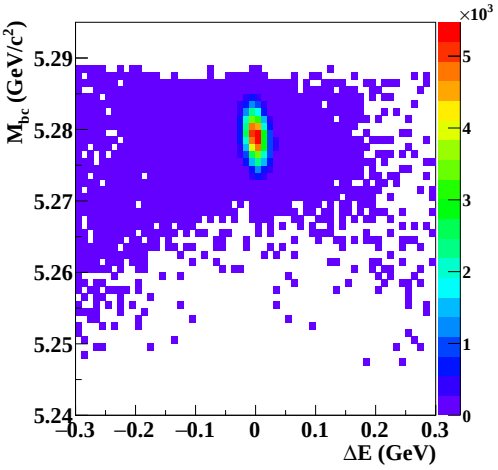
Type	Correlation Factor
True signal	0.8%
SCF	12.6%
Continuum MC	-
Off-res data	−1.5%
Generic B	−13.7%
Rare B	−15.2%
$B^+ \to K^+K^-K^+$	12.2%
$B^+ \to K^+\pi^+\pi^-$	24.5%

Table 3.8.: Summary of PDFs in the both overall and bin-by-bin fit

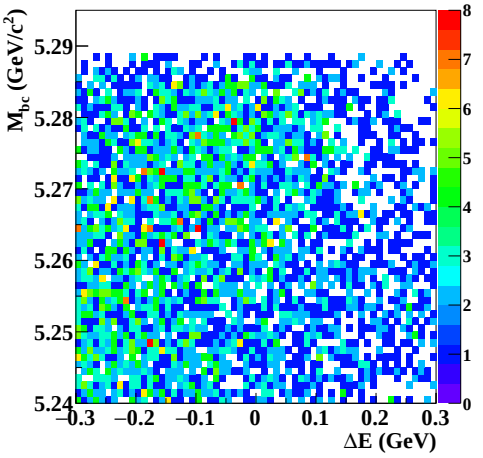
PDF	ΔE	M_{bc}
True signal	Triple Gaussian	Double Gaussian
SCF	2D smoothed histogram	
Continuum	2nd order Cheby. poly.	ARGUS
Generic B	2D smoothed histogram	
Rare B^0	2D smoothed histogram	
Rare $B^\pm$	2D smoothed histogram	
$B^+ \to K^+K^-K^+$	2D smoothed histogram	
$B^+ \to K^+\pi^+\pi^-$	2D smoothed histogram	

3.3.3. Ensemble Test for Fitting Method

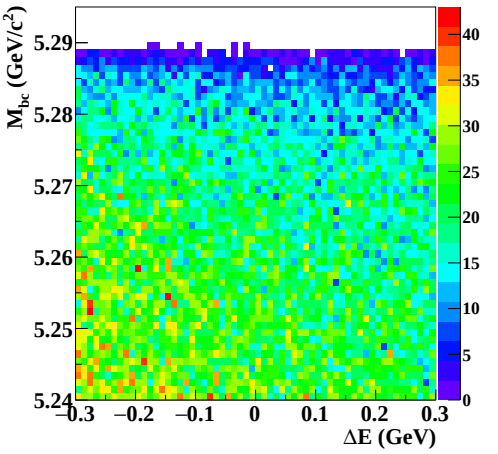
An ensemble test is performed to verify the unbiasedness and stability of the fitting method. Sets of one thousand toy MC samples are generated according to the expected numbers of events and PDF models for each component as



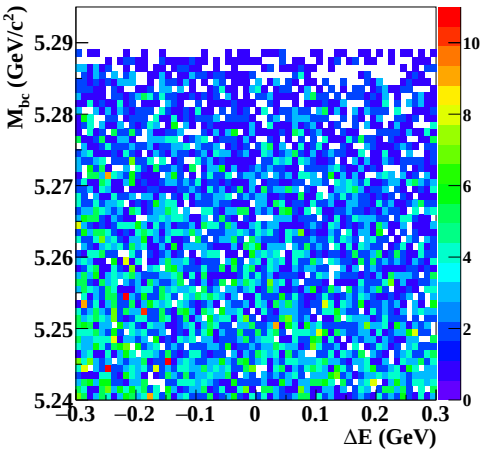
(a) True signal



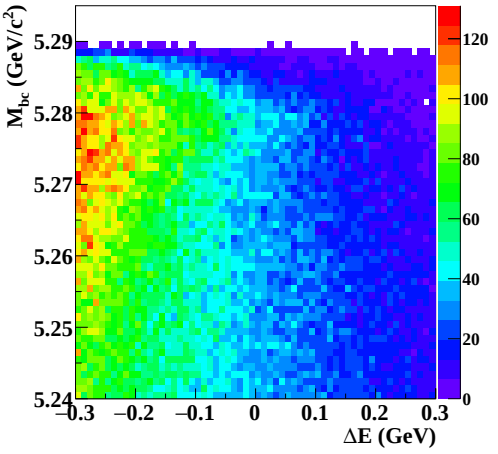
(b) Self cross feed



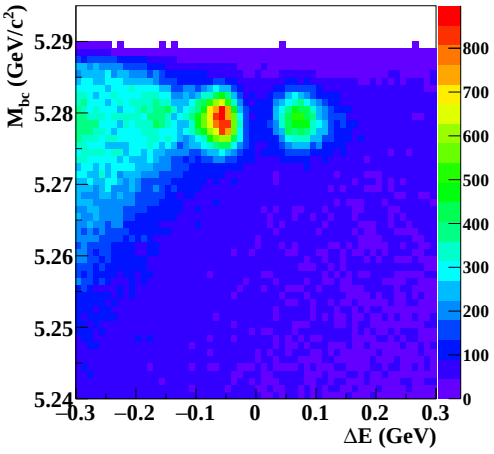
(c) Continuum MC



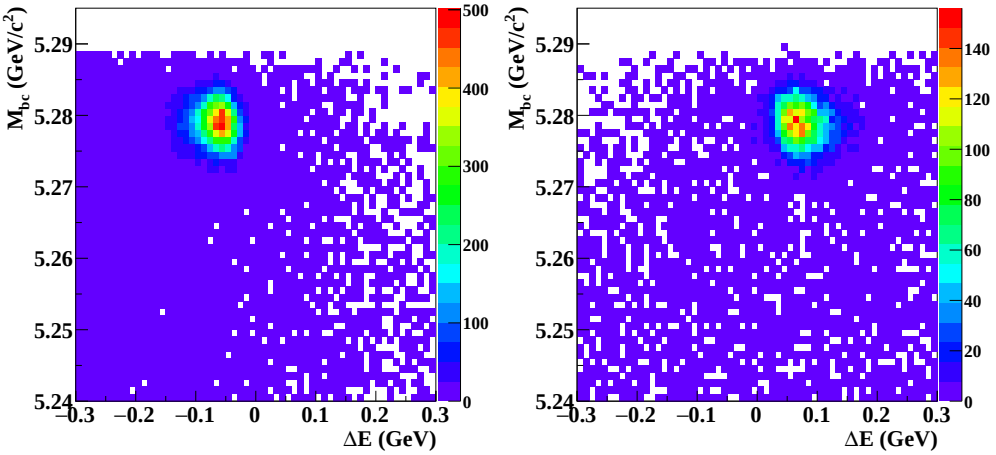
(d) Off-resonance data



(e) Generic B



(f) Rare B



(g) $B^+ \to K^+K^-K^+$ (h) $B^+ \to K^+\pi^+\pi^-$

Figure 3.19.: The two-dimentional scatter plots of $M_{\rm bc}$ and ΔE for all the fitting components. The plots are, from left to right, top to bottom, signal, SCF, continuum MC, off-resonance data, Generic B, Rare B, $B^+ \to K^+K^-K^+$, and $B^+ \to K^+\pi^+\pi^-$ events.

shown in Table 3.5. The result of the ensemble tests includes yield, fitting error, and pull value. The pull value is defined as

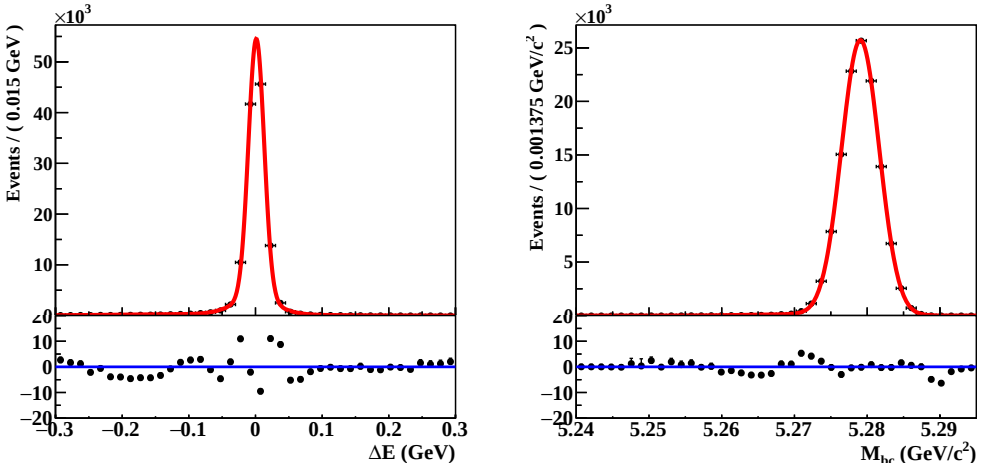
$$\text{Pull} = \frac{N_{\rm fit} - N_{\rm expect}}{\sigma_{\rm fit}}, \tag{3.25}$$

where $N_{\rm fit}$ and $\sigma_{\rm fit}$ are the signal yields and the errors obtained in each trial respectively, and $N_{\rm expect}$ is the expected number of signal events. If there is no bias in the test, the pull distribution should be a gaussian distribution with mean of 0 and standard deviation of 1.

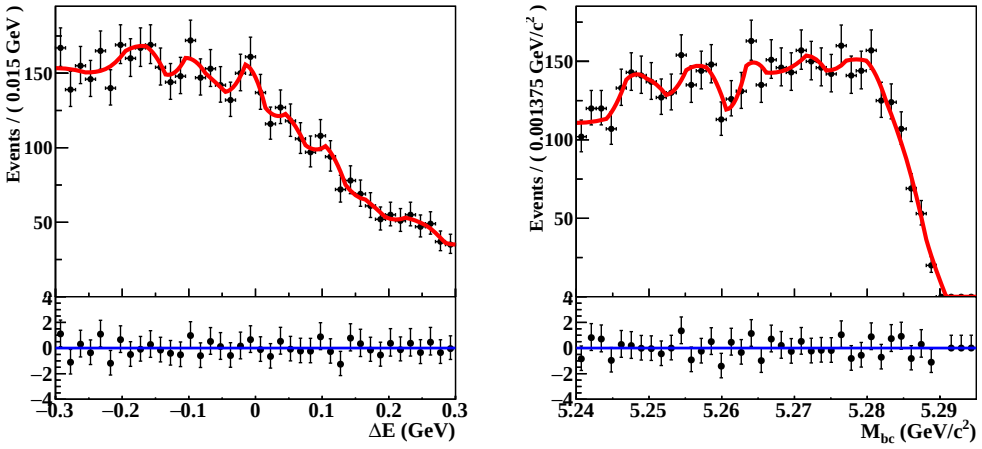
Firstly, the ensemble test with toy MC samples was performed to validate the stability of the fitting model. Later, another test with the signal GSIM sample was performed to check for possible bias raised by the small correlation between two kinematic variables which our signal model does not take into account.

Overall $M_{\rm KK}$ Fit

When fitting a sample $N_{\rm sig}$, $N_{\rm gen}$, and $N_{\rm qq}$ are floated, while the rest of the fitting variables are fixed. The ensemble test result and an example fit with the

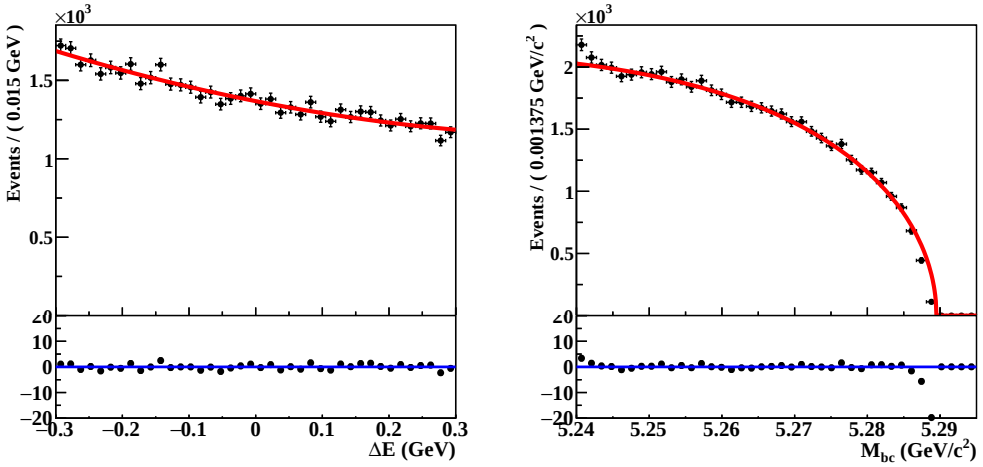


(a) true signal

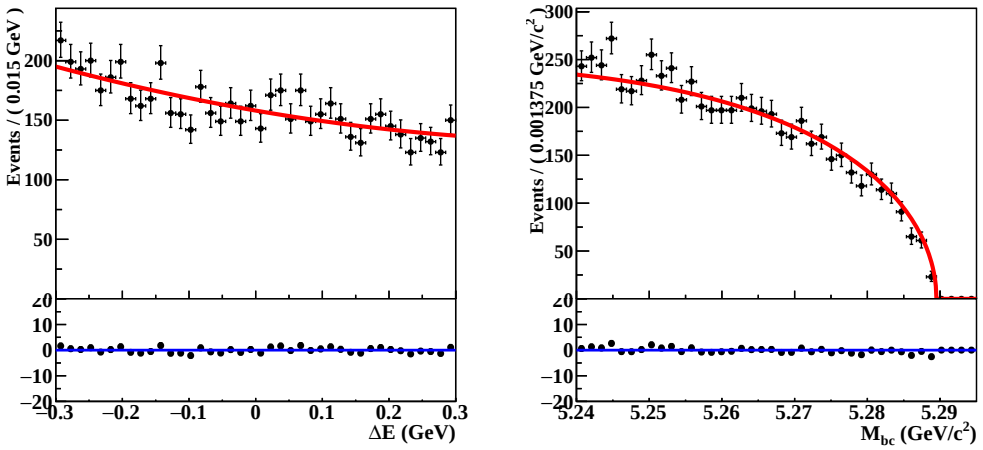


(b) self-cross-feed

Figure 3.20.: The projection plots of fitting results of true signal (a) and self-cross-feed (b) events from the signal MC sample. ΔE and $M_{\rm bc}$ of true signal events are modeled by triple Gaussian and double Gaussian, respectively. For the self-cross feed background, the ΔE and $M_{\rm bc}$ are modeled by a 2D smoothed histogram function. The projection regions in the same as the fitting region defined in Sec. 3.1.4. Also plotted beneath the histograms are the residual-pulls of the fit to the data.



(a) continuum MC



(b) off-resonance data

Figure 3.21.: The projection plots of fitting results of the continuum MC sample (a). ΔE and $M_{\rm bc}$ are modeled by 2nd order Chebyshev polynomial and ARGUS function, respectively. (b) shows the comparison of off-resonance data and PDFs obtained from continuum MC. The projection regions in the same as the fitting region defined in Sec. 3.1.4. Also plotted beneath the histograms are the residual-pulls of the fit to the data.

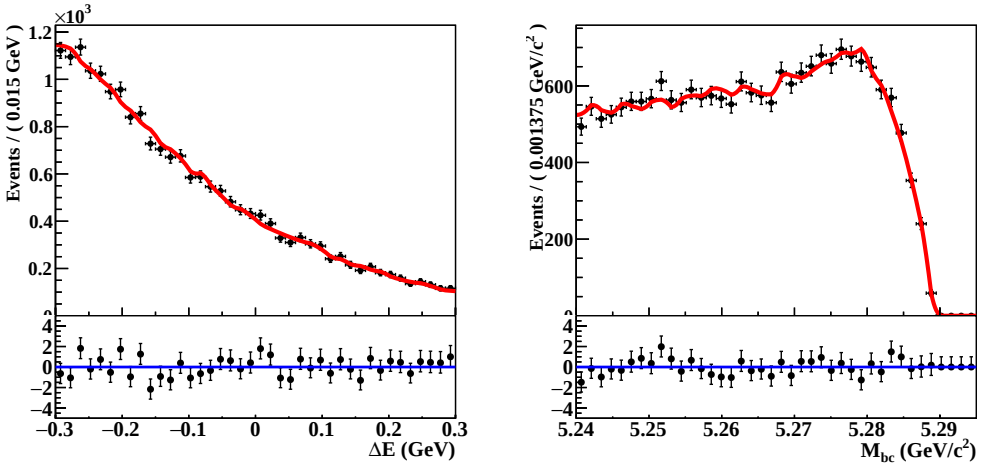
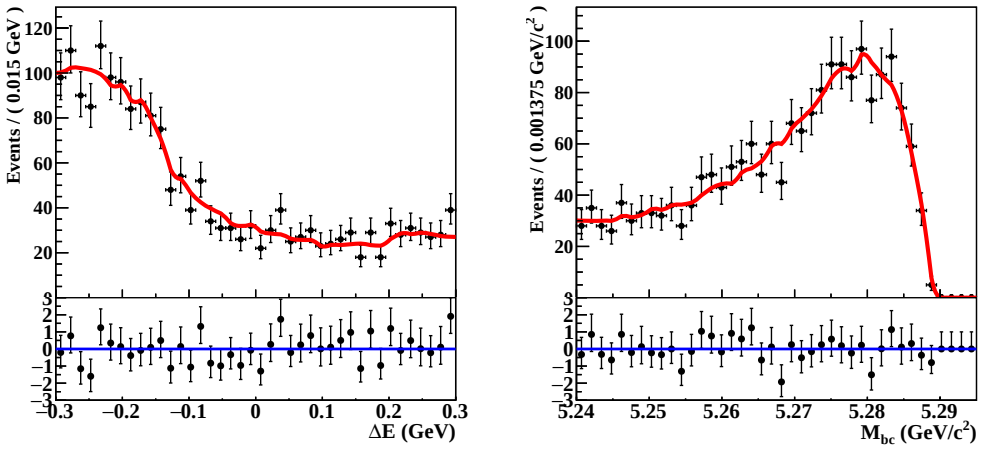


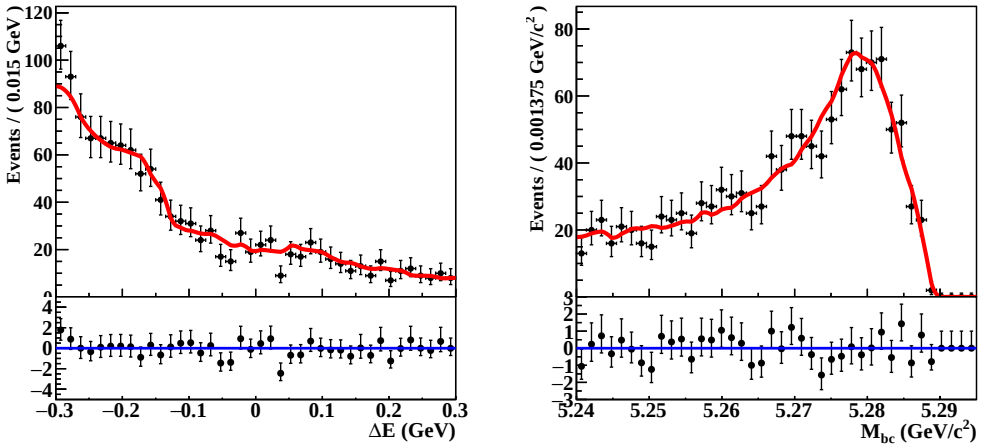
Figure 3.22.: The projection plots in ΔE and $M_{\rm bc}$ showing the 2D smoothed histogram fit for Generic B MC background. The data points are chosen from a subset of the total sum of events in the MC samples (to provide an independent comparison). Also plotted beneath the histograms are the residual-pulls of the fit to the data.

toy MC samples are shown in Fig. 3.28. We perform a linearity test where the ensemble tests are carried out with an assumed signal yield ranging from 14 to 914, and the result is presented in Fig. 3.29. Ideally the measured values should match the corresponding input values, so a first order polynomial should have a y-intercept of 0 and a gradient of 1. Any significant deviation from this indicates a bias. We also perform another ensemble test with the signal events from the GSIM signal MC sample. For each trial, events from signal MC are randomly sampled and employed from the MC simulation. Figure 3.30 shows the result of the GSIM ensemble test and a fitting result of one of the trials. A linearity test is performed with the GSIM ensemble tests as well, and the result is shown in Fig. 3.31. A bias of 2.3% is found in the signal yield in the GSIM test as the gradient is 2.3% below 1 while no significant y-intercept is found. Given the statistical error is expected to be 6.7%, the bias will be included in the systematic uncertainty (See Sec. 4.1).

To verify that the $\mathcal{A}_{CP}$ obtained from our fitter is correct, we perform the pull test with 1000 sets which are picked from the GSIM samples. By varying $\mathcal{A}_{CP}$ input values, we repeat the ensemble tests with $\mathcal{A}_{CP}$ from -0.5 to 0.5. The $\mathcal{A}_{CP}$ result at LHCb in this decay is -0.123 ± 0.017 ($stat$) $\pm$ 0.012 ($syst$) [20]. No

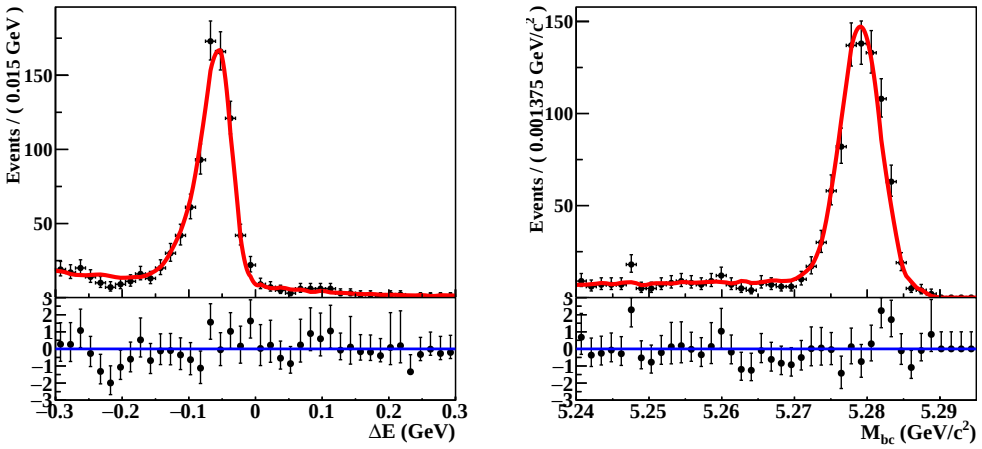


(a) Rare B^0

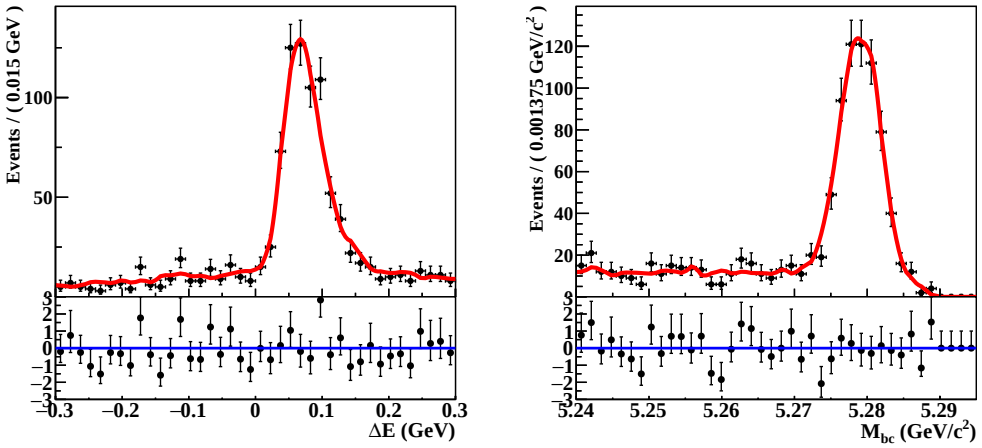


(b) Rare $B^{\pm}$

Figure 3.23.: The projection plots in ΔE and M_{bc} showing the 2D smoothed histogram fit for Rare B^0 (a) and Rare $B^{\pm}$ (b) MC samples. The data points are chosen from a subset of the total sum of events in the MC samples (to provide an independent comparison). Also plotted beneath the histograms are the residual-pulls of the fit to the data.

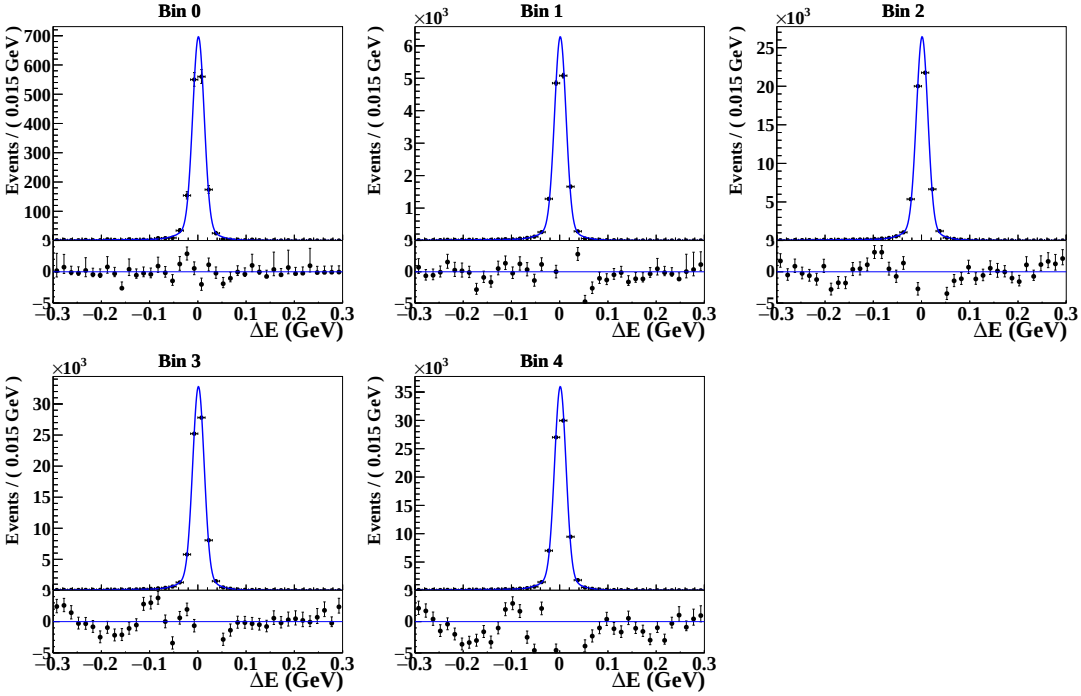


(a) $B^+ \to K^+K^-K^+$

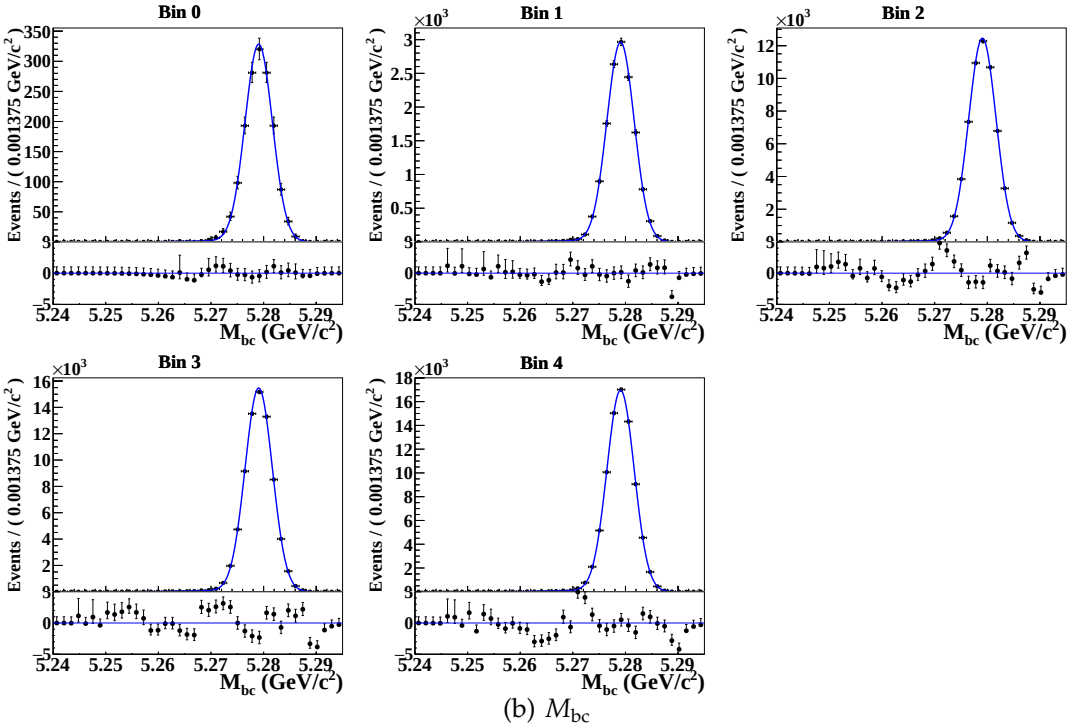


(b) $B^+ \to K^+\pi^+\pi^-$

Figure 3.24.: The projection plots in ΔE and $M_{\rm bc}$ showing the 2D smoothed histogram fit for $B^+ \to K^+K^-K^+$ and $B^+ \to K^+\pi^+\pi^-$ MC backgrounds. The ΔE and $M_{\rm bc}$ are modeled by a 2D smoothed histogram function. The data points are chosen from a subset of the total sum of events in the MC samples (to provide an independent comparison). Also plotted beneath the histograms are the residual-pulls of the fit to the data.

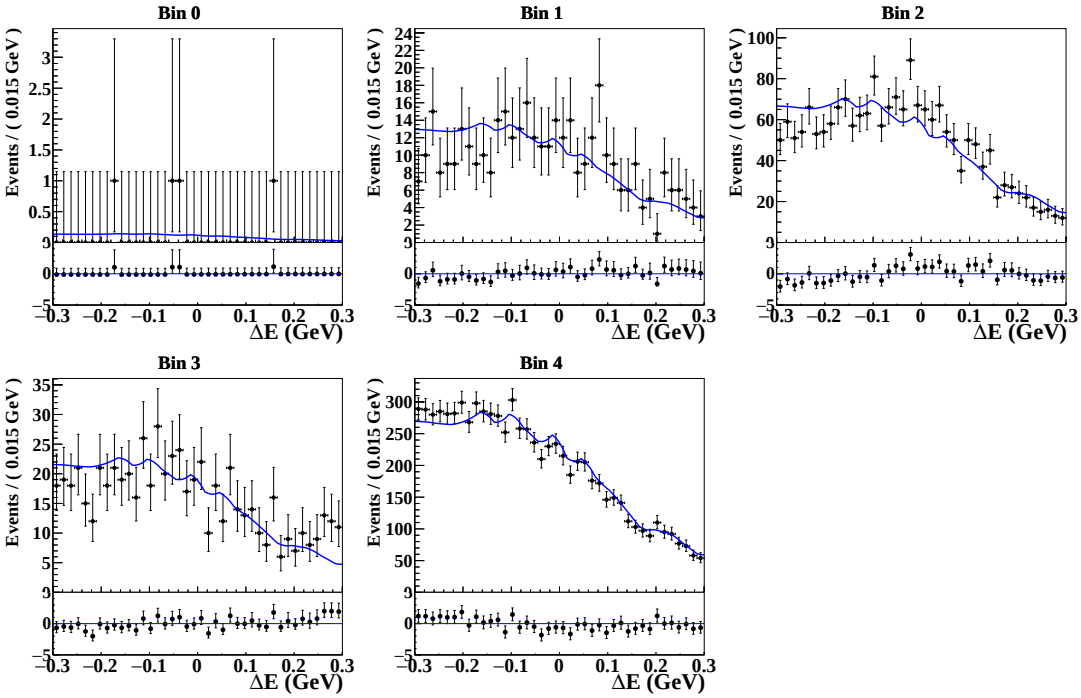


(a) ΔE

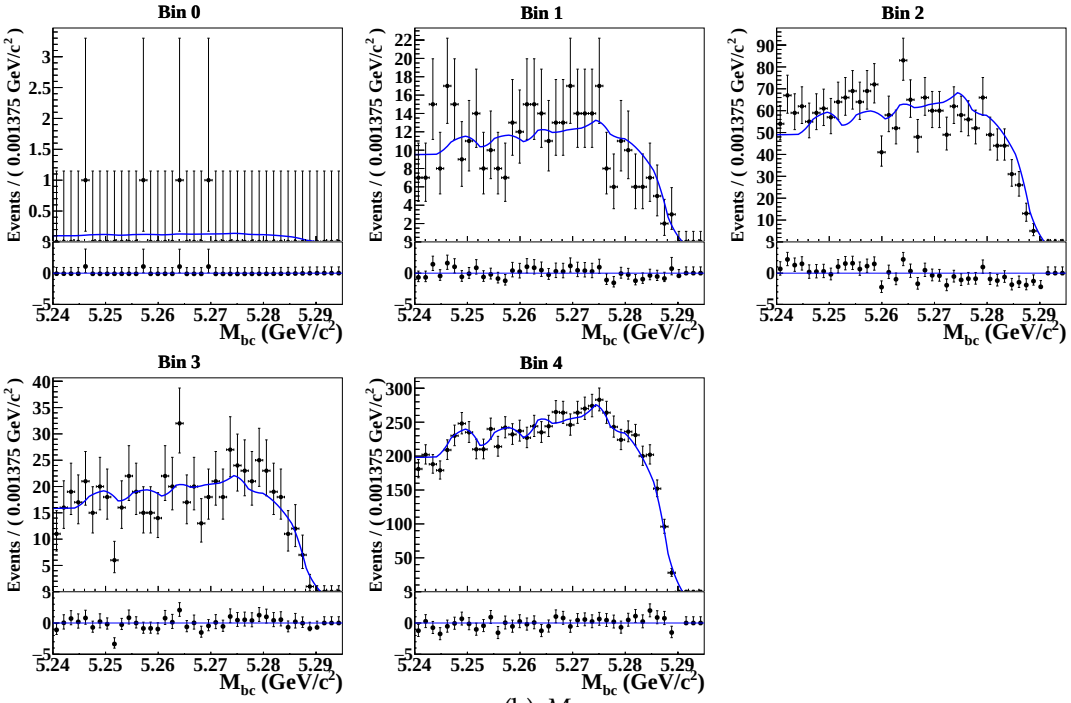


(b) $M_{\rm bc}$

Figure 3.25.: The ΔE (a) and $M_{\rm bc}$ (b) distributions of signal in different $M_{\rm KK}$ bins. The dots are MC sample event distribution, and the blue line indicates the PDF for overall $M_{\rm KK}$ region. Also plotted beneath the histograms are the residual-pulls of the fit to the data.

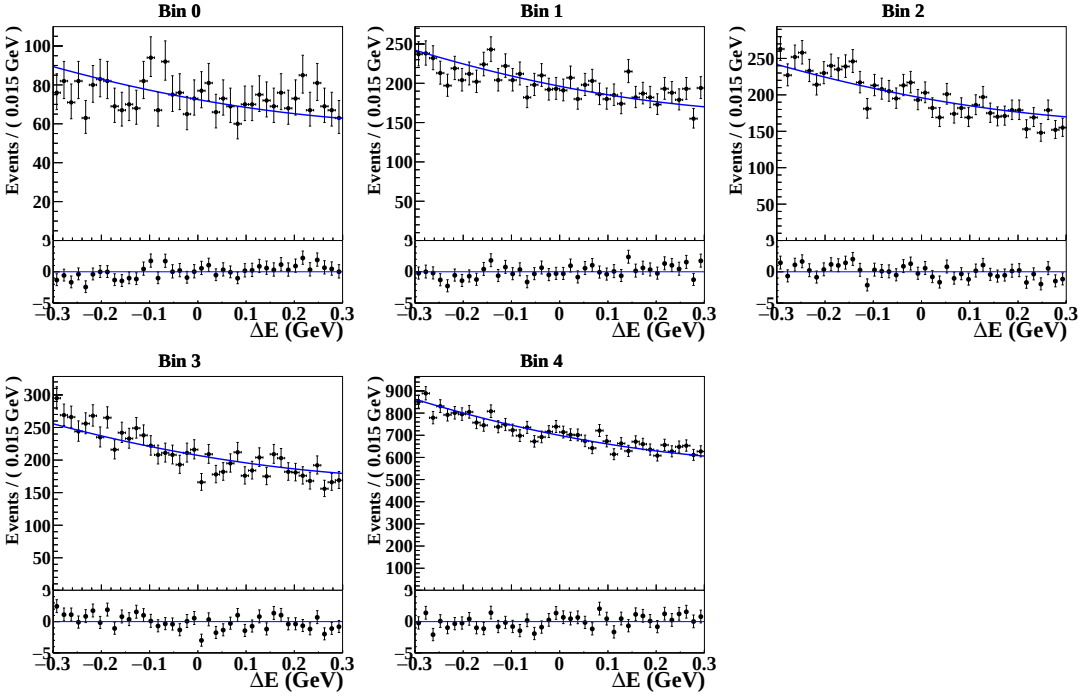


(a) ΔE

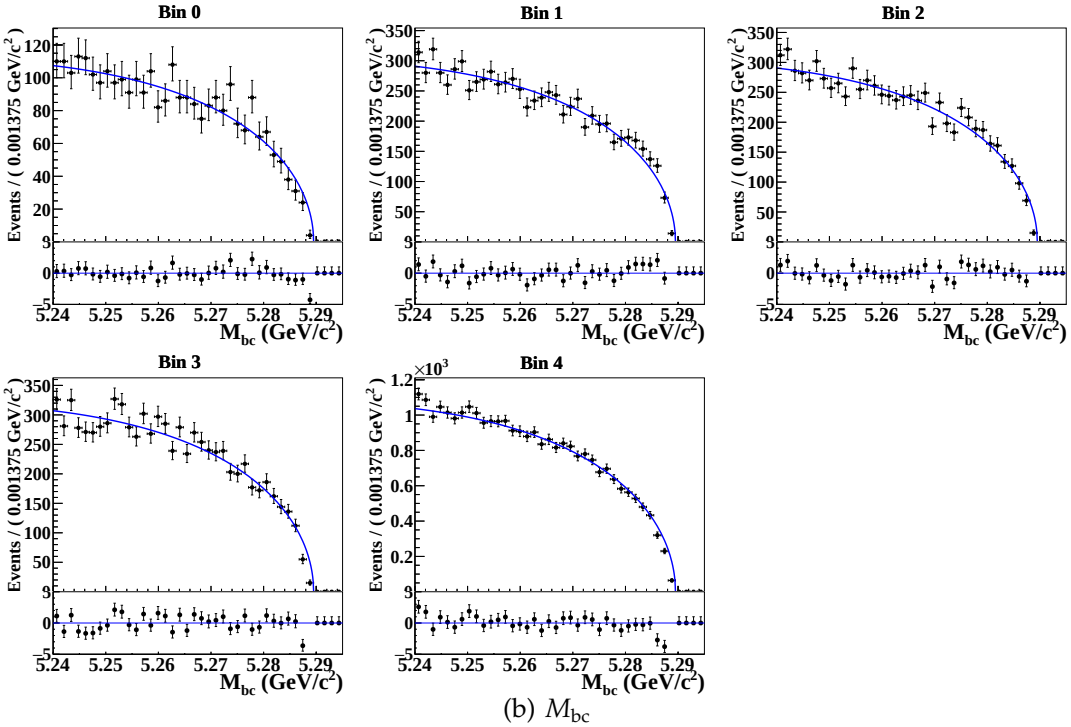


(b) $M_{\rm bc}$

Figure 3.26.: The ΔE (a) and $M_{\rm bc}$ (b) distributions of self-cross feed background in different $M_{\rm KK}$ bins. The dots are MC sample event distribution, and the blue line indicates the PDF for overall $M_{\rm KK}$ region. Also plotted beneath the histograms are the residual-pulls of the fit to the data.



(a) ΔE



(b) $M_{\rm bc}$

Figure 3.27.: The ΔE (a) and $M_{\rm bc}$ (b) distributions of continuum MC background in different $M_{\rm KK}$ bins. The dots are MC sample event distribution, and the blue line indicates the PDF for overall $M_{\rm KK}$ region. Also plotted beneath the histograms are the residual-pulls of the fit to the data.

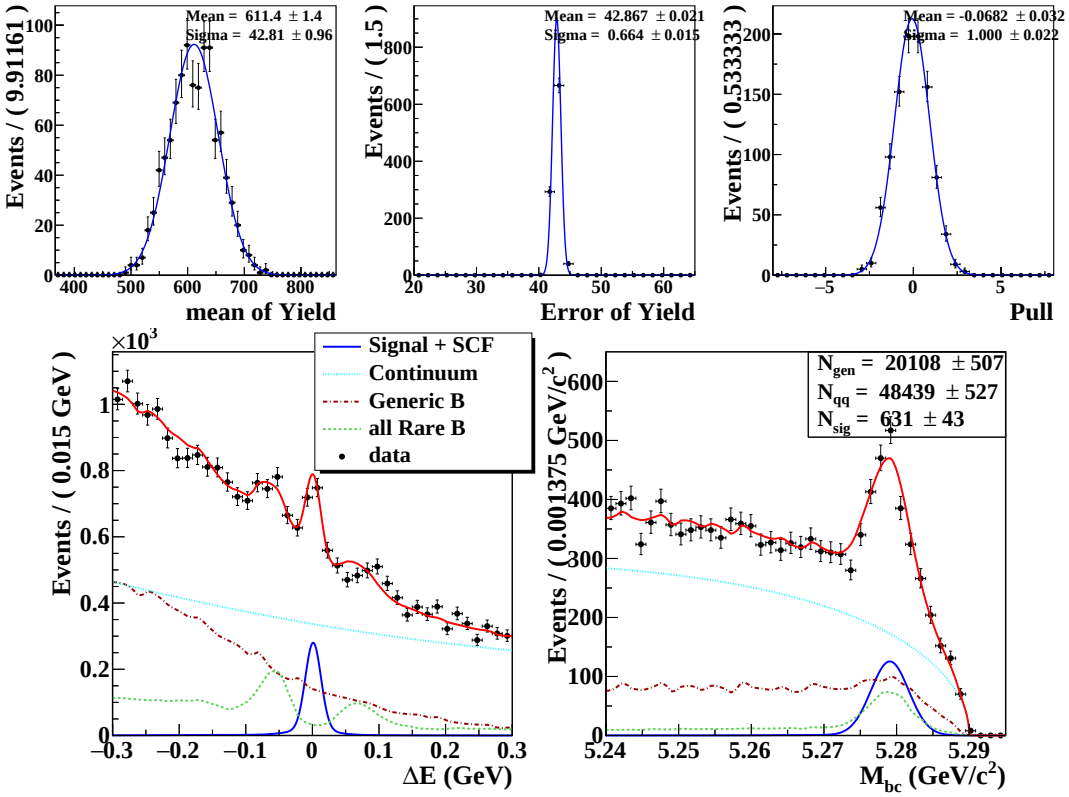


Figure 3.28.: The ensemble test result with expected number of events and toy MC samples. The upper plot shows the distributions of signal yields, fitting error, and pull value. The lower one shows the projection plot of the fitting result from one of the toy MC samples. The projection region is defined to be $5.27 < M_{\rm bc} < 5.29$ GeV/c^2 for ΔE and $-0.05 < \Delta E < 0.05$ GeV for $M_{\rm bc}$.

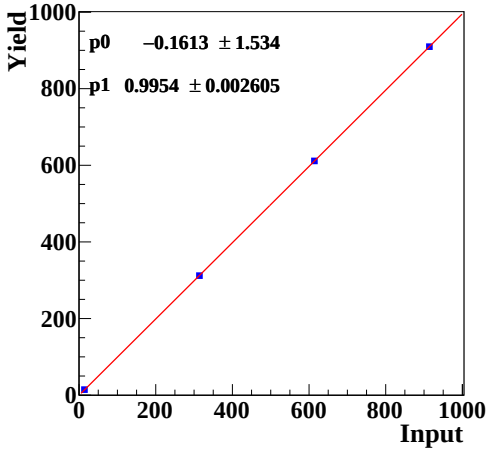


Figure 3.29.: The result of the linearity test for yield with toy MC samples. No bias is found in the ensemble test with toy MC samples.

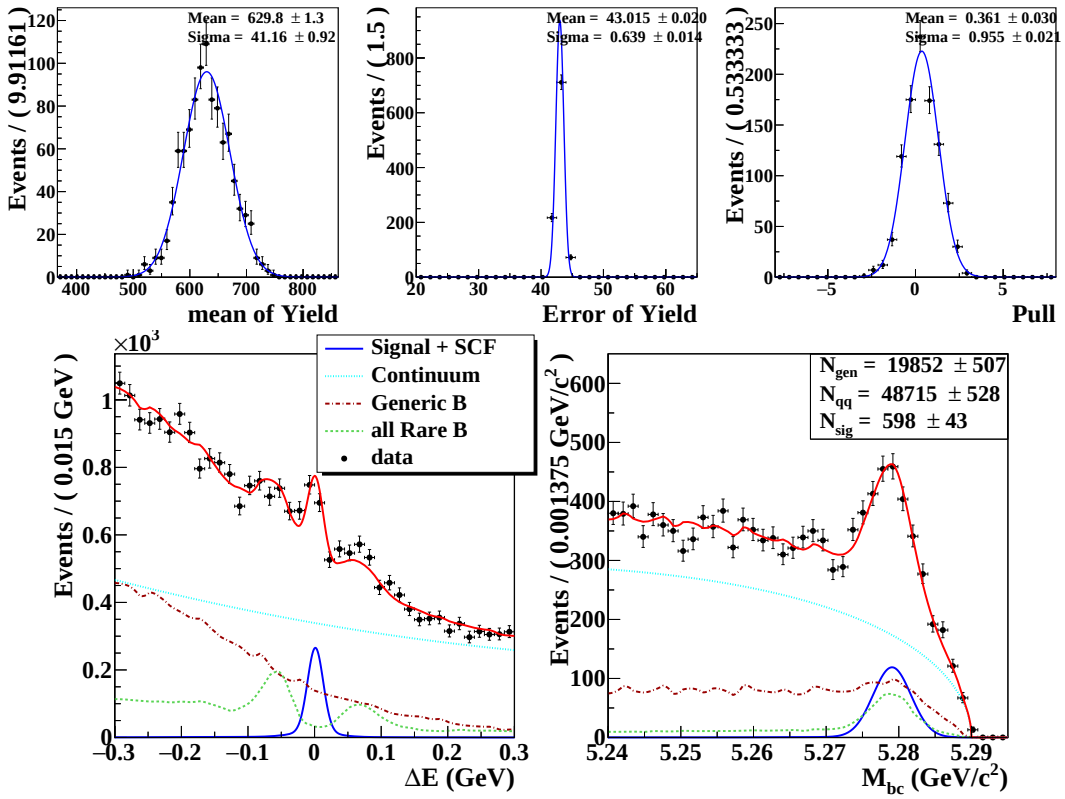


Figure 3.30.: The ensemble test result with expected number of events. The upper plot shows the distributions of signal yields, fitting error, and pull value. The lower one shows the projection plot of the fitting result from one of the trials. The signal and self-cross feed events are picked from GSIM MC samples while the other components are purely toy MC samples. The projection region is defined to be $5.27 < M_{\rm bc} < 5.29$ GeV/c^2 for ΔE and $-0.05 < \Delta E < 0.05$ GeV for $M_{\rm bc}$.

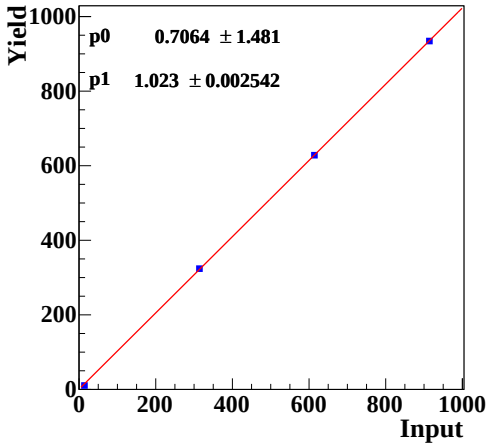


Figure 3.31.: The result of the linearity test for yield with GSIM samples. A bias of 2.3% is found, and will be quoted as a systematic uncertainty.

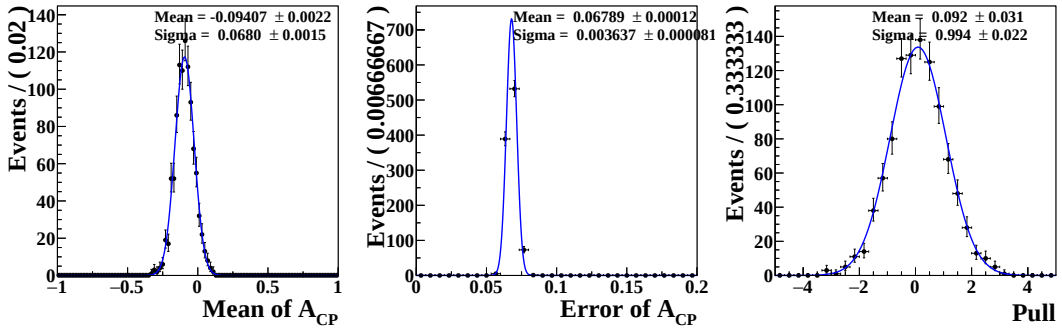
bias is found in the $\mathcal{A}_{CP}$ test, and the results are shown in Fig. 3.32 and Fig. 3.33. The expected error of $\mathcal{A}_{CP}$ is around 0.067.

Bin-by-bin Fit

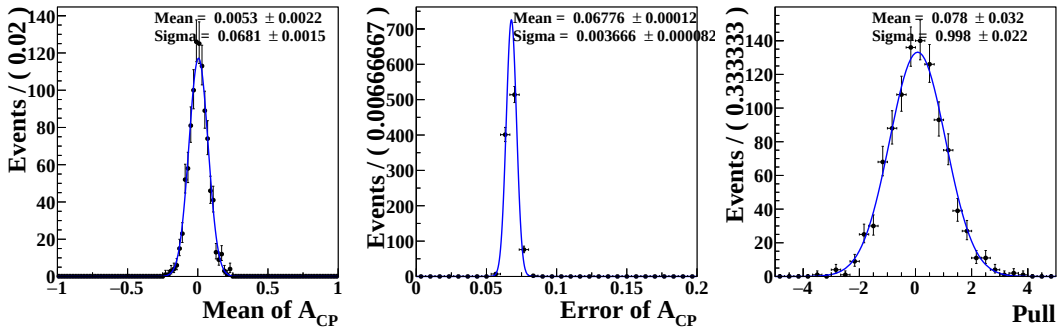
In the bin-by-bin fitting method, we also employ pull tests with toy MC and GSIM MC samples, and the expected number of events in each M_{KK} bin obtained from MC samples. The pull results show no bias on signal yields with toy MC samples (Fig. 3.34) and $\mathcal{A}_{CP}$ (Fig. 3.36), while 1.4 $\sim$ 2% bias is found on signal yields in GSIM tests (see Fig. 3.35) which is consistent with the overall fitting bias. Table 3.9 shows the bias on signal yields, which should be included in the systematic uncertainty of the bin-by-bin results. However, the error on the bias is large. To be more conservative, we will use the value of bias obtained from the overall fitting test (2.3%) as the systematic uncertainty (See Sec. 4.1).

3.3.4. Efficiency variation over M_{KK}

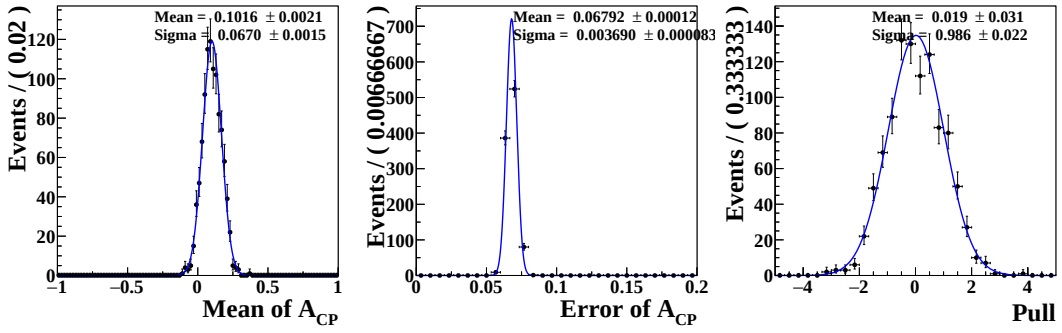
An anomalous excess was found in the low KK invariant mass region by Babar [30], and the LHCb result shows that the excess only occurs in B^+ (Fig. 3.37) [20]. Therefore, to validate both branching fraction and $\mathcal{A}_{CP}$



(a) $\mathcal{A}_{CP} = -0.1$



(b) $\mathcal{A}_{CP} = 0.0$



(c) $\mathcal{A}_{CP} = 0.1$

Figure 3.32.: The results of pull tests with $\mathcal{A}_{CP}$ equal to $-0.1,\ 0,\ 0.1$.

Table 3.9.: Fitter bias in each M_{KK} bin.

Index	Bias on signal yields (%)
0	3.4 ± 2.0
1	2.0 ± 0.5
2	1.8 ± 0.3
3	1.7 ± 0.2
4	1.4 ± 0.2

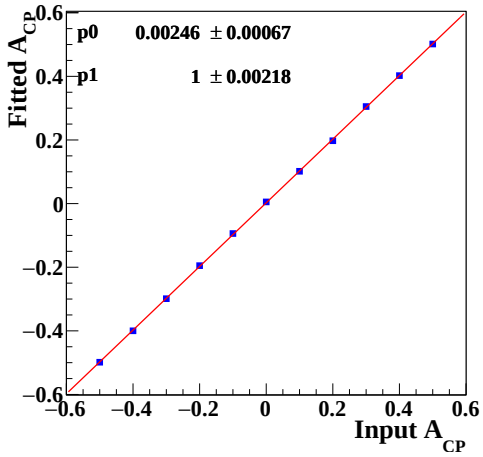


Figure 3.33.: The result of the linearity test for $\mathcal{A}_{CP}$ with GSIM samples. The input $\mathcal{A}_{CP}$ values are set to be -0.5, -0.4, -0.3, -0.2, -0.1, 0, 0.1, 0.2, 0.3, 0.4, 0.5. The red line indicates the linear fit to input and output $\mathcal{A}_{CP}$.

extraction as a function of M_{KK} and the efficiency map, we generate a MC sample to model the excess in M_{KK} by assuming an intermediate state X with a mass of 1.2 GeV/c^2 and width of 0.25 GeV/c^2. The branching fraction of $B^+ \to X(K^+K^-)\pi^+$ is set to be 10% of $B^+ \to K^+K^-\pi^+$. The obtained signal efficiency in each M_{KK} bin is shown in Table 3.10. The result using the MC with a resonance shows that the efficiency is consistent with the signal MC sample (using phase space model), which means that the signal efficiency variation over M_{KK} is model-independent. To further study the possible efficiency variation over M_{KK} due to different angular distributions for K and π, we generate a MC sample for $B^+ \to K^+\overline{K}^*(892)^0$. In this process, a scalar particle, B, decays to a vector particle, $K^*(892)^0$, and a scalar particle, K. This decay will form an non-uniform distribution in the Dalitz plane, as shown in Figure 3.40. Despite the significantly different event population compared to $B^+ \to K^+K^-\pi^+$, the efficiency distribution is similar to that obtained from signal MC sample. Please note that this is an extreme case as the upper limit on the branching fraction of this decay is 1.1×10^{-6} with 90% confidence level [32]. For the following calculation, we use the efficiency obtained from the signal MC sample.

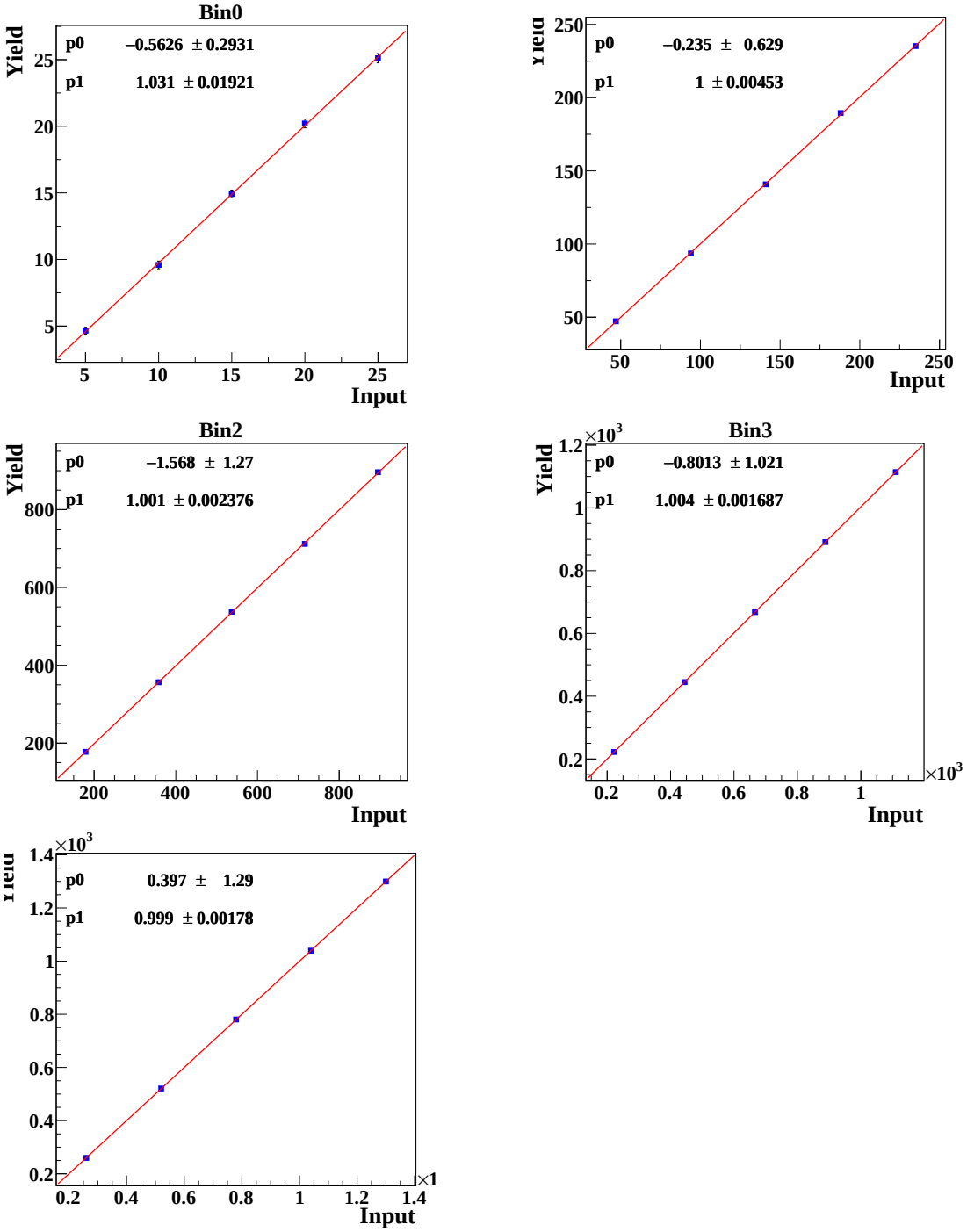


Figure 3.34.: The results of linearity tests for yield with toy MC ensemble in each M_{KK} bin. The result shows no bias is found in the toy MC test.

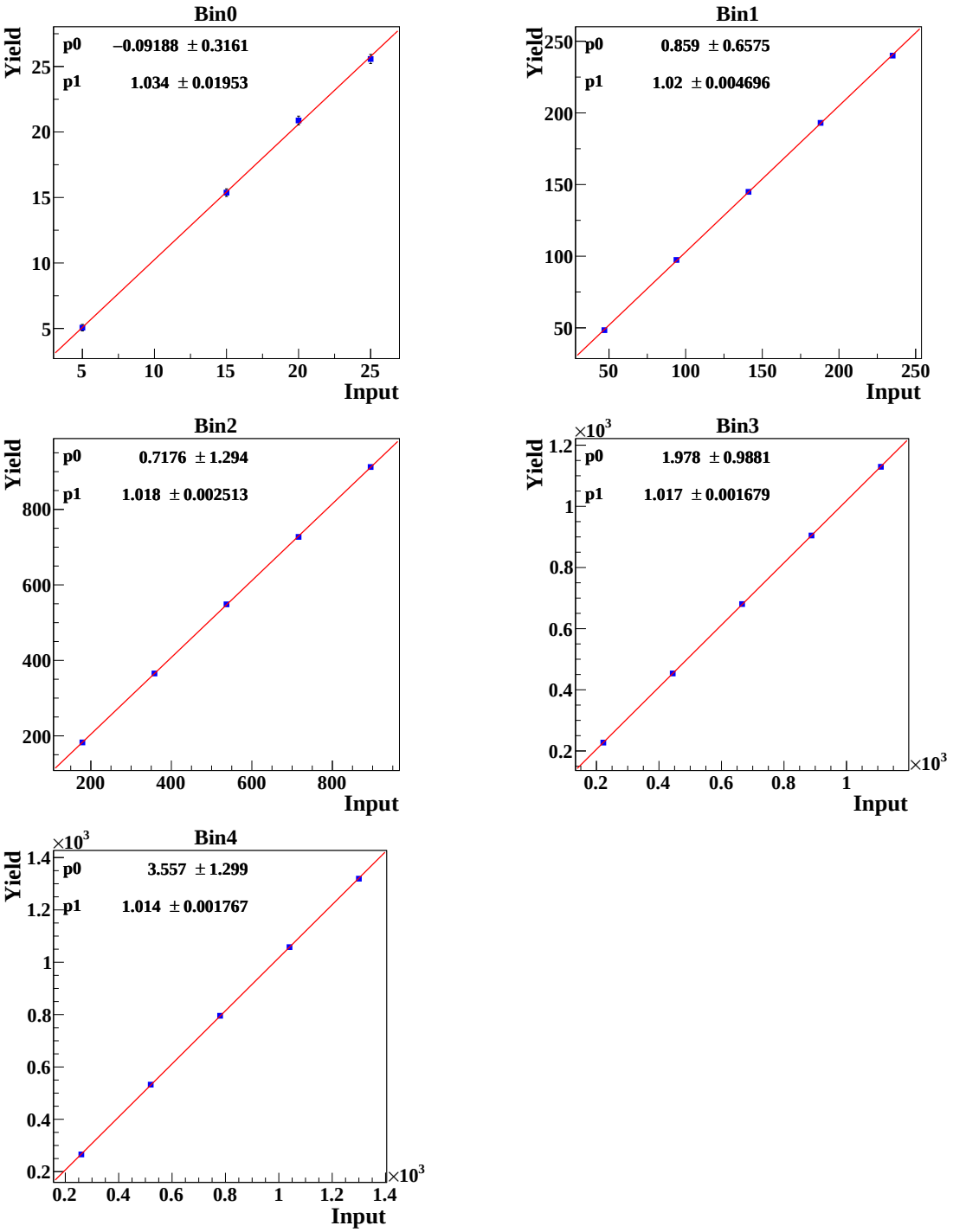


Figure 3.35.: The results of linearity tests for yield with GSIM ensemble in each M_{KK} bin. The bias in each bin is consistent with one found in the overall M_{KK} test.

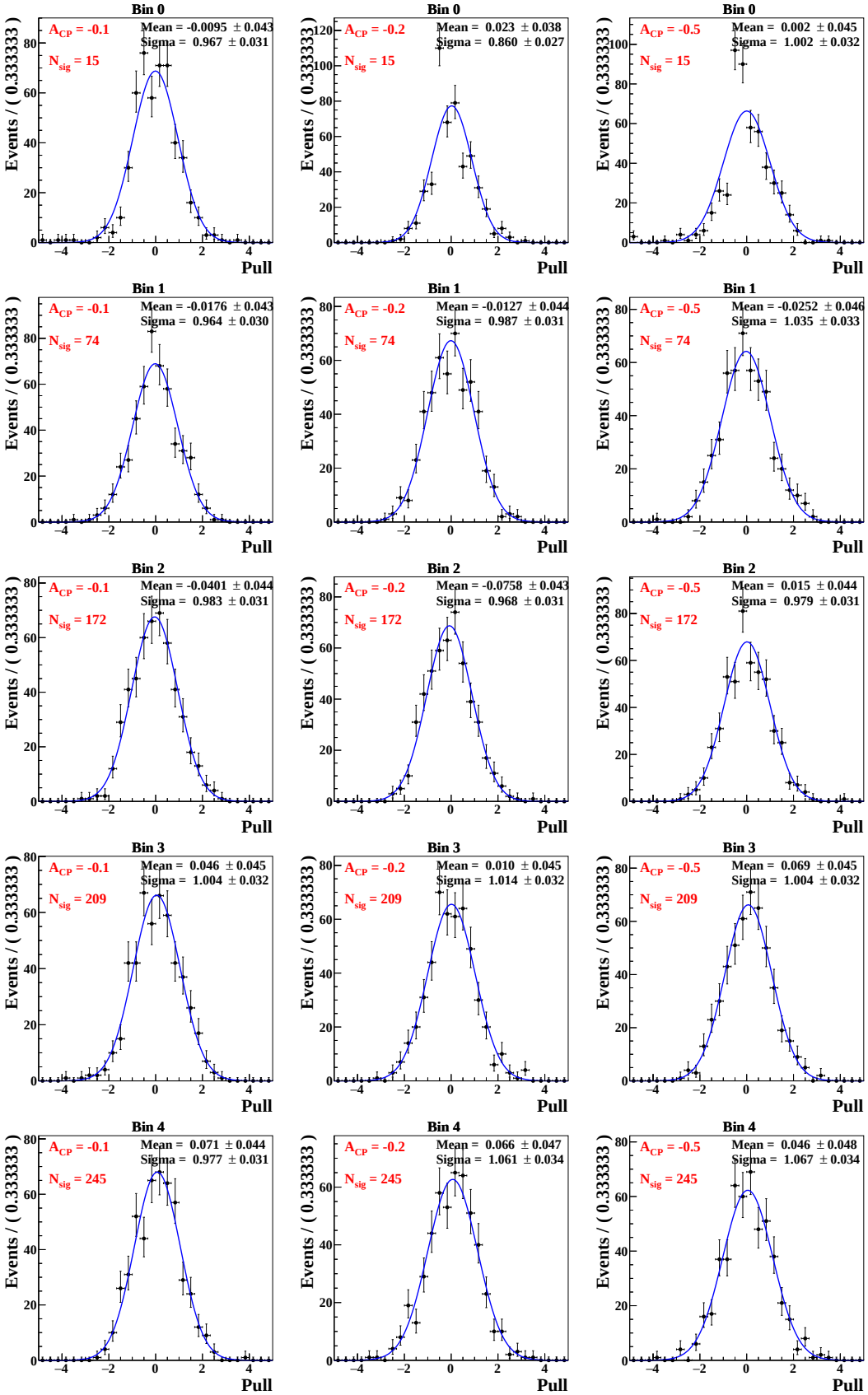


Figure 3.36.: The pull test results on $\mathcal{A}_{CP}$ in the bin-by-bin fit. The top row of plots corresponds to bin0, with subsequent rows of plots corresponding to subsequent bins.

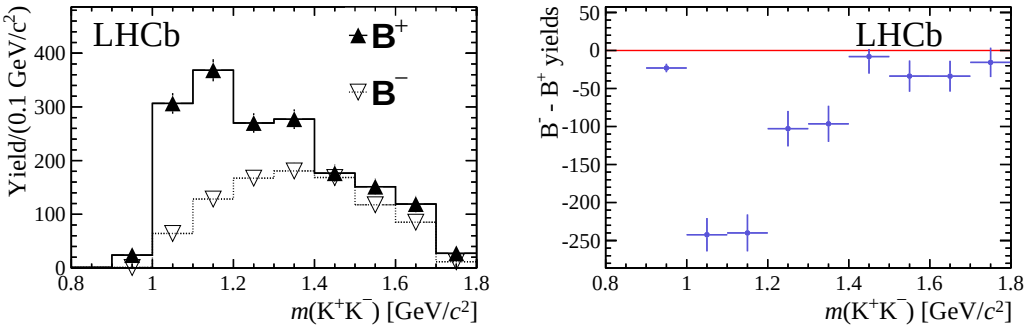


Figure 3.37.: Results from LHCb. Projections in bins of the M_{KK} variable of (a) the number of B^+ and B^- signal events and (b) their difference, for $B^+ \to K^+K^-\pi^+$ decays. The yields are acceptance-corrected and background-subtracted. A guide line for zero (horizontal red line) was included on plot (b). [20]

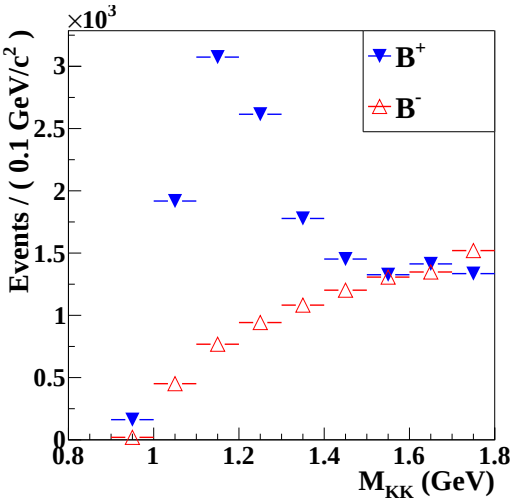


Figure 3.38.: M_{KK} distribution in signal MC with 10% $B^+ \to X(K^+K^-)\pi^+$. To compare with the LHCb result, only $M_{KK} < 1.8$ GeV/c^2 is shown.

Table 3.10.: Efficiency in each M_{KK} bin.

Bin	Phase Space Model	With Resonance
0	0.197 ± 0.006	0.200 ± 0.006
1	0.193 ± 0.002	0.191 ± 0.002
2	0.156 ± 0.001	0.158 ± 0.001
3	0.151 ± 0.001	0.152 ± 0.001
4	0.163 ± 0.001	0.163 ± 0.001

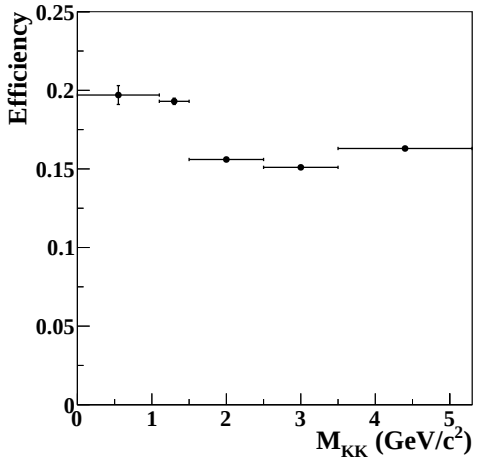
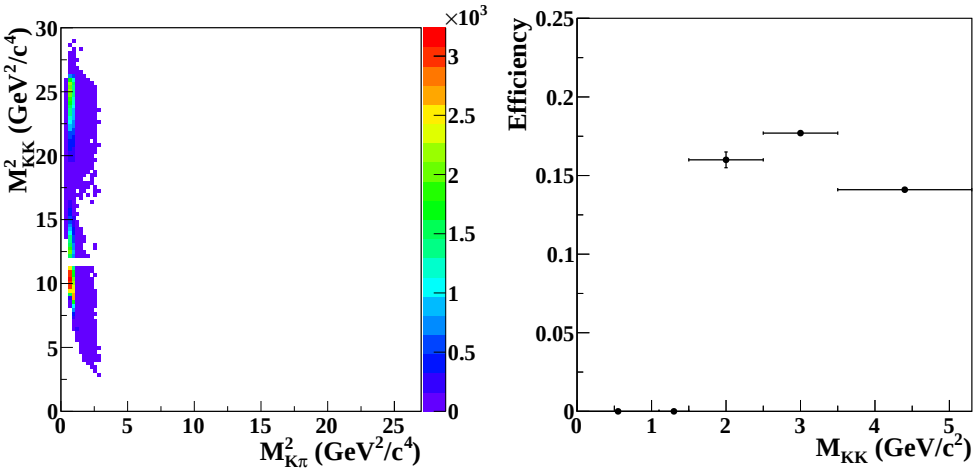


Figure 3.39.: The efficiency distribution of signal events over M_{KK}.



(a) (b)

Figure 3.40.: (a) shows the event distribution of $B^+ \to K^+\overline{K}^*(892)^0$ on the Dalitz plane with vetoes applied, while (b) shows the efficiency distribution over M_{KK}.

3.3.5. Inclusive Branching Fraction and $\mathcal{A}_{CP}$

There are two ways to obtain the inclusive results. One is to obtain the signal yield from the overall fit, and then apply the average efficiency evaluated by weighting of signal yields obtained in each bin. A better way to calculate branching fraction is to consider the signal efficiency (ϵ_{MC}) as a function of a parameter, such as masses or momenta of certain particles, instead of a constant. The formulae used to evaluate $\mathcal{B}$ and $\mathcal{A}_{CP}$ from the bin-by-bin fit are

$$\mathcal{B}(B^+ \to K^+K^-\pi^+) = \frac{\sum_i (N_{Yield,i}/\epsilon_{MC,i}) \pm \sqrt{\sum_i (\sigma_{stat,i}/\epsilon_{MC,i})^2} \pm \sum_i (\sigma_{sys,i} \cdot N_{Yield,i}/\epsilon_{MC,i})}{N_{B\overline{B}} \times C_{PID}} \tag{3.26}$$

and

$$\begin{aligned} \mathcal{A}_{CP} &= \frac{\sum_i [(N_{B^-,i} - N_{B^+,i})/\epsilon_{MC,i}]}{\sum_i [(N_{B^-,i} + N_{B^+,i})/\epsilon_{MC,i}]} \\ &= \frac{\sum_i N_{Yield,i} \cdot \mathcal{A}_{CP,i}/\epsilon_{MC,i}}{\sum_i N_{Yield,i}/\epsilon_{MC,i}} \end{aligned} \tag{3.27}$$

with its uncertainties defined as

$$\sigma_{CP,stat} = \sqrt{\sum_i \left\{ \left[\delta N_{Yield,i} \cdot \frac{\partial}{\partial N_{Yield,i}} (\mathcal{A}_{CP}) \right]^2 + \left[\delta \mathcal{A}_{CP,i} \cdot \frac{\partial}{\partial \mathcal{A}_{CP,i}} (\mathcal{A}_{CP}) \right]^2 \right\}} \tag{3.28}$$

$$\sigma_{CP,sys} = \frac{\sum_i N_{Yield,i} \cdot \sigma_{CP,sys,i}/\epsilon_{MC,i}}{\sum_i N_{Yield,i}/\epsilon_{MC,i}}, \tag{3.29}$$

and all the variables with subscript i denote the bin-dependent variable to be summed. Here σ_{stat} and σ_{sys} are the absolute statistical and the percentage systematic uncertainties (See Sec. 4.1) on the signal yield, and $\sigma_{CP,sys}$ is the absolute systematic uncertainty on $\mathcal{A}_{CP}$ (See Sec. 4.2). The statistical uncertainties should be independent between bins, thus the term is a quadratic sum. For the systematic uncertainties, some sources are correlated, as the modeling and fixed yield uncertainties in each bin can differ slightly, and some are not (e.g. $N_{B\overline{B}}$,). Contributions from bin-correlated sources will be combined by a simple sum; contributions from bin-uncorrelated sources will be combined as a quadratic sum. The statistical uncertainty on $\mathcal{A}_{CP}$ will propagate from $N_{Yield,i}$

and $\mathcal{A}_{CP,i}$. The systematic uncertainties in $\mathcal{A}_{CP}$ are also divided into two groups as bin-correlated and bin-uncorrelated, and summed accordingly.

3.4. Control sample study: $B^+ \to \overline{D}^0 \pi^+$

3.4.1. Monte Carlo and Event Selection

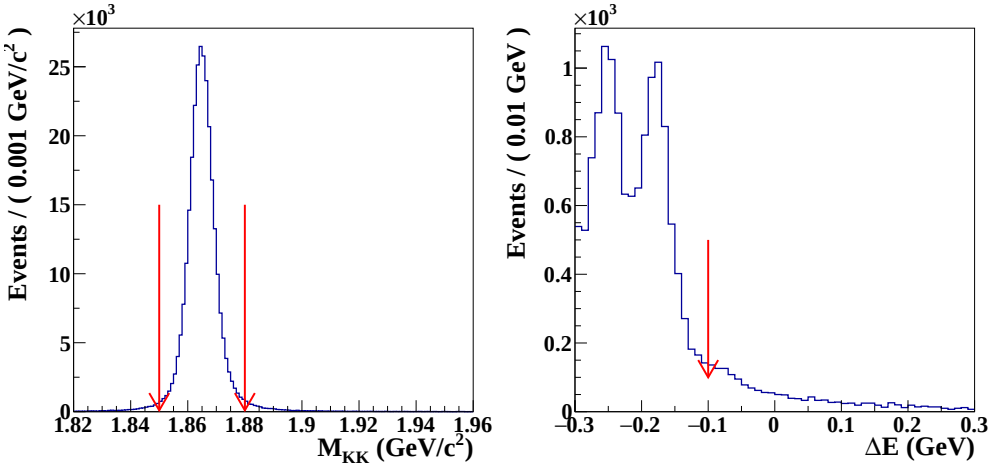
In order to study the calibration factors for the signal PDF from the data/MC discrepancy, and systematic uncertainty from continuum suppression, we choose a control channel $B^+ \to \overline{D}^0(K^+K^-)\pi^+$, which has higher statistics and the same final states as our signal analysis. The summary of event selection criteria for $B^+ \to \overline{D}^0\pi^+$ is shown in Table 3.11. To avoid the peaking background from D^* included decay modes, we choose an asymmetric windows in ΔE. After all the event selection criteria, the efficiency of true signal in MC is 16.9%, and the ratio between true signal and SCF events is 1.9%. The mass distribution of $M_{\rm KK}$ and the ΔE and $M_{\rm bc}$ distributions in signal MC are shown in Fig. 3.41 and Fig 3.42.

Table 3.11.: Summary of selection criteria for $B^+ \to \overline{D}^0\pi^+$

Particle Identification	$\mathcal{R}_{K/\pi} > 0.6$ for kaons; $\mathcal{R}_{K/\pi} < 0.4$ for pion
Impact parameters	$\lvert dr\rvert < 0.2$ cm and $\lvert dz\rvert < 5$ cm for all tracks
Continuum suppression	$NN > 0.88$
Candidate region	$5.2 < M_{\rm bc} < 5.295$ GeV/c^2 $-0.1 < \Delta E < 0.3$ GeV
D^0 selection	$1.85 < M_{\rm KK} < 1.88$ GeV/c^2
Multiple candidate selection	χ^2 of ExKFitter of the B vertex fit

3.4.2. Probability Density Function

We use a triple Gaussian to model the ΔE distribution and a double Gaussian for $M_{\rm bc}$. For continuum MC, we choose a 2nd order Chebyshev polynomial for



(a) D^0 distribution for in the signal MC (b) ΔE distribution in Generic B MC

Figure 3.41.: The D^0 mass distribution in signal MC for $B^+ \to \overline{D}^0\pi^+$ (a), and ΔE distribution in Generic B MC (b). The arrows show the selection criteria on D^0 mass and ΔE.

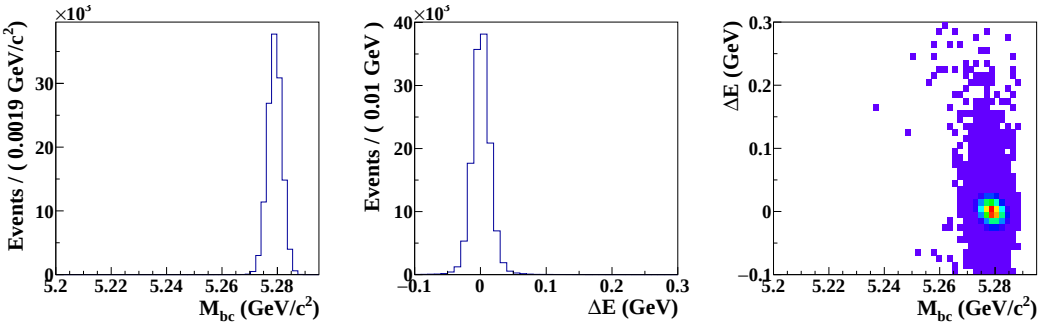


Figure 3.42.: The ΔE and $M_{\rm bc}$ distributions and the two-dimensional scatter plot for signal MC after applying all event selection.

ΔE and an ARGUS function for $M_{\rm bc}$. Because of the correlation between ΔE and $M_{\rm bc}$ in $B\overline{B}$ backgrounds, we use 2D smoothed histogram to characterize the ΔE and $M_{\rm bc}$ distributions for self-cross feed, Generic B, and Rare B backgrounds. The distribution of ΔE, $M_{\rm bc}$, and 2D plots are shown in Fig. 3.42-3.46.

3.4.3. Fitting Result and Branching Fraction

In order to study possible discrepancy between MC and data for the signal PDF, a shift in the mean and an expansion factor in the width of all Gaussian

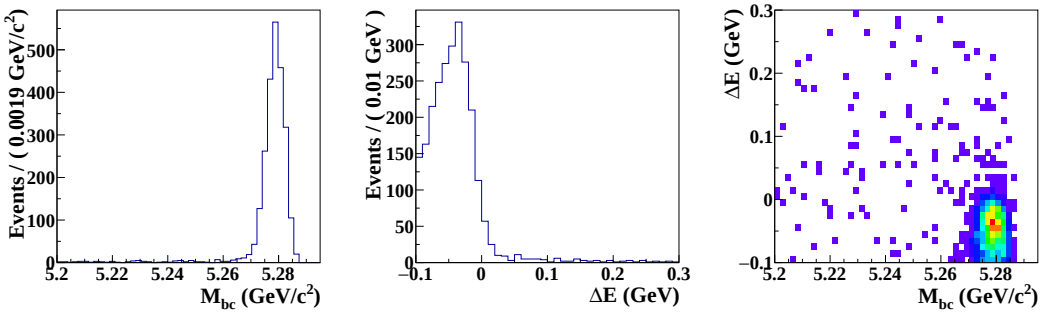


Figure 3.43.: The ΔE and $M_{\rm bc}$ distributions and the two-dimensional scatter plot for SCF events in the control study.

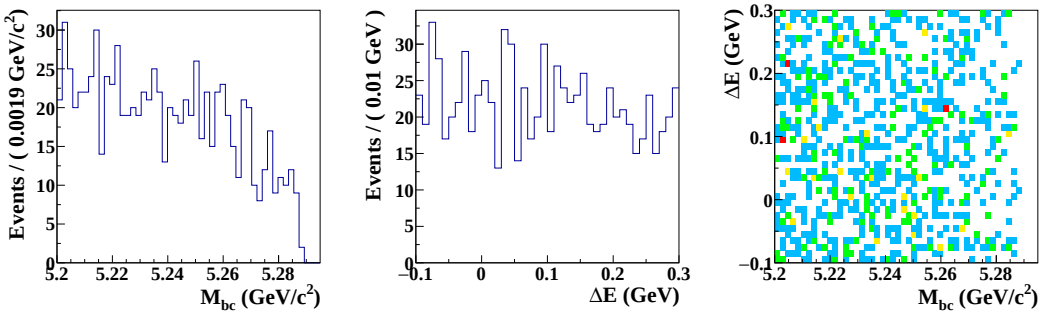


Figure 3.44.: The ΔE and $M_{\rm bc}$ distributions and the two-dimensional scatter plot for continuum MC.

functions are added to the fitting method as calibration factors. The fitting results of the control sample and the calibration factors for the signal PDFs are shown in Fig. 3.47 and Table. 3.12, respectively. The calibration factors will be applied to the signal PDFs in the fit of the $B^+ \to K^+K^-\pi^+$ analysis.

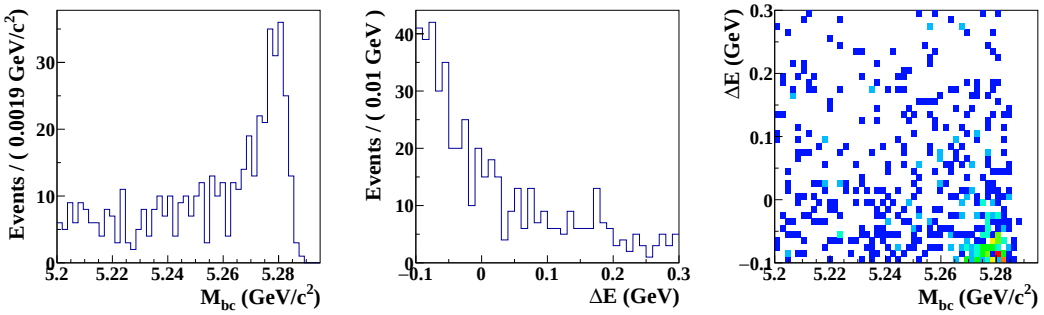


Figure 3.45.: The ΔE and $M_{\rm bc}$ distributions and the two-dimensional scatter plot for Generic B MC.

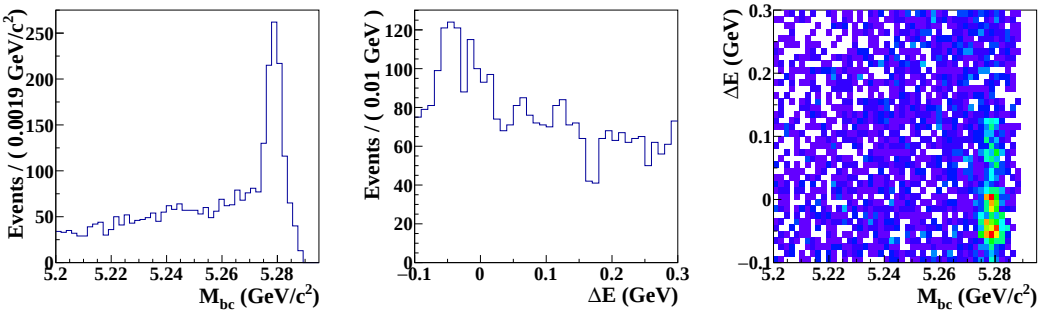


Figure 3.46.: The ΔE and $M_{\rm bc}$ distributions and the two-dimensional scatter plot for Rare B MC .

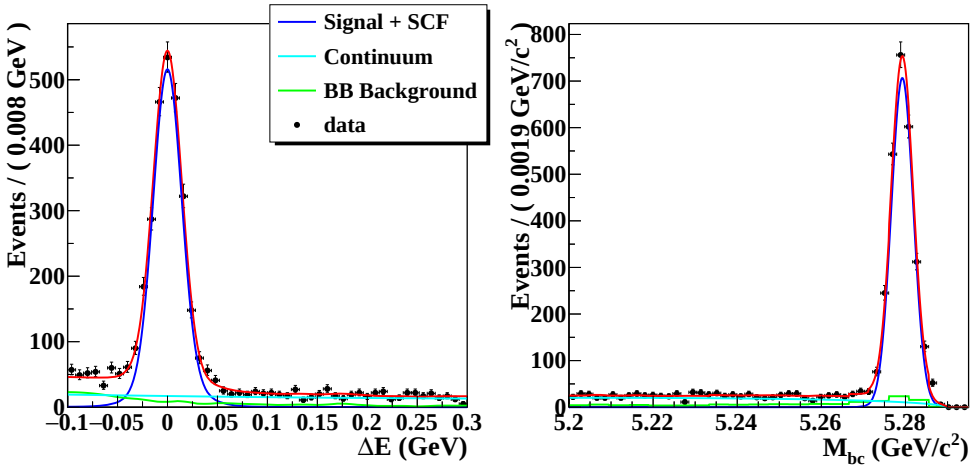


Figure 3.47.: The projection plots on ΔE (left) and $M_{\rm bc}$ (right) of the fit to data for $B^+ \to \overline{D}^0\pi^+$. The blue line is for signal, the cyan is for continuum, and green is the sum of $B\overline{B}$ backgrounds.The plots for ΔE require $5.27 < M_{\rm bc} < 5.29$ GeV/c^2. The plots for $M_{\rm bc}$ require $-0.05 < \Delta E < 0.05$ GeV.

The branching fraction of the control sample mode is evaluated by the formula:

$$\mathcal{B}(B^+ \to \overline{D}^0\pi^+) = \frac{N_{yield}}{N_{B\overline{B}} \times \epsilon_{MC}}, \tag{3.30}$$

Table 3.12.: Calibration factors for the signal PDF.

ΔE mean (MeV)	-0.8806 ± 0.3355
ΔE width	1.1976 ± 0.0230
$M_{\rm bc}$ mean (MeV/c^2)	0.1884 ± 0.0555
$M_{\rm bc}$ width	1.0289 ± 0.0163

where ϵ_{MC} is the signal efficiency obtained from MC. The obtained branching fraction and branching fraction product are $\mathcal{B}(B^+ \to \overline{D}^0\pi^+) = (4.75 \pm 0.14(stat.)) \times 10^{-3}$ and $\mathcal{B}(B^+ \to \overline{D}^0\pi^+) \times \mathcal{B}(D^0 \to K^+K^-) = (1.88 \pm 0.04(stat.)) \times 10^{-5}$, which are consistent with the PDG values $(4.81 \pm 0.15) \times 10^{-3}$ and $(1.9 \pm 0.07) \times 10^{-5}$, thus validating the signal efficiency estimation and the fitting method.

Chapter 4.

Systematic Uncertainties

In this chapter, the sources of systematic uncertainty are discussed. The main contributions to the uncertainty arise from the background PDF modeling used in the fit. The uncertainty on the branching fraction and $\mathcal{A}_{CP}$ measurements are discussed separately in this chapter.

4.1. Systematic Uncertainties of Branching Fraction Measurement

The systematic uncertainties for branching fractions come from the following items. The summary of systematic uncertainties is shown in Table 4.2.

4.1.1. Tracking

The tracking uncertainty will be evaluated from the previous study for high momentum tracks with $P_t > 200$ MeV/c [59]. The track finding efficiency is measured by comparing the number of partially and fully reconstructed D^* decays. We introduce a systematic uncertainty of 0.35% per track .

4.1.2. PID performance

As discussed in Sec 2.3.1, the PID performance for K/π separation is studied in data and MC using the decay:

$$\begin{aligned} D^{*+} \rightarrow & D^0 \pi_s^+ \\ & D^0 \rightarrow K^- \pi^+, \end{aligned} \tag{4.1}$$

where π_s^+ denotes a low momentum charged π meson. The slow pion allows the event to be reconstructed with good signal/noise ratio without relying on PID information. Since the PID efficiency is related to the momenta of the tracks, the efficiency and fake rate were studied in different momentum and polar angle bins. A look-up table is provided by the PID joint group to obtain the average correction of the efficiency ratio between MC and data. The factors are shown in Table 4.1. The uncertainty from the PID correction is 1.44%.

Table 4.1.: PID correction factors.

K^+	0.9953 ± 0.0084
K^-	0.9951 ± 0.0084
π	0.9506 ± 0.0082

4.1.3. Continuum Suppression

The study of the systematic uncertainty from continuum suppression is performed in the control sample study. We compare the branching fractions in the control sample study with and without applying continuum suppression. Figure 4.1 shows the fitting result without continuum suppression. There are $254,054.0 \pm 504.0$ signal events in MC and 4779.1 ± 76.4 in the fitted result without continuum suppression; there are $129,938.0 \pm 360.5$ events in MC and 2440.8 ± 50.5 in the fitted result with continuum suppression. Comparing the ratio of passing events in MC and data:

$$\text{Ratio}_{MC} = 51.1\% \times (1 \pm 0.0019) \tag{4.2}$$

$$\text{Ratio}_{Data} = 51.1\% \times (1 \pm 0.0131) \tag{4.3}$$

A systematic uncertainty of 1.33% is quoted for this term.

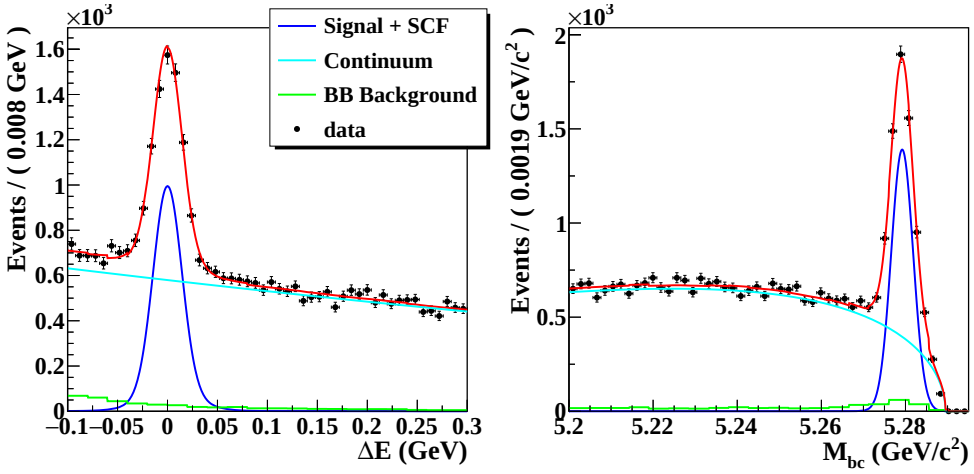


Figure 4.1.: The projection plots of ΔE (left) and $M_{\rm bc}$ (right) of the fit to data for $B^+ \to \overline{D}^0 \pi^+$ without continuum suppression. The blue line is for signal, cyan is for continuum, and green is for Generic B. The plots for ΔE require $5.27 < M_{\rm bc} < 5.29$ GeV/c^2. The plots for $M_{\rm bc}$ require $-0.05 < \Delta E < 0.05$ GeV.

4.1.4. Signal PDF Modeling

The uncertainty due to the calibration factors of the signal PDFs is also evaluated on the control sample study (See 3.4.3). We compare the signal yields with and without calibration factors . The signal yield difference of 1.77% is quoted as the systematic uncertainty for the signal PDF modeling.

4.1.5. Background PDF Modeling

The uncertainty due to the fixed PDF shapes is obtained by varying the binning of the 2D smooth histograms or by repeating the fit by varying the shape parameters by 1σ. To take the possible data/MC discrepancy on continuum samples into account, we compared the distribution between continuum MC and off-resonance data with loose ($NN > -0.28$) and tight ($NN > 0.88$) continuum suppression selection criteria, shown in Fig 4.2. We found a disagreement in

the ΔE distributions with loose and tight selections in continuum MC sample despite only a tiny correlation between ΔE and NN. In the off-resonance data, the statistics after the tight selection applied is too few to see the difference. We calculate the negative error of shape parameter x as

$$\sigma_{x_-} = |x_{\text{off-res loose}} - x_{\text{MC loose}}| + \sigma_{x_{\text{off-res loose}}} \tag{4.4}$$

to take the data/MC difference into account since the values of the parameters in the off-resonance data are 2-3 sigma smaller than MC. The difference in signal yield between the yields with varied parameters and the nominal fit result is quoted as the systematic uncertainty.

4.1.6. Fixed Yields of Rare B components

Since the normalizations of the backgrounds from Rare B decays are fixed in the fit, the uncertainty is estimated by varying the normalization by 1σ from the expected value and comparing the difference in signal yields.

4.1.7. Background $\mathcal{A}_{CP}$

The uncertainty due to non-zero $\mathcal{A}_{CP}$ of Rare B peaking components is estimated by fixing the $\mathcal{A}_{CP}$ values to the measured values and varying each $\mathcal{A}_{CP}$ value up or down by its error. The $\mathcal{A}_{CP}$ values are $-0.036 \pm 0.004 \pm 0.002 \pm 0.007$ [20] and $0.025 \pm 0.004 \pm 0.004 \pm 0.007$ [20] for $B^\pm \to B^+ \to K^+K^-K^+$ and $B^\pm \to B^+ \to K^+\pi^+\pi^-$, respectively. The difference between the yields of the nominal fit and of the fit with non-zero background $\mathcal{A}_{CP}$ is quoted as the systematic uncertainty.

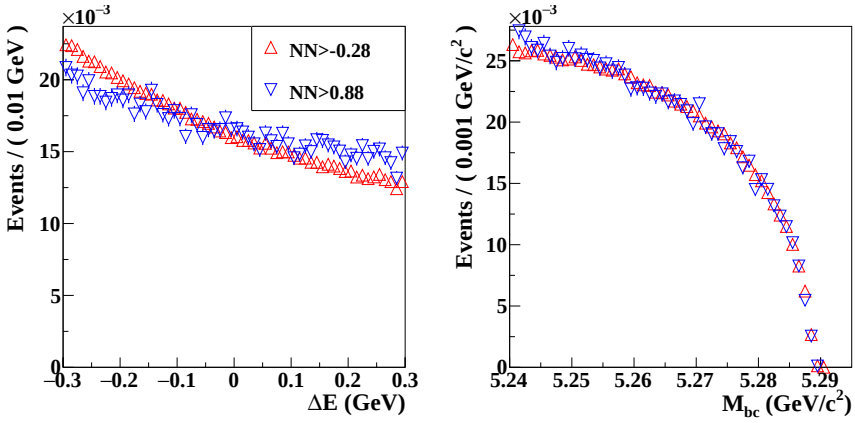
4.1.8. Other Systematics

Number of $B\overline{B}$:
The total number of $B\overline{B}$ events has an uncertainty associated with it. This adds a systematic uncertainty to our result. The uncertainty in the number of $B\overline{B}$ events is 1.37% based on the data of $(771.6 \pm 10.6) \times 10^6$ $B\overline{B}$ pairs [60].

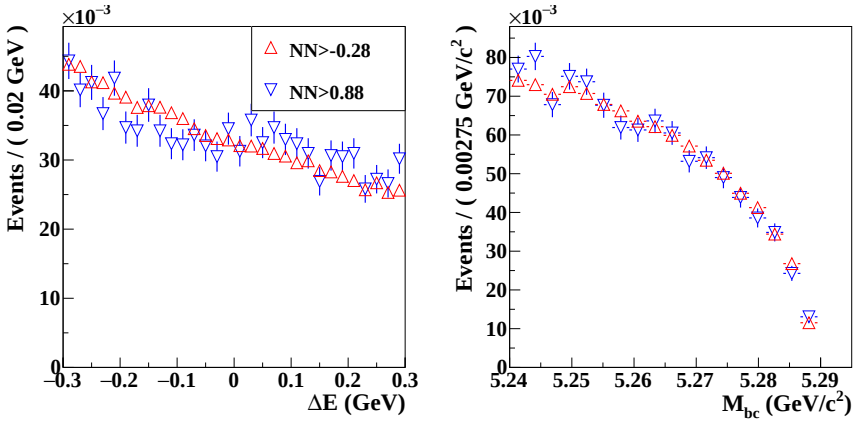
Fitting Bias:
The bias of 2.3% obtained from the GSIM ensemble test is quoted as a systematic uncertainty. For more details please see Sec. 3.3.3.

Table 4.2.: Systematic uncertainties in the measured branching fraction in the individual bins. The † sign indicates M_{KK} dependent uncertainties. The – sign indicates an uncertainty less than 0.05%.

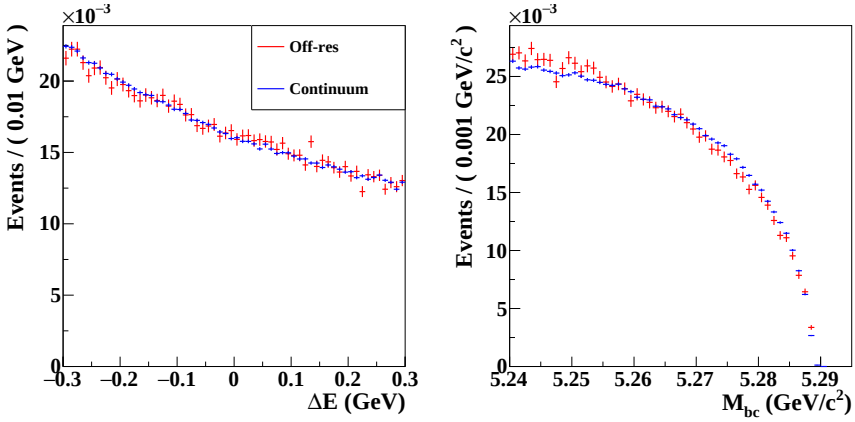
Source	Relative uncertainties in $\mathcal{B}$ (%)				
M_{KK}(GeV/c^2)	0.8-1.1	1.1-1.5	1.5-2.5	2.5-3.5	3.5-5.3
Number of $B\overline{B}$ pairs			1.37		
Tracking			1.05		
Particle identification			1.44		
Continuum suppression			1.33		
Signal PDF modeling			1.77		
Fit bias			2.30		
Background PDF modeling†	3.65	2.15	16.16	3.77	3.59
Fixed yields†	–	–	–	0.07	–
Background $\mathcal{A}_{CP}{}^\dagger$	0.23	0.28	1.46	0.80	0.36



(a) Continuum MC samples.



(b) Off-resonance data samples.



(c) Continuum MC and off-resonance data samples.

Figure 4.2.: Comparison of ΔE and M_{bc} between (a) continuum MC samples and (b) off-resonance data samples with loose (red) and tight (blue) continuum suppression selection, and (c) continuum MC (blue) and off-resonance (red) data with loose continuum suppression selection.

4.2. Systematic Uncertainties of $\mathcal{A}_{CP}$

The systematic uncertainties for $\mathcal{A}_{CP}$ come from the following items. In order to obtain the systematic uncertainty from the following sources, we apply the same procedure as in the previous section.

- **Background PDF Modeling:**
  The procedure is the same as the one for branching fraction. Instead of comparing the signal yield, we compare the $\mathcal{A}_{CP}$ value.

- **Fixed Yields of Rare B Backgrounds:**
  From the previous analyses, the fixed normalizations of Rare B components might also affect the measured $\mathcal{A}_{CP}$. We vary the normalizations by $\pm 1\sigma$ and then compare the $\mathcal{A}_{CP}$ difference.

- **Background $\mathcal{A}_{CP}$:**
  The procedure is the same as the one for branching fraction. Instead of comparing the signal yield, we compare the $\mathcal{A}_{CP}$ value.

- **Detector Asymmetry for π:**
  The detector bias is studied with the off-resonance data. The $\mathcal{A}_{CP}$ from the off-resonance data is evaluated as the following:

$$\mathcal{A}_{CP} = \frac{N_{K^+K^-\pi^-} - N_{K^+K^-\pi^+}}{N_{K^+K^-\pi^-} + N_{K^+K^-\pi^+}}, \tag{4.5}$$

  here we assume the number of B candidates is

$$T_{K^+K^-\pi^-} = \frac{N(K^+K^-\pi^-)}{\epsilon(K^+) \cdot \epsilon(K^-) \cdot \epsilon(\pi^-)}, \tag{4.6}$$

  where $N_{K^+K^-\pi^-}$ is the observed number of B^- candidates in the off-resonance data and $\epsilon(K^+)$ is the efficiency of K^+ detection. Then the equation becomes

$$\mathcal{A}_{CP} = \frac{\epsilon(K^+)\epsilon(K^-)\epsilon(\pi^-)T_{K^+K^-\pi^-} - \epsilon(K^+)\epsilon(K^-)\epsilon(\pi^+)T_{K^+K^-\pi^+}}{\epsilon(K^+)\epsilon(K^-)\epsilon(\pi^-)T_{K^+K^-\pi^-} + \epsilon(K^+)\epsilon(K^-)\epsilon(\pi^+)T_{K^+K^-\pi^+}}. \tag{4.7}$$

Due to the K^+K^- pair appearing in both B^+ and B^- candidates, the efficiencies they contribute can be cancelled at the first order to give

$$\mathcal{A}_{CP} = \frac{\epsilon(\pi^-)T_{K^+K^-\pi^-} - \epsilon(\pi^+)T_{K^+K^-\pi^+}}{\epsilon(\pi^-)T_{K^+K^-\pi^-} + \epsilon(\pi^+)T_{K^+K^-\pi^+}}. \tag{4.8}$$

The asymmetry of detector efficiency between particles and anti-particles depends on momentum and polar angle of particles in the laboratory frame. In this analysis, K^+ and K^- in the decay product have similar distribution as shown in Fig. 4.3. In addition to this, since we have a pair of K^+ and K^- in the final state, the detector bias for kaons will be canceled at least at the first order.

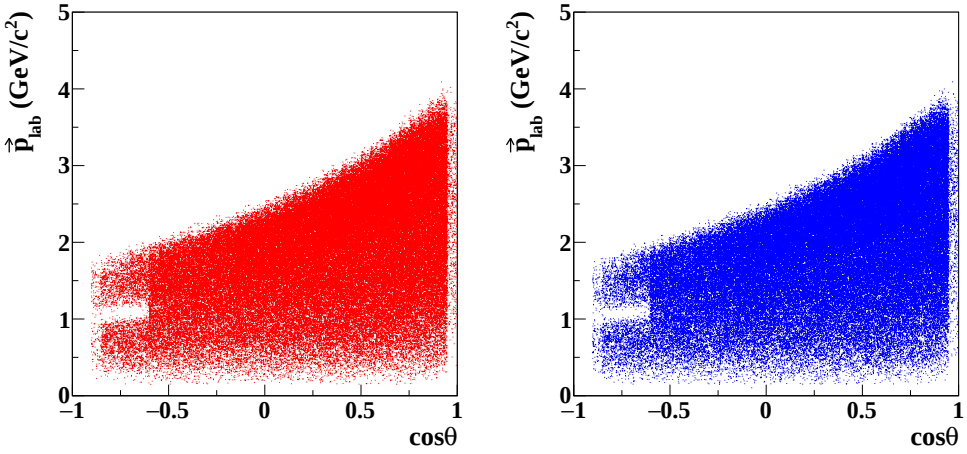
To test this, we reweighted the signal MC sample with following asymmetric K^+ and K^- efficiencies:

$$\epsilon(\vec{p}) = \begin{cases} 0.8 + 0.04 \times |\vec{p}|, & \text{for } K^+ \\ 1.0 - 0.04 \times |\vec{p}|, & \text{for } K^-, \end{cases} \tag{4.9}$$

and the total reweighted factor is $\epsilon(K^+) \cdot \epsilon(K^-)$. The original $\mathcal{A}_{CP}$ value of the testing sample is 0.0017 ± 0.0019 (stat.) and that of the reweighted sample is 0.0021 ± 0.0019 (stat.). The conclusion is that the $\mathcal{A}_{CP}$ values from the original and reweighted samples are consistent and hence any asymmetric efficiency due to kaons has a negligible contribution to the systematic uncertainty.

Therefore, for this analysis we only need to consider the detector asymmetry from charged pions. To reduce the statistical uncertainty, the continuum suppression selection is not applied to the off-resonance data for the systematic uncertainty study. The $\mathcal{A}_{CP}$ obtained from the off-resonance data is 0.0024 ± 0.0014, the central shift plus 1σ statistical error are quoted as the systematic uncertainty for the detector asymmetry.

The summary of the systematic uncertainties on $\mathcal{A}_{CP}$ is shown in Table 4.3.



(a) K^+ (b) K^-

Figure 4.3.: Momentum verse polar angle distribution of K^+ and K^-.

Table 4.3.: Systematic uncertainties on $\mathcal{A}_{CP}$ in the individual bins. The † sign indicates M_{KK} dependent uncertainties, and the – sign indicates an uncertainty less than 0.001.

Source	Absolute uncertainties on $\mathcal{A}_{CP}$				
M_{KK}(GeV/c^2)	0.8-1.1	1.1-1.5	1.5-2.5	2.5-3.5	3.5-5.3
Background PDF modeling†	0.036	0.005	0.028	0.006	0.003
Fixed yields†	–	–	–	0.002	-
Background $\mathcal{A}_{CP}$†	0.015	0.004	0.009	0.005	0.002
Detector bias			0.004		

4.3. Summary of Systematics of Inclusive Results

The modeling, fixed yield, and background $\mathcal{A}_{CP}$ uncertainties are bin-dependent in the bin-by-bin fit, while other terms are bin-independent. As mentioned in Sec. 3.3.4, the systematic uncertainty will be divided into two parts: bin-dependent and bin-independent terms. The bin-dependent term is a simple sum, and the bin-independent term is a quadratic sum.

Chapter 5.

Result and Conclusion

5.1. Result

5.1.1. Overall signal extraction

Using the fit strategy described above, we obtain a signal yield of $714.6^{+48.3}_{-47.6}$ with 18σ significance, and the fitted results are shown in Fig. 5.1.

The significance for signal yield and $\mathcal{A}_{CP}$ is calculated by:

$$\mathcal{S} = \sqrt{-2\ln\frac{\mathcal{L}_0}{\mathcal{L}_{max}}}. \tag{5.1}$$

The $\mathcal{A}_{CP}$ obtained from the simultaneous fit to the data is -0.177 ± 0.067 with 3σ significance. Figure 5.2 shows the projection plots on $M_{\rm bc}$ and ΔE for the B^+ and B^- samples.

5.1.2. Bin-by-bin signal extraction in $M_{\rm KK}$

The signal yields and $\mathcal{A}_{CP}$ over $M_{\rm KK}$ are obtained by the extraction method described in Sec. 3.3.5. Table 5.1 and Figures 5.3-5.4 show the results in each bin. To investigate whether the fit describes the data adequately, we use the binned χ^2 test, which was introduced by Pearson [61] as an estimator for the "goodness

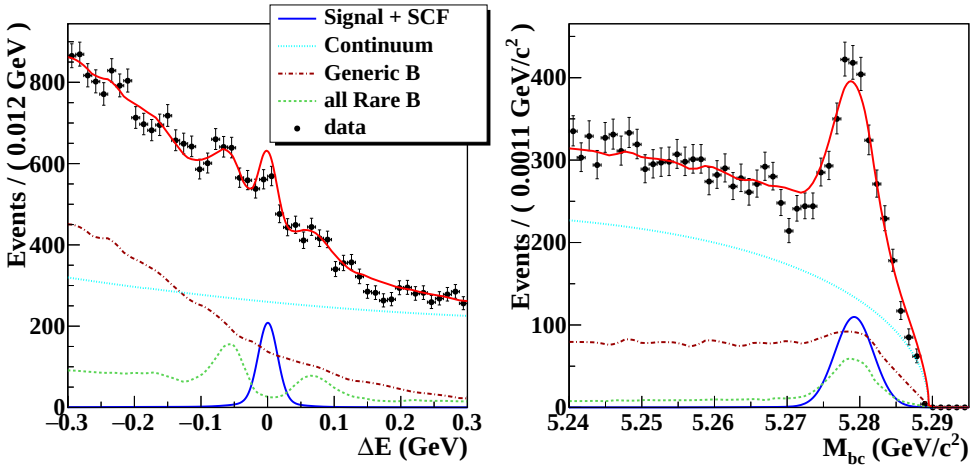
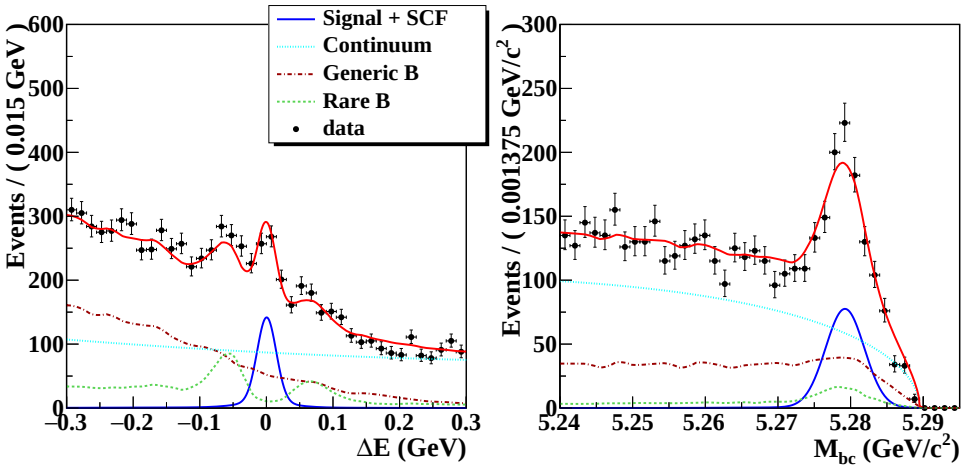


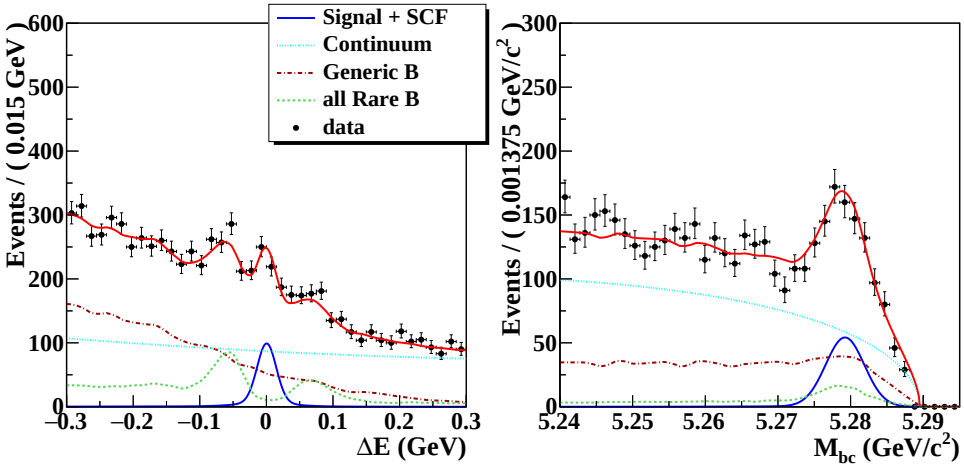
Figure 5.1.: Projection plots showing the fit results of $B^+ \to K^+K^-\pi^+$. The black points with error bars are data, the red line is total PDF, the blue line is the combination of signal and SCF events, the green line is the total Rare B components, the brown line is the Generic B component, and the cyan line is for continuum background. The plots for ΔE require $5.27 < M_{\rm bc} < 5.29$ GeV/c^2. The plots for $M_{\rm bc}$ require $-0.05 < \Delta E < 0.05$ GeV.

Table 5.1.: Signal yield, significance of signal yield, $\mathcal{A}_{CP}$, significance of $\mathcal{A}_{CP}$, and differential branching fraction for individual bins. The first error is statistical error, and the second is systematic error.

$M_{\rm KK}$	N_{sig}	$\mathcal{S}(\sigma)$	$\mathcal{A}_{CP}$	$\mathcal{S}(\sigma)$	$d\mathcal{B}/dM$ ($\times 10^{-7}$)
0.8-1.1	$59.8 \pm 11.4 \pm 2.6$	6.7	$-0.896 \pm 0.166 \pm 0.039$	4.8	$14.0 \pm 2.7 \pm 0.8$
1.1-1.5	$212.4 \pm 21.3 \pm 6.7$	12.5	$-0.157 \pm 0.098 \pm 0.007$	1.6	$37.8 \pm 3.8 \pm 1.9$
1.5-2.5	$113.5 \pm 26.7 \pm 18.6$	3.7	$-0.135 \pm 0.231 \pm 0.030$	0.6	$10.0 \pm 2.3 \pm 1.7$
2.5-3.5	$110.1 \pm 17.6 \pm 4.9$	7.4	$-0.092 \pm 0.158 \pm 0.009$	0.6	$10.0 \pm 1.6 \pm 0.6$
3.5-5.3	$172.6 \pm 25.7 \pm 7.4$	7.4	$-0.053 \pm 0.147 \pm 0.006$	0.4	$8.1 \pm 1.2 \pm 0.5$



(a) B^+



(b) B^-

Figure 5.2.: The projection plots on ΔE (left) and M_{bc} (right) for the B^+ (a) and B^- (b) samples. The projections on M_{bc} require $-0.035 < \Delta E < 0.035$ GeV, while the plots of ΔE require $5.275 < M_{\mathrm{bc}} < 5.285$ GeV/c^2.

of fit". The χ^2 is defined as

$$\chi^2 = \sum_{c=1}^{n_c} \frac{(o_c - e_c)^2}{e_c}, \tag{5.2}$$

where n_c is the number of cells and o_c (e_c) is the number of observed (expected) events in the c^{th} cell. To avoid cells with too few events, a multidimensional adaptive binning algorithm, TKDTreeBinning , implemented in ROOT [58] is used. The data set in each M_{KK} bin is divided into cells containing approximately 25 events each. Figure 5.5 shows an example of adaptive binning in the 2D $M_{\text{bc}} - \Delta E$ plane in the region of $0.8 < M_{\text{KK}} < 1.1$ GeV/c^2. The reduced χ^2 and the number of degrees of freedom are listed in Table 5.2.

The systematic uncertainty is taken into account in the significance calculation by convolving the likelihood function with a Gaussian whose width equals the systematic uncertainty [62]:

$$\mathcal{L}_{smear}(n) = \int_{\infty}^{-\infty} \mathcal{L}(n') \frac{e^{-\frac{(n-n')^2}{2\Delta n^2}}}{\sqrt{2\pi}\Delta n} dn'. \tag{5.3}$$

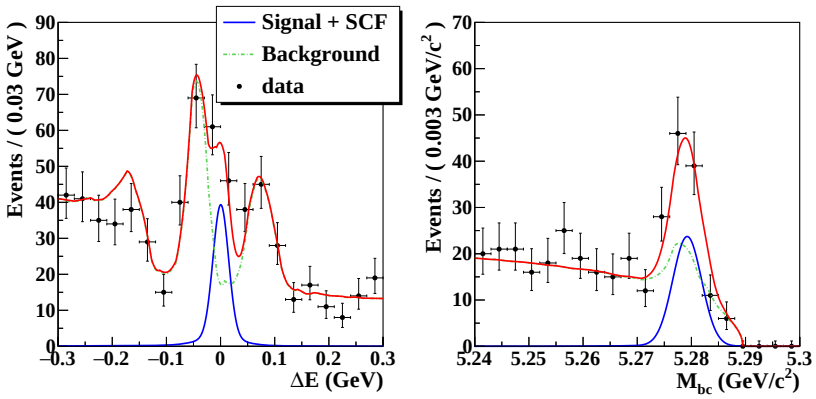
The total signal yield, 668.4 ± 54.6, is consistent with the overall fit result. We integrate the differential branching fraction and the measured $\mathcal{A}_{CP}$ in each M_{KK} by Eq. 3.26 and Eq. 3.27 to obtain the inclusive branching fraction and $\mathcal{A}_{CP}$:

$$\mathcal{B}(B^+ \to K^+K^-\pi^+) = (5.38 \pm 0.40 \pm 0.35) \times 10^{-6} \tag{5.4}$$

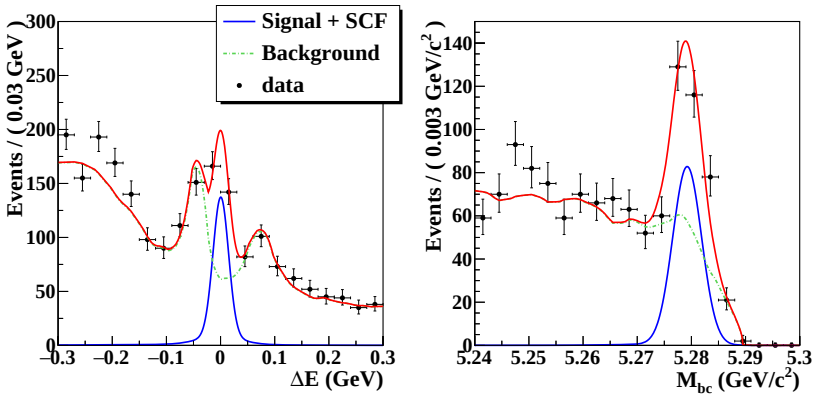
$$\mathcal{A}_{CP} = -0.170 \pm 0.073 \pm 0.017, \tag{5.5}$$

where the first uncertainties are statistical and the second are systematic.

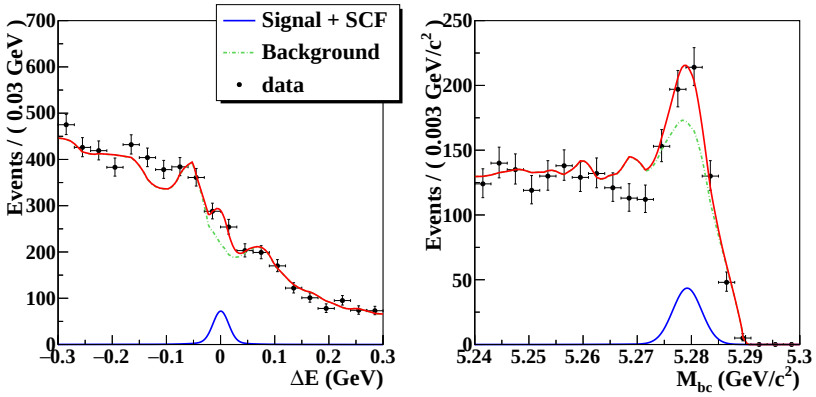
The results of signal yields and $\mathcal{A}_{CP}$ as functions of M_{KK} are shown in Fig. 5.6. We obtain a large $\mathcal{A}_{CP}$ of -0.896 with 4.8σ significance in the region $M_{\text{KK}} < 1.1\text{GeV}/c^2$.



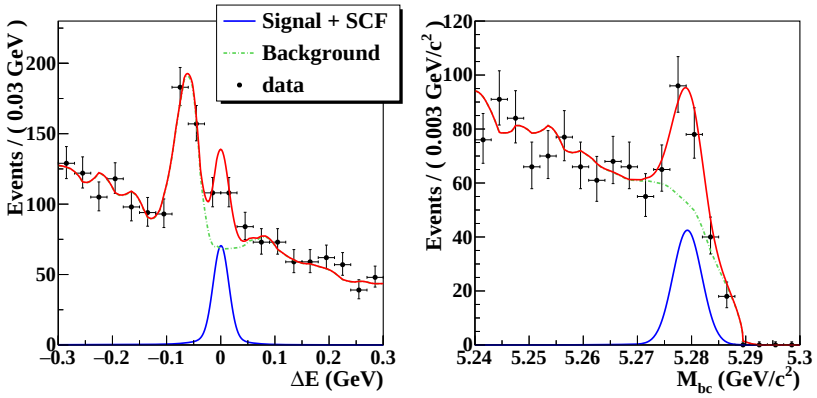
(a) $0.8 < M_{KK} < 1.1$ GeV/c^2



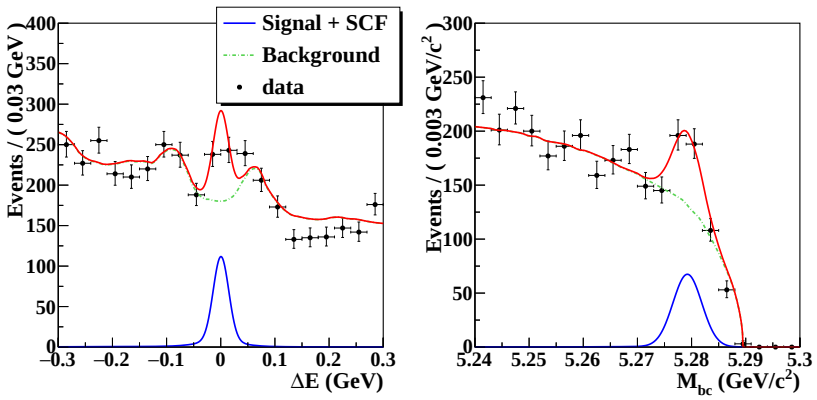
(b) $1.1 < M_{KK} < 1.5$ GeV/c^2



(c) $1.5 < M_{KK} < 2.5$ GeV/c^2



(d) $2.5 < M_{\mathrm{KK}} < 3.5$ GeV/c^2

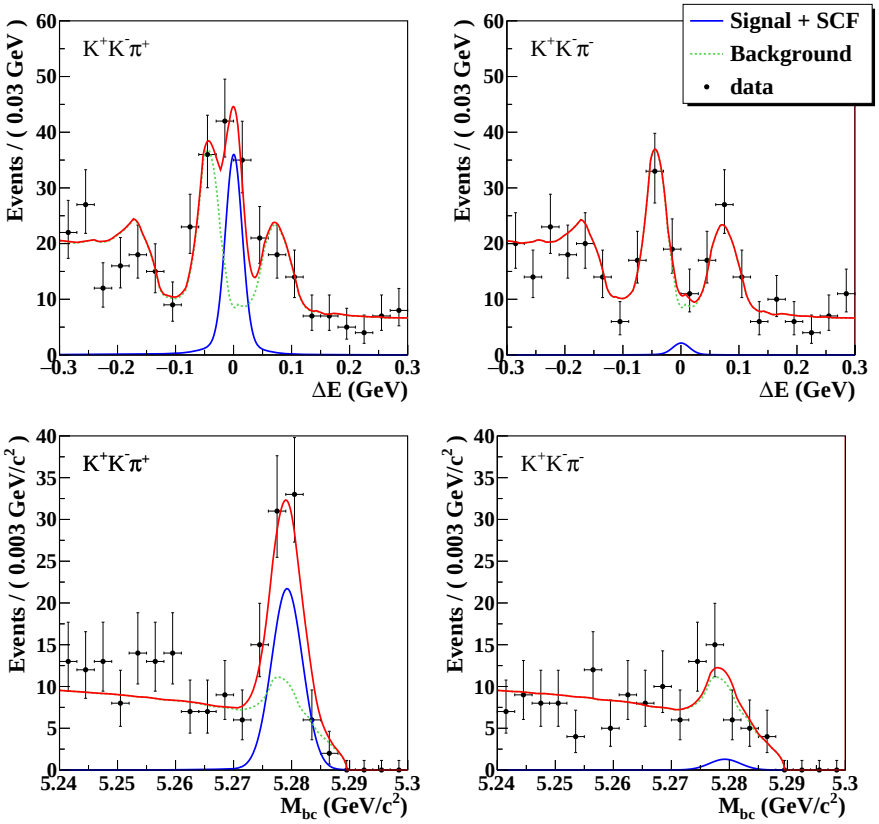


(e) $3.5 < M_{\mathrm{KK}} < 5.3$ GeV/c^2

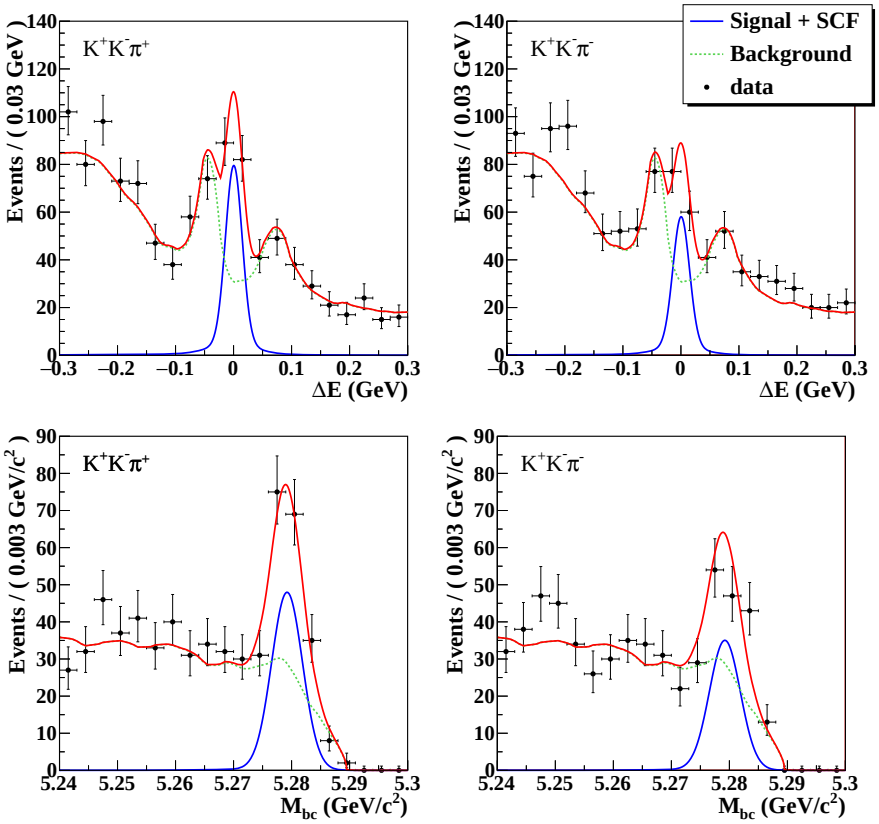
Figure 5.3.: The projection plots on ΔE and M_{bc} in individual M_{KK} bin. The plots for ΔE require $5.27 < M_{\mathrm{bc}} < 5.29$ GeV/c^2. The plots for M_{bc} require $-0.05 < \Delta E < 0.05$ GeV.

Table 5.2.: Summary of reduced χ^2 and number of degrees of freedom used in each M_{KK} bin.

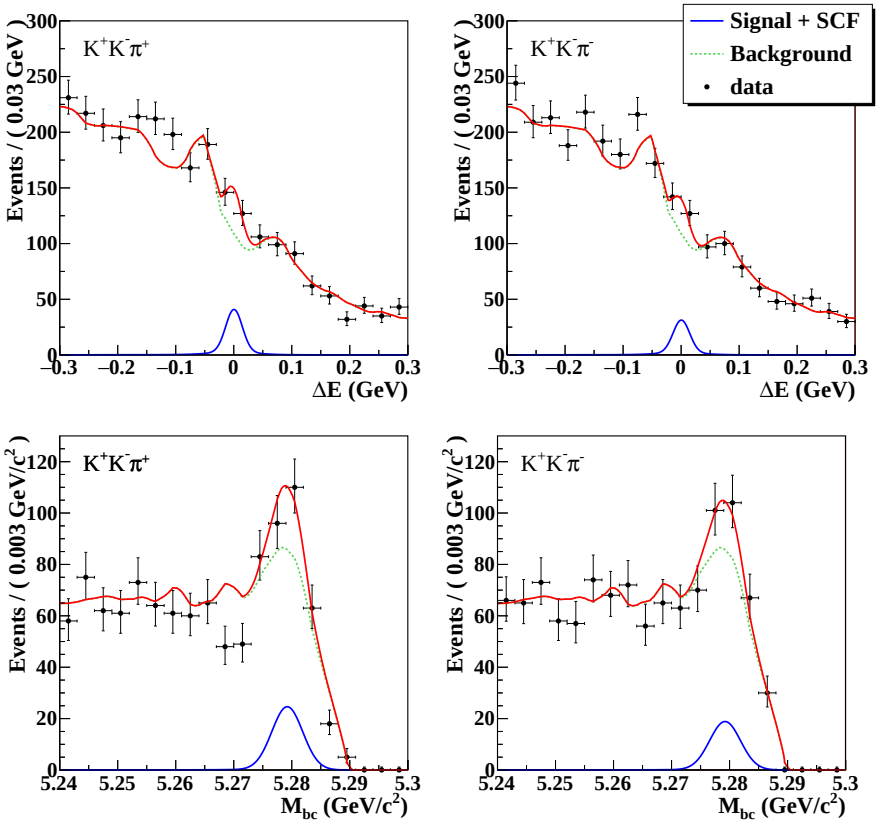
M_{KK}	χ^2/ndof	ndof
0.8-1.1	1.03	112
1.1-1.5	1.16	438
1.5-2.5	1.05	973
2.5-3.5	1.06	448
3.5-5.3	1.16	1089



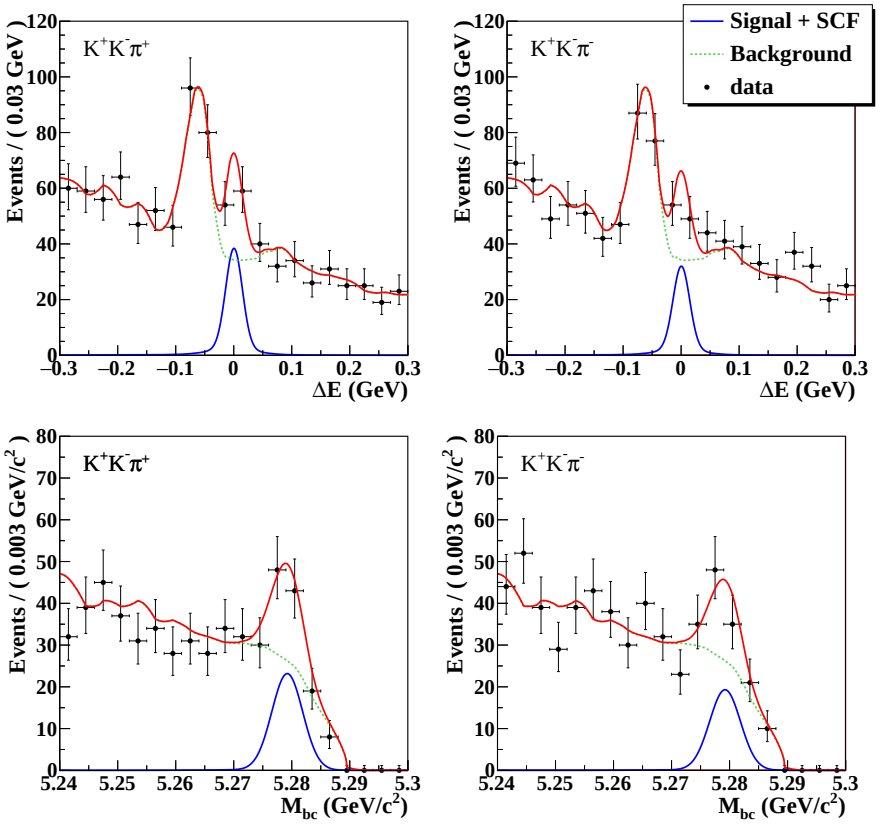
(a) $0.8 < M_{KK} < 1.1$ GeV/c^2



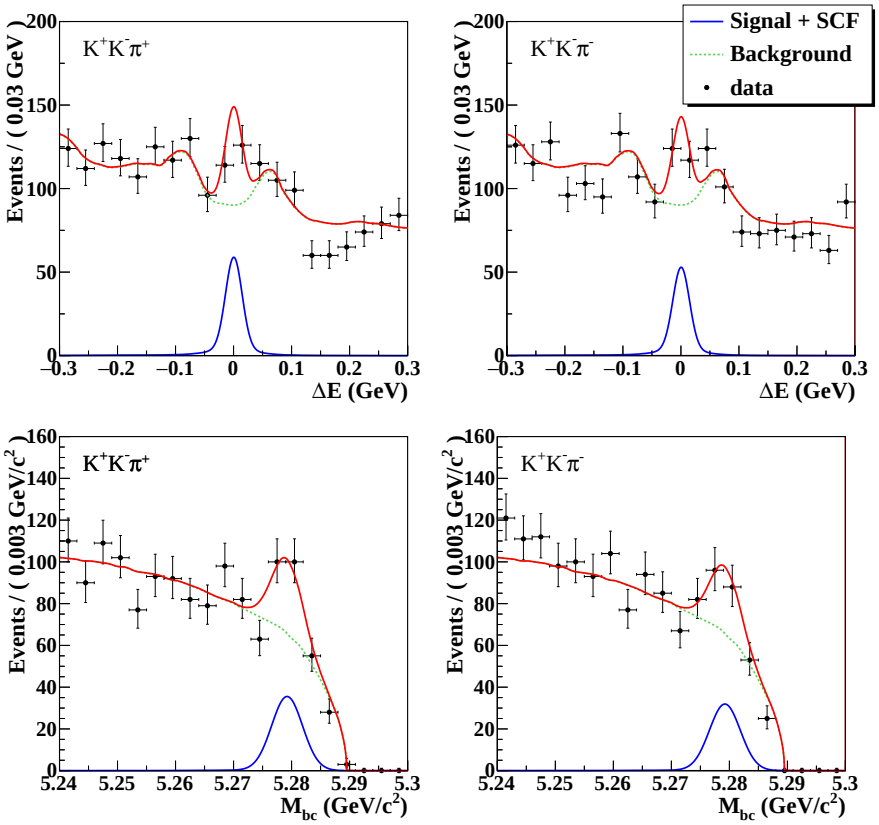
(b) $1.1 < M_{KK} < 1.5$ GeV/c^2



(c) $1.5 < M_{KK} < 2.5$ GeV/c^2

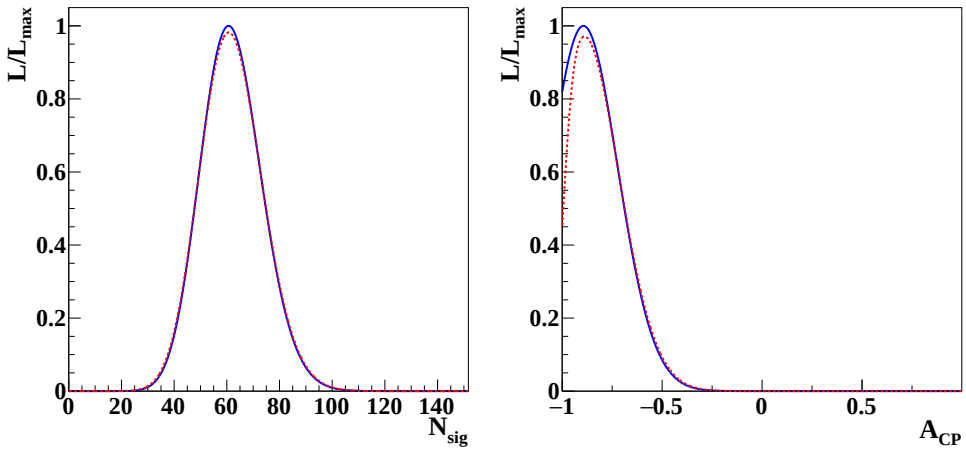


(d) $2.5 < M_{KK} < 3.5$ GeV/c^2

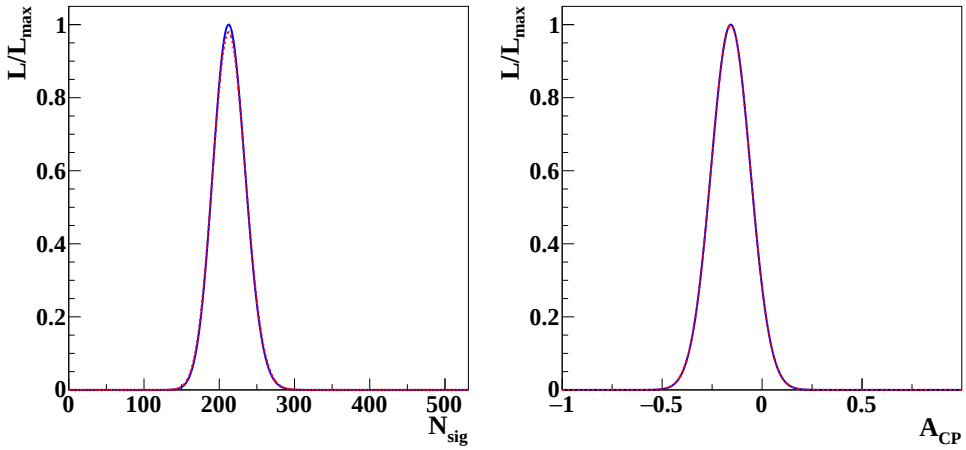


(e) $3.5 < M_{KK} < 5.3$ GeV/c^2

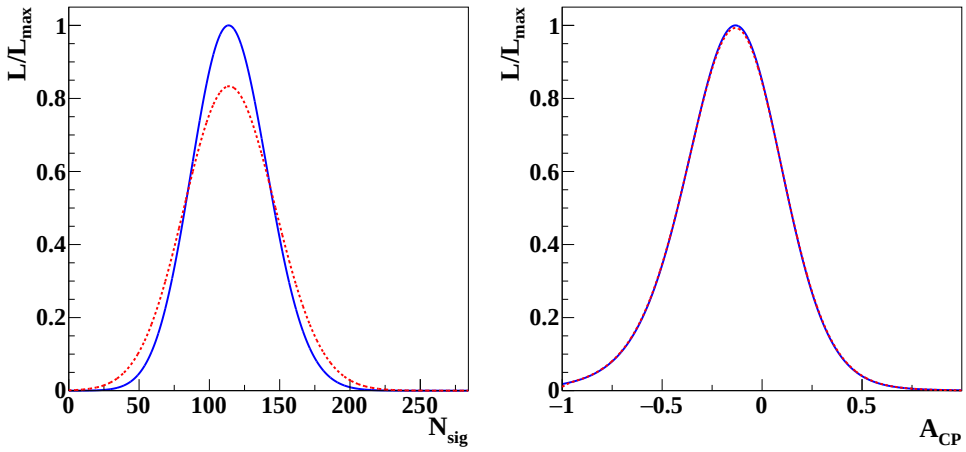
Figure 5.4.: Projection plots showing the fit results on ΔE and M_{bc} of B^+/B^- in each M_{KK} bin. The plots for ΔE require $5.275 < M_{bc} < 5.2835$ GeV/c^2. The plots for M_{bc} require $-0.03 < \Delta E < 0.03$ GeV.



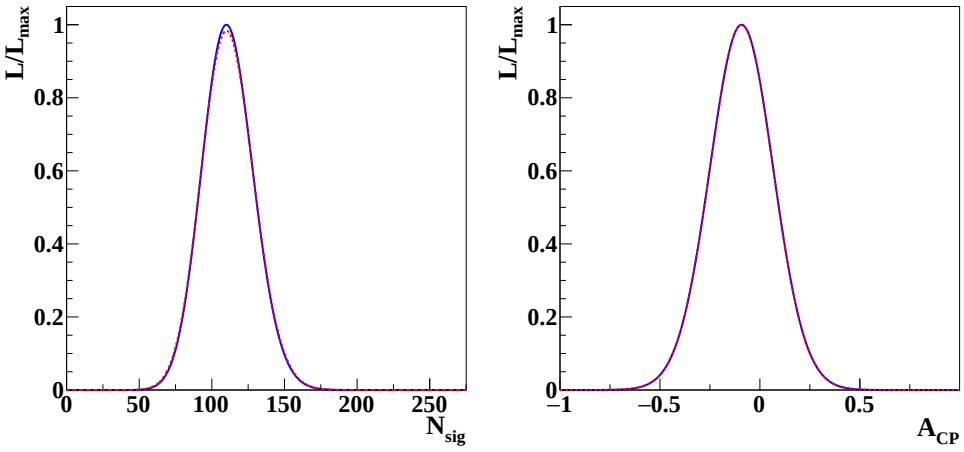
(a) $0.8 < M_{KK} < 1.1$ GeV/c^2



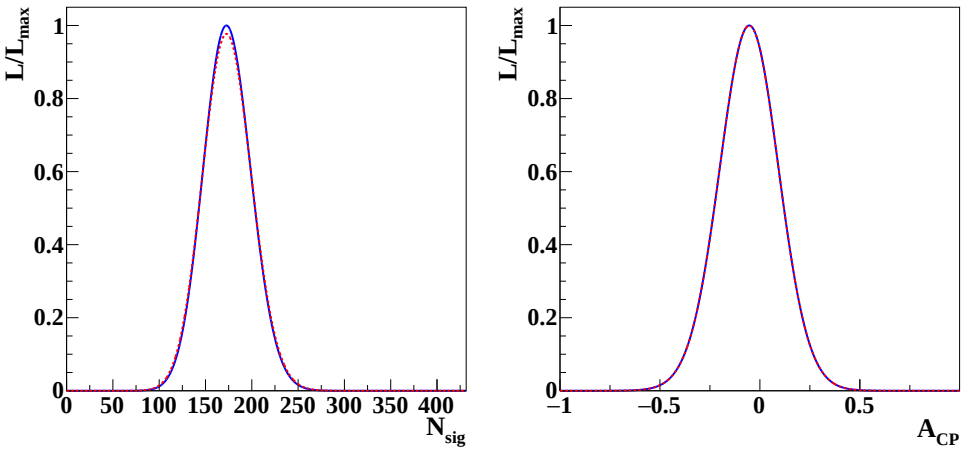
(b) $1.1 < M_{KK} < 1.5$ GeV/c^2



(c) $1.5 < M_{KK} < 2.5$ GeV/c^2



(d) $2.5 < M_{\mathrm{KK}} < 3.5\ \mathrm{GeV}/c^2$



(e) $3.5 < M_{\mathrm{KK}} < 5.3\ \mathrm{GeV}/c^2$

Figure 5.4.: Likelihood ratio distribution for signal yields(left) and $\mathcal{A}_{CP}$(right) in individual bins. The blue histogram shows plain likelihood ratio, and the red histogram shows convoluted likelihood ratio.

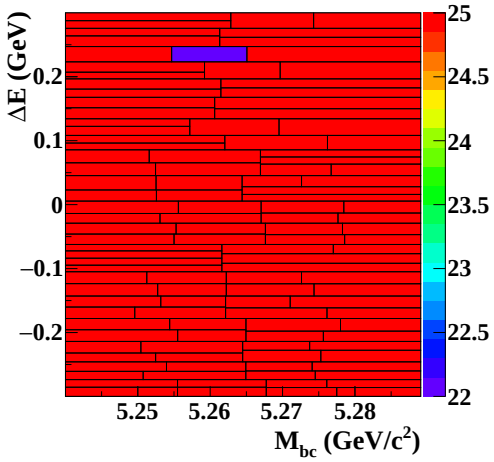


Figure 5.5.: Example of the adaptive binning in the region of $0.8 < M_{\mathrm{KK}} < 1.1$ GeV/c^2. Each cell contains approximately 25 events.

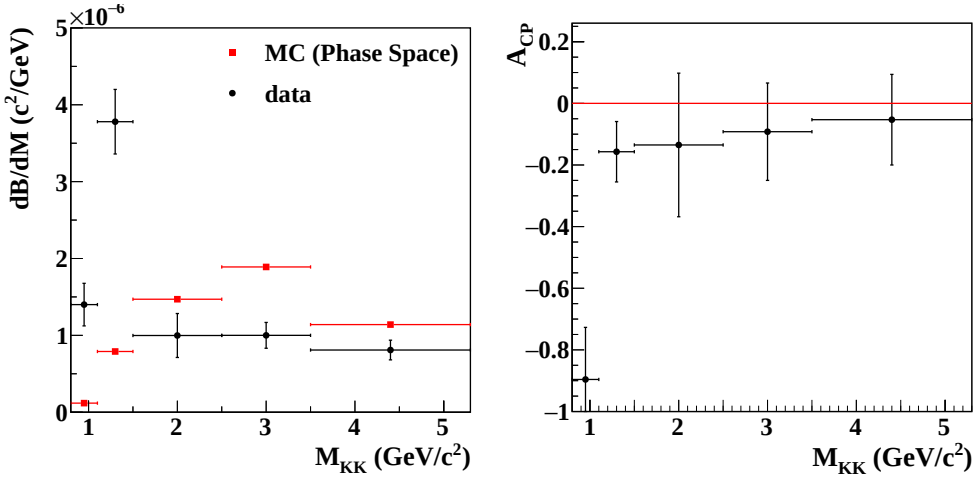


Figure 5.6.: Differential branching fractions (left) and measured $\mathcal{A}_{CP}$ (right) as a function of M_{KK}. Each point is obtained from a two-dimensional fit with systematic uncertainty included. Red squares with error bars in the left figure show the expected signal distribution in a three-body phase space MC. Note that the phase space hypothesis is rescaled to the total observed total $B^+ \to K^+K^-\pi^+$ signal yield.

5.2. Conclusion

We present the measured branching fraction and direct *CP* asymmetry for the suppressed decay $B^+ \to K^+K^-\pi^+$ using the full $\Upsilon(4S)$ data sample collected with the Belle detector at the KEKB asymmetric e^+e^- collider. We employ a two-dimensional fit to determine the signal yield and $\mathcal{A}_{CP}$ as a function of $M_{\rm KK}$.

We measure an inclusive branching fraction and direct *CP* asymmetry of:

$$\mathcal{B}(B^+ \to K^+K^-\pi^+) = (5.38 \pm 0.40(\text{stat.}) \pm 0.35(\text{syst.})) \times 10^{-6} \tag{5.6}$$

and

$$\mathcal{A}_{CP} = -0.170 \pm 0.073(\text{stat.}) \pm 0.017(\text{syst.}), \tag{5.7}$$

which are consistent with the world average branching fraction and the LHCb $\mathcal{A}_{CP}$ measurement,

$$\mathcal{B}(B^+ \to K^+K^-\pi^+) = (5.0 \pm 0.5(\text{stat.}) \pm 0.5(\text{syst.})) \times 10^{-6} \tag{5.8}$$

and

$$\mathcal{A}_{CP} = -0.123 \pm 0.017(\text{stat.}) \pm 0.012 \pm 0.007(\text{syst.}). \tag{5.9}$$

We confirm the excess and local $\mathcal{A}_{CP}$ in the low $M_{\rm KK}$ region reported by LHCb, and quantify the differential branching fraction and $\mathcal{A}_{CP}$ in each K^+K^- invariant mass bin.

We find 4.8σ evidence for a negative *CP* asymmetry in the region $M_{\rm KK} < 1.1$ GeV/c^2. The measurement challenges the conventional description of direct *CP* violation since it requires large contributions from separate weak tree and strong penguin amplitudes in the same small region of phase space in order to simultaneously enhance both the yield and provide the cancellation required for such a large *CP* effect. For example, if the enhancement were due to a large final state resonance in a strong penguin diagram, there would have to be an accompanying tree-level process of the same magnitude and opposite phase to provide the almost complete cancellation observed in the measurement. Moreover, this weak contribution is only present in a very small region of phase

space, which, for example, can be provided by a narrow resonance, that couples strongly to a tree diagram. It must couple to a tree level process, via $b \to u$ transition, in order to provide the necessary change of phase between B^+ and B^-. It is not obvious what this weak process could be. To understand the origin of the low-mass dynamics, a full Dalitz analysis from experiments with a sizeable data set, such as LHCb and Belle II, will be needed in the future.

Appendix A.

${}_s\mathcal{P}lot$ Technique

The ${}_s\mathcal{P}lot$ [63] technique is a powerful tool to unfold the contribution of different components to the data distributions. ${}_s\mathcal{P}lot$ provides a covariance-weighted quantity, sWeight, for each event in the fit data set. This can be utilized to recover the marginal distribution in the fit. The sWeight is calculated according to the maximum likelihood from the fit. A fit is performed first to determine the shape parameters as well as yields, and then the subsequent fit in ${}_s\mathcal{P}lot$ fixes all the parameters except the yields. For the regular sWeight, the inverted covariance matrix and sWeight are defined as:

$$\mathbf{V}_{nj}^{-1} = \frac{\partial^2(\mathcal{L})}{\partial N_n \partial N_j} = \sum_{e=1}^{N} \frac{f_n(y_e) f_j(y_e)}{\sum_{k=1}^{N_s} N_k f_k(y_e)} \tag{A.1}$$

and

$${}_s\mathcal{P}_n(y_e) = \frac{\sum_{j=1}^{N_s} \mathbf{V}_{nj} f_j(y_e)}{\sum_{k=1}^{N_s} N_k f_k(y_e)}, \tag{A.2}$$

where

- N is the total number of events in the data sample,
- N_s is the number of species in the data sample,
- N_j is the expected number of events on average for the j^{th} species,
- f_j is the PDF of the discriminating variables for the j^{th} species,
- $f_j(y_e)$ is the value taken by f_j for event e, the later being associated with a set of values y_e of the discriminating variables.

In this analysis, the yields of Rare B components are fixed, hence the corresponding quantities become:

$$\mathbf{V}_{nj}^{-1} = \sum_{e=1}^{N} \frac{f_n f_j}{(\sum_{k=1}^{N_s} N_k f_k + N_0 f_0)^2} \tag{A.3}$$

and

$${}_s\mathcal{P}_n(y_e) = \frac{\sum_{j=1}^{N_s} \mathbf{V}_{nj} f_j(y_e)}{\sum_{j=1}^{N_s} N_k f_k(y_e) + N_0 f_0}, \tag{A.4}$$

where "0" denotes the fixed species in the fit. For the marginal distribution of a fixed species which is well known, the species dependent coefficient,

$$c_n = N_n - \sum_j \mathbf{V}_{nj}, \tag{A.5}$$

is the weighting factor for the fixed species which can be used to deduct the contribution of the fixed species from the sWeight distribution. On the other hand, if the fixed species is poorly known, the sWeight for species '0' is defined as:

$${}_s\mathcal{P}_0 \equiv 1 - \sum_i {}_s\mathcal{P}_i, \tag{A.6}$$

and the Extended sWeight is redefined as

$${}_{es}\mathcal{P}_n \equiv {}_s\mathcal{P}_n + \frac{N_i - \sum_j \mathbf{V}_{ij}}{N - \sum_{i,j} \mathbf{V}_{ij}} {}_s\mathcal{P}_0. \tag{A.7}$$

Because the ${}_s\mathcal{P}lot$ package in RooStats is not able to calculate the sWeight with fixed species, we developed a patch, called mySPlot, to take the fixed species into account.

A.1. mySPlot Test

We performed a toy MC test to check mySPlot. There are three components (two floated and one fixed) with two variables (ΔE and $M_{\rm bc}$) in the toy MC data set. ΔE is used in the fit, and $M_{\rm bc}$ is recovered by the sWeight distribution provided by mySPlot. Figure A.1 shows the result of the toy MC test, the $M_{\rm bc}$ distribution is recovered by the sWeights successfully.

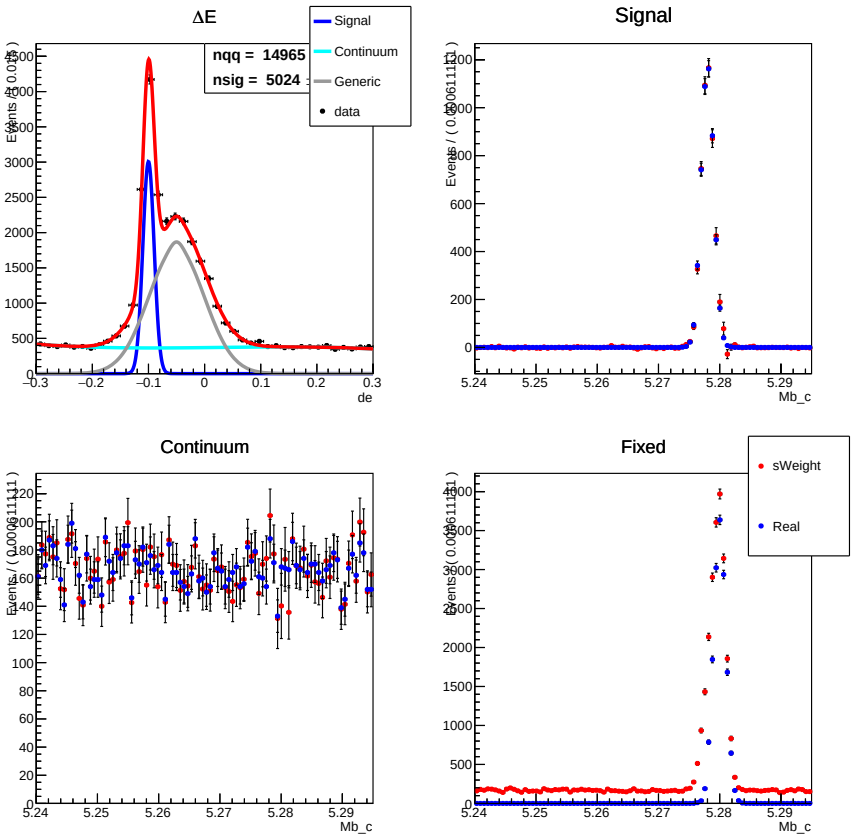


Figure A.1.: The results of mySPlot test. From left to right, top to bottom, the plots are: the fitted result of a 1-D fit to ΔE, the comparison of the sWeight distribution (red dot) and the real M_{bc} distribution (blue dot) one of of the floated species (Signal), the comparison for the other floated species (Continuum), and the real M_{bc} distribution of the fixed species (Generic) and the total M_{bc} distribution for the data set.

A.2. Correlation between variables

Although mySPlot can deal with the fixed species, the results on the GSIM MC test are not as expected. Figure A.2 shows the test results with the GSIM MC subset, the sWeight distribution on M_{KK} doesn't match the real distribution. However, if we generate a mock Generic B data set with M_{KK} uncorrelated with the fitting variables, the sWeight distribution is consistent with the M_{KK} distribution, which is shown in Fig. A.3. According to the the ${}_s\mathcal{P}lot$ paper, in the case significant correlations are observed, we may still use the sWeights. However, in our test the distribution obtained with ${}_s\mathcal{P}lot$ cannot be compared directly with the marginal distribution, so the result cannot be verified. Therefore we do not use the ${}_s\mathcal{P}lot$ method.

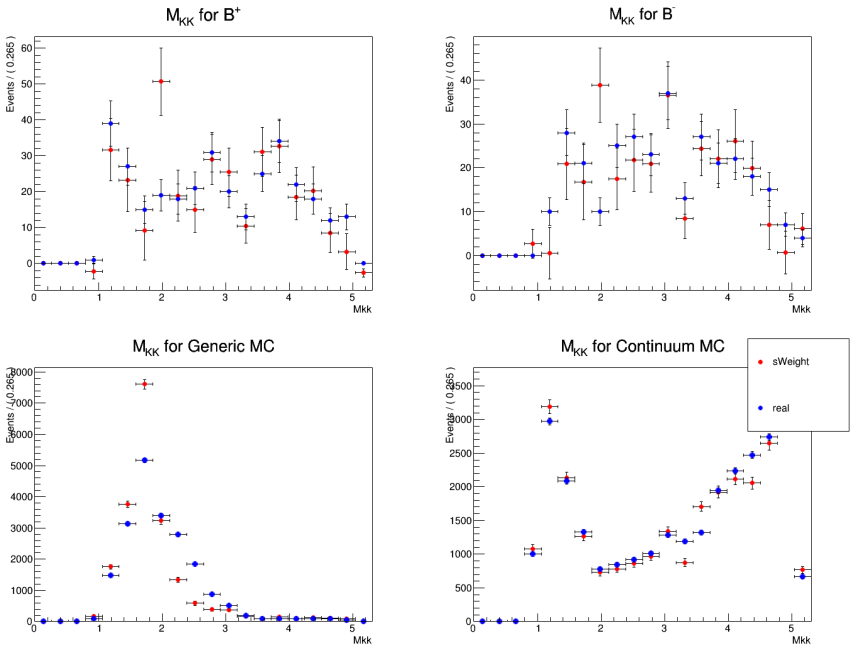


Figure A.2.: The GSIM results of ${}_s\mathcal{P}lot$ with three floated components. The sWeight distribution does not match the M_{KK} distribution in the data set.

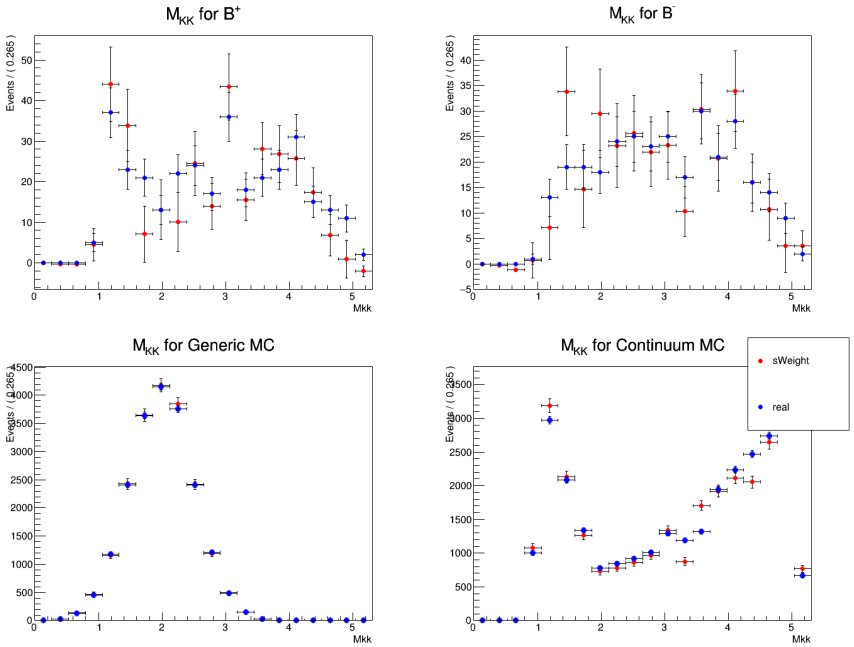


Figure A.3.: The results of $_s\mathcal{P}lot$ with a mock data set in which M_{KK} is uncorrelated with fitting variables. The sWeight distribution is consistent with the M_{KK} distribution.

Bibliography

1. Hsu, C. L., Dossett, D., Sevior, M. E., *et al.* Measurement of branching fraction and direct *CP* asymmetry in charmless $B^+ \to K^+K^-\pi^+$ decays at Belle. *Phys. Rev. D* **96,** 031101 (3 2017).
2. Patrignani, C. *et al.* Review of Particle Physics. *Chinese Physics C* **40,** 100001 (2016).
3. Sakharov, A. Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe. *Pisma Zh.Eksp.Teor.Fiz.* **5,** 32–35 (1967).
4. Lee, T. D. & Yang, C. N. Question of Parity Conservation in Weak Interactions. *Phys. Rev.* **104,** 254–258 (1 1956).
5. Wu, C. S., Ambler, E., Hayward, R. W., Hoppes, D. D. & Hudson, R. P. Experimental Test of Parity Conservation in Beta Decay. *Phys. Rev.* **105,** 1413–1415 (4 1957).
6. Christenson, J. H., Cronin, J. W., Fitch, V. L. & Turlay, R. Evidence for the 2π Decay of the K_2^0 Meson. *Phys. Rev. Lett.* **13,** 138–140 (4 1964).
7. Cabibbo, N. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.* **10,** 531–533 (1963).
8. Glashow, S. L., Iliopoulos, J. & Maiani, L. Weak Interactions with Lepton-Hadron Symmetry. *Phys. Rev. D* **2,** 1285–1292 (7 1970).
9. Kobayashi, M. & Maskawa, T. *CP*-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics* **49,** 652–657 (1973).
10. Wolfenstein, L. Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.* **51,** 1945–1947 (21 1983).
11. Herb, S. W. *et al.* Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions. *Physical Review Letters* **39,** 252–255 (Aug. 1977).
12. Bigi, I. & Sanda, A. *CP Violation* <https://books.google.com.au/books?id=ZUANziLsV-MC> (Cambridge University Press, 2000).
13. Cheng, H.-Y., Chua, C.-K. & Soni, A. Final state interactions in hadronic *B* decays. *Phys. Rev. D* **71,** 014030 (1 2005).
14. Lin, S.-W., Unno, Y., Hou, W.-S., Chang, P., *et al.* Difference in direct charge-parity violation between charged and neutral B meson decays. *Nature* **452,** 332–335 (2008).

15. *Penguin diagram* <https://en.wikipedia.org/wiki/Penguin_diagram>.
16. Bediaga, I. *et al.* On a *CP* anisotropy measurement in the Dalitz plot. *Phys. Rev. D* **80,** 096006 (9 2009).
17. Bediaga, I. *et al.* Second generation of "Miranda procedure" for *CP* violation in Dalitz studies of B (and D and τ) decays. *Phys. Rev. D* **86,** 036005 (3 2012).
18. Aaij, R. *et al.* Measurement of *CP* Violation in the Phase Space of $B^{\pm} \to K^{\pm}\pi^{+}\pi^{-}$ and $B^{\pm} \to K^{\pm}K^{+}K^{-}$ Decays. *Phys. Rev. Lett.* **111,** 101801 (10 2013).
19. Aaij, R. *et al.* Measurement of CP violation in the phase space of $B^{\pm} \to K^{+}K^{-}\pi^{\pm}$ and $B^{\pm} \to \pi^{+}\pi^{-}\pi^{\pm}$ decays. *Phys. Rev. Lett.* **112,** 011801 (2014).
20. Aaij, R. *et al.* Measurements of *CP* violation in the three-body phase space of charmless $B^{\pm}$ decays. *Phys. Rev. D* **90,** 112004 (11 2014).
21. Okubo, S. ϕ-meson and unitary symmetry model. *Physics Letters* **5,** 165 –168 (1963).
22. Zweig, G. *An SU_3 model for strong interaction symmetry and its breaking; Version 1* tech. rep. CERN-TH-401 (CERN, 1964). <http://cds.cern.ch/record/352337>.
23. Zweig, G. An SU_3 model for strong interaction symmetry and its breaking; Version 2, 80 p (1964).
24. Iizuka, J. A Systematics and Phenomenology of Meson Family*. *Progress of Theoretical Physics Supplement* **37-38,** 21–34 (1966).
25. Beneke, M. & Neubert, M. QCD factorization for $B \to PP$ and $B \to PV$ decays. *Nuclear Physics B* **675,** 333 –415 (2003).
26. Li, Y., Lü, C.-D. & Wang, W. Revisiting $B \to \phi\pi$ decays in the standard model. *Phys. Rev. D* **80,** 014024 (1 2009).
27. Gronau, M. & Rosner, J. L. B decays dominated by $\omega - \phi$ mixing. *Physics Letters B* **666,** 185 –188 (2008).
28. Grossman, Y., Ligeti, Z., Nir, Y. & Quinn, H. SU(3) relations and the CP asymmetries in B decays to $\eta' K_S$, φK_S, and $K^{+}K^{-}K_S$. *Phys. Rev. D* **68,** 015004 (1 2003).
29. Engelhard, G. & Raz, G. Using $SU(3)$ relations to bound the *CP* asymmetries in $B \to KKK$ decays. *Phys. Rev. D* **72,** 114017 (11 2005).
30. Aubert, B., Bona, M., Boutigny, D., Karyotakis, Y., Lees, J. P., *et al.* Observation of the Decay $B^{+} \to K^{+}K^{-}\pi^{+}$. *Phys. Rev. Lett.* **99,** 221801 (22 2007).
31. Aaij, R *et al.* Measurement of the charge asymmetry in $B^{\pm} \to \phi K^{\pm}$ and search for $B^{\pm} \to \phi\pi^{\pm}$ decays. *Phys.Lett.* **B728,** 85–94 (2014).
32. Aubert, B. *et al.* Search for the decay $B^{+} \to \bar{K}^{*}\,0(892)K^{+}$. *Phys.Rev.* **D76,** 071103 (2007).

33. Bhattacharya, B., Gronau, M. & Rosner, J. L. *CP* asymmetries in three-body decays to charged pions and kaons. *Physics Letters B* **726,** 337 –343 (2013).
34. Bediaga, I., Frederico, T. *CP* violation and *CPT* invariance in $B^{\pm}$ decays with final state interactions. *Phys. Rev. D* **89,** 094013 (9 2014).
35. Kurokawa, S. & Kikutani, E. Overview of the KEKB accelerators. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **499,** 1 –7 (2003).
36. Abe, T. *et al.* Achievements of KEKB. *Progress of Theoretical and Experimental Physics* **2013,** 03A001 (2013).
37. Palmer, R. B. Energy Scaling, Crab Crossing, and the Pair Problem. *SLAC-PUB-4707* (1988).
38. *Belle Internal Home Page* <http://belle.kek.jp/belle/slides/slide_index.html> (1999).
39. *Crab Cavity* <https://en.wikipedia.org/wiki/Crab_cavity>.
40. Abashian, A. *et al.* The Belle detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **479,** 117 –232 (2002).
41. Natkaniec, Z. *et al.* Status of the Belle silicon vertex detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **560.** Proceedings of the 13th International Workshop on Vertex Detectors, 1 –4 (2006).
42. Tsai, Y.-S. Pair production and bremsstrahlung of charged leptons. *Rev. Mod. Phys.* **46,** 815–851 (4 1974).
43. Ushiroda, Y. *et al.* Development of the central trigger system for the BELLE detector at the KEK B-factory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **438,** 460 –471 (1999).
44. Hojo, T., Hazumi, M. & Itoh, R. Level 3 trigger at Belle. *Belle Note* **421** (2001).
45. Nishida, S. Study of kaon and pion identification using inclusive D^* sample. *Belle Note* **779** (2005).
46. *PID Joint Group* <http://belle.kek.jp/group/pid_joint/>.
47. Lange, D. J. The EvtGen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **462,** 152–155 (2001).
48. Brun, R. GEANT 3.21. *CERN Report DD/EE/84-1* (1984).
49. Olive, K. *et al.* Review of Particle Physics. *Chinese Physics C* **38,** 090001 (2014).
50. Feindt, M. A Neural Bayesian Estimator for Conditional Probability Densities. arXiv: physics/0402093 (2004).

51. Fox, G. C. & Wolfram, S. Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes. *Phys. Rev. Lett.* **41,** 1581–1585 (23 1978).

52. Lee, S. H., Suzuki, K., *et al.* Evidence for $B^0 \to \pi^0\pi^0$. *Phys. Rev. Lett.* **91,** 261801 (26 2003).

53. Fisher, R. A. The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics* **7,** 179–188 (1936).

54. Kakuno, H., Hara, K., Casey, B., Chen, K.-F., *et al.* Neutral flavor tagging for the measurement of mixing-induced violation at Belle. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **533,** 516 –531 (2004).

55. Dalitz, R. CXII. On the analysis of τ-meson data and the nature of the τ-meson. *Philosophical Magazine Series 7* **44,** 1068–1080 (1953).

56. Verkerke, W. & Kirkby, D. The RooFit toolkit for data modeling. eprint: arXiv:physics/0306116 (2003).

57. Albrecht, H. *et al.* Search for $b \to s\gamma$ in exclusive decays of B mesons. *Physics Letters B* **229,** 304 –308 (1989).

58. Antcheva, I. *et al.* ROOT-A C++ framework for petabyte data storage, statistical analysis and visualization. *Computer Physics Communications* **180.** 40 YEARS OF CPC: A celebratory issue focused on quality software for high performance, grid and novel computing architectures, 2499 –2512 (2009).

59. Bhuyan, B. High PT Tracking Efficiency Using Partially Reconstructed D^* Decays. *Belle Note* **1165** (2010).

60. Brodzicka, J., Browder, T., Chang, P., *et al.* Physics achievements from the Belle experiment. *Progress of Theoretical and Experimental Physics* **2012,** 04D001 (2012).

61. F.R.S., K. P. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **50,** 157–175 (1900).

62. Demortier, L. A Convolution Method for Folding Systematic Uncertainties into Likelihood Functions (2005).

63. Pivk, M. & Diberder, F. L. : A statistical tool to unfold data distributions. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **555,** 356 –369 (2005).