國立臺灣大學理學院物理學研究所

碩士論文

Department of Physics

College of Science

National Taiwan University

Master Thesis

測量 Belle II 中 $B^0 \to K^0_S\pi^0$ 之衰變率及宇稱不對稱性

Time-Integrated Measurement of the Branching Fraction and Direct *CP* Asymmetry of $B^0 \to K^0_S\pi^0$ Decay at Belle II

林紓平

Shu-Ping Lin

指導教授: 張寶棣 博士

Advisor: Pao-Ti Chang, Ph.D.

中華民國 112 年 5 月

May, 2023

國立臺灣大學碩士學位論文
口試委員會審定書
MASTER'S THESIS ACCEPTANCE CERTIFICATE
NATIONAL TAIWAN UNIVERSITY

測量 Belle II 中 $B^0 \to K^0_S\pi^0$ 之衰變率及宇稱不對稱性

Time-integrated measurement of the branching fraction and direct CP asymmetry of $B^0 \to K^0_S\pi^0$ Decay at Belle II

本論文係 林紓平 (R09222092) 在國立臺灣大學 物理學研究所 完成之碩士學位論文，於民國112年3月27日承下列考試委員審查通過及口試及格，特此證明。

The undersigned, appointed by the Department of Physics on 27, March, 2023 have examined a Master's thesis entitled above presented by Shu-Ping Lin (R09222092) candidate and hereby certify that it is worthy of acceptance.

口試委員 Oral examination committee:

張寶棣

（指導教授Advisor）

徐靜戈　　張敏娟

系主任/所長Director: 張寶棣

摘要

本論文紀錄了 Belle II 實驗中 $B^0 \to K^0_S\pi^0$ 之衰變率及宇稱不對稱性的測量，透過非訊號端之 B 介子 (B_{tag}) 衰變的風味，使用時間積分方法來測量直接宇稱不對稱性參數。本測量使用 Belle II 於 2019-2022 年測量的數據，由 e^+e^- 對撞器 SuperKEKB 在 $\Upsilon(4S)$ 能量碰撞產生，總光度為 362 fb^{-1}，得到

$$\mathcal{B}(B^0 \to K^0\pi^0) = (10.16 \pm 0.65 \pm 0.65) \times 10^{-6},$$

$$\mathcal{A}(B^0 \to K^0\pi^0) = -0.06 \pm 0.15 \pm 0.05,$$

$$I_{K\pi} = -0.03 \pm 0.13 \pm 0.05.$$

其中誤差項第一項為統計誤差，第二項為系統誤差。加法原理之計算合併其他 Belle II 實驗中 $B \to K\pi$ 衰變率及宇稱不對稱性的測量，計算結果符合標準模型之預測 ($I_{K\pi} = 0$)。

關鍵字： Belle II 實驗、B 介子、B 衰變、量子色動力學

Abstract

We report updated measurements of the branching ratio and the direct *CP*-violating asymmetry in $B^0 \to K^0_S\pi^0$ decay at Belle II. The flavour information of the decay is obtained from the pair-produced neutral B partner B_{tag}, and the direct *CP* asymmetry parameter is determined with the time-integrated method. We analyse the 2019-2022 Belle II data from e^+e^- collisions produced by the asymmetric-energy SuperKEKB collider at the $\Upsilon(4S)$ resonance. The data analysed contain 387 million $B\overline{B}$ pairs and correspond to an integrated luminosity of $362\,\text{fb}^{-1}$. We obtain the branching fraction and the direct *CP* asymmetry

$$\mathcal{B}(B^0 \to K^0\pi^0) = (10.16 \pm 0.65 \pm 0.65) \times 10^{-6},$$

$$\mathcal{A}(B^0 \to K^0\pi^0) = -0.06 \pm 0.15 \pm 0.05,$$

where the first uncertainties are statistical and the second systematic. With other Belle II $B \to K\pi$ measurements of branching fractions and direct *CP* asymmetries, we compute the $K\pi$ isospin sum rule to be

$$I_{K\pi} = -0.03 \pm 0.13 \pm 0.05,$$

which is in agreement with the standard model prediction of zero.

Keywords: Belle II, B meson, B decay, quantum chromodynamics

Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 The Standard Model

The standard model (SM) describes the elementary particles given in Tab. 1.1 and their interactions. The SM is mathematically self-consistent and has continuously succeeded in predicting experimental results. However, there are some physical phenomena left unexplained. In particular, baryogenesis, the physical process leading to the predominance of matter over antimatter. The baryogenesis incoporated in the SM does not sufficiently account for the observed baryon asymmetry. *CP* violation is one of the conditions that must be satisfied for matter and antimatter to be produced at different rates [1].

Table 1.1: Elementary particles of the SM.

	fermions			bosons
quarks	u	c	t	g
	d	s	b	γ
leptons	e	μ	τ	Z
	ν_e	ν_μ	ν_τ	W

In the SM, *CP* violation is based on the complex phase in the Cabibbo-Kobayashi-Maskawa (CKM) matrix [2], [3], which describes the mismatch of quark mass eigenstates and weak eigenstates, the quantum states for propagating freely in time and participating in weak interactions, respectively. The CKM matrix relates the mass eigenstates to the weak eigenstates,

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1.1}$$

where q' denotes the weak eigenstates, and q denotes the mass eigenstates. It requires four parameters to define the CKM matrix. The standard parameterisation uses three Euler

angles ($\theta_{12}, \theta_{13}, \theta_{23}$) and one *CP* violation phase (δ_{13}) [4],

$$\begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix}, \tag{1.2}$$

where c_{ij} and s_{ij} denotes the cosine and sine of the angle θ_{ij}, respectively. Although the SM is able to explain the observed *CP* violation of quarks, it is not sufficient to account for the baryon asymmetry. There might yet be undiscovered *CP* violating processes beyond the SM.

1.2 *CP* Violation in the B Meson System

The first experimental evidence of *CP* violation was observed in the neutral kaon system [5]. Whilst neutral B and K systems are described by the same underlying theory, the experimental features are quite different due different lifetime differences of the mass eigenstates. The lifetimes of the two mass eigenstates are practically identical for neutral B mesons, whilst for neutral K mesons they are almost three orders of magnitudes different [6].

In experiments, B mesons are produced via strong interactions, that is, in flavour eigenstates B^0 and $\overline{B}^0$. The Hamiltonian of the time evolution of B^0 and $\overline{B}^0$ can be written as

$$\mathcal{H} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} = M - \frac{i}{2}\Gamma = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}, \tag{1.3}$$

where both M and Γ are Hermitian. The heavy and light mass eigenstates are

$$|B_{H,L}\rangle = p|B^0\rangle \mp q|\overline{B}^0\rangle, \tag{1.4}$$

with

$$\left(\frac{q}{p}\right)^2 = \frac{M_{12}^* - i/2\Gamma_{12}^*}{M_{12} - i/2\Gamma_{12}}. \tag{1.5}$$

The time evolution of $|B_{H,L}\rangle$ is described by

$$|B_{L,H}\rangle(t) = e^{-i\mathcal{H}_{H,L}t}|B_{H,L}\rangle(0) = e^{-\frac{1}{2}\Gamma_{H,L}t}e^{-iM_{H,L}t}|B_{H,L}\rangle(0), \tag{1.6}$$

where $\mathcal{H}_{H,L} = H_{11} \pm \sqrt{H_{12}H_{21}}$ under CPT symmetry. This can be interpreted as two neutral mesons with decay constants $\Gamma_{H,L}$. In general, *CP* violation can be observed as three distinct effects:

- direct *CP* violation resulting from interference between various terms contributing to the decay amplitude,
- mixing-induced *CP* violation from interference between different oscillation paths of the neutral mesons in the decay process,
- *CP* violation in the interference between decays to a common final state, with and without mixing.

In the SM, the effect of *CP* violation due to $B^0 - \overline{B}^0$ mixing is small, however, *CP* violating effects in the interference between decays to the same *CP* eigenstate can be relatively large.

Table 1.2: Properties of the B meson [6].

Mean mass m_{B^0}	$(5279.66 \pm 0.12)\,\text{MeV}/c^2$
$m_{B^0_H} - m_{B^0_L}$	$(3.334 \pm 0.013) \times 10^{-10}\,\text{MeV}/c^2$
Mean lifetime τ_{B^0}	$(1.519 \pm 0.004)\,\text{ps}$
$\tau_{B^0}/\tau_{\overline{B}^0}$	1.000 ± 0.012

1.3 $K\pi$ Isospin Sum Rule

The direct *CP* asymmetry and branching fraction of $B^0 \to K^0\pi^0$ contribute to an isospin sum rule ($I_{K\pi}$) test [7] for a set of $B \to K\pi$ decays, including $B^0 \to K^+\pi^-$, $B^+ \to K^0\pi^+$, $B^+ \to K^+\pi^0$, and $B^0 \to K^0\pi^0$,

$$\begin{aligned} I_{K\pi} =& \mathcal{A}_{K^+\pi^-} + \mathcal{A}_{K^0\pi^+} \cdot \frac{\mathcal{B}_{K^0\pi^+}}{\mathcal{B}_{K^+\pi^-}} \frac{\tau_{B^0}}{\tau_{B^+}} \\ & - 2\mathcal{A}_{K^+\pi^0} \cdot \frac{\mathcal{B}_{K^+\pi^0}}{\mathcal{B}_{K^+\pi^-}} \frac{\tau_{B^0}}{\tau_{B^+}} - 2\mathcal{A}_{K^0\pi^0} \cdot \frac{\mathcal{B}_{K^0\pi^0}}{\mathcal{B}_{K^+\pi^-}}, \end{aligned} \tag{1.7}$$

where $\mathcal{A}_{K\pi}$ and $\mathcal{B}_{K\pi}$ denote the direct *CP* asymmetry and branching fraction of a $B \to K\pi$ decay, and τ_{B^+} and τ_{B^0} are the lifetimes of the charged and neutral B mesons. Derivation of the sum rule is briefly discussed in App. A.1. The sum rule linearly combines the direct *CP* asymmetries and branching fractions and is predicted to be zero in the SM with a precision of better than 1% [8], thus providing a way to probe the SM [9]. The current world average of the sum rule is $I_{K\pi} = -0.13 \pm 0.11$, evaluated with averages of current measurements [10], and the Belle result is $I_{K\pi} = -0.270 \pm 0.132 \pm 0.060$ [11]. The precision of the isospin sum rule test is limited by the uncertainty of $\mathcal{A}_{K^0\pi^0}$.

By convention, the direct *CP* asymmetry is defined as

$$\mathcal{A} = \frac{\Gamma(\overline{B} \to \bar{X}) - \Gamma(B \to X)}{\Gamma(\overline{B} \to \bar{X}) + \Gamma(B \to X)}, \tag{1.8}$$

where Γ is the decay widths to some final state X of a B meson decay. For $B^0 \to K^0\pi^0$, the decay rate is described a time-dependent function

$$\mathcal{P}(\Delta t, q) = \frac{e^{-|\Delta t|/\tau_B^0}}{4\tau_B^0}\{1 + q \times [\mathcal{S}\sin(\Delta m \Delta t) + \mathcal{A}\cos(\Delta m \Delta t)]\}, \tag{1.9}$$

where $\mathcal{S}$ and $\mathcal{A}$ are the mixing-induced and direct *CP* asymmetries, respectively, τ_B^0 is the B^0 lifetime, Δm is the mass difference between the two B^0 mass eigenstates, Δt is the decay time difference between the signal- and tag-side B^0, and $q = +1(-1)$ when the tag-side B meson is a $B^0(\overline{B}^0)$. Integration over the decay time in Eqn. 1.9 results in

$$\mathcal{P}(q) = \frac{1}{2}[1 + (1 - 2\chi_d) \times q\mathcal{A}], \tag{1.10}$$

where χ_d is the time-integrated B^0 mixing parameter. The asymmetry $\mathcal{A}$ is diluted by the time-integrated mixing probability $(1 - 2\chi_d)$.

1.4 $B^0 \to K_S^0\pi^0$ Decay

In $B^0 \to K_S^0\pi^0$ decay, both B^0 and $\overline{B}^0$ decay to the same *CP* eigenstate $K_S^0\pi^0$. K_S^0 is a good approximation of the $CP = 1$ eigenstate of the neutral kaon system with spin-parity $J^P = 0^-$, neglecting *CP* violation in neutral kaon mixing. Since both B^0 and π^0 have $J^P = 0^-$ as well, the decays $B^0 \to K_S^0\pi^0$ and $\overline{B}^0 \to K_S^0\pi^0$ must result in an $l = 0$ state to conserve angular momentum. The overall parity of the final state is therefore

$$P(K_S^0\pi^0) = P(K_S^0) \times P(\pi^0) \times (-1)^l = (-1)(-1)(-1)^0 = +1. \tag{1.11}$$

The flavour wavefunction of π^0 is

$$|\pi^0\rangle = \frac{1}{\sqrt{2}}(u\overline{u} - d\overline{d}), \tag{1.12}$$

which means that π^0 has $C = +1$, and $CP = -1$. Therefore,

$$C(K_S^0\pi^0) = C(K_S^0) \times C(\pi^0) = (-1)(+1) = -1, \tag{1.13}$$

and the *CP* state of the final state is

$$CP(K_S^0\pi^0) = CP(K_S^0) \times CP(\pi^0) \times (-1)^l = (+1)(-1)(-1)^0 = -1. \tag{1.14}$$

The $B^0/\overline{B}^0$ first decays to the flavour eigenstates, $B^0 \to K^0\pi^0$ and $\overline{B}^0 \to \overline{K}^0\pi^0$, and the neutral kaon evolves as a linear combination of the mass eigenstates K_S^0 and K_L^0, which then decay as K_S^0 and K_L^0.

In Tab. 1.3 we summarise previous measurements of the branching fraction and direct

CP asymmetry by BaBar [12], [13] and Belle [11], [14], as well as the Belle II preliminary result at an integrated luminosity of $62.8\,\text{fb}^{-1}$, using also a time-integrated method [15].

Table 1.3: Belle and BaBar measurements of the branching fraction and direct *CP* asymmetry of $B^0 \to K^0\pi^0$.

	$\mathcal{B}_{K^0\pi^0}$ $[10^{-6}]$	$\mathcal{A}_{K^0\pi^0}$
BaBar	$10.1 \pm 0.6 \pm 0.4$	$-0.13 \pm 0.13 \pm 0.03$
Belle	$9.68 \pm 0.46 \pm 0.50$	$0.14 \pm 0.13 \pm 0.06$
Belle II	$8.5 \pm 1.7 \pm 1.2$	$-0.40 \pm 0.46 \pm 0.04$

1.5 Thesis Outline

In this study, we simultaneously extract the direct *CP* asymmetry and the branching fraction of $B^0 \to K^0\pi^0$ by performing a fit on ΔE-C'-q space, where ΔE is the difference between the beam energy and the B meson candidate, C' is the transformed output classifier of the multivariate analysis (MVA) for continuum background suppression, and q is the B decay flavour information provided by the flavour tagging algorithm [16]. Before unblinding the Moriond 2023 LS1 dataset, we first optimise the candidate selection and fitting strategy with Monte-Carlo (MC) simulation. We test the fitter with toys, linearity tests, and partial unblinding to validate the direct *CP* asymmetry and branching fraction measurements and to check for bias before the full unblinding.

The reconstruction and selection is documented in Ch. 3, the continuum background suppression in Ch. 4, the extraction of the direct *CP* asymmetry and the branching fraction in Ch. 5, and the data-simulation discrepancy and calibration is studied with control modes in Ch. 6. The result of the measurement is shown in Ch. 7, where we also calculate the isospin sum rule with measurements of the direct *CP* asymmetries and branching fractions of other $B \to K\pi$ decays [17]. In Ch. 8 we show the estimation of the systematic uncertainties.

Chapter 2 Belle II Experiment

Belle II [18], the first super B-Factory experiment, is designed to make precise measurements of weak interaction parameters and find New Physics (NP) beyond the SM of particle physics. It is located at Tsukuba, Ibaraki Prefecture, Japan. Its data taking commenced in 2019 and has reached an integrated luminosity of $362\,\text{fb}^{-1}$ at the $\Upsilon(4S)$ resonance. Its instantaneous luminosity set a world record of $4.7 \times 10^{34}\ \text{cm}^{-2}\text{s}^{-1}$ in June 2022. Compared to its predecessor, KEKB and the Belle detector [19], the Belle II detector has been upgraded, introducing new devices to provide better precision and competency on several measurements.

2.1 SuperKEKB Accelerator

SuperKEKB [20] is the particle collider for Belle II. It consists of two storage rings, the high energy ring for electron beam and the low energy ring for positron beam. Electron ($7\,\text{GeV}$) and positron ($4\,\text{GeV}$) beams collide asymmetrically at an angle of $83\,\text{mrad}$. The centre-of-mass energy is at the $\Upsilon(4S)$ resonance ($10.58\,\text{GeV}/c^2$). The small mass difference between the $\Upsilon(4S)$ resonance and the sum of the two B mesons signifies that the B mesons are produced almost at rest in the centre-of-mass frame, with very small relative velocity, and only the lightest B mesons, $B^+(\bar{b}u)$ and $B^0(\bar{b}d)$. At this threshold, the production of B mesons is clean and efficient.

The asymmetric beam energies result in a Lorentz boost along "forward" direction, determined by the beam with higher energy. SuperKEKB targets $8 \times 10^{35}\ \text{cm}^{-2}\text{s}^{-1}$ instantaneous luminosity, which is about 40 times larger than that of KEKB. The improvement mainly arises from the "nano-beam" scheme [21]. The beam size is reduced to 50 nm, with a large crossing angle, at the collision point to increase luminosity. The nano-beam scheme also allows for detection much closer to the interaction point (IP). To cope with the high rate of electron-positron collision provided by SuperKEKB, an upgrade of the detector is also necessary. Schematic diagrams of the SuperKEKB accelerator and the Belle II detector are shown in Fig. 2.1 and Fig. 2.2, respectively.

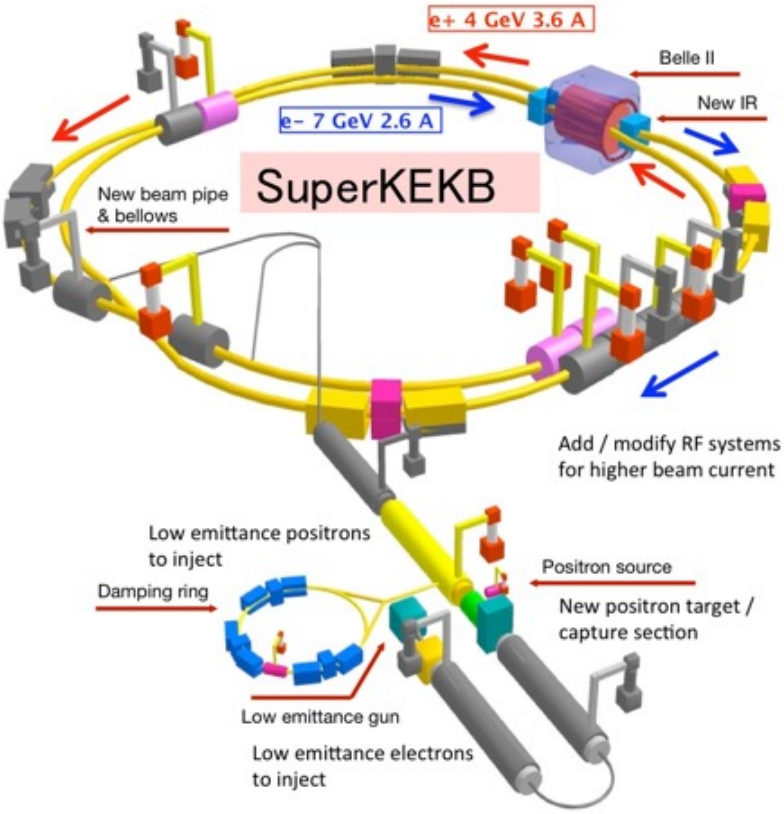


Figure 2.1: Schematic diagram of the SuperKEKB accelerator.

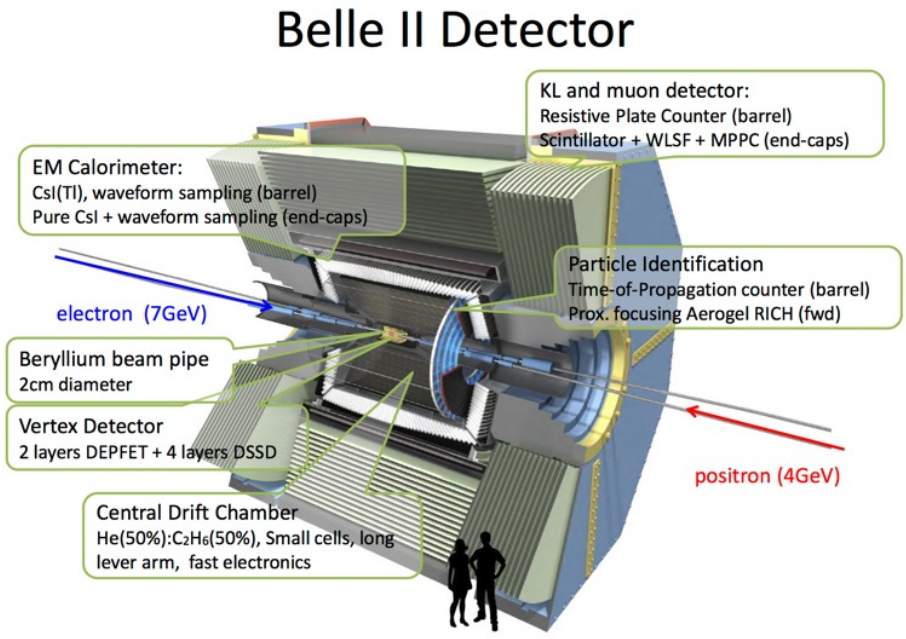


Figure 2.2: Schematic diagram of the Belle II detector.

2.2 Vertex Detector

The Belle II vertex detector [22] consists of two components: the two-layer silicon pixel detector (PXD) and the four-layer silicon vertex detector (SVD), making up a total of six layers around the beryllium beam pipe. The PXD inner layers at $r = 14\,\text{mm}$ and $r = 22\,\text{mm}$ implement DEPFET (depleted field effect transistor) pixelated sensors to provide precise charged particle trajectory measurement. The SVD outer layers at radii of $38\,\text{mm}$, $80\,\text{mm}$, $115\,\text{mm}$, and $140\,\text{mm}$ are equipped with double-sided silicon strip sensors, arranged cylindrically, with an inclined endcap in the forward direction. The layout of the PXD and the SVD are shown in Fig. 2.3. The beam pipe and the first two PXD layers are considerably closer to the IP than Belle, and the outermost layer covers a much larger range, significantly improving the vertex resolution and the reconstruction efficiency for $K^0_S \to \pi^+\pi^-$ decays. Most of the typical K^0_S flight distance is covered.

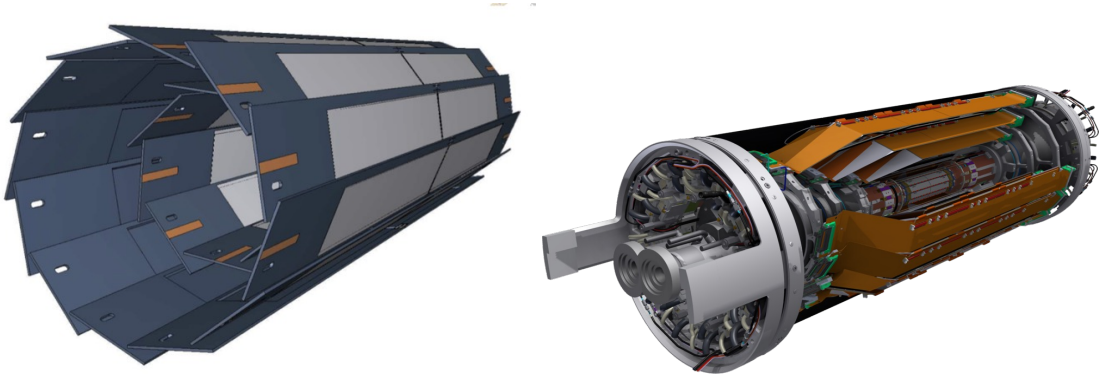


Figure 2.3: Schematic view of the PXD (left) and SVD (right) layouts [23].

2.3 Central Drift Chamber

The Central Drift Chamber (CDC) provides charged track reconstruction and momentum measurement, particle identification through measurement of energy loss within its volume, and trigger information for charged particles. The CDC is a large volume of drift chamber with small drift cells, extending to a larger radius compared to Belle due to a more condensed particle identification device in the barrel region. The CDC contains 56 layers in either axial or stereo orientations, which are aligned with the solenoidal magnetic field or skewed with respect to the axial wires, respectively. Combined information from such a configuration allows reconstruction of 3D helix tracks. The chamber gas is a He-C_2H_6 1:1 mixture.

2.4 Particle Identification

The particle identification device at a Belle II consists of two subsystems: the time-of-propagation (TOP) counter for the barrel region and the aerogel ring-imaging Cherenkov (ARICH) detector for the forward region [24]. The TOP counter provides particle identification in the barrel region, with a polar angle coverage of $32°$ to $120°$. Cherenkov photons collected at one end of a 2.6 m quartz bar give the timing and two-dimensional spatial information of the Cherenkov ring image. Cherenkov photons are emitted at different angles by different particles of the same momentum, resulting in different path lengths, and thus a different times of propagation for different particles. The TOP counter achieves single photon time resolution of about 100 ps through 16-channel micro-channel plate photomultiplier tubes. A focusing system [25] for dividing the ring image according to the Cherenkov photon wavelength is introduced to minimise the chromatic effect. A custom-made waveform sampling read-out electronics is used to improve the timing precision, achieving a precision of about 50 ps for the starting time.

For charged particle identification in the forward end-cap region, ARICH is employed with a polar angle coverage of $17°$ to $35°$. It provides a good separation of pions and kaons with momentum ranging from 0.4 GeV/c to 4 GeV/c, which covers most of their spectrum, and offers pion, muon, and electron discrimination ability for momentum lower than 1 GeV/c^2. ARICH comprises of two 2 cm layers of aerogel radiator of different refractive indices at the front for proximity focusing, and an array of photon detectors at the end. The schematic views of the TOP counter and the ARICH detector is shown in Fig. 2.4.

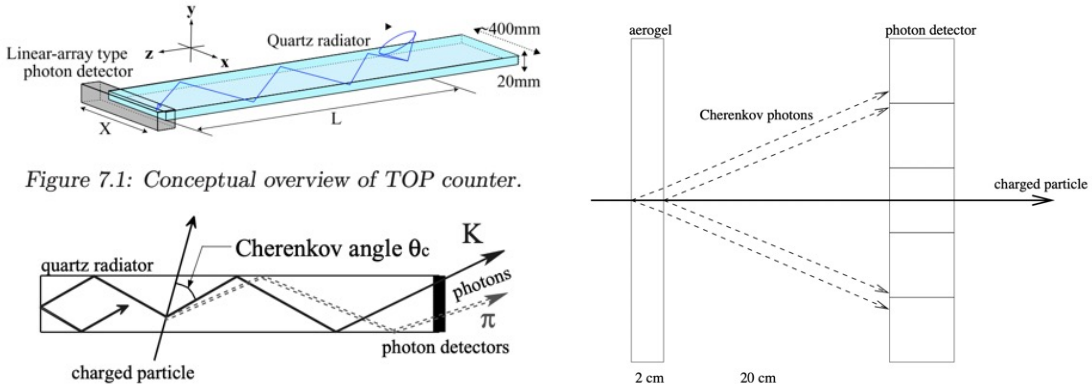


Figure 7.1: Conceptual overview of TOP counter.

Figure 2.4: Left: Schematic side-view of TOP counter and internal reflecting Cherenkov photons. Right: Proximity focusing ARICH [23].

2.5 Electromagnetic Calorimeter

The main purpose of the Electromagnetic Calorimeter (ECL) is to measure the energy of photons from π^0 decays, which could range from a few MeV to a few GeV. Fine resolution is required in order to separate two photons from a high-momentum π^0. The ECL is built of highly-segmented array of thallium-doped caesium iodide CsI(Tl) scintillator crystals, arranged in a projective geometry and covers the polar angle from $12°$ to $155°$. The ECL has been reused from Belle, but with upgraded readout electronics and reconstruction software to accommodate the high event rate of Belle II. The CsI(Tl) crystals, which has a long scintillation decay time, will be replaced with pure CsI crystals in the forward region to cope with the high background and occupancy. The layout of the ECL is shown in Fig. 2.5.

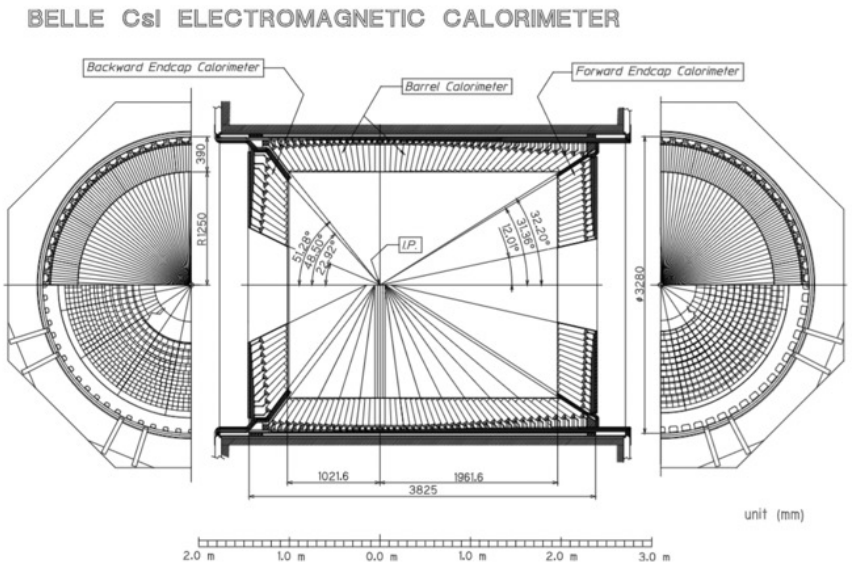


Figure 2.5: Configuration of Belle ECL [26].

2.6 K_L^0 and Muon Detector

The K_L and muon detector (KLM) provides identification of K_L^0 and muon particles with momentum $> 600\,\text{MeV}$. The detector is made of alternating layers of 4.7 cm thick iron plates, which serve as the magnetic flux return for the solenoid, and active detector elements based on glass-electrode resistive plate chambers (RPC). The iron plates also provide room for the K_L^0 mesons to shower hadronically. The KLM detector is adopted from Belle as well, with the RPC replaced by layers of scintillator strips with wavelength shifting fibres in the two innermost barrel layers. This is because the long dead time of RPC is not suitable for the high background rates expected at Belle II, caused by neutrons produced in electromagnetic showers.

Chapter 3 Reconstruction and Selections

3.1 Data Samples

In this study, we study simulated MC data of signal-only and generic events in order to optimise the selections and fitting models. A large amount of signal-only sample is generated centrally for each reconstructed decay channel. Generic simulation consists of centrally produced, run-independent `BGx1 MC15ri` generic MC samples that include $B^0\overline{B}^0$, B^+B^-, $u\overline{u}$, $d\overline{d}$, $s\overline{s}$, and $c\overline{c}$ processes, corresponding to an integrated luminosity of $1\,\text{ab}^{-1}$ at the $\Upsilon(4S)$ resonance. We analyse the 2019-2022 Belle II data from e^+e^- collisions produced by the asymmetric-energy SuperKEKB collider at the $\Upsilon(4S)$ resonance. The data analysed contain 387 million $B\overline{B}$ pairs and correspond to an integrated luminosity of $362\,\text{fb}^{-1}$. We use all good LS1 $\Upsilon(4S)$ runs of the latest official processing proc13 and prompt processings. The LS1 dataset has an integrated luminosity of $362 \pm 2\,\text{fb}^{-1}$ and consists of

- Exp 7 (proc13),
- Exp 8 (proc13),
- Exp 10 (proc13),
- Exp 12 (proc13),
- Exp 14 (proc13),
- Exp 16 (proc13),
- Exp 17 (proc13),
- Exp 18 (proc13),
- Exp 20 (prompt),

- Exp 22 (prompt),
- Exp 24 (prompt),
- Exp 26 (prompt).

All candidates meet the high level trigger (HLT) hadron skim selection criteria, which requires more than 3 tracks in the event and absence of any track pair consistent with the topology of a Bhabha event. The HLT performs a rough skimming, assigning events that satisfy the HLT conditions into categories such as Bhabha, hadronic, tau, etc. This categorisation is for quickly selecting and sending the events that require calibration to the calibration centre.

3.2 Reconstructed Modes

We reconstruct the following signal and control modes in this analysis:

- $B^0 \to K^0_S\pi^0$
- $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$

The signal-only MC consists of four million events for $B^0 \to K^0_S\pi^0$, and two million events for $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$. The signal mode $B^0 \to K^0_S\pi^0$ is used to study the reconstruction, selection, continuum suppression, and fitting procedures before the analysis is performed on real data. We use the control mode $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$, which decays to similar particles as the signal mode, to study the discrepancy between data and simulation in continuum suppression and fitting. Charge conjugation is implied throughout this document unless otherwise stated, and quantities denoted by a star symbol ($*$) are estimated in the centre-of-mass frame.

3.3 Monte-Carlo Simulation

We generate simulation samples in order to compare the result with real data from the detector, to ensure that what we observe is not some artefact of the complex experiment. Simulated events should behave just as detector events. In high-energy physics we use the MC random sampling algorithm [27] for simulations.

The simulation procedure consists of two parts: the event generation and the detector simulation. In the event generation we generate the positions and momenta of the particles resulting from the electron-positron collision using `KKMC` [28] for continuum events,

PYTHIA8 [29] to simulate hadronisation, and EvtGen [30], [31] to simulate decays of hadrons. In the detector simulation we use the simulation software GEANT4 [32]–[34] to model the detector response. Interaction of the final-state particles with a virtual Belle II detector results in deposited energy and particles produced from interaction with the subdetectors. Our simulation includes beam backgrounds [35].

In this the analysis, the signal-only sample is simulated for the signal and control channels. For the signal mode $B^0 \to K^0_S\pi^0$, the $\Upsilon(4S)$ first decay to $B^0\overline{B}^0$ with mixing enabled, then the signal-side B^0 and $\overline{B}^0$ mesons decay to $K^0_S\pi^0$. Further decays of K^0_S and π^0 are simulated according to the branching ratios of the Particle Data Group (PDG) [6]. The non-signal B^0 or $\overline{B}^0$ decays follow the PDG. For the control mode $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$, the $\Upsilon(4S)$ first decays to B^+B^-, then the signal-side B^+ (B^-) meson decays to $\overline{D}^0\pi^+$ ($D^0\pi^-$), where the neutral D mesons decay to $K^0_S\pi^0$. Further decays of K^0_S and π^0 are the same as the signal channel. The non-signal $B^\pm$ decays according to the PDG. The signal decay files are given in App. A.2.

3.4 Reconstruction and Baseline Selections

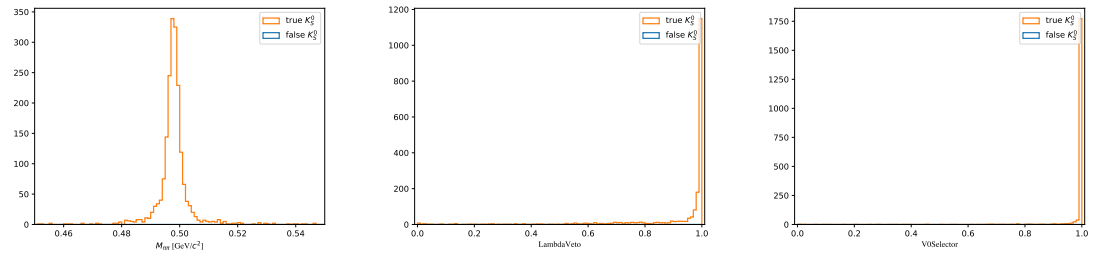
Throughout this analysis, we use the Belle II analysis framework (basf2) with release light-2210-devonrex for both signal and control channels. We use the grid-based basf2 (gbasf2) for reconstruction of MC and data samples distributed in storage sites across the world. B meson candidates are reconstructed by applying loose baseline selection criteria and combining final-state particle candidates in kinematic fits. Background events consist of continuum background events $e^+e^- \to q\bar{q}$, where q is a light quark of u, d, s, or c, and the $B\overline{B}$ background from B meson decays. The dominant background for $B^0 \to K^0_S\pi^0$ is the former. Signal-only and background MC simulation are studied to optimise the selections to be applied on data. In Tab. 4.2 we summarise the final selections applied in this analysis.

3.4.1 K^0_S Selections

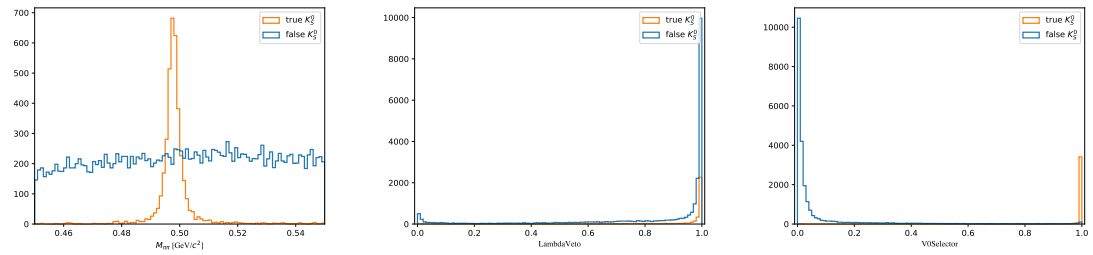
K^0_S candidates are reconstructed from the K_S0:merged standard list [36], which combines two charged tracks to form a K^0_S candidate. The charged tracks are not required to have originated from the IP. The invariant mass of the K^0_S candidates fall within the mass window $0.450 < m(\pi\pi) < 0.550$ GeV/c^2 and the K^0_S candidates are required to pass a TreeFit [37] vertex fit. The mass of the K^0_S candidates is constrained to the nominal value in the B^0 vertex fit. We then tighten the selection by applying cuts on the invariant mass, VOSelector, and LambdaVeto. The latter two are both outputs of the ksSelector [38],

which is an MVA-based module for K_S^0 selection, consisting of two selections through MVA. `V0Selector` is the output classifier of an MVA trained to discriminate real K_S^0 particles from short-lived fake ones coming from the IP, using K_S^0 and pion kinematics as input, whereas `LambdaVeto` is the output classifier of an MVA trained to suppress background from Λ baryons, with proton ID, pion kinematics, and mass as input.

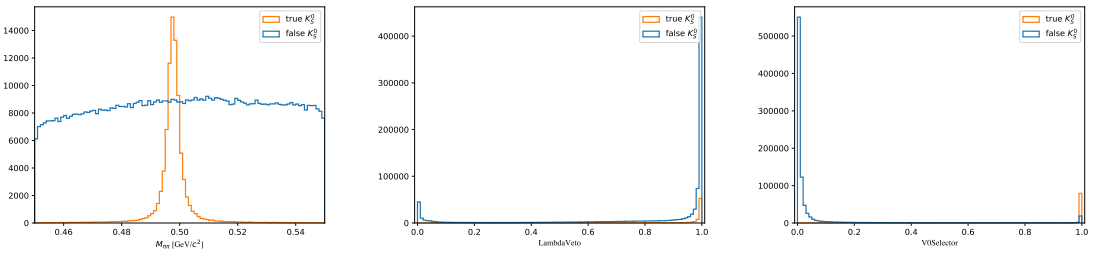
The cuts on these variables are optimised by maximising the significance in the signal-enhanced region defined by $-0.13 < \Delta E < 0.10\,\text{GeV}$. The significance is defined as $S/\sqrt{S+B}$, where S and B are the number of signal and background events, respectively. We calculate the significance with varying cuts on the K_S^0 variables, and the optimised cuts for K_S^0 are $0.480 < m(\pi\pi) < 0.510\,\text{GeV}/c^2$, `V0Selector` > 0.2, and `LambdaVeto` > 0.06. The significance is increased from 0.63 to 1.50 with these selections. Fig. 3.1 shows the distributions of K_S^0 variables of signal-only and background MC. In Tab. 3.1 we list the K_S^0 reconstruction and selection criteria.



(a) Signal MC



(b) $B\overline{B}$ background MC



(c) Continuum background MC

Figure 3.1: Distributions of K_S^0 variables. Left: K_S^0 invariant mass; middle: `LambdaVeto`; right: `V0Selector`.

Table 3.1: K_S^0 reconstruction and selection criteria.

Reconstruction	Selection
$0.450 < m(\pi\pi) < 0.550\,\text{GeV}/c^2$	$0.480 < m(\pi\pi) < 0.510\,\text{GeV}/c^2$
–	$\texttt{LambdaVeto}(K_S^0) > 0.06$
–	$\texttt{V0Selector}(K_S^0) > 0.2$

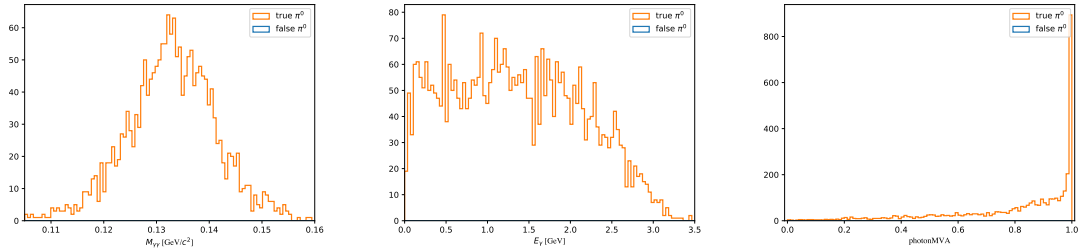
3.4.2 π^0 Selections

γ candidates are reconstructed from the `gamma:pi0eff60_May2020` standard list [39], which consists of ECL clusters that are within the polar angle range $0.2967 < \theta < 2.6180$ [rad] and that are not associated to a track. The clusters must satisfy $E > 0.0225$ GeV for γ candidates in the forward region, and $E > 0.020$ GeV for those in the barrel and backward regions. Two γ candidates are then combined to form a π^0 candidate. We constrain the invariant mass of the π^0 candidates to the nominal value through a mass-constrained fit using `kFit` [40].

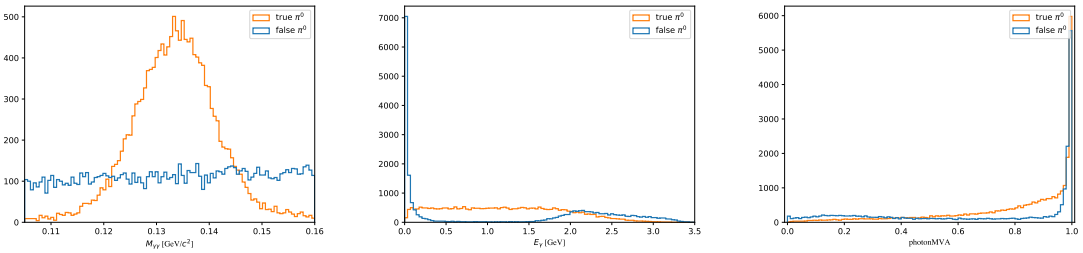
After the reconstruction, we apply the π^0 selections developed for $B^0 \to \pi^0\pi^0$ analysis [41], where the π^0 candidates exhibit similar kinematic properties as those from $B^0 \to K_S^0\pi^0$. We drop the π^0 momentum, `abs(cosHelicityAngleMomentum)`, and `abs(daughterDiffofPhi)` requirements, which have only negligible impact on the $B^0 \to \pi^0\pi^0$ analysis. We require the invariant mass of the π^0 candidate to be within the mass window $0.115 < m(\gamma\gamma) < 0.150\,\text{GeV}/c^2$, the γ cluster energy to be greater than 0.03 GeV, the cluster timing to be within 200 ns of the hardware trigger time ($|\texttt{clusterTiming}| < 200$ ns), the weights of all crystals within an ECL cluster to be above 1.5 ($\texttt{clusterNHits} > 1.5$), the angle between the two clusters used in reconstructing π^0 to be below 0.4 [rad] ($\texttt{daughterAngle} < 0.4$), and that the output of `photonMVA` to be greater than 0.20. `photonMVA` is trained for $B^0 \to \pi^0\pi^0$ to suppress hadronic interactions and photons from non-signal sources and to distinguish between real and mis-reconstructed photons using ECL variables [41]. The significance is further increased to 1.71 with these selections. Fig. 3.2 shows the distributions of π^0 variables of signal-only and background MC. In Tab. 3.2 we list the π^0 reconstruction and selection criteria.

Table 3.2: π^0 reconstruction and selection criteria.

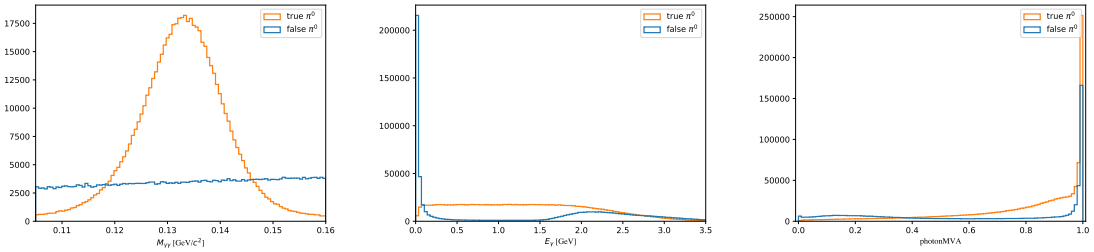
Reconstruction	Selection
$E(\gamma) > 0.0225\,\text{GeV}$ (forward)	$E(\gamma) > 0.03\,\text{GeV}$
$E(\gamma) > 0.0200\,\text{GeV}$ (barrel)	$\lvert\texttt{clusterTiming}(\gamma)\rvert < 200\,\text{ns}$
$E(\gamma) > 0.0200\,\text{GeV}$ (backward)	$\texttt{photonMVA}(\gamma) > 0.20$
$0.02967 < \theta(\gamma) < 2.6180$ [rad]	$0.115 < m(\gamma\gamma) < 0.150\,\text{GeV}/c^2$
$\texttt{clusterNHits}(\gamma) > 1.5$	$\texttt{daughterAngle}(\pi^0) < 0.4$ [rad]



(a) Signal MC



(b) $B\overline{B}$ background MC



(c) Continuum background MC

Figure 3.2: Distributions of π^0 variables. Left: π^0 invariant mass; middle: γ energy; right: `photonMVA`.

3.4.3 B^0 Selections

Finally, K^0_S and π^0 candidates are combined to form B^0 candidates, on which we perform a vertex fit. Two important variables in B factories are the difference between the beam energy and the B meson candidate ΔE, and the beam-constrained mass M_{bc}, defined as

$$\Delta E = E^*_B - \sqrt{s}/2, \tag{3.1}$$

$$M_{\text{bc}} = \sqrt{s/(4c^4) - (p^*_B/c)^2}, \tag{3.2}$$

where p^*_B and E^*_B are the three-momentum and energy of the B meson in the centre-of-mass frame, and s is the square of centre-of-mass energy. Channels that contain high-energy neutral particles have highly-correlated ΔE and M_{bc}, as both variables have the energies of the neutral particles as input. The energy loss in the ECL due to hadronic shower results in tails in both ΔE and M_{bc} distributions. In order to reduce this correlation,

we replace the B meson momentum $p_B = p_{K_S^0} + p_{\pi^0}$ in the calculation of M_{bc} with

$$p_B = p_{K_S^0} + \sqrt{(E_{\text{beam}} - E_{K_S^0})^2 - M_{\pi^0}^2} \cdot \frac{p_{\pi^0}}{|p_{\pi^0}|}, \tag{3.3}$$

where $p_{K_S^0}$ and $E_{K_S^0}$ are the momentum and energy of the K_S^0 candidate, p_{π^0} and M_{π^0} are the π^0 momentum and its nominal mass, and E_{beam} is the beam energy [11]. The π^0 momentum is scaled by the deduced value from the beam energy and K_S^0 energy, which are more precisely measured, effectively mitigating the lossy effect on M_{bc} and thus decreasing its corelation with ΔE. In Fig. 3.3 we show the distributions of M_{bc}-ΔE and modified M_{bc}-ΔE for the signal, $B\overline{B}$ background, and continuum background simulations. We will use the modified M_{bc} instead of the classical M_{bc} in the following analysis. For convenience, we will refer to the modified M_{bc} as M_{bc} for the rest of this document.

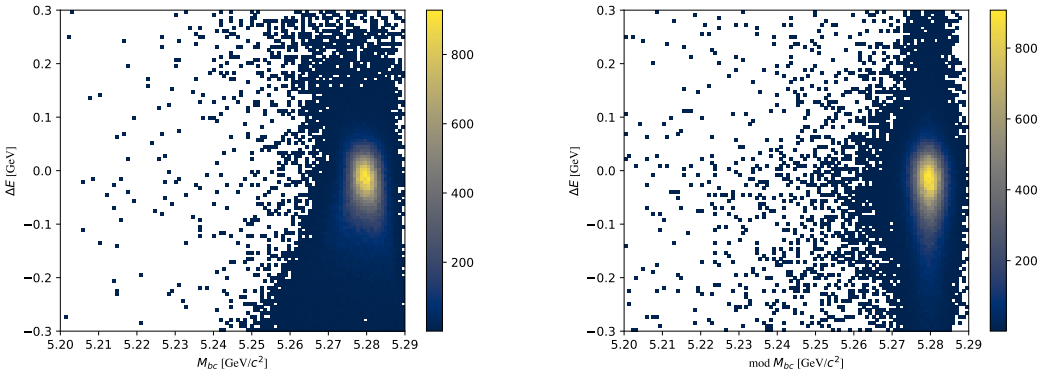
The reconstructed B^0 candidates are required to fall within the mass window $5.272 < M_{\text{bc}} < 5.288\,\text{GeV}/c^2$, which corresponds to a 3σ window around the nominal B meson mass, as shown in Fig. 3.4. We apply a loose requirement on the energy difference $-0.3 < \Delta E < 0.3\,\text{GeV}$, due to the large radiative tail in the signal ΔE distribution in decays that contain a π^0. The vertex of the accompanying tag-side B mesons are reconstructed using the `kFit` algorithm [40] with the `Btube` constraint, and we apply tag track selection of the most recent recommendations by the TDCPV subgroup [42], summarised in Tab. 3.3.

Table 3.3: The standard TDCPV tag track selection.

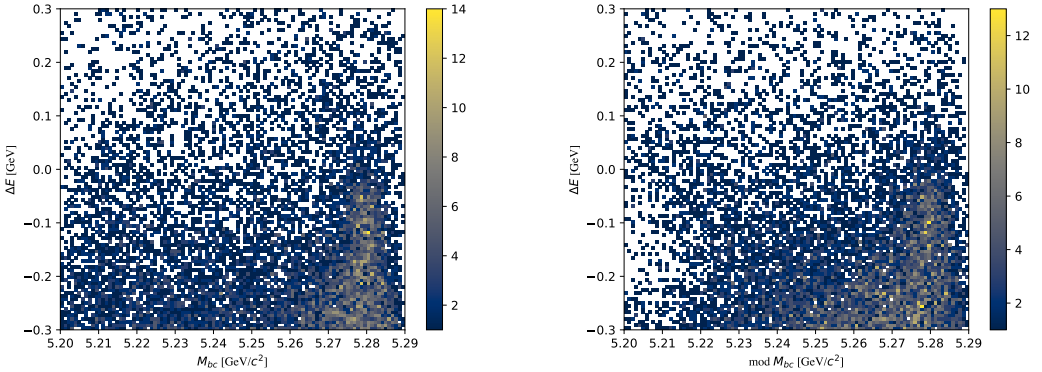
$dr < 0.5$ cm
$\lvert dz\rvert < 2$ cm
`nCDCHits` > 0
`nSVDHits` > 0
$p \geq 0.05\,\text{GeV}/c$

3.4.4 Control Channel Selections

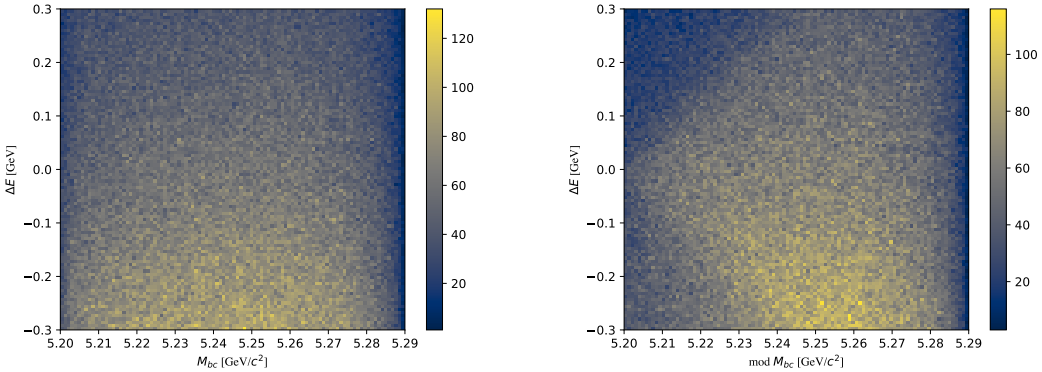
We reconstruct the control channel $B^+ \to \overline{D}^0(\to K_S^0\pi^0)\pi^+$ in order to study the discrepancy between data and simulation. The K_S^0 and π^0 candidates have the same selections as those in the signal channel $B^0 \to K_S^0\pi^0$. The charged pions are reconstructed from the `pi+:all` list while requiring that `thetaInCDCAcceptance` $=$ `true`, `nCDCHits` > 20, $dr < 0.5$ cm, and $|dz| < 2$ cm. The D^0 candidates are reconstructed by combining K_S^0 and π^0 candidates with a kinematic fit. Only candidates with an invariant mass within a 3σ window around the nominal D^0 mass ($1.853 < \texttt{InvM} < 1.876\,\text{GeV}/c^2$) are kept. We follow the same procedure in reconstructing the B meson as in the signal channel and apply the same M_{bc} and ΔE requirements.



(a) Signal MC

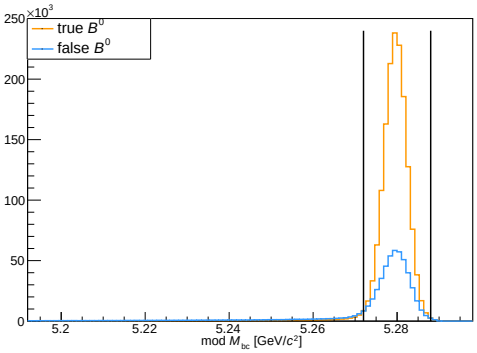


(b) $B\overline{B}$ background MC

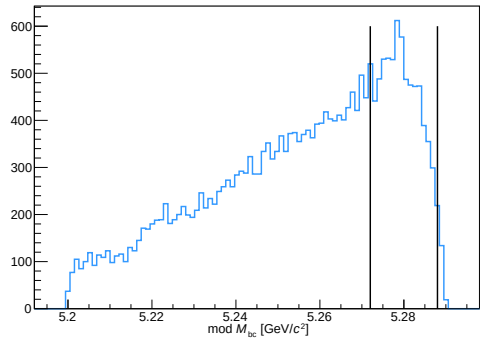


(c) Continuum background MC

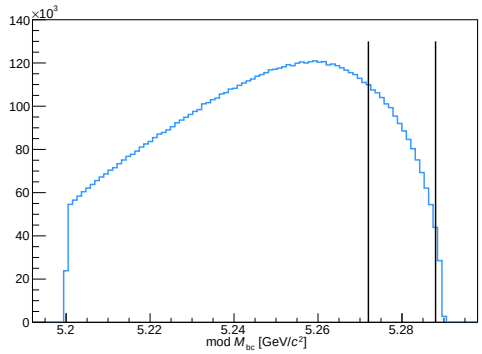
Figure 3.3: 2D Histograms of ΔE and M_{bc} (modified M_{bc}).



(a) Signal MC



(b) $B\overline{B}$ background MC



(c) Continuum background MC

Figure 3.4: Distributions of modified M_{bc}. The black solid lines denote the mass window.

Chapter 4 Continuum Suppression

The dominant background in $B^0 \to K^0_S\pi^0$ decay is the continuum background that arises from non-resonant $e^+e^- \to q\overline{q}$ hadronic events. A large amount of continuum background events still remains after the baseline selection. We use the fast boosted decision tree (FBDT) algorithm [43] to train a discriminator as a means to separate the signal events from the predominant continuum background events.

In the centre-of-mass frame, final-state particles in a continuum background event have a jet-like distribution as the decay to light hadrons leaves large available momentum for them to be collimated. On the other hand, $B\overline{B}$ events are uniformly distributed since the mass difference between the $\Upsilon(4S)$ resonance and the sum of the B meson pair is small. The B mesons are produced almost at rest in the centre-of-mass frame. The event shape for continuum and $B\overline{B}$ events are depicted in Fig. 4.1. Event shape variables describing the overall shape of the event are often used to train a classifier for suppressing the continuum background. These variables themselves may not have good separating powers. The signal efficiency would be reduced too greatly if strict cuts are directly applied to them.

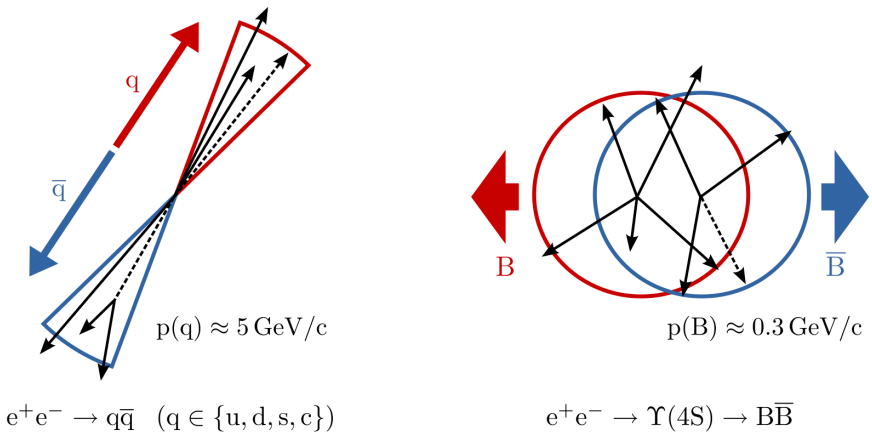


Figure 4.1: Event shape of continuum and $B\overline{B}$ events [44].

4.1 FastBDT Algorithm

In the multivariate analysis, the probability of an event being signal is calculated, using a set of characterising variables and a class label that labels signal as +1 and background as -1. The analysis proceeds in two steps, the fitting phase and the application phase. In the former a dataset with known labels is used, and in the latter the fitted classifier is applied to datasets with unknown labels.

In this analysis, we use the stochastic gradient boosted decision trees [45] developed for the Belle II experiment. Multivariate classification is used extensively in `basf2`, and the default multivariate algorithm is FastBDT, which is fast during fitting and application, robust for training in an automated environment, and generates an interpretable model with a good performance. We briefly introduce the stochastic gradient-boosted decision tree algorithm in this section.

A decision tree [46] performs classification with consecutive cuts determined in the fitting phase on the characterising variables. The maximum number of consecutive cuts is called the depth of the decision tree. At each node a binary decision is made, and the fraction of signal events in the node of all training events is taken as the signal probability. Only events that pass the preceding cuts are considered. A cumulative probability histogram for each variable at each node is calculated for signal and background. The characterising variable and the cut at each node are the ones that locally maximise the separation gain. The training result of a decision tree is often dominated by statistical fluctuation in the training dataset, leading to over-training and poor performance on other datasets.

In order to avoid over-training, a stochastic gradient boosted decision tree uses shallow decision trees instead of a single deep decision tree, the depth of each tree strongly limited. Signal and background are only roughly separated by each decision tree. A subsample is randomly drawn from the training sample for each tree and reweighted. The randomisation increases the robustness against over-training since statistical fluctuations would be averaged out in the sum over all trees [45]. The reweighting increases the influence of events hard to classify, often located near the optimal separating cuts [47]. The output classifier is constructed with multiple trees. The boost-weights calculated for the terminal node is the output of each decision tree, and the probability of an event being signal is the sigmoid-transformed sum of the boost-weights.

4.2 Training Variables

We use 10,000 correctly reconstructed signal events (`isSignal` $= 1$) and the same amount of continuum background events that survive the selections to train a CS output classifier. Another set of events of the same amount is used for testing and to check for over-training. We use the variable `isNotContinuumEvent` as the known label, and the input variables used for training the discriminator include M_{bc}, event shape variables, and vertex variables:

- M_{bc} - modified beam-constrained mass,
- `R2` - normalised second Fox-Wolfram moment,
- `cosTBTO` - cosine of the angle between the signal B thrust axis and the rest of event (ROE) thrust axis,
- `cosTBz` - cosine of the angle between the signal B thrust axis and the z-axis,
- `DeltaZ` - spacial difference between signal and tag-side B decay vertices in the beam direction,
- `DeltaZErr` - error of `DeltaZ`,
- `thrustAxisCosTheta` - cosine of the polar angle component of the thrust axis,
- `thrustOm` - magnitude of the ROE thrust axis,
- `CMScosTheta` - polar angle of the signal B in the $\Upsilon(4S)$ frame,
- dr - transverse distance with respect to the IP,
- `dcosTheta` - cosine of the polar angle with respect to the IP,
- `dphi` - azimuthal angle with respect to the IP,
- `KSFWVariables` - Kakuno-Super-Fox-Wolfram moments,
- `CLEOConeCS(ROE)` - CLEO cones for rest-of-event (ROE) particles.

Variables with a high correlation to ΔE are excluded from the input so as to prevent the trained continuum suppression (CS) output, which we employ as one of the fitting variables, from being correlated to ΔE. This enables us to model the distributions of the fitting variables, ΔE and the transformed CS output C' with a simpler approach. We apply a tight cut on M_{bc} and include it as an input training variable for continuum suppression, as using it as a fit variable does not help to improve the precision on the branching

fraction or the *CP* asymmetry. We test the fitting variables by performing fits to $400\,\text{fb}^{-1}$ of simulation sample using M_{bc}, ΔE, and the CS output, transformed with both the μ-transformation, described in Sec. 4.3 and the log-transformation. They yielded similar statistical uncertainties for the physics parameters, as summarised in Tab. 4.1.

We compare the CS output and ΔE of the training and testing set to check for over-training or if applying a cut on the output classifier sculpts the ΔE distribution, as shown in Fig. 4.2. We observe no significant over-training or sculpting.

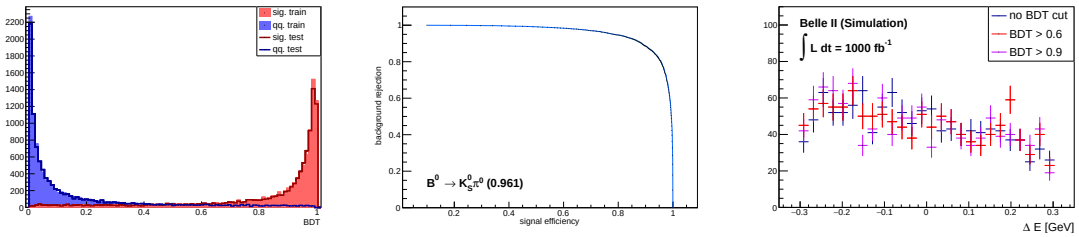


Figure 4.2: Left: over-training check by comparing training and testing datasets. Middle: receiver operating characteristic (ROC) of the CS classifier trained for $B^0 \to K^0_S\pi^0$ obained on the test sample. The area-under-curve is 0.961. Right: the $q\bar{q}$ ΔE distribution (scaled to the number of events corresponding to the tightest cut) before and after placing a requirement on the CS output classifier.

We require the CS output to be greater than 0.6, retaining about 85% of signal events and suppresses more than 90% of continuum background. This cut is chosen to optimise the sensitivity of the branching fraction and the direct *CP* asymmetry, for which we estimate the fitting uncertainty by evaluating large sets of toy pseudo-experiments with varying cuts around the PDG values of the physics parameters., as shown in Fig. 4.3.

Table 4.1: Statistical uncertainty estimated with $400\,\text{fb}^{-1}$ simulation sample using 2D and 3D fitting with μ-transformed CS output and 3D fitting with log-transformed CS output. The log-transformation is defined as $\log C = \ln\left([\text{CS} - \min(\text{CS})]/[1 - \text{CS}]\right)$.

	ΔE-C'	M_{bc}-ΔE-C'	M_{bc}-ΔE-$\log C$
$\mathcal{B}_{K^0\pi^0}$ uncert. [%]	6.0	5.6	5.8
$\mathcal{A}_{K^0\pi^0}$ uncert.	0.15	0.15	0.15

4.3 μ-Transformation

After the cut, we transform the CS output with the μ-transformation, or the probability integral transformation [48], to better model its distribution. The transformed classifier $\mathcal{R}(y)$ of classifier y is given by

$$\mathcal{R}(y) = \int_0^y \hat{y}_S(y')dy', \tag{4.1}$$

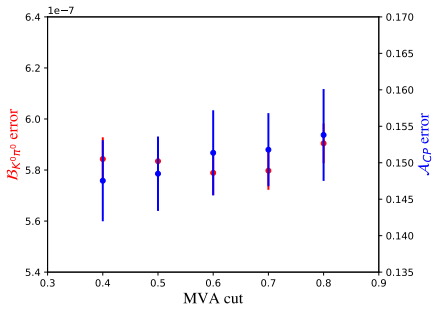


Figure 4.3: Mean uncertainties of $\mathcal{B}_{K^0\pi^0}$ and $\mathcal{A}_{K^0\pi^0}$ for varying cuts on the CS classifier variable, evaluated on 100 sets of pseudo-experiments.

where $\hat{y}_S(y')$ is the signal classifier distribution. This transformation is defined such that the transformed output for signal events $\mathcal{R}(y_S)$ is distributed uniformly between 0 and 1 within statistical fluctuation, whereas continuum events cluster towards 0. This allows for simple descriptions of the probability density functions (PDF) with analytical functions. We perform the μ-transformation of the CS output independently for each flavour tagger quality output interval, such that by construction the signal component would have a uniform distribution for each individual interval as well as for the total distribution. The CS output and the transformed output C' between data and simulation are compared using the control mode $B^+ \to \overline{D}(\to K^0_S\pi^0)\pi^+$ in Fig. 4.4. The data-MC comparison of the input variables is given in App. A.3, and they are listed in order of feature importance in Fig. 4.5.

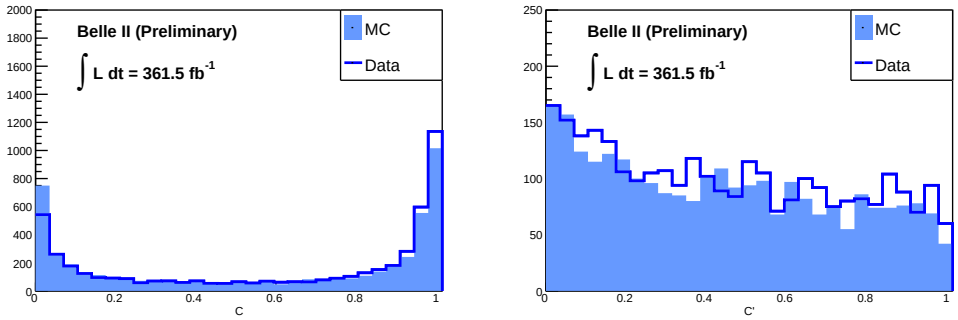


Figure 4.4: Left: comparison of the CS output between data and simulation in the $B^+ \to \overline{D}(\to K^0_S\pi^0)\pi^+$ control channel for the BDT trained for $B^0 \to K^0_S\pi^0$. Right: distribution of the transformed CS output after the cut $C > 0.6$. The number of simulated events before the cut is normalised to the total number of data events.

4.4 Summary of Final Selections

In this analysis we do not apply a best-candidate selection since the sample multiplicity is low (1.01%) after all the selection, and in general, applying a best-candidate selection

might introduce bias to the result. We assign this as the systematic uncertainty, as described in Ch. 8. In the signal-only simulation sample, the signal efficiency and the self-cross-feed efficiency are defined as

$$\epsilon = \frac{S}{G} \quad \text{and} \quad \text{SCF} = \frac{B}{S+B}, \tag{4.2}$$

where S is the number of correctly reconstructed signal events (`isSignal`$= 1$), G is the total number of generated signal events and B is the number of misreconstructed events in the signal-only MC sample. We summarise the selection requirements in Tab. 4.2 and the cumulative signal efficiencies determined from abundant signal-only simulation for subsequent selection steps in Tab. 4.3.

Table 4.2: Summary of selection requirements.

K_S^0 selection	$0.480 < m(\pi\pi) < 0.510\,\text{GeV}/c^2$	
	`LambdaVeto`$(K_S^0) > 0.06$	
	`VOSelector`$(K_S^0) > 0.2$	
π^0 selection	$0.115 < m(\gamma\gamma) < 0.150\,\text{GeV}/c^2$	
	`daugtherAngle`$(\pi^0) < 0.4$	
	γ selections	`clusterNHits`$(\gamma) > 1.5$
		$\lvert$`clusterTiming`$(\gamma)\rvert < 200$ ns
		$E(\gamma) > 0.03\,\text{GeV}$
		$0.2967 < \theta(\gamma) < 2.6180$ [rad]
		`photonMVA`$(\gamma) > 0.20$
B^0 selection	$5.272 < M_{\text{bc}} < 5.288\ \text{GeV}/c^2$	
	$-0.3 < \Delta E < 0.3$	
	CS classifier > 0.6	

Table 4.3: Signal efficiencies after subsequent analysis steps [%].

Operation	$\varepsilon(B^0 \to K_S^0\pi^0)$
Reconstruction	38.9
Selection	33.3
M_{bc} & ΔE	32.4
CS classifier	27.5

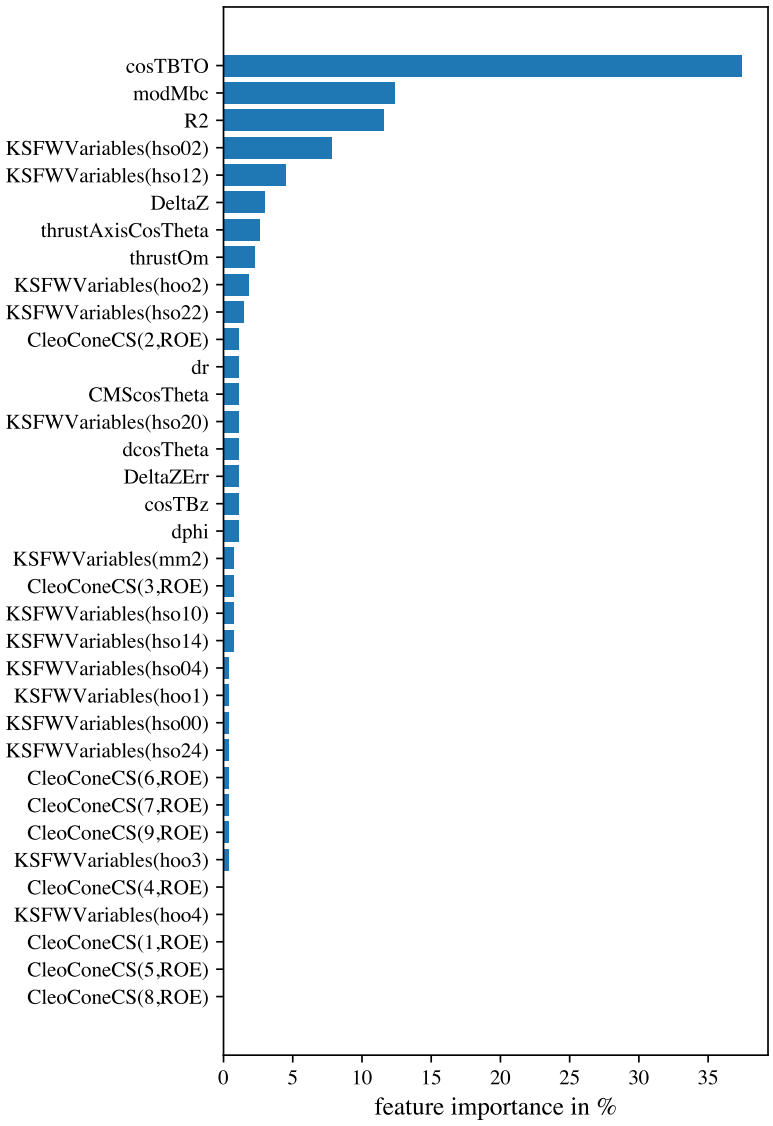


Figure 4.5: Variables used in BDT training for the continuum suppression in order of feature importance.

Chapter 5 Extraction of Physics Parameters

5.1 Flavour Tagger

The flavour tagger [16] is used to determine the flavour of neutral B mesons in events with a pair of neutral B meson, where one of them decays into a CP eigenstate, that is, both B^0 and $\overline{B}^0$ decay to the same final state, and the other decays through a flavour-specific channel. The event can be separated into the signal-side B meson (B_{sig}), which decays to a CP eigenstate, and the tag-side B meson (B_{tag}), which is the rest-of-event of B_{sig}. The flavour of B_{tag} is determined at the time of its decay with multivariate analysis, which is possible since many decay modes of neutral B mesons are flavour-specific.

There are 13 categories developed for grouping the different signatures of flavour-specific decays, as listed in Tab. 5.1. The charge of the final state particle tags the flavour of the B meson. The flavour tagger operates at the event level and the combiner level. The multivariate method is based on FastBDT [43]. The category-based flavour tagger [49] is inspired by the flavour tagger developed by Belle [50] and BaBar [51], first identifying the neutral B decay products and then discerning its flavour with combined information.

At the event level, particle lists are created for different types of reconstructed ROE tracks, including electrons, muons, kaons, pions, and protons. The ROE tracks are fitted with mass hypotheses corresponding to the target particle for determination of the flavour signature. Various flavour tagging input variables for category-specific multivariate analysis are calculated for each track, implicitly involving information from ECL and KLM clusters. For each category, and an output `RightCategory` is trained with the flavour tagger inputs to give the probability of the track being the target particle of the category.

At the combiner level, the track with the highest `RightCategory` value is selected as the target for that category. The input for the combiner is the product of the charge and `RightCategory`, resulting in 13 inputs at for the combiner, one for each category. The product qr of the flavour (q) and the dilution factor (r) is then trained for each B_{tag}.

Table 5.1: Flavour tagger categories of flavour specific decays. The target particles for determining the flavour signatures are marked in colour.

Categories	Decays
Electron	$b \to X e^-$
IntermediateElectron	$b \to X_c X \to X e^+$
Muon	$b \to X \mu^-$
IntermediateMuon	$b \to X_c X \to X \mu^+$
KinLepton	$b \to X l^-$
IntermediateKinLepton	$b \to X_c X \to X l^+$
Kaon	$b \to X K^-$
KaonPion	$b \to X D^{*+} \to X D^0 \pi^+ \to X K^- \pi^+$
SlowPion	$b \to X D^{*+} \to X D^0 \pi^+$
FastHadron	$b \to X^+ \pi^- (K^-)$
MaximumP	The particle with the highest CMS momentum.
FastSlowCorrelated	Slow pions from $D^{*\pm}$ and high-momentum primary particles.
Lambda	$b \to X \Lambda_c^* \to X \Lambda \to X p \pi^-$

5.2 Sample Composition

We identify three fit components: signal, $B\overline{B}$ background, and continuum background. The self-cross-feed efficiency amounts to SCF $< 2.5\%$. We include these candidates in the signal component for they peak at the same place in ΔE as signal events, as shown in Fig 5.1. Fig. 5.2 shows the distribution of different components making up the total sample in the generic simulation in ΔE, M_{bc}, and the product of the tagged flavour q and the dilution factor r of the flavour tagger, as shown in App. A.4.

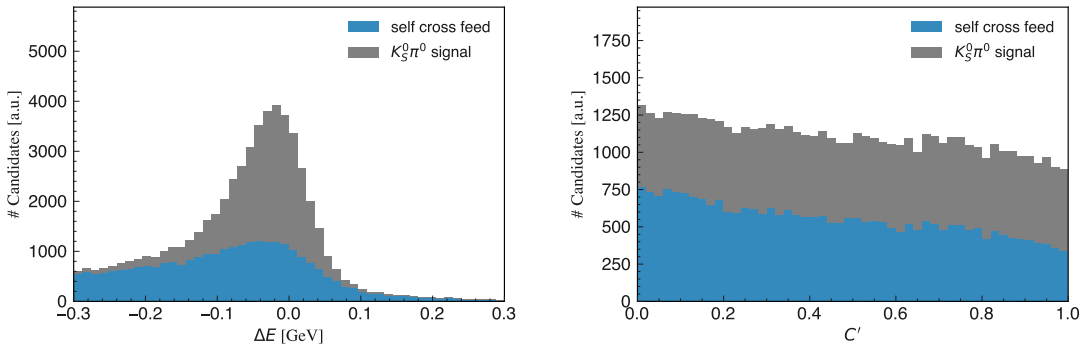


Figure 5.1: The ΔE and C' distributions of correctly and mis-reconstructed $B^0 \to K^0_S \pi^0$ signal candidates, scaled to the same number of events, after the final selection. The mis-reconstructed events amounts to 2.5% and is included in the signal component.

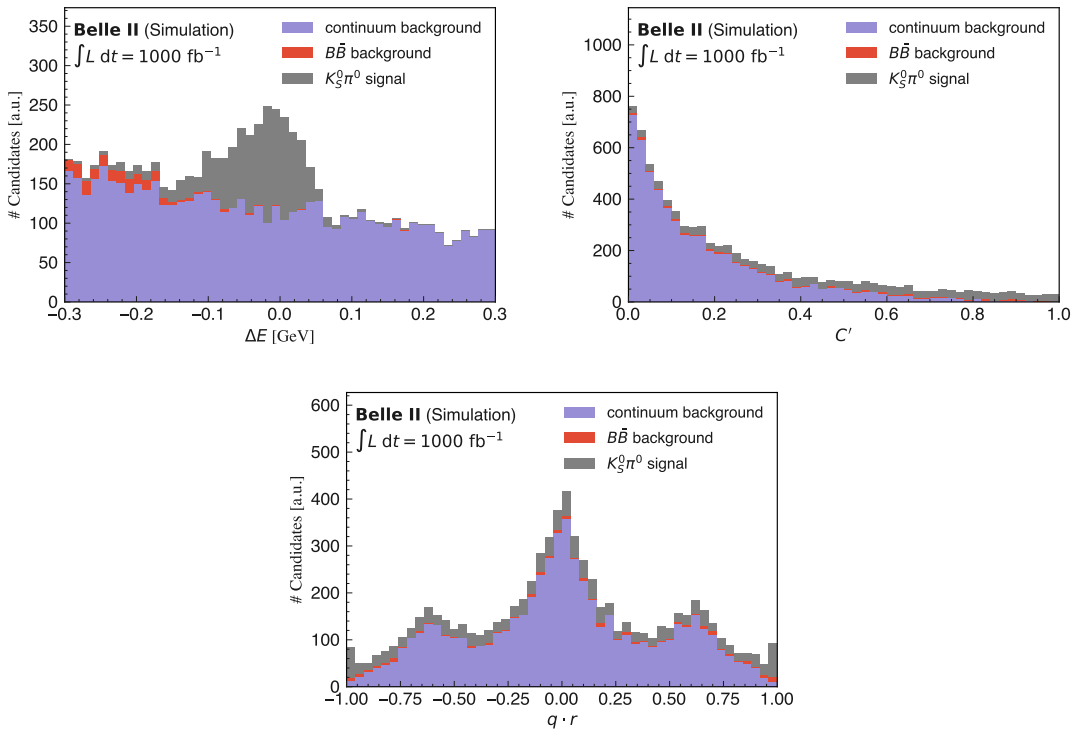


Figure 5.2: Final sample composition of ΔE, C', and $q \cdot r$ for 1 ab^{-1} of simulation sample.

5.3 Fitter Details

We extract the physics parameters, the branching fraction and the direct *CP* asymmetry $\mathcal{A}_{K^0\pi^0}$, with an extended two-dimensional unbinned maximum likelihood fit on the distributions of energy difference ΔE and the transformed continuum suppression classifier C'. Inclusion of M_{bc} as a fit variable complexes the fitting machinery and makes it prone to mismodelling while showing no improvement in the statistical precision. We therefore exclude M_{bc} from the fitting variables and use it as an input in the continuum suppression after a tight cut ($5.272 < M_{\text{bc}} < 5.288\,\text{GeV}/c^2$).

To measure the direct *CP* violation parameter, a category-based flavour tagger [16] algorithm is used to determine the the tag-side flavour. A positively tagged quark flavour $q = 1$ ($q = -1$) corresponds to the tag-side B meson being tagged as B^0 ($\overline{B}^0$). The flavour tagging algorithm yields a dilution factor r for each candidate, which is related to the wrong-tag fraction w by $w = 1 - 2r$. The events are categorised into seven r-bins according to the flavour tagger output quality. A set of flavour parameters is associated to each r-bin:

- partial efficiency (ε_r),
- difference of partial efficiencies between B^0 and $\overline{B}^0$ (μ_r),

- wrong-tag fraction (w_r),
- difference of wrong-tag fractions between B^0 and $\overline{B}^0$ (Δw_r).

We extract the branching fraction and direct *CP* asymmetry from a simultaneous fit over all seven r-bins, with the flavour parameters fixed to values determined from $B^0 \to D^{*-}\pi^+$ [52]. They are Gaussian constrained in the final fit with means and widths specified in Tab. 5.2 and Tab. 5.3. We gain additional statistical signal-to-background separation in the fit since the r distribution is different for signal and continuum. The $q \cdot r$ distribution of the generic MC sample is shown in Fig. 5.2. For the $B\overline{B}$ background component, we use the same flavour parameters as for the signal component, and for the continuum background component the partial efficiencies are fixed to values obtained from simulation.

The signal yield can be computed from the branching fraction by

$$N^{sig}_{K^0_S\pi^0} = 2 \times N_{B\overline{B}} \times \mathcal{B}_{K^0\pi^0} \times f_{K^0\to K^0_S} \times \epsilon_{K^0_S\pi^0}, \tag{5.1}$$

where $N^{sig}_{K^0_S\pi^0}$ is the signal yield, $N_{B\overline{B}}$ is the number of produced $B\overline{B}$ pairs, $\mathcal{B}_{K^0\pi^0}$ is the branching fraction (fit parameter), $f_{K^0\to K^0_S}$ is the fraction of K^0_S in K^0, and $\epsilon_{K^0_S\pi^0}$ is the signal reconstruction efficiency (estimated from simulation). The sub-branching fractions of $K^0_S \to \pi^+\pi^-$ and $\pi^0 \to \gamma\gamma$ are accounted for in the signal reconstruction efficiency $\epsilon_{K^0_S\pi^0}$. The signal and $B\overline{B}$ background PDF shapes are fixed from integrated r-bin fits to dedicated simulation samples, and the following parameters are free to float when fitting the real data:

- The direct *CP* asymmetry of $B^0 \to K^0\pi^0$ (1 parameter),
- The branching fraction of $B^0 \to K^0\pi^0$ (1 parameter),
- The continuum background yield (1 parameter),
- The $B\overline{B}$ background yield (1 parameter),
- The continuum background ΔE and C' shape parameters (3 parameters).

5.4 Fit Shapes

The PDF shapes and the correlation coefficients between the fitting variables are summarised in Tab. 5.4. Since the correlations are small, we assume the fitting variables ΔE and C' to be uncorrelated, and that the two-dimensional PDFs are simply the products of

Table 5.2: r-interval definition of each r-bin and associated flavor parameters for `MC15ri`. The partial efficiencies for the continuum background are calculated from simulated sample.

r-bin	r-interval	μ_r	w_r	Δw_r	ϵ_r	$\epsilon_{r,q\bar{q}}$
0	0.000 – 0.100	-0.0143 $\pm$ 0.0071	0.4725 $\pm$ 0.0031	-0.0008 $\pm$ 0.0063	0.1655	0.2513
1	0.100 – 0.250	-0.0099 $\pm$ 0.0073	0.4137 $\pm$ 0.0032	0.0119 $\pm$ 0.0065	0.1527	0.2116
2	0.250 – 0.450	-0.0026 $\pm$ 0.0071	0.3220 $\pm$ 0.0031	-0.0062 $\pm$ 0.0062	0.1585	0.1626
3	0.450 – 0.600	0.0003 $\pm$ 0.0075	0.2291 $\pm$ 0.0031	-0.0007 $\pm$ 0.0062	0.1385	0.1456
4	0.600 – 0.725	0.0185 $\pm$ 0.0079	0.1638 $\pm$ 0.0031	-0.0040 $\pm$ 0.0062	0.1203	0.1318
5	0.725 – 0.875	0.0281 $\pm$ 0.0078	0.0930 $\pm$ 0.0027	-0.0141 $\pm$ 0.0055	0.1134	0.0713
6	0.875 – 1.000	-0.0228 $\pm$ 0.0064	0.0181 $\pm$ 0.0017	-0.0039 $\pm$ 0.0034	0.1511	0.0257

Table 5.3: r-interval definition of each r-bin and associated flavor parameters for Moriond 2023 data.

r-bin	r-interval	μ_r	w_r	Δw_r	ϵ_r
0	0.000 – 0.100	-0.0240 $\pm$ 0.0129 $\pm$ 0.0057	0.4804 $\pm$ 0.0070 $\pm$ 0.0036	-0.0068 $\pm$ 0.0103 $\pm$ 0.0070	0.1580
1	0.100 – 0.250	0.0138 $\pm$ 0.0134 $\pm$ 0.0086	0.4240 $\pm$ 0.0071 $\pm$ 0.0042	0.0381 $\pm$ 0.0107 $\pm$ 0.0052	0.1553
2	0.250 – 0.450	-0.0115 $\pm$ 0.0130 $\pm$ 0.0049	0.3410 $\pm$ 0.0065 $\pm$ 0.0040	-0.0188 $\pm$ 0.0105 $\pm$ 0.0039	0.1652
3	0.450 – 0.600	0.0088 $\pm$ 0.0129 $\pm$ 0.0061	0.2362 $\pm$ 0.0069 $\pm$ 0.0048	-0.0070 $\pm$ 0.0107 $\pm$ 0.0053	0.1392
4	0.600 – 0.725	0.0368 $\pm$ 0.0140 $\pm$ 0.0055	0.1675 $\pm$ 0.0065 $\pm$ 0.0037	0.0197 $\pm$ 0.0109 $\pm$ 0.0050	0.1161
5	0.725 – 0.875	-0.0202 $\pm$ 0.0140 $\pm$ 0.0066	0.1073 $\pm$ 0.0059 $\pm$ 0.0054	-0.0005 $\pm$ 0.0097 $\pm$ 0.0072	0.1145
6	0.875 – 1.000	-0.0123 $\pm$ 0.0115 $\pm$ 0.0048	0.0274 $\pm$ 0.0032 $\pm$ 0.0033	0.0024 $\pm$ 0.0063 $\pm$ 0.0018	0.1517

one-dimensional PDFs. The PDF shape parameters are obtained by fitting MC samples, using four million signal-only and 1 ab^{-1} background MC events. The fits to dedicated simulation samples are shown in Fig. 5.3, and the shape parameters are listed in Tab. 5.5.

We derive the flavour PDF by incorporating the flavour parameters into Eqn. 1.9, which then becomes

$$\begin{aligned}\mathcal{P}(\Delta t, q) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}}\{\, 1 - q\Delta w + q\mu(1-2w) + [q(1-2w) + \mu(1-q\Delta w)] \\ \times[\mathcal{A}\cos(\Delta m_d \Delta t) + \mathcal{S}\sin(\Delta m_d \Delta t)]\,\}.\end{aligned} \tag{5.2}$$

Integrating over the decay time, we obtain the time-integrated flavour PDF shapes described by

$$\begin{aligned}\mathcal{P}_{sig}(q) =& \frac{1}{2}\,[1 - q\Delta w_r + q\mu_r(1-2w_r) + (1-2\chi_d)\{q(1-2w_r) + \mu_r(1-q\Delta w_r)\} \times \mathcal{A}_{K^0\pi^0}], \\ \mathcal{P}_{B\bar{B}}(q) =& \frac{1}{2}\,[1 - q\Delta w_r + q\mu_r(1-2w_r) + (1-2\chi_d)\{q(1-2w_r) + \mu_r(1-q\Delta w_r)\} \times \mathcal{A}_{B\bar{B}}], \\ \mathcal{P}_{q\bar{q}}(q) =& \frac{1}{2}(1 - q \times \mathcal{A}_{q\bar{q}}),\end{aligned} \tag{5.3}$$

where χ_d is the time-integrated B^0 mixing parameter. The background asymmetries $A_{B\bar{B}}$

and $A_{q\bar{q}}$ are fixed to zero in the final fit to decrease the estimated uncertainty and bias on $A_{K^0\pi^0}$. The total PDF $\mathcal{P}(\Delta E, C', q)$ is the product of $\mathcal{P}(\Delta E)$, $\mathcal{P}(C')$, and $\mathcal{P}(q)$,

$$\mathcal{P}_{r,j}(\Delta E, C', q) = \mathcal{P}_j(\Delta E) \cdot \mathcal{P}_j(C') \cdot \mathcal{P}_{r,j}(q), \tag{5.4}$$

where j denotes the fit components and r denotes flavour tagger r-bins. The ΔE and C' distributions are modelled with the same PDF over different r-bins. We use the extended maximum unbinned likelihood method to determine the branching fraction and the direct *CP* asymmetry, respectively, which are encoded in the signal PDF and signal yield. The likelihood is given by

$$\mathcal{L}(\Delta E, C', q) = \frac{e^{-N}}{N!} \prod_r \prod_i^{N_r} \sum_j \epsilon_{r,j} N_j \mathcal{P}^i_{r,j}(\Delta E, C', q), \tag{5.5}$$

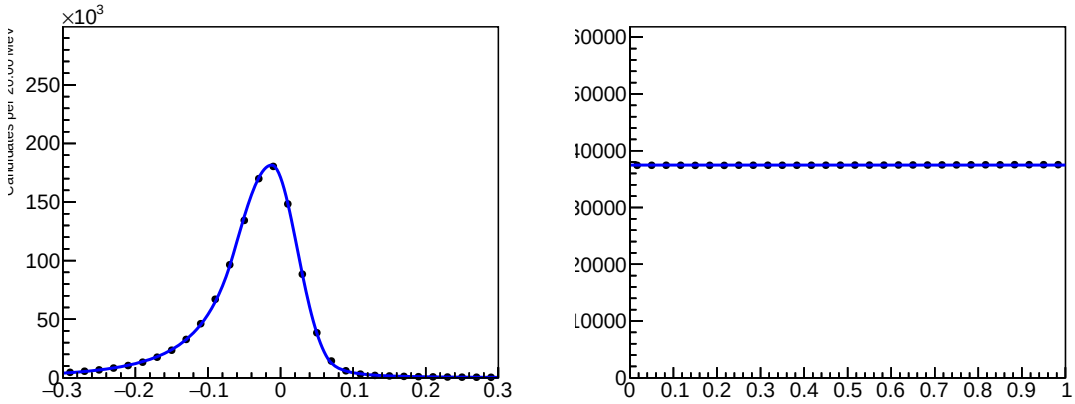
where N is the number of all the events, N_j is the number of event for each fit component, N_r is the number of events in each r-bin, i represents the i-th event, and $\epsilon_{r,j}$ is the partial efficiency for fit component j in each r-bin, as given in Tab. 5.2 and Tab. 5.3.

Table 5.4: $B^0 \to K^0_S\pi^0$ PDF shapes and the correlation coefficients between ΔE and the μ-transformed continuum suppression output. ΔE is modelled with a double-sided Crystal Ball lineshape, a Gaussian, and a first-order polynomial for the signal, $B\overline{B}$ background, and continuum background, respectively. C' is modelled with first-order polynomials for signal and $B\overline{B}$ background, and an exponential function for the continuum background.

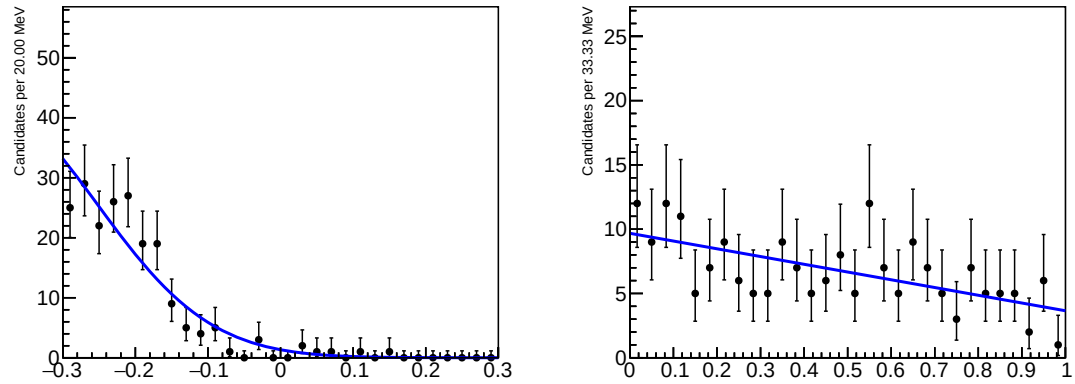
Component	ΔE	C'	corr.
$B^0 \to K^0_S\pi^0$ signal	Double CB	1st-order polynomial	0.0033
$B\overline{B}$ background	Gaussian	1st-order polynomial	0.0400
Continuum background	1st-order polynomial	exp($ax + bx^2$)	0.0034

Table 5.5: Shape parameters for $B^0 \to K^0_S\pi^0$ that are fixed in the final fit.

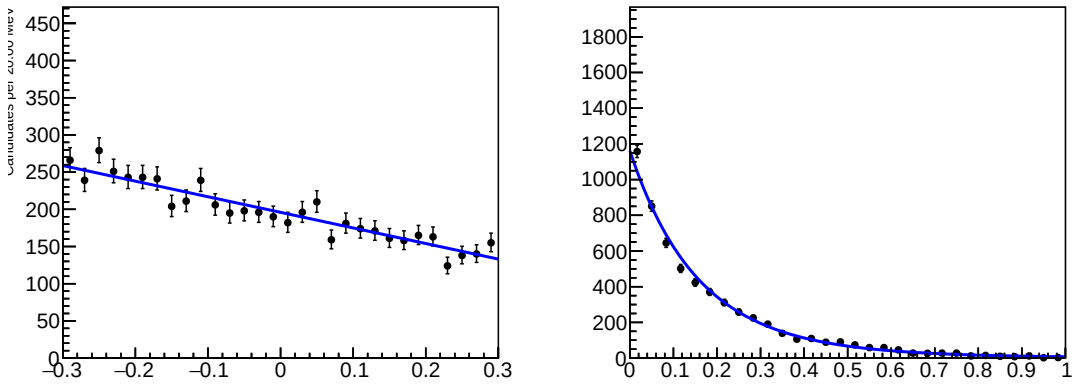
Component	Variable	Parameter	Fit Result
Signal	ΔE	μ	-0.01216 ± 0.00017
		σ_L	0.05077 ± 0.00027
		σ_R	0.03519 ± 0.00012
		n_L	4.499 ± 0.093
		n_R	1.773 ± 0.032
		α_L	1.0477 ± 0.0082
		α_R	2.147 ± 0.011
$B\overline{B}$ background	ΔE	μ	-0.41 ± 0.11
		σ	0.155 ± 0.031
	C'	a_1	-0.62 ± 0.11



(a) Signal component



(b) $B\overline{B}$ background component



(c) Continuum background component

Figure 5.3: Left: the ΔE distribution (black dots with error bars) for signal, $B\overline{B}$ background, and continuum background in the $B^0 \to K^0_S\pi^0$ sample. The result of a fit is overlaid as blue solid lines. Right: same but for the μ-transformed continuum suppression output C' distribution.

5.5 Fitter Validation

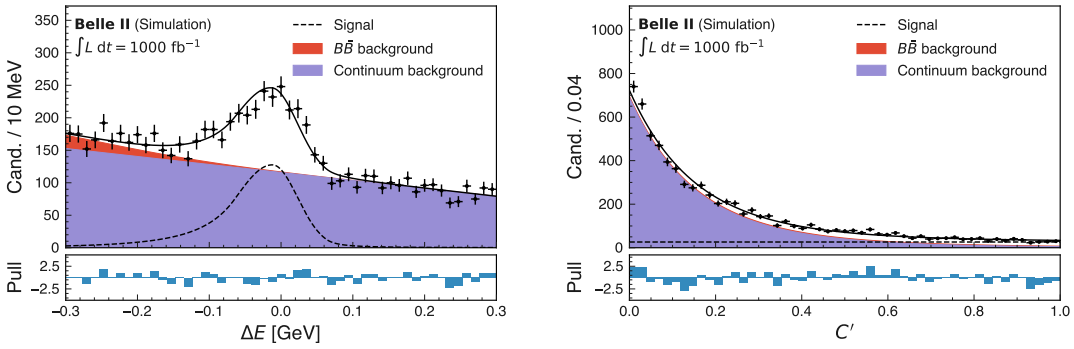
Before performing the fit to data, we first validate the fitter on $1\,\text{ab}^{-1}$ of simulated sample and pseudo experiments of ToyMC and gsim datasets. The fit result agree with the simulation truth and we observe no bias from the ensemble tests. We also perform partial unblinding with $62.8\,\text{fb}^{-1}$ of Moriond data to check that the result are in agreement with the previous Belle II measurement and the world average reported by the PDG [6].

5.5.1 Fit to Simulated Data

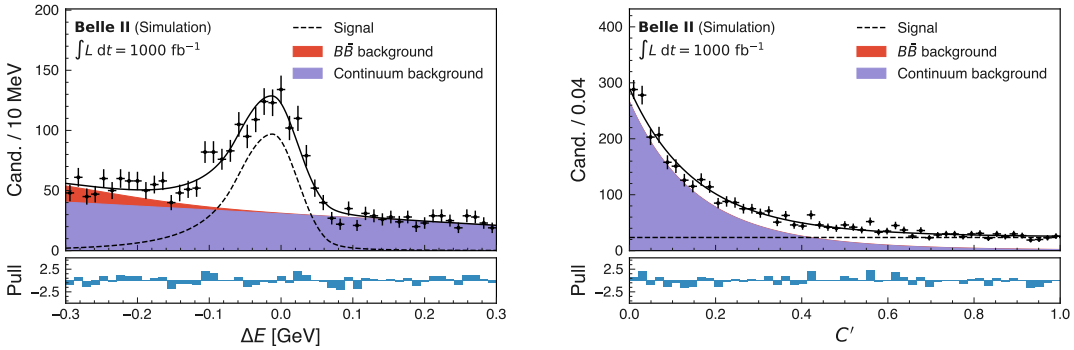
We validate the fitter on a simulated sample corresponding to an integrated luminosity of $1\ \text{ab}^{-1}$. The result of the fit is shown in Fig. 5.4 to Fig. 5.6, and the fitted parameters and the simulation truth are listed in Tab. 5.6. The overlaid fit match the MC distribution, and the fitted physics parameters are in agreement with the simulation truth.

Table 5.6: Fit result for the fit to the $1\,\text{ab}^{-1}$ simulation sample for $B^0 \to K^0_S\pi^0$.

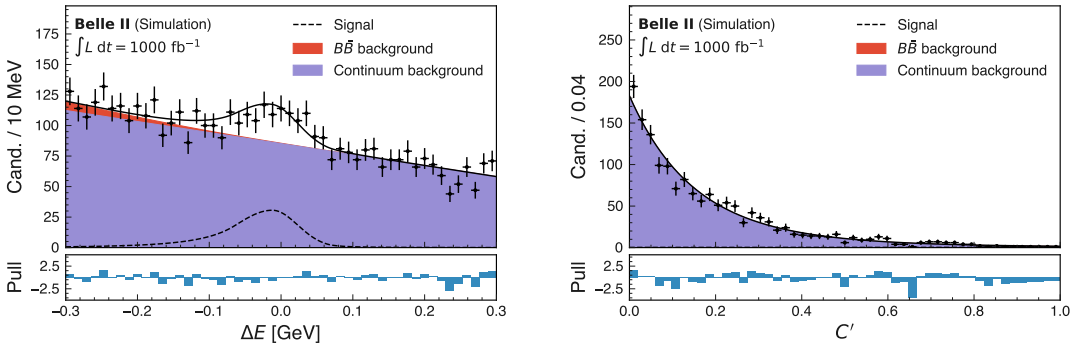
Parameter	Fit Result	Simulation Truth
$\mathcal{B}_{K^0\pi^0}$	$(8.84 \pm 0.35) \times 10^{-6}$	9.09×10^{-6}
$\mathcal{A}_{K^0\pi^0}$	0.036 ± 0.092	0
$B\overline{B}$ yield	209 ± 47	200
$q\overline{q}$ yield	5911 ± 99	5906
$q\overline{q}$ ΔE 1st param	-1.060 ± 0.082	–
$q\overline{q}$ C' a	-6.48 ± 0.22	–
$q\overline{q}$ C' b	1.59 ± 0.33	–



(a) Fitting region, where $-0.3 < \Delta E < 0.3$ and $0 < C' < 1$

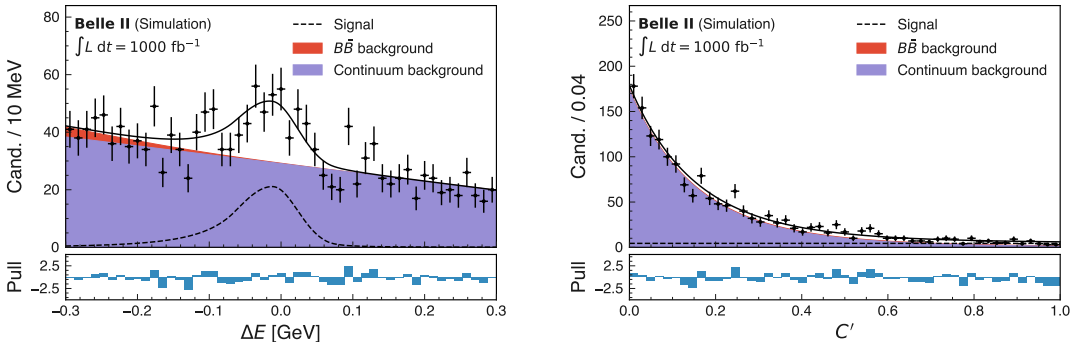


(b) Signal-enhanced region, defined by $-0.13 < \Delta E < 0.1$ and $0.24 < C' < 1$

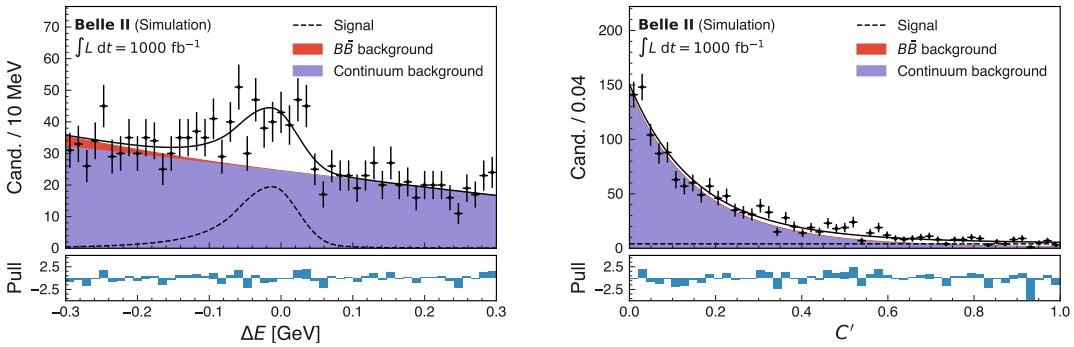


(c) Background-enhanced region, defined by $0.1 < \Delta E < 0.3$ and $0 < C' < 0.24$

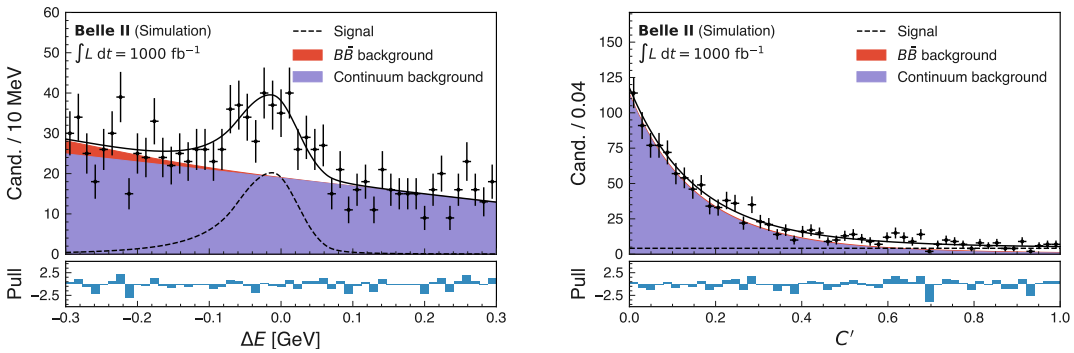
Figure 5.4: The ΔE and C' distributions of $1\,\text{ab}^{-1}$ generic `MC15ri` simulation in (a) the fitting region, (b) the signal-enhanced region, and (c) the background-enhanced region. The fit result to the sample is overlaid with a solid black line, and the fit components are shown as black dashed line (signal), red shaded area ($B\overline{B}$ background), and purple shaded area (continuum background).



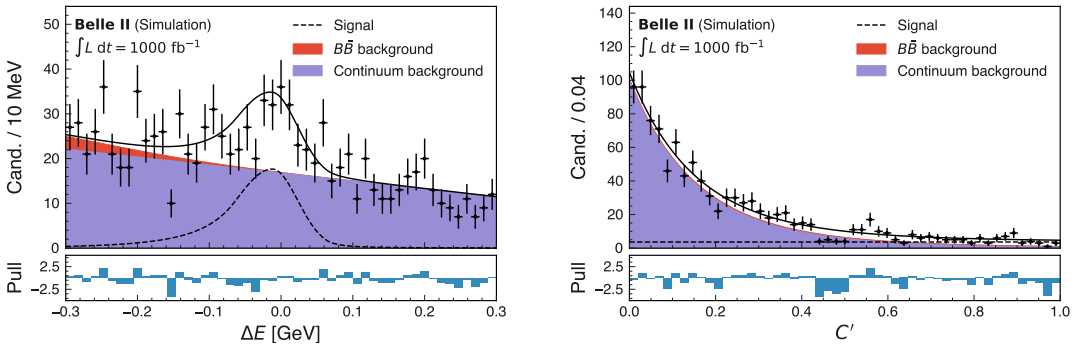
(a) 0th r-bin



(b) 1st r-bin

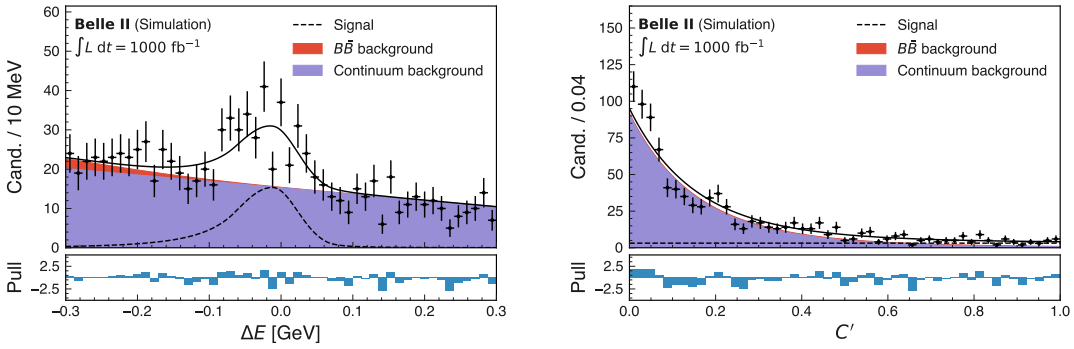


(c) 2nd r-bin

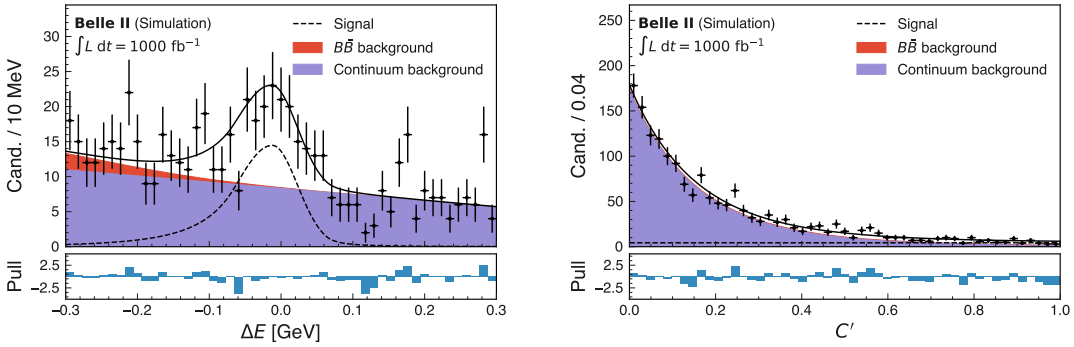


(d) 3rd r-bin

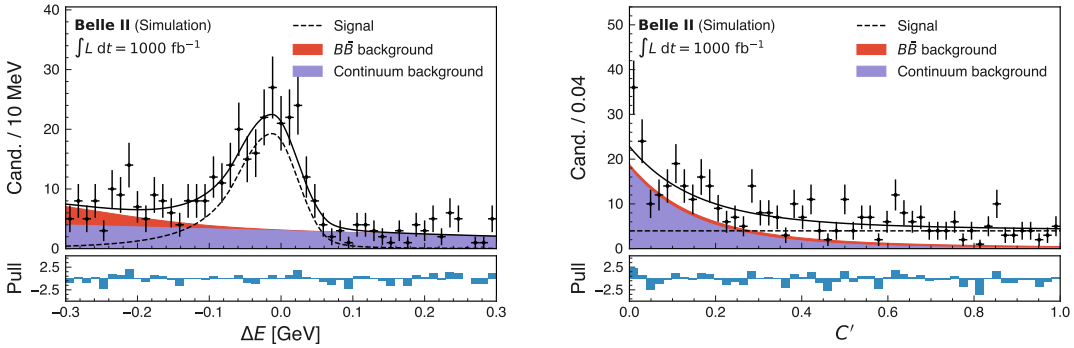
Figure 5.5: The ΔE and C' distributions of 1 ab^{-1} generic `MC15ri` simulation projected to each r-bin.



(a) 4th r-bin



(b) 5th r-bin



(c) 6th r-bin

Figure 5.6: The ΔE and C' distributions of 1 ab^{-1} generic `MC15ri` simulation projected to each r-bin.

5.5.2 Pseudo-Experiments

We also validate the fitter with a large ensemble of ToyMC datasets, drawn from the nominal fit shapes. We generate events for each component corresponding to the integrated luminosity of the LS1 dataset ($362\,\text{fb}^{-1}$). The yields for each component are drawn according to their expected proportions, randomly fluctuated from a Poisson distribution around the value indicated by the simulation truth of the 1 ab^{-1} generic MC sample. We then fit these ToyMC datasets and calculate the pull for each fit, which is defined as

$$\text{pull} = \frac{x_{fit} - x_{true}}{\sigma_x}, \tag{5.6}$$

where x_{fit} is the measured fit parameter, x_{true} is true input value of this parameter and σ_x is the uncertainty on x_{fit}. Fig. 5.7 shows the pull distributions of the physics parameters. We observe no biases in any of the fitted physics parameters or yields. We estimate the statistical uncertainties of the physics parameters by calculating the mean of the fit uncertainties in this ensemble of ToyMC datasets.

$$\langle \sigma_{\mathcal{A}(K^0\pi^0)} \rangle = 0.14$$

$$\langle \sigma_{\mathcal{B}(K^0\pi^0)} \rangle = 6.2\,\%$$

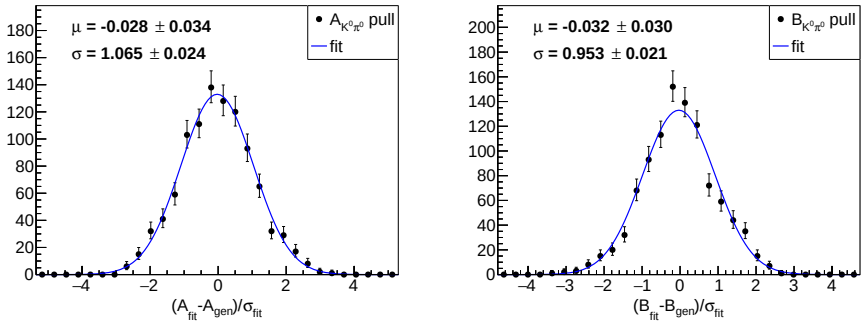


Figure 5.7: Pull distributions of the physics parameters $\mathcal{A}_{K^0_S\pi^0}$ and $\mathcal{B}_{K^0_S\pi^0}$ obtained by fits to 1000 pseudo-experiments. Projections of fits with a Gaussian function are shown along with the numeric fit results.

In addition, we perform linearity tests in order to check for biases when the physics parameters deviate from the nominal values announced by the PDG. We follow the same approach as described above. Large ensembles of ToyMC datasets are generated and subsequently fitted. The input value of the branching fraction is varied from about 80% to 120% of its nominal value, and that of the direct *CP* asymmetry is varied from -0.5 to 0.5. The results of the linearity tests are shown in Fig. 5.8.

We also perform linearity tests on pseudo-experiments sampled from MC simula-

tion. Large ensembles of Gsim datasets are bootstrapped around the nominal values of the branching fraction and the direct *CP* asymmetry. The pseudo-experiment for the direct *CP* asymmetry are generated from `MC15ri` generic and signal-only samples, and for the branching fraction the signal component is generated from signal-only simulation samples, whereas the $q\bar{q}$ and $B\overline{B}$ background components are ToyMC datasets in order to avoid bias originating from bootstrapping from a small sample. The results are shown in Fig. 5.9. We observe no biases in either of the physics parameters.

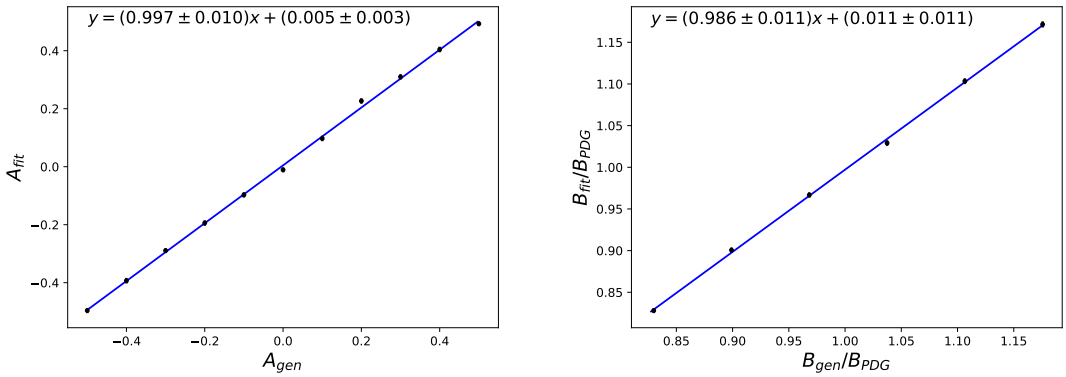


Figure 5.8: Result of the toyMC linearity test for the direct *CP* asymmetry (right) and branching fraction (left).

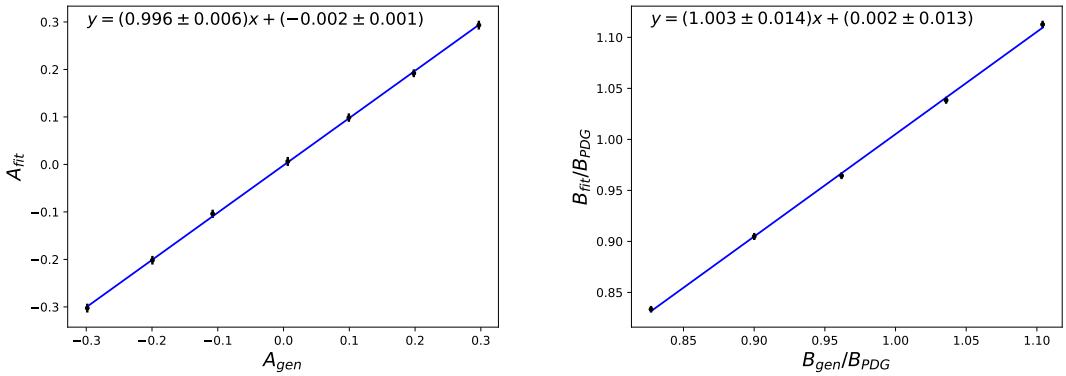
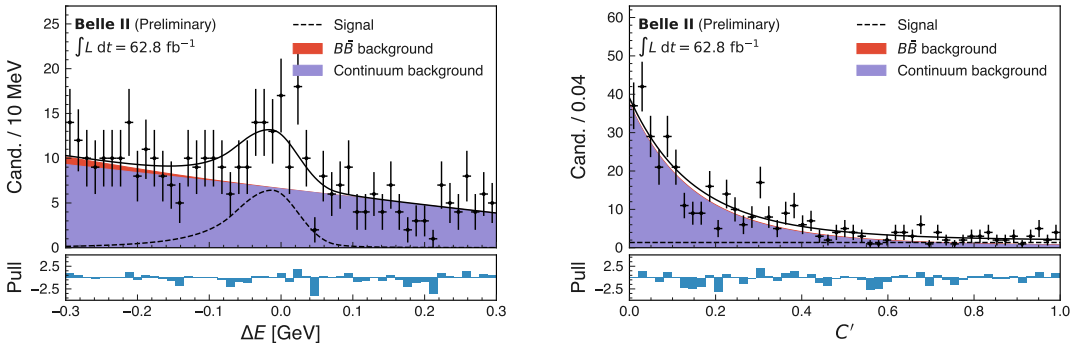


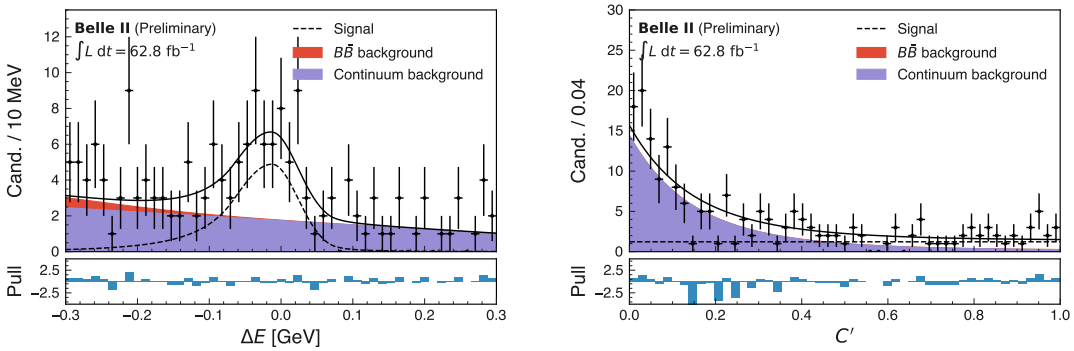
Figure 5.9: Result of the bootstrapping linearity test for the direct *CP* asymmetry (right) and branching fraction (left).

5.5.3 Partial Unblinding

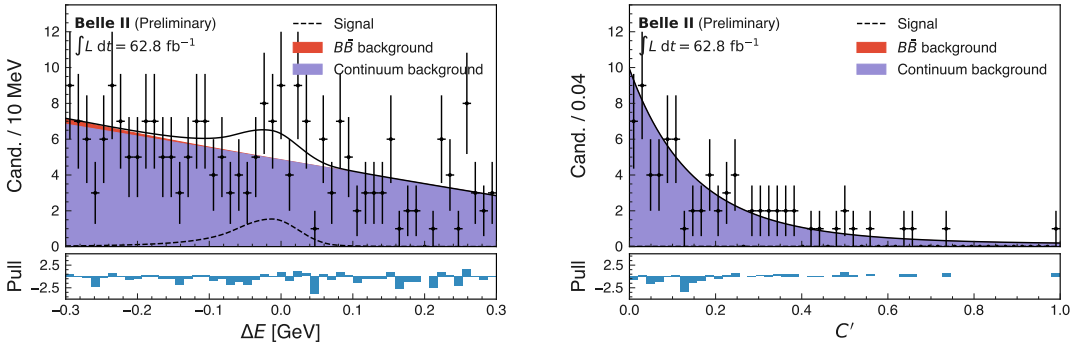
We partially unblind the Moriond data up to Exp 12 (Sec. 3.1) as the final check of the fitting machinery before the full unblinding. The integrated luminosity of partial unblinding is $62.8\,\text{fb}^{-1}$, which is the same as the previous Belle II measurement [15]. The fit result is shown in Fig. 5.10, and the fitted parameters, the PDG values, and the previous Belle II measurement are listed in Tab. 5.7. The result agrees with the PDF values and the previous measurement considering uncertainties.



(a) Fitting region, where $-0.3 < \Delta E < 0.3$ and $0 < C' < 1$



(b) Signal-enhanced region, defined by $-0.13 < \Delta E < 0.1$ and $0.24 < C' < 1$



(c) Background-enhanced region, defined by $0.1 < \Delta E < 0.3$ and $0 < C' < 0.24$

Figure 5.10: The ΔE and C' distributions of $62.8\,\text{fb}^{-1}$ data in (a) the fitting region, (b) the signal-enhanced region, and (c) the background-enhanced region.

Table 5.7: Fit result for the fit to the $62.8\,\text{fb}^{-1}$ data sample for $B^0 \to K^0_S\pi^0$.

Parameter	Fit Result	PDG [6]	Belle II ($62.8\,\text{fb}^{-1}$) [15]
$\mathcal{B}_{K^0\pi^0}$ $[10^{-6}]$	8.0 ± 1.4	9.9 ± 0.5	8.5 ± 1.7
$\mathcal{A}_{K^0\pi^0}$	-0.70 ± 0.45	0.00 ± 0.13	-0.40 ± 0.45
$B\overline{B}$ yield	9 ± 14	–	–
$q\overline{q}$ yield	336 ± 27	–	–
$q\overline{q}$ ΔE 1st param	-1.38 ± 0.35	–	–
$q\overline{q}$ C' a	-6.66 ± 0.88	–	–
$q\overline{q}$ C' b	2.7 ± 1.2	–	–

Chapter 6 Control Channels

We used abundant B meson charmed decays as control channels for testing the agreement between data and simulation, validate the fitter, and extract calibration parameters or fudge factors for the signal channel $B^0 \to K^0_S\pi^0$. The shift parameter Δ_μ and the scaling parameter s_σ are used to shift the mean μ and scale the width σ of the signal PDFs to provide a better description of the real data.

$$\begin{aligned} \text{simulation} &\to \text{real data} \\ \mu &\to \mu + \Delta_\mu \\ \sigma &\to \sigma \cdot s_\sigma. \end{aligned} \tag{6.1}$$

We reconstruct the $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$ control mode to assess the continuum suppression efficiency in data and simulation, and we apply the ΔE shift and scaling parameters obtained from $B^0 \to \overline{D}^0(\to K^+\pi^-)\pi^0$ [17]. In Tab. 6.1 we summarise the extracted continuum suppression efficiencies as well as the shift and scaling parameters.

Table 6.1: Calibration parameters obtained from control modes.

CS eff. ratio [%]	96.79 ± 1.68
ΔE shift	(-0.0005 ± 0.0017) GeV
ΔE scale	1.034 ± 0.046

6.1 Continuum Suppression Efficiency

We perform a simultaneous fit on ΔE with the $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$ samples divided into two disjoint subsamples: one passing the continuum suppression trained for $B^0 \to K^0_S\pi^0$ and the other failing it. The PDF shapes are the same as those for $B^0 \to K^0_S\pi^0$ except for the BB background component, where kernel estimation [53] is used for ΔE instead of a Gaussian for the control mode. The shape parameters are obtained by fitting dedicated simulation samples (two million signal-only events and $1\,\text{ab}^{-1}$ generic background). In the fit to data, only the $q\bar{q}$ shape parameters and the yields are free-floating. The fit result is shown in Fig. 6.1. We scale the branching ratio by the data-MC ratio and assign the

uncertainty of the ratio as the systematic uncertainty. The measured branching ratio of $B^+ \to \overline{D}^0\pi^+$ is $\mathcal{B}_{\overline{D}^0\pi^+} = 4.81 \pm 0.14$, which agrees with the world average [6].

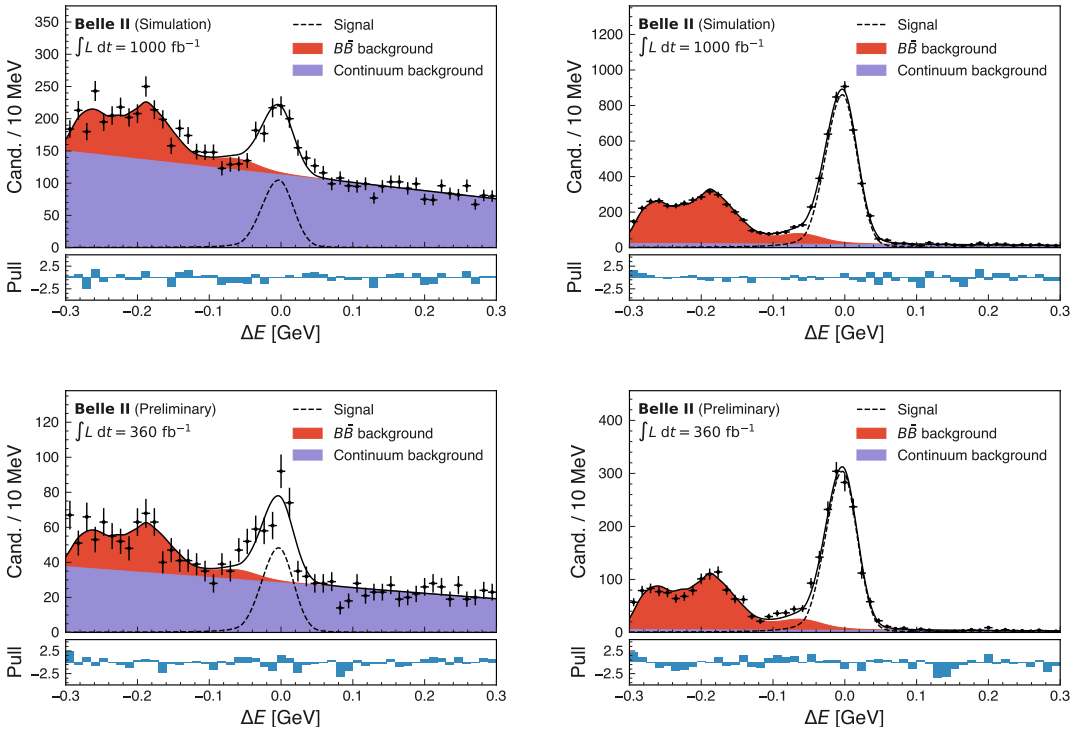


Figure 6.1: Upper (lower): Result of the fit to the simulation (data) ΔE distribution of $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$ for the extraction of the efficiency of the continuum suppression cut for $B^0 \to K^0_S\pi^0$. Left (right): the fit to the sample failing (passing) the continuum suppression cut is shown.

Table 6.2: Summary of fit parameters for $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$ ΔE fit.

Parameter	$1\,\text{ab}^{-1}$ MC	$362\,\text{fb}^{-1}$ data
$\mathcal{B}_{\overline{D}^0\pi^+}[\times 10^{-3}]$	4.47 ± 0.08	4.81 ± 0.14
$B\overline{B}$ yield	4668 ± 157	1587 ± 84
$q\overline{q}$ yield	6678 ± 184	1649 ± 95
Signal CS eff. [%]	89.14 ± 0.83	86.28 ± 1.27
$B\overline{B}$ CS eff. [%]	77.26 ± 2.00	78.80 ± 3.14
$q\overline{q}$ CS eff. [%]	13.85 ± 0.77	12.35 ± 1.62

6.2 ΔE Shift and Scaling Parameters

We use the ΔE shift and scaling parameters extracted from $B^0 \to \overline{D}^0(\to K^+\pi^-)\pi^0$ [17] for the ΔE shape of the signal component. The π^0 momentum distribution in this control channel is similar enough to that of $B^0 \to K^0_S\pi^0$ for using it as the control channel, since the ΔE shape depends primarily on π^0. The shift and scaling parameters are extracted

Table 6.3: Continuum suppression efficiencies of simulation and data. The data-MC ratio is incompatible with unity.

	CS efficiency [%]
data	86.28 ± 1.27
simulation	89.14 ± 0.83
ratio	96.79 ± 1.68

from a 2D fit in ΔE and C' (trained for $B^+ \to h^+\pi^0$). The PDF shapes of different fit components (signal, $B\overline{B}$ background, and $q\overline{q}$ background) are determined by fitting dedicated simulation samples. In the fit to data, only the $q\overline{q}$ shape parameters, the yields, and the ΔE shift and scaling parameters are free-floating. The extracted shift parameter is (-0.0005 ± 0.0017) GeV, and the scaling parameter is 1.034 ± 0.046.

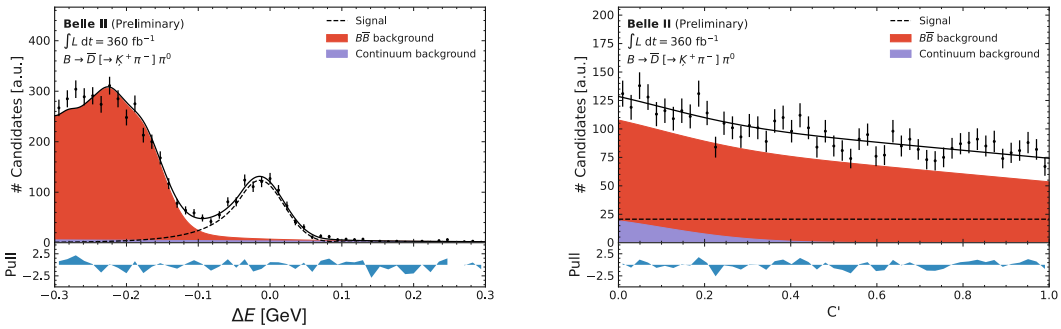


Figure 6.2: ΔE (left) and C' (right) distributions for $B^0 \to \overline{D}^0(\to K^+\pi^-)\pi^0$ [17]. The fit result is shown as a solid black line. Individual fit components are shown as black dashed line (signal), red shaded area ($B\overline{B}$ background) and purple shaded area ($q\overline{q}$ background).

Chapter 7 Result

7.1 Fit to Data

The ΔE and C' distributions of the $362\,\text{fb}^{-1}$ Moriond 2023 LS1 dataset are shown in Fig. 7.1 to Fig. 7.6. Flavour-specific projections in each r-bin are given in App. A.5. The measured branching fraction and direct *CP* asymmetry are

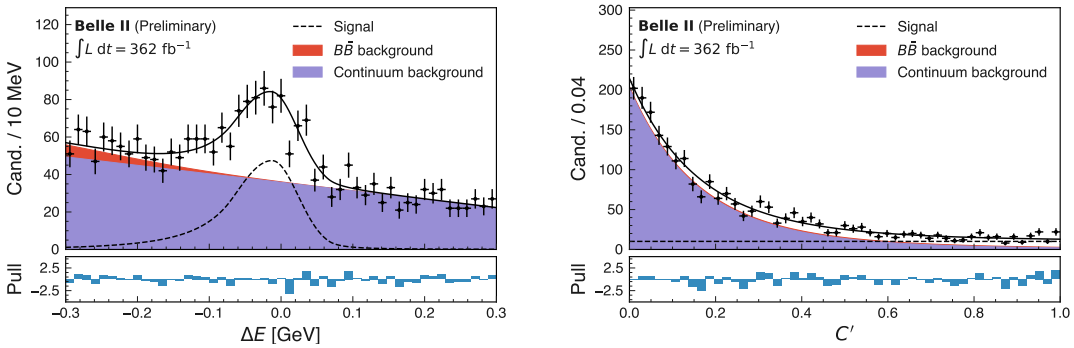
$$\mathcal{A}_{K^0\pi^0} = -0.06 \pm 0.15 \pm 0.05,$$

$$\mathcal{B}_{K^0\pi^0} = 10.16 \pm 0.65 \pm 0.67,$$

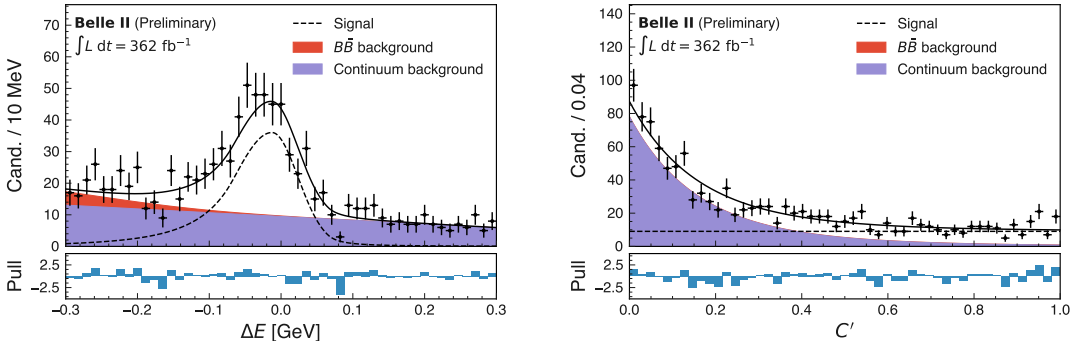
where the first uncertainty is statistical and the second is systematic. Estimation of the systematic uncertainty is described in detail in Ch. 8. The fitted parameters and data-MC corrections are summarised in Tab. 7.1, and the correlation matrix of the fitted parameters is given in App. A.6.

Table 7.1: Summary of fit parameters and data-MC corrections

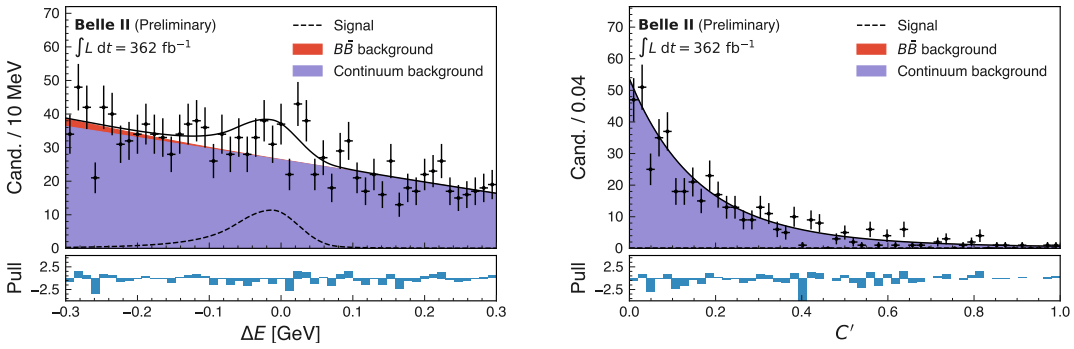
Parameter	Moriond 2023 data	PDG value
$\mathcal{A}_{K^0\pi^0}$	-0.06 ± 0.15	0 ± 0.13
$\mathcal{B}_{K^0\pi^0}[\times 10^{-6}]$	10.16 ± 0.65	9.9 ± 0.5
$N_{B\bar{B}}$	66 ± 28	–
$N_{q\bar{q}}$	1828 ± 57	–
$q\bar{q}\ \Delta E\ a_1$	-1.26 ± 0.15	–
$q\bar{q}\ C'\ a$	-6.20 ± 0.40	–
$q\bar{q}\ C'\ b$	1.61 ± 0.60	–
Corrections		
K^0_S efficiency	0.9384 ± 0.0198	
π^0 efficiency	1.018 ± 0.0501	
CS efficiency	0.9679 ± 0.0168	
ΔE shift	-0.0005 ± 0.0017	
ΔE scale	1.034 ± 0.046	



(a) Fitting region, where $-0.3 < \Delta E < 0.3$ and $0 < C' < 1$

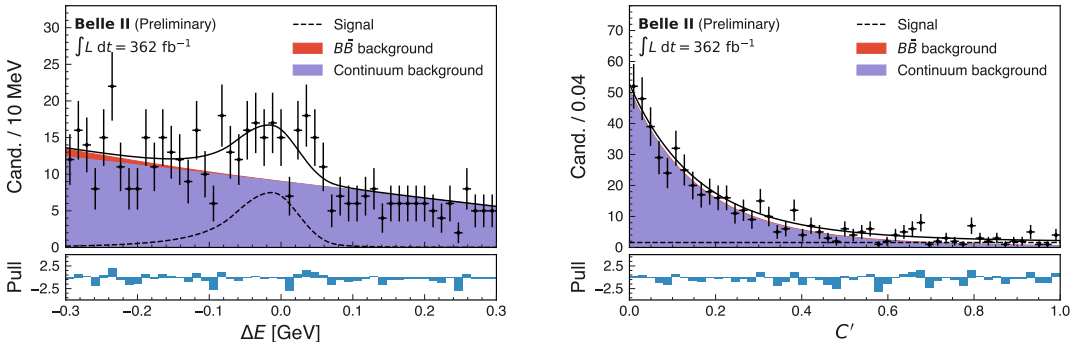


(b) Signal-enhanced region, defined by $-0.13 < \Delta E < 0.1$ and $0.24 < C' < 1$

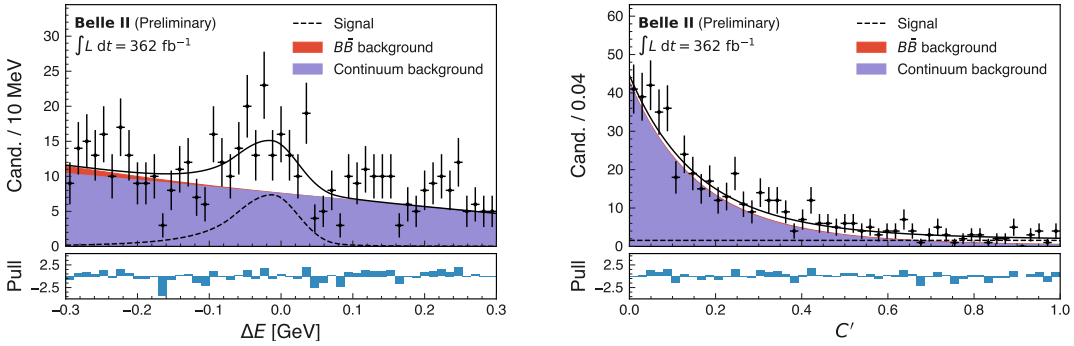


(c) Background-enhanced region, defined by $0.1 < \Delta E < 0.3$ and $0 < C' < 0.24$

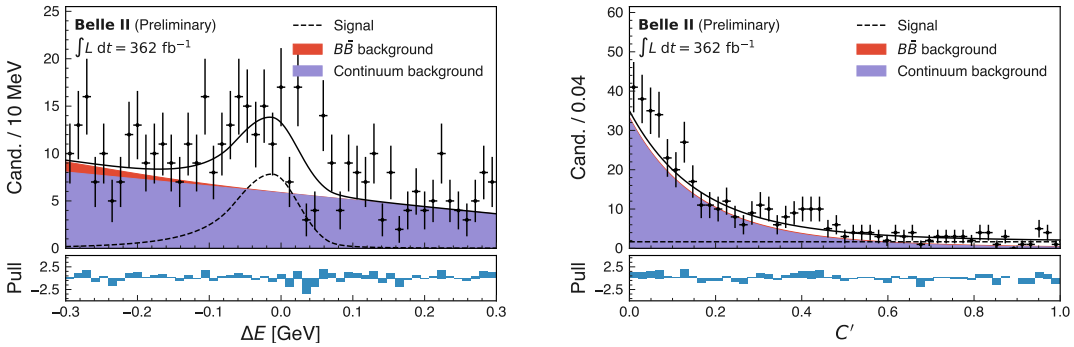
Figure 7.1: The ΔE and C' distributions of $362\,\mathrm{fb}^{-1}$ data in (a) the fitting region, (b) the signal-enhanced region, and (c) the background-enhanced region. The fit result to the sample is overlaid with a solid black line, and the fit components are shown as black dashed line (signal), red shaded area ($B\overline{B}$ background), and purple shaded area (continuum background).



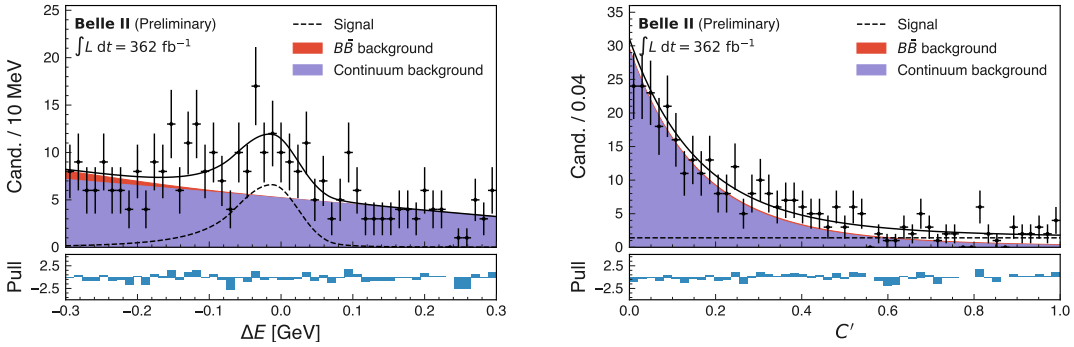
(a) 0th r-bin



(b) 1st r-bin

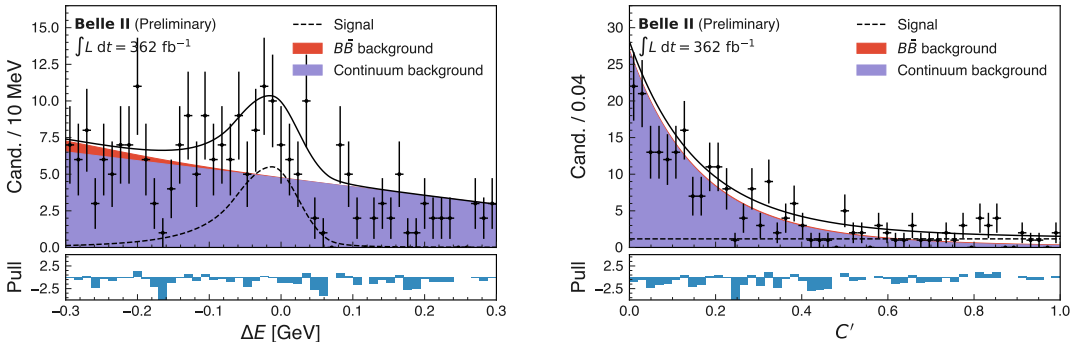


(c) 2nd r-bin

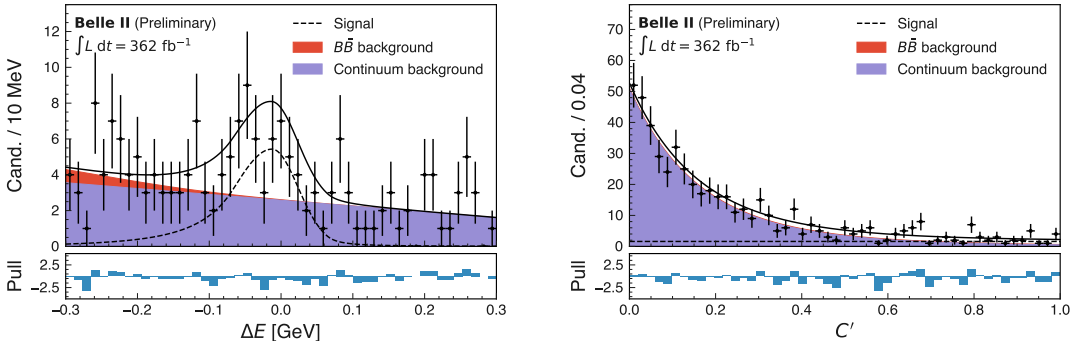


(d) 3rd r-bin

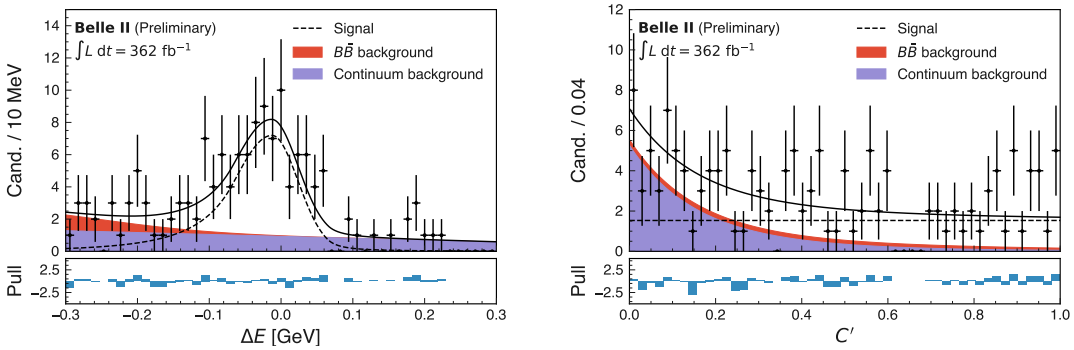
Figure 7.2: The ΔE and C' distributions of $362\,\mathrm{fb}^{-1}$ data projected to each r-bin.



(a) 4th r-bin

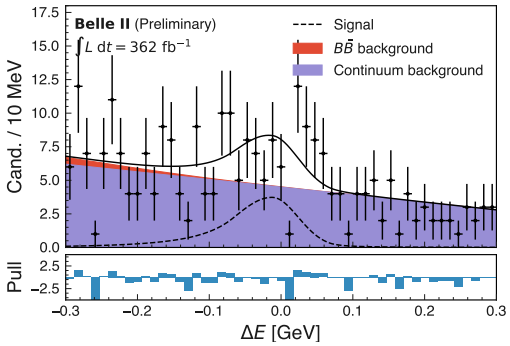


(b) 5th r-bin

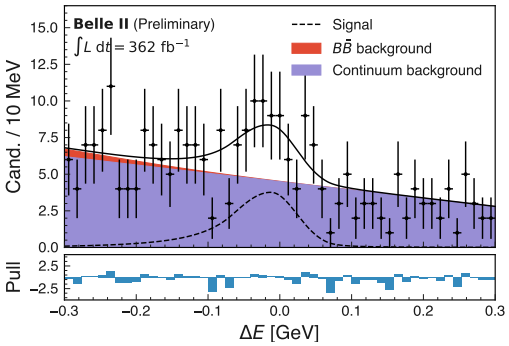


(c) 6th r-bin

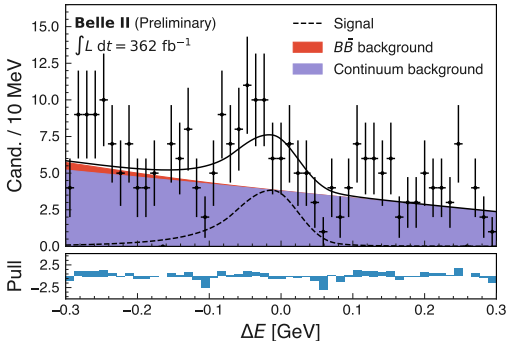
Figure 7.3: The ΔE and C' distributions of $362\,\mathrm{fb}^{-1}$ data projected to each r-bin.



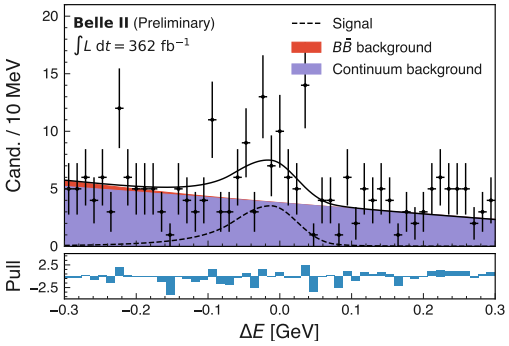
(a) 0th r-bin, $q = -1$



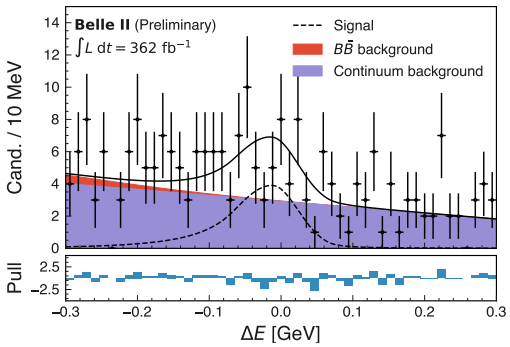
(b) 0th r-bin, $q = 1$



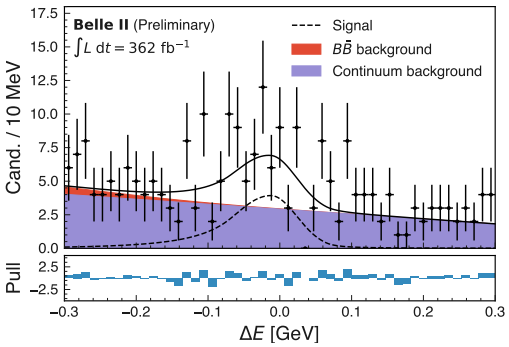
(c) 1st r-bin, $q = -1$



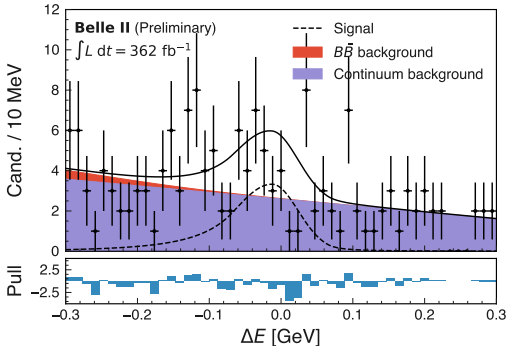
(d) 1st r-bin, $q = 1$



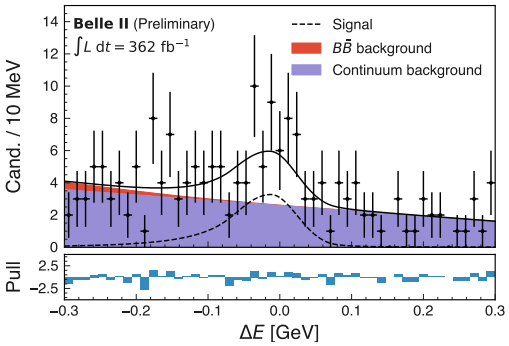
(e) 2nd r-bin, $q = -1$



(f) 2nd r-bin, $q = 1$

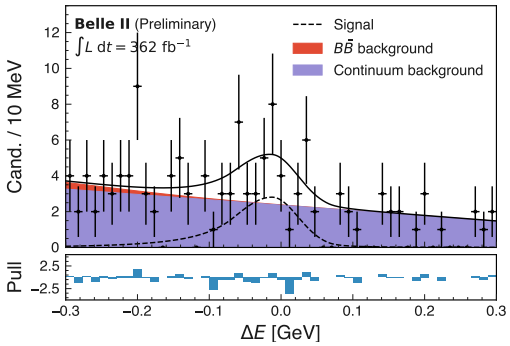


(g) 3rd r-bin, $q = -1$

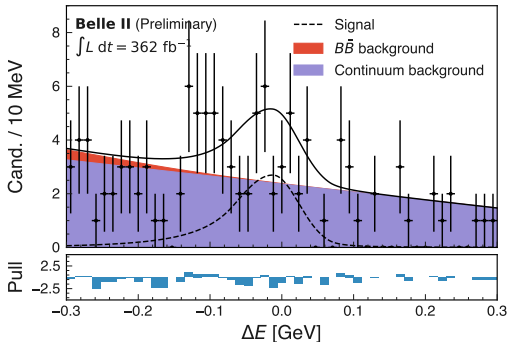


(h) 3rd r-bin, $q = 1$

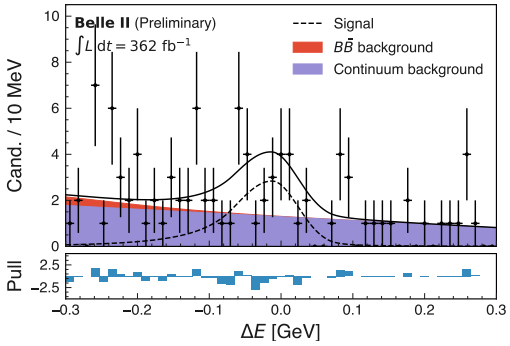
Figure 7.4: Flavour-specific ΔE distributions of $362\,\text{fb}^{-1}$ data projected to each r-bin.



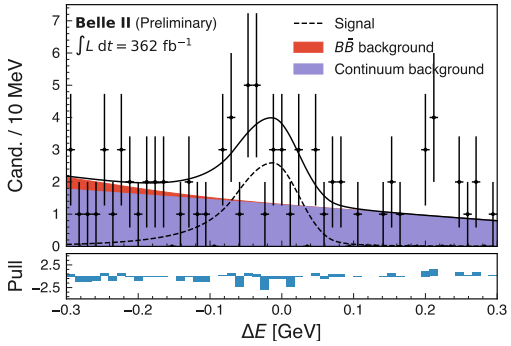
(a) 4th r-bin, $q = -1$



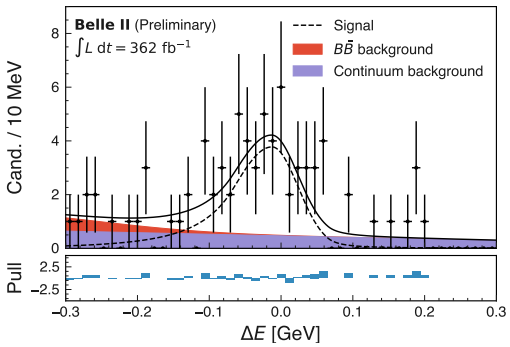
(b) 4th r-bin, $q = 1$



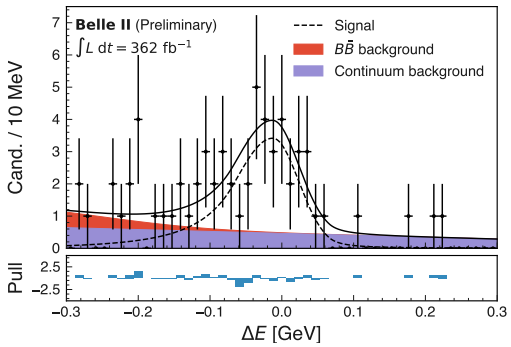
(c) 5th r-bin, $q = -1$



(d) 5th r-bin, $q = 1$

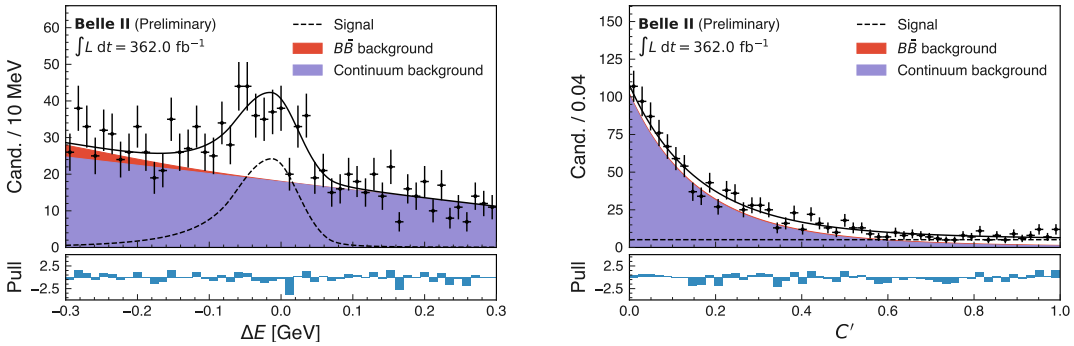


(e) 6th r-bin, $q = -1$

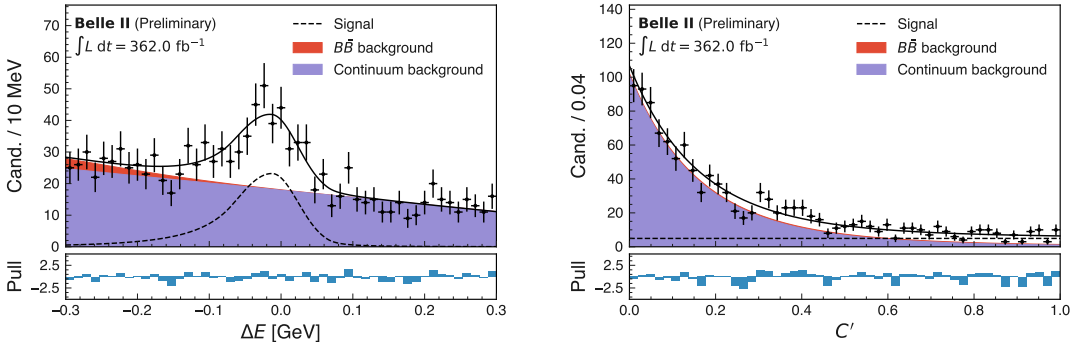


(f) 6th r-bin, $q = 1$

Figure 7.5: Flavour-specific ΔE distributions of $362\ \mathrm{fb}^{-1}$ data projected to each r-bin.



(a) $B_{tag} = \overline{B}^0$, $q = -1$



(b) $B_{tag} = B^0$, $q = 1$

Figure 7.6: Flavour-specific ΔE and C' projections of $362\ \mathrm{fb}^{-1}$ data.

7.2 Combination with Time-Dependent Analysis

Besides our time-integrated (TI) measurement, the the direct *CP* asymmetry and the branching fraction of $B^0 \to K^0\pi^0$ have also been measured in a decay-time-dependent (TD) analysis [54], featuring a different selection scheme from that of this analysis. The signal reconstruction efficiency of Ref. [54] is 25% lower, and the fraction of common events between the two analyses is 46% of those used here. This fraction increases to 53% in the signal-enhanced region, defined as $-0.13 < \Delta E < 0.1$ and $C' > 0.24$.

To compute the correlations of the measurements, we separate the data sample into (a) overlapping events that are analysed in both analyses, (b) events that are only in the TD analysis, and (c) events that are only in the TI analysis (this analysis). We randomly bootstrap 1000 samples from (a), (b), and (c), allowing for replacement of candidates, and create 1000 (a)+(b) and (a)+(c) samples, where the same overlapping bootstrapped sample (a) is used. By rerunning the two analyses, TD for (a)+(b) and TI for (a)+(c), we compute the statistical correlation between the direct *CP* asymmetry to be 21%, and the branching fraction to be 76%. The 2D distributions of $\mathcal{A}_{K^0\pi^0}$ and $\mathcal{B}_{K^0\pi^0}$ are shown in Fig. 7.7.

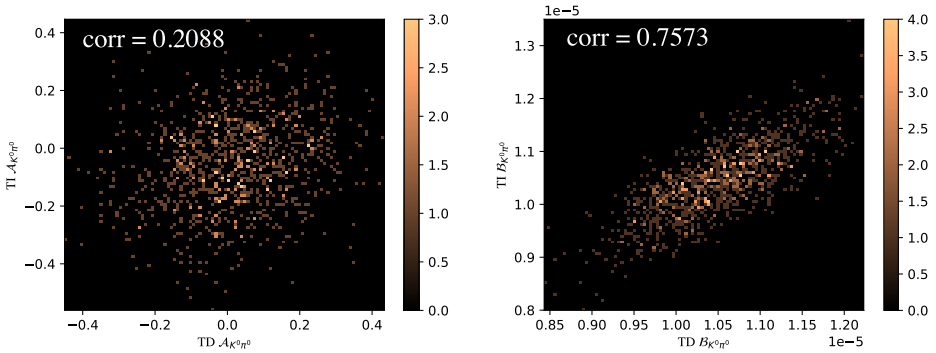


Figure 7.7: 2D distribution of the TI and TD measurements of $\mathcal{A}_{K^0\pi^0}$ (left) and $\mathcal{B}_{K^0\pi^0}$ (right).

The measurements are in agreement considering uncertainties, thus we combine them with the best linear unbiased estimator [55] to enhance precision. The results of both analyses, the correlations between the measurements, and the combination are listed in Tab. 7.2. Taking into account the systematic uncertainties, the combined branching fraction and direct *CP* asymmetry measurements are

$$\mathcal{B}_{K^0\pi^0} = 10.50 \pm 0.62 \pm 0.69,$$

$$\mathcal{A}_{K^0\pi^0} = -0.01 \pm 0.12 \pm 0.05,$$

which are the Belle II result of $B^0 \to K^0_S\pi^0$ for the Moriond 2023 dataset. We assume the systematic uncertainties of the two analyses to be 100% correlated since the dominant contributions are the same, that is, π^0 and K^0_S efficiencies for the branching fraction and the $B\overline{B}$ background asymmetry for the direct *CP* asymmetry. Therefore, we simply assign the systematic uncertainties estimated in Ch. 8.

Table 7.2: Branching ratio and direct *CP* asymmetry measurements of this analysis and the TD analysis [54], the combined results, and the correlation of the measurements between the two analyses.

	TI	TD	combination	corr. [%]
$\mathcal{A}_{K^0\pi^0}$	-0.06 ± 0.15	0.04 ± 0.15	-0.01 ± 0.12	20.88
$\mathcal{B}_{K^0\pi^0}$ $[\times 10^{-6}]$	10.16 ± 0.65	11.00 ± 0.67	10.50 ± 0.62	75.73

7.3 Isospin Sum Rule

We calculate the value of the isospin sum rule described by Eqn. 1.7 with Belle II measurements of the branching ratios and direct *CP* asymmetries [17], and the ratio of τ_{B^0}/τ_{B^+} being 0.9273 ± 0.0033 [6]. The Belle II $K\pi$ measurements using the Moriond 2023 dataset are summarised in Tab. 7.3, from which ratios of branching fractions are computed for the

calculation of the sum rule, as shown in Tab. 7.4. Common systematic uncertainties are cancelled out in ratios of branching fractions, including those originating from the tracking efficiency, the number of produced B-mesons, and $f^{+-/00}$. We calculate the value of the isospin sum rule to be

$$I_{K\pi} = -0.03 \pm 0.13 \pm 0.05,$$

where the correlations between uncertainties are accounted for. This result is in accordance with the SM prediction considering uncertainties. To be noted, the dominant uncertainty on $I_{K\pi}$ is the statistical uncertainty of $A_{K^0\pi^0}$. The update of the $A_{K^0\pi^0}$ measurement at Belle II is therefore crucial to improving the precision of the isospin sum rule test.

Table 7.3: Branching fractions and direct *CP* asymmetry measurements using Moriond 2023 dataset [17].

Modes	$\mathcal{B}[\times 10^{-6}]$	$\mathcal{A}$
$B^0 \to K^+\pi^-$	$20.67 \pm 0.37 \pm 0.62$	$-0.072 \pm 0.019 \pm 0.007$
$B^+ \to K^+\pi^0$	$14.21 \pm 0.38 \pm 0.85$	$0.013 \pm 0.027 \pm 0.005$
$B^+ \to K^0\pi^+$	$24.4 \pm 0.71 \pm 0.86$	$0.046 \pm 0.029 \pm 0.007$
$B^0 \to K^0\pi^0$	$10.50 \pm 0.62 \pm 0.69$	$-0.01 \pm 0.12 \pm 0.05$

Table 7.4: Ratios of branching fractions used as input for the calculation of $I_{K\pi}$. The ratios are calculated from the branching fractions in Tab. 7.3.

Modes	Ratio
$\mathcal{B}_{K^0\pi^+}/\mathcal{B}_{K^+\pi^-}$	$1.180 \pm 0.040 \pm 0.063$
$\mathcal{B}_{K^+\pi^0}/\mathcal{B}_{K^+\pi^-}$	$0.687 \pm 0.022 \pm 0.050$
$\mathcal{B}_{K^0\pi^0}/\mathcal{B}_{K^+\pi^-}$	$0.508 \pm 0.031 \pm 0.030$

Chapter 8 Systematic Uncertainties

We summarise the systematic uncertainties of the measured branching fraction $\mathcal{B}_{K^0\pi^0}$ and direct *CP* asymmetry $\mathcal{A}_{K^0\pi^0}$ in Tab. 8.1. The total systematic uncertainties are 6.4% and 0.05 for the $\mathcal{B}_{K^0\pi^0}$ and $\mathcal{A}_{K^0\pi^0}$, respectively. For the branching fraction, the systematic and statistical uncertainties are of the same magnitude, and for the direct *CP* asymmetry the precision is dominated by the statistical uncertainty. Each listed contribution is documented in the following sections.

Table 8.1: Summary of the fractional uncertainties on the branching ratio and the absolute uncertainties on the direct *CP* asymmetry.

Source	$\mathcal{B}_{K^0\pi^0}$ [%]	$\mathcal{A}_{K^0\pi^0}$
Tracking	0.5	–
$N_{B\bar{B}}$	1.5	–
f^{+-}/f^{00}	2.5	–
π^0 Efficiency	5.0	–
K^0_S Efficiency	2.0	–
CS efficiency	1.7	–
ΔE shift and scale	1.7	<0.01
Signal model	0.1	<0.01
$B\overline{B}$ model	0.3	<0.01
Flavour model	<0.1	<0.01
$B\overline{B}$ background asymmetry	<0.1	0.05
$q\overline{q}$ background asymmetry	<0.1	0.02
χ_d	<0.1	<0.01
Multiple Candidates	0.3	<0.01
Syst. uncert.	6.6	0.05
Stat. uncert.	6.4	0.15

8.1 Tracking

We assign an uncertainty of 0.24% on the branching fraction per track in the final state on the branching fraction to accommodate for tracking efficiency uncertainties, as described in Ref. [56].

8.2 $B\overline{B}$ pair counting

The total number of produced $B\overline{B}$ pairs is $(387\pm6)\times10^6$. As provided by the performance group, we assign a systematic uncertainty of 1.5% on the branching fraction to account for the uncertainty of produced $B\overline{B}$ pairs.

8.3 f^{+-}/f^{00}

The uncertainty on the fraction f^{+-}/f^{00} of $\Upsilon(4S) \to B^+B^-$ and $B^0\overline{B}^0$ is assigned as systematic uncertainty on the branching fraction [57].

8.4 π^0 Reconstruction Efficiency

The π^0 reconstruction efficiency is assessed using $D^0 \to K^+\pi^-\pi^0$ and $\tau^+ \to \pi^+\pi^-\pi^+\pi^0\nu_\tau$ decays. The full documentation of this systematic uncertainty can be found in Ref. [58]. Fig. 8.1 shows the π^0 momentum distribution of correctly reconstructed signal events in $B^0 \to K^0_S\pi^0$ signal-only sample. Tab. 8.2 shows the π^0 data-MC ratios of reconstruction efficiency for π^0 of different momentum ranges. The averaged value is $(1.018\pm0.050)\%$. The ratio is compatible with unity. The uncertainty of this ratio is assigned as the systematic uncertainty on the branching fraction. This is the leading component for the branching fraction.

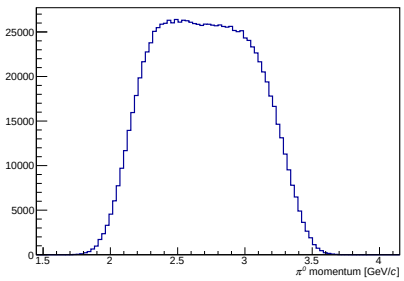


Figure 8.1: Distribution of π^0 momentum of correctly reconstructed signal-only simulation sample.

Table 8.2: Data-MC ratio of π^0 efficiency. The first uncertainty is the statistical uncertainty, the second the systematic uncertainty, and the third is the uncertainty originated from the difference of D and τ measurements. All uncertainties except the first are correlated across the momentum bins.

p_{π^0} [GeV/c]	data-MC ratio of ϵ_{π^0}
$1.0 - 1.5$	$1.047 \pm 0.003 \pm 0.043 \pm 0.039$
$1.5 - 2.0$	$1.044 \pm 0.003 \pm 0.041 \pm 0.051$
$2.0 - 3.0$	$1.018 \pm 0.003 \pm 0.039 \pm 0.031$

8.5 K^0_S Reconstruction Efficiency

The K^0_S reconstruction efficiency is evaluated using $D^{*+} \to \overline{D}^0(\to K^0_S\pi^+\pi^-)\pi^+$ and $D^{*+} \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$ decays. Fig. 8.2 shows the flight distance and $\cos\theta$ distribution of correctly reconstructed signal events in $B^0 \to K^0_S\pi^0$ signal-only sample. Tab. 8.3 shows the K^0_S data-MC ratios of reconstruction efficiency for K^0_S of different flight distance and $\cos\theta$ ranges. An additional systematic uncertainty is assigned for normalisation of the first K^0_S flight distance bin, calculated by adding in quadrature weighted by the fraction of events per $\cos\theta$ bin the statistical uncertainties of the ratios in the first K^0_S flight distance bin. The systematic uncertainty due to the K^0_S selection is computed in three bins, with corrections shown in Tab. 8.4. The ratio is 0.9955 ± 0.0025, which is compatible with unity. The overall averaged value of ratio of data-MC K^0_S efficiency is $(0.938 \pm 0.020)\%$. The ratio is incompatible with unity. The uncertainty of this ratio is assigned as the systematic uncertainty on the branching fraction.

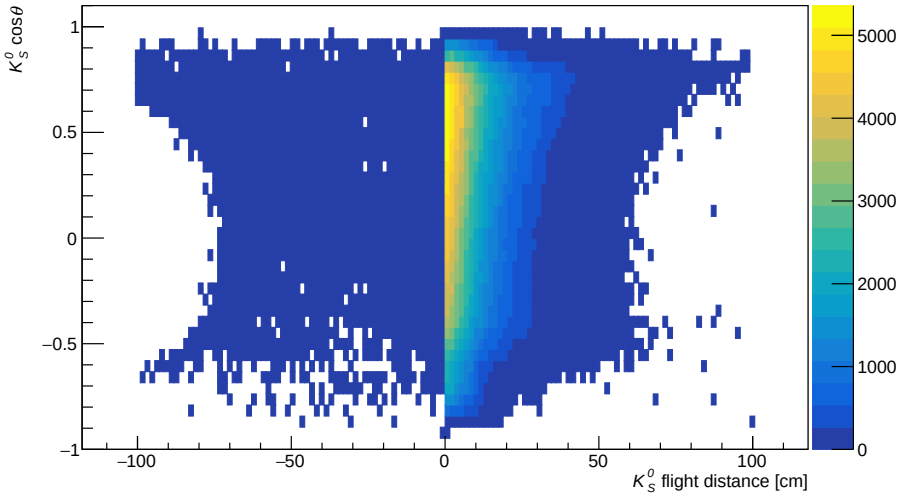


Figure 8.2: Distribution of K^0_S flight distance and $\cos\theta$ of correctly reconstructed signal-only simulation sample.

Table 8.3: Data-MC ratio of K^0_S efficiency. The first uncertainty is the statistical uncertainty and the second the systematic uncertainty, which is correlated across the K^0_S distance and cos θ bins.

K^0_S flight	$\cos\theta_{K^0_S}$		
distance [cm]	-1 – 0.3	0.3 – 0.7	0.7 – 1
0.5 – 1.4	$1.000 \pm 0.019 \pm 0.000$	$1.000 \pm 0.018 \pm 0.000$	$1.000 \pm 0.022 \pm 0.000$
1.4 – 2.4	$0.995 \pm 0.018 \pm 0.024$	$0.973 \pm 0.016 \pm 0.001$	$1.050 \pm 0.021 \pm 0.017$
2.4 – 3.6	$0.985 \pm 0.016 \pm 0.024$	$0.972 \pm 0.015 \pm 0.010$	$1.009 \pm 0.018 \pm 0.011$
3.6 – 5.2	$1.022 \pm 0.016 \pm 0.018$	$0.966 \pm 0.014 \pm 0.011$	$1.038 \pm 0.016 \pm 0.019$
5.2 – 7.0	$1.012 \pm 0.016 \pm 0.020$	$0.980 \pm 0.014 \pm 0.011$	$1.006 \pm 0.016 \pm 0.012$
7.0 – 9.0	$0.945 \pm 0.017 \pm 0.006$	$0.958 \pm 0.014 \pm 0.012$	$1.006 \pm 0.016 \pm 0.017$
9.0 – 13.0	$0.917 \pm 0.015 \pm 0.012$	$0.923 \pm 0.012 \pm 0.013$	$1.000 \pm 0.013 \pm 0.017$
13.0 –22.0	$0.863 \pm 0.014 \pm 0.014$	$0.906 \pm 0.011 \pm 0.005$	$0.960 \pm 0.013 \pm 0.002$
22.0 – 35.0	$0.766 \pm 0.024 \pm 0.080$	$0.854 \pm 0.014 \pm 0.000$	$0.940 \pm 0.018 \pm 0.027$
35.0 – 55.0	$0.795 \pm 0.083 \pm 0.189$	$0.854 \pm 0.034 \pm 0.022$	$0.964 \pm 0.032 \pm 0.004$

Table 8.4: Data-MC raio of K^0_S efficiency due to K^0_S selection.

	data-MC ratio of $\epsilon_{K^0_S}$
$\cos\theta_{K^0_S} > 0.4$ $\texttt{flightDistance} \geq 10$	0.9997 ± 0.0020
$\cos\theta_{K^0_S} > 0.4$ $\texttt{flightDistance} < 10$	0.9996 ± 0.0049
$\cos\theta_{K^0_S} < 0.4$	0.9927 ± 0.0038

8.6 Continuum Suppression Efficiency

The data-MC continuum suppression (CS) efficiency ratio obtained in Sec. 6.1 with the control mode $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$ is $(96.79 \pm 1.68)\%$. We assign the uncertainty of the ratio as the systematic uncertainty on the branching fraction.

8.7 Fudge Factors

To estimate the systematic uncertainty associated to the ΔE shift and scaling parameters, we generate 100 toy datasets around the nominal fit shape. Each dataset is fitted with the nominal fit shape and an alternative fit shape where the values of the fudge factors are drawn randomly within their uncertainties. The difference between the two results are compared and the width σ of the differences is assigned as the systematic uncertainty. The distributions of the fit differences are shown in Fig. 8.3. The contribution to the systematic

uncertainty of $\mathcal{A}_{K^0\pi^0}$ is insignificant.

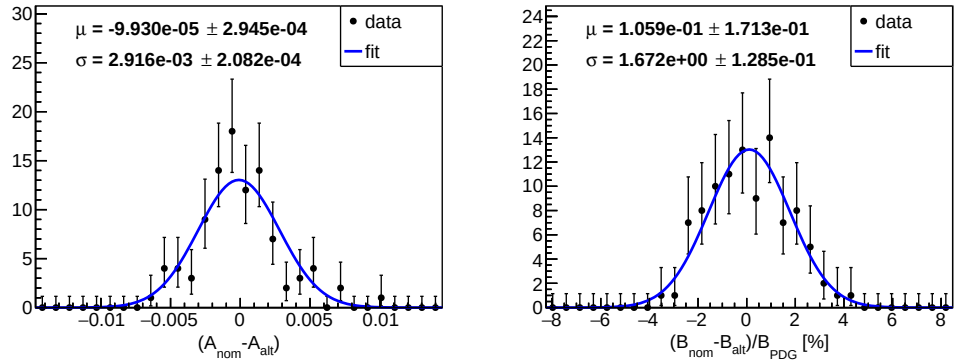


Figure 8.3: Distribution of $\mathcal{A}_{K^0\pi^0}$ (left) and $\mathcal{B}_{K^0\pi^0}$ (right) fit differences obtained using nominal and alternative ΔE shift and scaling parameters.

8.8 Signal Modelling

To obtain systematic uncertainties originating from imperfect modelling, we generate 500 toy datasets around the nominal fit shape. Each dataset is fitted with the nominal fit shape and an alternative fit shape where the shape parameters of the signal component are varied randomly within their uncertainties, extracted from the fit to large dedicated simulated signal sample. We then quote the width σ of the differences in measured physics parameters of the nominal and alternative fits as the systematic uncertainty. The distributions of the fit differences are shown in Fig. 8.4. The contribution to the systematic uncertainty of $\mathcal{A}_{K^0\pi^0}$ is insignificant, the distribution of $\mathcal{A}_{nom} - \mathcal{A}_{alt}$ clustered toward zero.

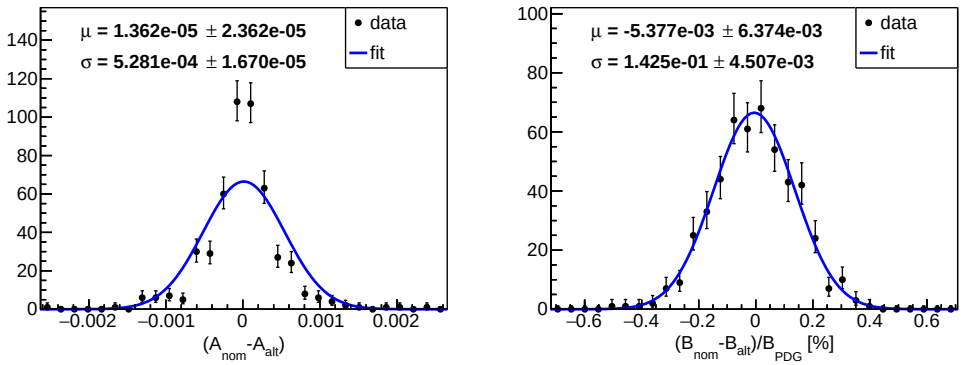


Figure 8.4: Distribution of $\mathcal{A}_{K^0\pi^0}$ (left) and $\mathcal{B}_{K^0\pi^0}$ (right) fit differences obtained using nominal and alternative signal shape parameters.

8.9 $B\overline{B}$ Background Modelling

To assess a systematic uncertainty for the $B\overline{B}$ shape, we replace the Gaussian PDF used to describe the ΔE shape with a kernel PDF. In addition, we vary the $B\overline{B}\,C'$ shape within its uncertainties. We generate 100 toy datasets around this alternative fit model and fit the datasets with the alternative and nominal fit model. The mean μ of the differences between the fit results is assigned as systematic uncertainty. The distributions of the fit differences are shown in Fig. 8.5. The contribution to the systematic uncertainty of $\mathcal{A}_{K^0\pi^0}$ is insignificant.

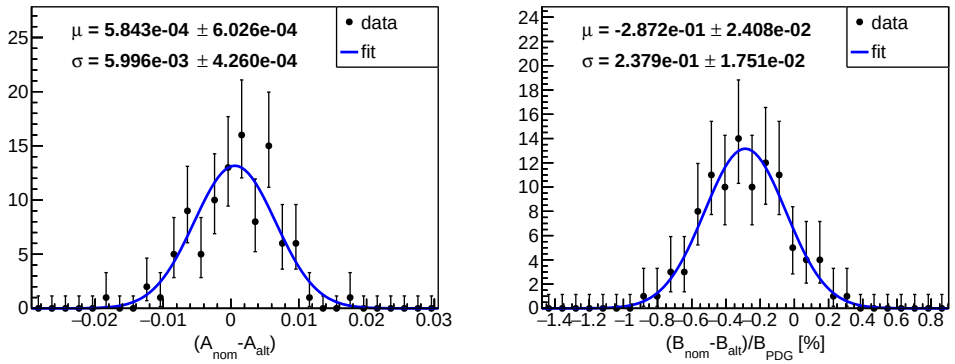


Figure 8.5: Distributions of $\mathcal{A}_{K^0\pi^0}$ (left) and $\mathcal{B}_{K^0\pi^0}$ (right) fit differences obtained using nominal and alternative $B\overline{B}$ models and shape parameters.

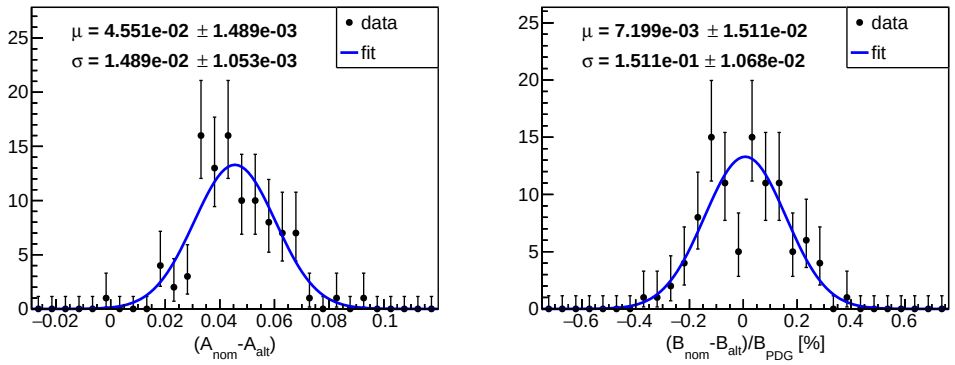
8.10 Flavour Tagging Modelling

The uncertainties associated with the flavour parameters are already accounted for since in the simultaneous fit the flavour parameters are Gaussian-constrained by their uncertainties. To estimate their contribution, we perform the fit with and without Gaussian constraints on the flavour parameters, and quote the difference in quadrature between the uncertainty of $\mathcal{A}_{K^0\pi^0}$ as the systematic uncertainty. The contributions to the systematic uncertainties of both the branching fraction and the *CP* asymmetry are insignificant.

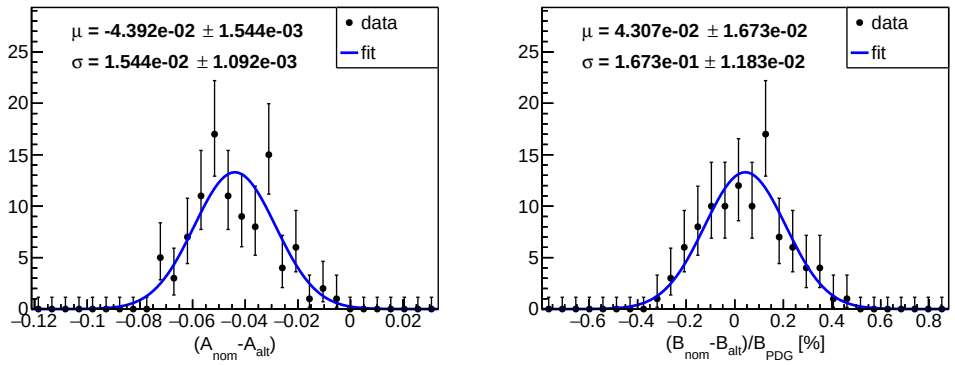
8.11 $B\overline{B}$ Background Asymmetry

The $B\overline{B}$ background is assumed to be *CP*-symmetric in the nominal fit, but it is possible for there to be a non-trivial asymmetry. We generate two sets of 100 toy datasets, where the $B\overline{B}$ background asymmetry is fixed to ± 1. Then we perform two fits on each toy dataset: the nominal fit with $\mathcal{A}_{B\overline{B}}$ fixed to zero, and the alternative fit with $\mathcal{A}_{B\overline{B}} = \pm 1$.

The distributions of the fit differences are shown in Fig. 8.6. We quote the larger of the mean μ of the differences between the nominal and alternative fit results as the systematic uncertainty. This is the leading component for $\mathcal{A}_{K^0\pi^0}$, though setting the $B\overline{B}$ background asymmetry to extreme values in the evaluation may result in overestimation of the systematic uncertainty. The $B\overline{B}$ background asymmetry has insignificant contribution on the branching fraction.



(a) $\mathcal{A}_{B\overline{B}} = +1$



(b) $\mathcal{A}_{B\overline{B}} = -1$

Figure 8.6: Distributions of $\mathcal{A}_{K^0\pi^0}$ (left) and $\mathcal{B}_{K^0\pi^0}$ (right) fit differences obtained with $\mathcal{A}_{B\overline{B}} = \pm 1$.

8.12 Continuum Background Asymmetry

We generate 100 toy datasets from the nominal PDF and perform alternative fits with the $q\overline{q}$ background asymmetry is fixed to the value $\mathcal{A}_{q\overline{q}} = -0.021$, obtained from the difference between the numbers of tagged B and $\overline{B}$ in the side-band data ($\Delta E > 0.1\,\text{GeV}$). The mean μ of the differences between the nominal and alternative fit results is assigned as the systematic uncertainty. The distributions of the fit differences are shown in Fig. 8.7. The contribution of the $q\overline{q}$ background asymmetry is insignificant for the branching fraction.

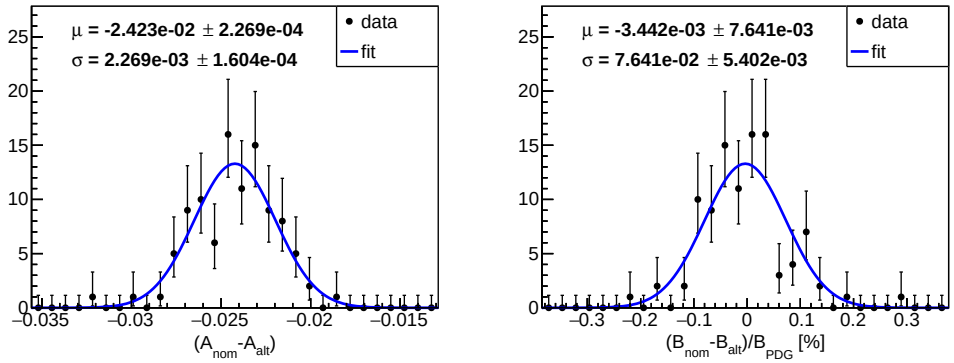


Figure 8.7: Distributions of $\mathcal{A}_{K^0\pi^0}$ (left) and $\mathcal{B}_{K^0\pi^0}$ (right) fit differences obtained with fixed $\mathcal{A}_{q\bar{q}}$.

8.13 B^0 Mixing Parameter χ_d

The value of χ_d is fixed in the nominal fit, where in the alternative fit we vary χ_d around the nominal value within its uncertainties (0.1858 ± 0.0011). We generate 100 toy datasets around the nominal fit model, and the mean μ of the differences between the nominal and alternative fit results is assigned as the systematic uncertainty. Its contributions to both $\mathcal{B}_{K^0\pi^0}$ and $\mathcal{A}_{K^0\pi^0}$ are insignificant.

8.14 Multiple Candidates

We apply no best-candidate selection in this analysis since the multiplicity is low. To assess the systematic uncertainty associated with multiple candidates, we repeat the fit to the data with a random best-candidate selection after unblinding. The fit result is shown in Fig. 8.8. The difference from the nominal fit result is taken as the systematic uncertainty. The contribution to the systematic uncertainty of $\mathcal{A}_{K^0\pi^0}$ is insignificant.

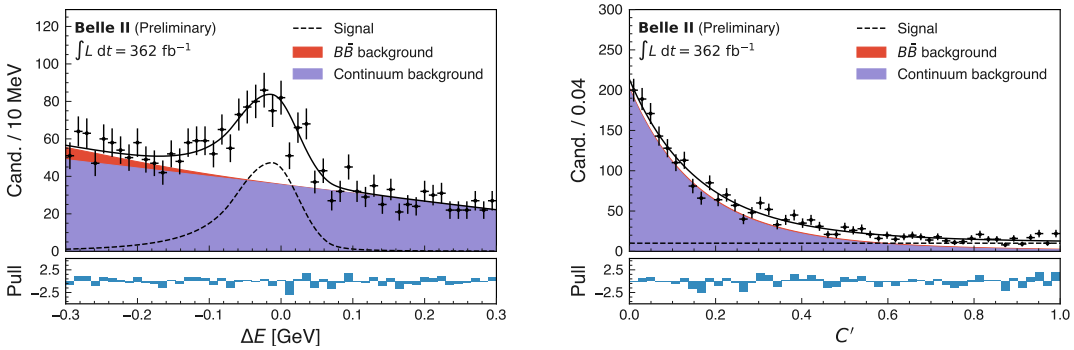


Figure 8.8: The ΔE and C' distributions of $362\,\text{fb}^{-1}$ data with a best-candidate selection.

Chapter 9 Conclusions

We report an update of the measurement of the direct *CP* asymmetry and the branching fraction of $B^0 \to K^0\pi^0$ using the data collected by Belle II at the $\Upsilon(4S)$ resonance from 2019 to 2022, which corresponds to an integrated luminosity of $362\,\text{fb}^{-1}$. A total of 502 ± 32 signal candidates are reconstructed and passed the candidate selection, optimised with simulation, and from which we perform a decay-time-integrated fit to extract the direct *CP* asymmetry and the branching fraction,

$$\mathcal{A}_{K^0\pi^0} = -0.06 \pm 0.15 \pm 0.05,$$

$$\mathcal{B}_{K^0\pi^0} = 10.16 \pm 0.65 \pm 0.65,$$

which are in agreement with current world averages [6]. Our measurements are comparable in precision to previous Belle and BaBar measurements, in spite of using a smaller dataset. We expect the π^0 and K^0_S efficiency ratios, which are the leading components of the systematic uncertainty of the branching fraction, to be reduced with future accumulation of data. We combine the result of this analysis with that of a decay-time-dependent study [54] on $B^0 \to K^0\pi^0$ in order to reduce statistical uncertainties and provide the official Belle II measurements for Moriond 2023. With the other Belle II $B \to K\pi$ measurements [17], we calculate the isospin sum rule to be

$$I_{K\pi} = -0.03 \pm 0.13 \pm 0.05,$$

which agrees with the SM expectation, with a precision limited by the statistical uncertainty whose dominant contribution is the statistical uncertainty of $\mathcal{A}_{K^0\pi^0}$. We obtain a result with a better precision than that of Belle, even though the Belle II dataset amounts to only half the size, due to the improvement of $\mathcal{A}_{K^0\pi^0}$ precision. With future accumulation of Belle II data, the precision of $\mathcal{A}_{K^0\pi^0}$, and therefore $I_{K\pi}$ can be further improved. In Fig. 9.1 we estimate the statistical uncertainty of $\mathcal{A}_{K^0\pi^0}$ at increasing integrated luminosities up to $10\,\text{ab}^{-1}$ using toy datasets drawn from the PDF, as well as the uncertainty of $I_{K\pi}$ computed with these estimations of $\mathcal{A}_{K^0\pi^0}$ uncertainty and current Belle II uncertainties of other $B \to K\pi$ decays, as given in Tab. 7.3.

If we assume the isospin sum rule to be consistent with the SM, that is, $I_{K\pi} = 0$,

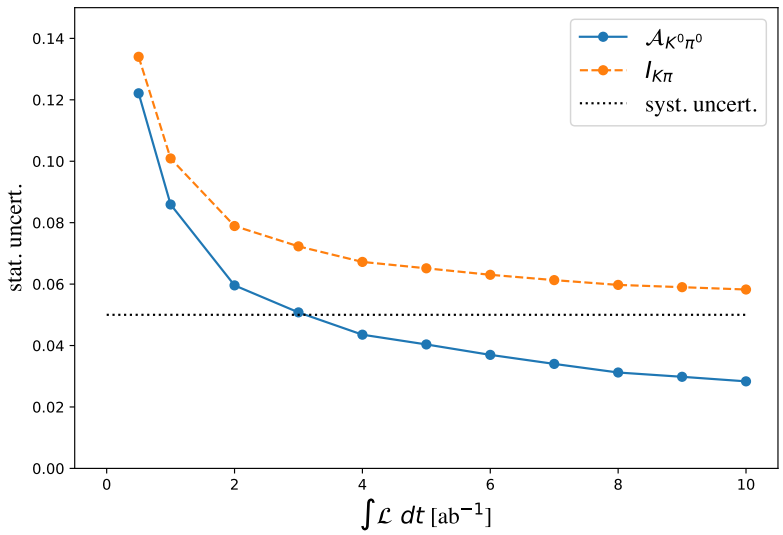


Figure 9.1: Estimated Belle II statistical uncertainties of $\mathcal{A}_{K^0\pi^0}$ and $I_{K\pi}$.

we may estimate the expected value of $\mathcal{A}_{K^0\pi^0}$, the input parameter with the highest uncertainty, using the world averages of other $K\pi$ branching fractions and asymmetry parameters [6]. The result is $\mathcal{A}_{K^0\pi^0} = -0.15 \pm 0.03$. The statistical uncertainty of the direct *CP* asymmetry approaches the systematic uncertainty at an integrated luminosity of $3\,\text{ab}^{-1}$. However, the leading component of the systematic uncertainty is overestimated (Sec. 8.9), and could be lowered with improved modelling of the $B\overline{B}$ background asymmetry. We expect Belle II to be able to claim a 5σ discovery for $\mathcal{A}_{K^0\pi^0}$ when it reaches an integrated luminosity of approximately $9\,\text{ab}^{-1}$, at which point the estimated uncertainty for $I_{K\pi}$ obtained from Belle II can be reduced to 0.06, as shown in Fig. 9.1.

References

[1] A. D. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*, Sov. Phys. Usp., **34**, 392 (1991).

[2] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett., **10**, 531 (1963).

[3] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys., **49**, 652 (1973).

[4] L.-L. Chau and W.-Y. Keung, *Comments on the parametrization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett., **53**, 1802 (1984).

[5] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the* 2π *decay of the* K_2^0 *meson*, Phys. Rev. Lett., **13**, 138 (1964).

[6] R. L. Workman *et al.* (Particle Data Group), *Review of Particle Physics*, Prog. Theor. Exp. Phys., **2022**, 083C01 (2022).

[7] M. Gronau, *A precise sum rule among four* $B \to K\pi$ *CP asymmetries*, Phys. Lett. B, **627**, 82 (2005).

[8] M. Gronau and J. L. Rosner, *Combining CP asymmetries in* $B \to K\pi$ *decays*, Phys. Rev. D, **59**, 113002 (1999).

[9] T. Gershon and A. Soni, *Null tests of the standard model at an international super B factory*, J. Phys. G: Nucl. Part. Phys., **34**, 479 (2007).

[10] Y. Amhis *et al.* (HFLAV Collaboration), *Averages of* b*-hadron,* c*-hadron, and* τ*-lepton properties as of 2021* (2022), arXiv:2206.07501.

[11] Y.-T. Duh *et al.* (Belle Collaboration), *Measurements of branching fractions and direct CP asymmetries for* $B \to K\pi$*,* $B \to \pi\pi$ *and* $B \to KK$ *decays*, Phys. Rev. D, **87**, 031103 (2013).

[12] B. Aubert *et al.* (BaBar Collaboration), *Measurement of time dependent CP asymmetry parameters in* B^0 *meson decays to* ωK_S^0*,* $\eta' K^0$*, and* $\pi^0 K_S^0$, Phys. Rev. D, **79**, 052003 (2009).

[13] J. P. Lees *et al.* (BaBar Collaboration), *Measurement of CP asymmetries and branching fractions in charmless two-body* B*-meson decays to pions and kaons*, Phys. Rev. D, **87**, 052009 (2013).

[14] M. Fujikawa *et al.* (Belle Collaboration), *Measurement of CP asymmetries in* $B^0 \to K^0\pi^0$ *decays*, Phys. Rev. D, **81**, 011101 (2010).

[15] P.-T. Chang, Y.-C. Chen, and C.-H. Li (Belle II Collaboration), *Preliminary measurement of yield and branching fraction and direct CP violation in* $B^0 \to K^0\pi^0$ *decay at Belle II* (2020), BELLE2-NOTE-PH-2020-046.

[16] F. Abudinén *et al.* (Belle II Collaboration), *B-flavor tagging at Belle II*, Eur. Phys. J. C, **82**, 283 (2022).

[17] T. Humair, H.-G. Moser, J. Skorupa, B. Wach, P.-T. Chang, and S.-P. Lin (Belle II Collaboration), *Measurement of the branching ratio and direct CP asymmetry of* $B^0 \to K^+\pi^-$, $B^0 \to \pi^+\pi^-$, $B^+ \to \pi^+\pi^0$, $B^+ \to K^+\pi^0$, $B^+ \to K^0_S\pi^+$, *and* $B^0 \to K^0_S\pi^0$ *using the moriond 2023 dataset* (2022), BELLE2-NOTE-PH-2022-062.

[18] E. Kou *et al.* (Belle II Collaboration), *The Belle II physics book*, Prog. Theor. Exp. Phys., **2019**, 123C01 (2019).

[19] A. J. Bevan *et al.*, *The physics of the B factories*, Eur. Phys. J. C, **74**, 3026 (2014).

[20] Y. Ohnishi *et al.*, *Accelerator design at SuperKEKB*, Prog. Theor. Exp. Phys., **2013**, 03A011 (2013).

[21] S. Collaboration, *SuperB: A high-luminosity asymmetric e+ e- super flavor factory. conceptual design report* (2007), arXiv:0709.045.

[22] G. Casarosa, *Inner tracking devices at the Belle II experiment*, Proc. Sci., **EPS-HEP2015**, 255 (2015).

[23] T. Abe *et al.*, *Belle II technical design report* (2010), arXiv:1011.0352.

[24] R. Pestotnik, *Particle identification devices at the Belle II experiment*, Proc. Sci., **EPS-HEP2015**, 256 (2016).

[25] B. N. Ratcliff, *Imaging rings in ring imaging Cherenkov counters*, Nucl. Instrum. Methods Phys. Res. A, **502**, 211 (2003).

[26] A. Abashian *et al.*, *The Belle detector*, Nucl. Instrum. Methods Phys. Res. A, **479**, 117 (2002).

[27] S. Weinzierl, *Introduction to Monte Carlo methods* (2000), arXiv:hep-ph/0006269.

[28] S. Jadach, B. Ward, and Z. Waş, *The precision Monte Carlo event generator for two-fermion final states in collisions*, Comput. Phys. Commun., **130**, 260 (2000).

[29] T. Sjöstrand *et al.*, *An introduction to PYTHIA 8.2*, Comput. Phys. Commun., **191**, 159 (2015).

[30] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Methods Phys. Res. A, **462**, 152 (2001).

[31] T. Ferber and P. Urquijo, *Overview of the Belle II physics generators* (2015), BELLE2-NOTE-PH-2015-006.

[32] J. Allison *et al.*, *Recent developments in Geant4*, Nucl. Instrum. Meth. A, **835**, 186 (2016).

[33] J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci., **53**, 270 (2006).

[34] S. Agostinelli *et al.*, *Geant4 - a simulation toolkit*, Nucl. Instrum. Meth. A, **506**, 250 (2003).

[35] Z. Liptak *et al.*, *Measurements of beam backgrounds in superkekb phase 2*, Nucl. Instrum. Methods Phys. Res. A, **1040**, 167168 (2022).

[36] Belle II Collaboration, *Basf2 documentation of the standard K^0_S list*, https://b2-master.belle2.org/software/development/sphinx/analysis/doc/StandardParticles.html?highlight=stdkshorts#stdV0s.stdKshorts.

[37] B. Collaboration, *Basf2 documentation of tree fitter*, https://software.belle2.org/sphinx/release-05-02-17/analysis/doc/TreeFitter.html.

[38] Belle II Collaboration, *Basf2 documentation of the K^0_S selector*, https://b2-master.belle2.org/software/development/sphinx/analysis/doc/KsSelector.html?highlight=ksselector#ksselector.

[39] Belle II Collaboration, *Basf2 documentation of the standard photon list*, https://software.belle2.org/development/sphinx/analysis/doc/StandardParticles.html?highlight=stdphoton#stdPhotons.stdPhotons.

[40] P. Avery, *Applied fitting theory I, IV, VI*, (1999).

[41] F. Pham (Belle II Collaboration), *Measurement of $B^0 \to \pi^0\pi^0$ branching fraction and $\mathcal{A}_{CP}$* (2021), BELLE2-NOTE-PH-2021-050.

[42] TDCPV subgroup (Belle II Collaboration), *Confluence page: TDCPV recipes*, https://confluence.desy.de/display/BI/Recipes+for+TDCPV+tools, (2022).

[43] T. Keck, *FastBDT: A speed-optimized multivariate classification algorithm for the Belle II experiment*, Comput. Softw. Big Sci., **1**, 2 (2017).

[44] B. Collaboration, *Basf2 documentation of continuum suppression*, https://software.belle2.org/development/sphinx/online_book/basf2/cs.html?highlight=continuum%20suppression.

[45] J. H. Friedman, *Stochastic gradient boosting*, Comput. Stat. Data Anal., **38**, 367 (2002).

[46] L. Breiman, *Classification and regression trees (1st edition)*, Routledge (1984).

[47] J. H. Friedman, *Greedy function approximation: A gradient boosting machine.* Ann. Stat., **29**, 1189 (2001).

[48] Belle II Collaboration, *TMVA users guide*, https://root.cern.ch/download/doc/tmva/TMVAUsersGuide.pdf.

[49] F. Abudinén, C. Kiesling, and T. Kuhr (Belle II Collaboration), *Development of a b0 flavor tagger and performance study of a novel time-dependent CP analysis of the decay $B^0 \to \pi^0\pi^0$ at Belle II*, Ph.D. dissertation, Munich, Ludwig-Maximilians-Universität, Munich, (2018).

[50] H. Kakuno *et al.* (Belle II Collaboration), *Neutral b flavor tagging for the measurement of mixing-induced CP violation at Belle*, Nucl. Instrum. Methods Phys. Res. A, **533**, 516 (2004).

[51] A. J. Bevan *et al.* (BaBar Collaboration), *The physics of the B factories*, Eur. Phys. J. C, **74**, 3026 (2014).

[52] T. Humair, R. Mehta, and P. Stavroulakis (Belle II Collaboration), *Calibration of the decay-time resolution and flavor tagger using hadronic B decays reconstructed in moriond 2023 data* (2023), BELLE2-NOTE-TE-2023-002.

[53] K. Cranmer, *Kernel estimation in high-energy physics*, Comput. Phys. Commun., **136**, 198 (2001).

[54] S. Hazra, A. B. Kaliyar, and G. B. Mohanty (Belle II Collaboration), *Measurements of branching fraction and time-dependent CP asymmetry in $B^0 \to K^0_S\pi^0$ decays with moriond 2023 dataset* (2022), BELLE2-NOTE-PH-2022-065.

[55] A. Valassi, *Combining correlated measurements of several different physical quantities*, Nucl. Instrum. Methods Phys. Res. A, **500**, 391 (2003).

[56] B. Collaboration, *Conference readiness confluence page*, https://confluence.desy.de/display/BI/Conference+readiness.

[57] S. Choudhury *et al.* (Belle Collaboration), *Measurement of the B^+/B^0 production ratio in e^+e^- collisions at the $\Upsilon(4S)$ resonance using $B \to J/\psi(\ell\ell)K$ decays at Belle*, Phys. Rev. D, **107**, L031102 (2023).

[58] T. Koga (Belle II Collaboration), *Momentum dependent π^0 efficiency measurement with D decays* (2020), BELLE2-NOTE-PH-2020-061.

[59] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, Rev. Mod. Phys., **68**, 1125 (1996).

[60] M. Beneke, G. Buchalla, M. Neubert, and C. Sachrajda, *QCD factorization in $B \to \pi K$, $\pi\pi$ decays and extraction of Wolfenstein parameters*, Nucl. Phys. B, **606**, 245 (2001).

[61] M. Gronau, D. Pirjol, and T.-M. Yan, *Model-independent electroweak penguin amplitudes in b decays to two pseudoscalars*, Phys. Rev. D, **60**, 034021 (1999).

[62] M. Neubert and J. L. Rosner, *New bound on gamma from $B^\pm \to \pi K$ decays*, Phys. Lett. B, **441**, 403 (1998).

[63] A. Ryd *et al.*, *EvtGen: A Monte Carlo Generator for B-Physics* (2005), EVTGEN-V00-11-06.

Appendix

A.1 Derivation of the Isospin Sum Rule

Following the derivation of Ref. [7], we give a brief derivation of the isospin sum rule. In this derivation, no relation between $B \to K\pi$ and $B \to \pi\pi$ decays are assumed, in general. The effective Hamiltonian [59] for $B \to K\pi$ decays is

$$\mathcal{H}_{eff} = -\frac{4G_F}{\sqrt{2}} \left[\sum_{U=u,c} \lambda_U (c_1 O_1^U + c_2 O_2^U) - \lambda_t \sum_{i=3}^{10} c_i O_i \right], \tag{A.1}$$

where $\lambda_U = V_{Ub}^* V_{Us}$, $\lambda_t = V_{tb}^* V_{ts}$, O_j^U and O_i are four-quark operators, and c_i are the real Wilson coefficients. We may decompose the $B \to K\pi$ amplitudes with the effective Hamiltonian. Recall the unitarity condition of the CKM matrix $\lambda_u + \lambda_c + \lambda_t = 0$ and defining $P_{tc} = P_t - P_c$, $P_{uc} = P_u - P_c$, we may write the $K\pi$ amplitudes as

$$-A(K^+\pi^-) = \lambda_u(P_{uc} + T) + \lambda_t(P_{tc} + \frac{2}{3}P_{EW}^C), \tag{A.2}$$

$$-\sqrt{2}A(K^+\pi^0) = \lambda_u(P_{uc} + T + C + A) + \lambda_t(P_{tc} + P_{EW} + \frac{2}{3}P_{EW}^C), \tag{A.3}$$

$$A(K^0\pi^+) = \lambda_u(P_{uc} + A) + \lambda_t(P_{tc} - \frac{1}{3}P_{EW}^C), \tag{A.4}$$

$$\sqrt{2}A(K^0\pi^0) = \lambda_u(P_{uc} - C) + \lambda_t(P_{tc} - \frac{1}{3}P_{EW}^C), \tag{A.5}$$

where the amplitudes P_u, T, C, A and P_c are contributions from the first sum in Eq. A.1, with $U = u$ and $U = c$, respectively. P, T, C, A denotes penguin, color-allowed tree, color- suppressed tree and annihilation topologies, and penguin contributions are substituted with

$$\lambda_u C \to \lambda_u C + \lambda_t P_{EW}, \;\; \lambda_u T \to \lambda_u T + \lambda_t P_{EW}^C, \;\; \lambda_u P_{uc} \to \lambda_u P_{uc} - \frac{1}{3}P_{EW}^C. \tag{A.6}$$

Colour suppression is not assumed for C and P_{EW}^C, and neither is P_{uc} assumed to be smaller than T or C. This means that T, C, and P_{uc}, may be of comparable magnitudes. The same may be said for P_{EW} and P_{EW}^C.

The $K\pi$ amplitudes may also be decomposed according to the isospin symmetry. The isosinglet part of $\mathcal{H}_{eff}$ contributes to $B_{1/2}$, with $I(K\pi) = 1/2$, and the isotriplet part of $mathcal H_{eff}$ contributes to $A_{1/2}$, with $I(K\pi) = 1/2$, and $A_{3/2}$, with $I(K\pi) = 3/2$. The $K\pi$ amplitudes can be written as

$$-A(K^+\pi^-) = B_{1/2} - A_{1/2} - A_{3/2}, \tag{A.7}$$

$$-\sqrt{2}A(K^+\pi^0) = B_{1/2} + A_{1/2} - 2A_{3/2}, \tag{A.8}$$

$$A(K^0\pi^+) = B_{1/2} + A_{1/2} + A_{3/2}, \tag{A.9}$$

$$\sqrt{2}A(K^0\pi^0) = B_{1/2} - A_{1/2} + 2A_{3/2}. \tag{A.10}$$

Therefore we have

$$B_{1/2} = \lambda_u[P_{uc} + \frac{1}{2}(T + A)] + \lambda_t[P_{tc} + \frac{1}{6}P^C_{EW}], \tag{A.11}$$

$$A_{1/2} = -\frac{1}{6}\lambda_u(T - 2C - 3A) + \frac{1}{3}\lambda_t(P_{EW} - \frac{1}{2}P^C_{EW}), \tag{A.12}$$

$$A_{3/2} = -\frac{1}{3}\lambda_u(T + C) - \frac{1}{3}\lambda_t(P_{EW} + P^C_{EW}). \tag{A.13}$$

Eq. A.11-A.13 may be used to provide a basis for calculation of $B \to K\pi$ amplitudes [60]. Using approxmate SU(3) relations between P_{EW}, P^C_{EW} and T, C [61], [62] and the approximation $(c_9 - c_{10})/(c_1 - c_2) \approx (c_9 + c_{10})/(c_1 + c_2)$ [59], we may write

$$P_{EW} + P^C_{EW} \approx -\frac{3}{2}\frac{c_9 + c_{10}}{c_1 + c_2}(T + C), \tag{A.14}$$

$$P^C_{EW} \approx -\frac{3}{2}\frac{c_9 + c_{10}}{c_1 + c_2}C, \tag{A.15}$$

Direct *CP* asymmetries in $B \to K\pi$ decays are caused by interference of terms in the amplitudes involving λ_t and λ_u and different strong phases. By defining $\Delta(B \to f) = \Gamma(\overline{B} \to \bar{f}) - \Gamma(B \to f)$, we may write for $B \to K\pi$ decays

$$\Delta(K^+\pi^-) = \mathfrak{Im}\left[(P_{tc} + \frac{2}{3}P^C_{EW})(T + P_{uc})^*\right] I, \tag{A.16}$$

$$2\Delta(K^+\pi^0) = \mathfrak{Im}\left[(P_{tc} + P_{EW} + \frac{2}{3}P^C_{EW})(TC + A + P_{uc})^*\right] I, \tag{A.17}$$

$$\Delta(K^0\pi^-) = \mathfrak{Im}\left[(P_{tc} - \frac{1}{3}P^C_{EW})(A + P_{uc})^*\right] I, \tag{A.18}$$

$$2\Delta(K^0\pi^0) = \mathfrak{Im}\left[(P_{tc} - P_{EW} - \frac{1}{3}P^C_{EW})(-C + P_{uc})^*\right] I, \tag{A.19}$$

where $I = 4\mathfrak{Im}(\lambda_t \lambda_u^*)$. We may define $\delta_{K\pi}$ as

$$\delta_{K\pi} = \Delta(K^+\pi^-) + \Delta(K^0\pi^-) - 2\Delta(K^+\pi^0) - 2\Delta(K^0\pi^0), \tag{A.20}$$

which yields

$$\delta_{K\pi} = -\mathfrak{Im}\left[(P_{EW} + P_{EW}^C)\{(T+C)^* + A^*\} + (P_{EW}C^* - P_{EW}^C T^*)\right] I. \tag{A.21}$$

We find that all terms involving P_{tc} cancel. The remaining terms involve electroweak penguin amplitudes and are about an order of magnitude smaller relative to P_{tc}. Assuming that the annihilation amplitude is suppressed relative to a colour-allowed tree amplitude, and that $\arg(C/T)$ is suppressed, all terms are doubly suppressed relative to $\Delta(K^+\pi^-)$, each by about an order of magnitude. The ratio $\delta_{K\pi}/\Delta(K^+\pi^-)$ can be predicted to be of the magnitude $\sim 1\%$. Thus we may write $\delta_{K\pi} \approx 0$, or

$$\Delta(K^+\pi^-) + \Delta(K^0\pi^-) \approx 2\Delta(K^+\pi^0) + 2\Delta(K^0\pi^0). \tag{A.22}$$

In terms of *CP* asymmetries defined in Eq. 1.8, with the *CP*-averaged branching ratios and the B meson lifetime ratio τ_{B^+}/τ_{B^0}, Eq. A.22 becomes

$$\begin{aligned} \mathcal{A}_{K^+\pi^-} + \mathcal{A}_{K^0\pi^+} \cdot \frac{\mathcal{B}_{K^0\pi^+}}{\mathcal{B}_{K^+\pi^-}} \frac{\tau_{B^0}}{\tau_{B^+}} \\ \approx 2\mathcal{A}_{K^+\pi^0} \cdot \frac{\mathcal{B}_{K^+\pi^0}}{\mathcal{B}_{K^+\pi^-}} \frac{\tau_{B^0}}{\tau_{B^+}} + 2\mathcal{A}_{K^0\pi^0} \cdot \frac{\mathcal{B}_{K^0\pi^0}}{\mathcal{B}_{K^+\pi^-}}, \end{aligned} \tag{A.23}$$

which is the relation that we set out to prove.

In summary, we have demonstrated that $B^0 \to K^+\pi^-$, $B^+ \to K^+\pi^0$, $B^+ \to K^0\pi^+$, and $B^0 \to K^0\pi^0$ decays obey the isospin sum rule with precision $\sim 1\%$. Violation of the sum rule signifies NP in $b \to sq\overline{q}$ transitions.

A.2 Signal MC EvtGen Decay Files

- $B^0 \to K^0_S\pi^0$ decay file

```
Decay     Upsilon(4S)
   1.0    B0     anti-B0   B0sig   anti-B0sig  VSS_BMIX  dm  dgog;
Enddecay

Decay     B0sig
   1.0    K_S0     pi0    SSS_CP  beta   dm   -1  1  0   1    0;
Enddecay
CDecay    anti-B0sig

End
```

- $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$ decay file

```
Decay     Upsilon(4S)
   0.5    B+sig          B-     VSS;
   0.5    B+           B-sig    VSS;
Enddecay

Decay     B+sig
   1.0    anti-D0sig    pi+   PHOTOS  PHSP;
Enddecay
CDecay    B-sig

Decay     anti-D0sig
   1.0    K_S0          pi0   PHOTOS  PHSP;
Enddecay
CDecay    D0sig

End
```

`Decay` and `Enddecay` indicate the start and end of an decay entry for a particle, where different decay channels are separated by ";". The decay channels of the conjugate particle can be specified with the syntax `CDecay`. The $B^0 \to K^0_S\pi^0$ decay file generates either `B0 anti-B0`, `B0sig anti-B0`, `B0 anti-B0sig`, or `B0sig anti-B0sig` from the $\Upsilon(4S)$. The

particle aliases are summarised in Tab. A.1, and those with a suffix `-sig` are the signal-side particles whose decays are specified in the decay file. The decay models and variables are listed in Tab. A.2. Detailed documentation of EvtGen can be found in Ref. [63].

Table A.1: Particle aliases in EvtGen.

Particle	Alias	Particle	Alias
B^0	`B0`	$\overline{B}^0$	`anti-B0`
B^+	`B+`	B^-	`B-`
K_S^0	`K_S0`	π^0	`pi0`
D^0	`D0`	$\overline{D}^0$	`anti-D0`
π^+	`pi+`	π^-	`pi-`

Table A.2: Decay models of EvtGen. Variables of each decay model are listed according to the input order.

`VSS_BMIX`	`dm`	Mass difference of the two B eigenstates.
($B^0\overline{B}^0$ mixing)	`dgog`	$\Delta\Gamma/\Gamma$
	`beta`	Relevant CKM angle.
	`dm`	Mass difference of the two B eigenstates.
`SSS_CP`	`CP`	*CP* eigenvalue of the final state.
(*CP* violation)	`|A|`	
	`arg(A)`	Complex amplitudes for B^0 and $\overline{B}^0$ to
	`|Abar|`	produce the final state.
	`arg(Abar)`	
`VSS`	Decays a vector particle into two scalars.	
`PHOTOS`	Incorporates final-state radiation.	
`PHSP`	Generic phase space to n-bodies .	

A.3 Data-MC Comparison of Continuum Suppression Training Variables Using $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$

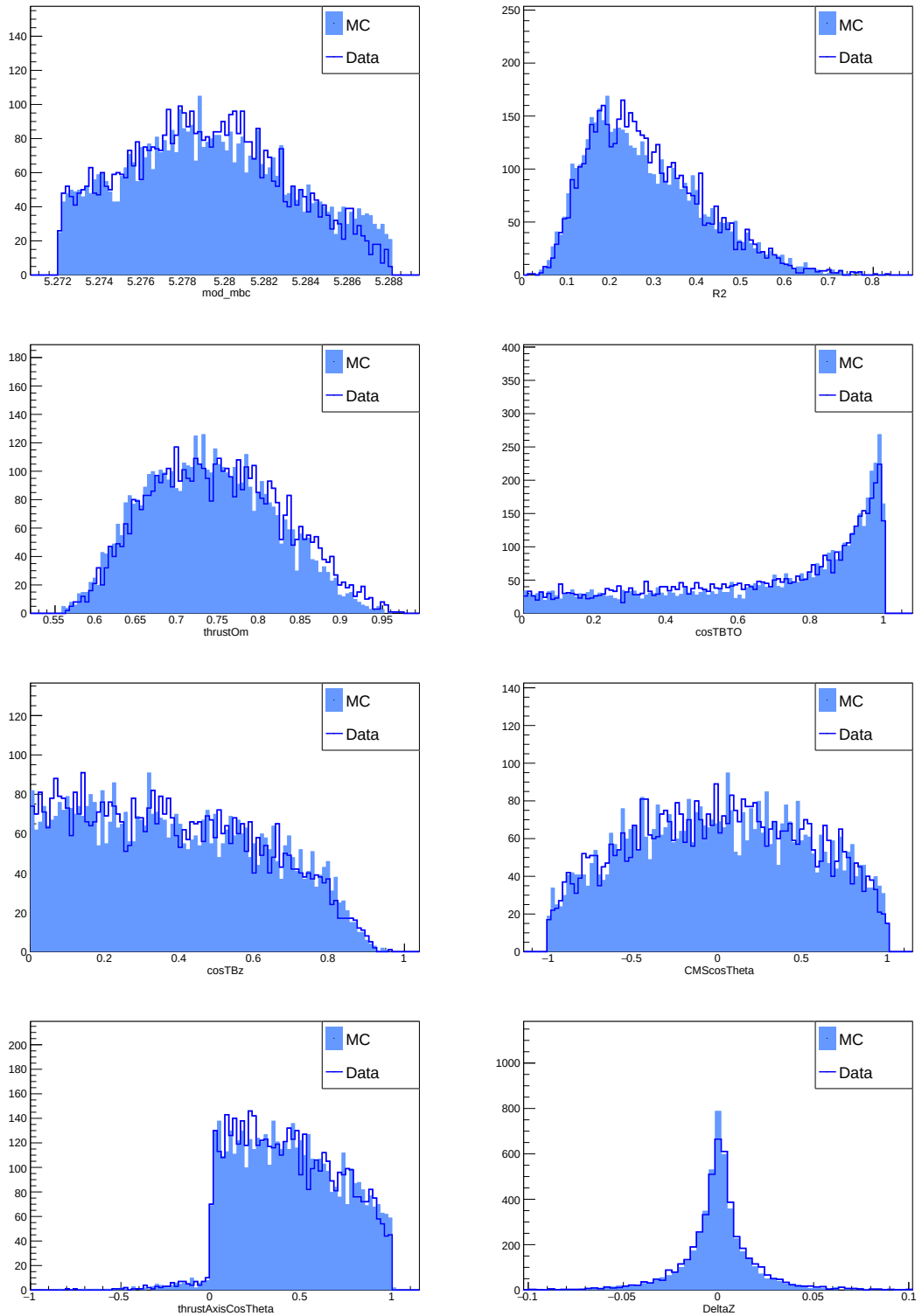


Figure A.2: Comparison between data and simulation for the BDT input variables in the $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$ control channel.

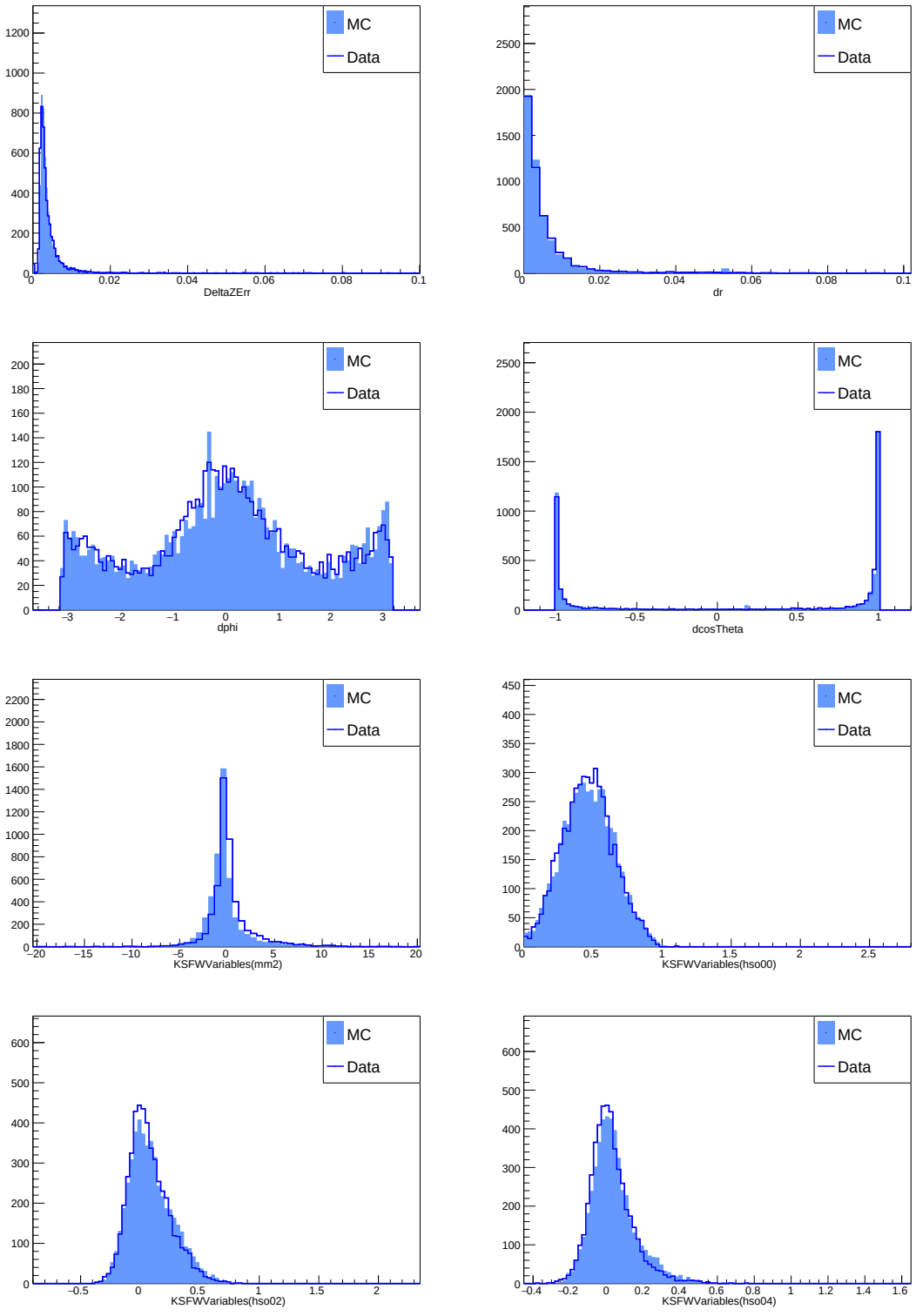


Figure A.3: Comparison between data and simulation for the BDT input variables in the $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$ control channel.

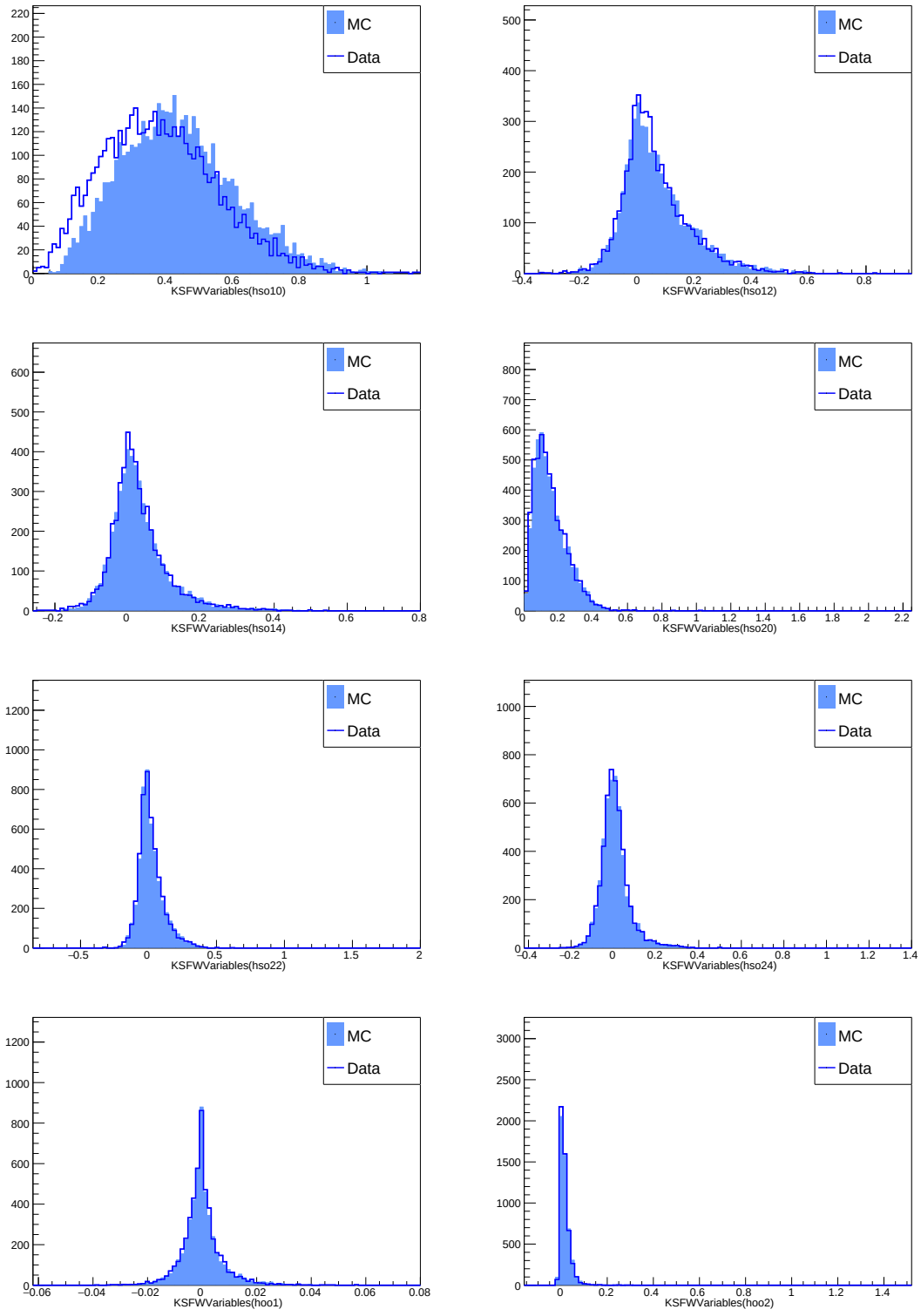


Figure A.4: Comparison between data and simulation for the BDT input variables in the $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$ control channel.

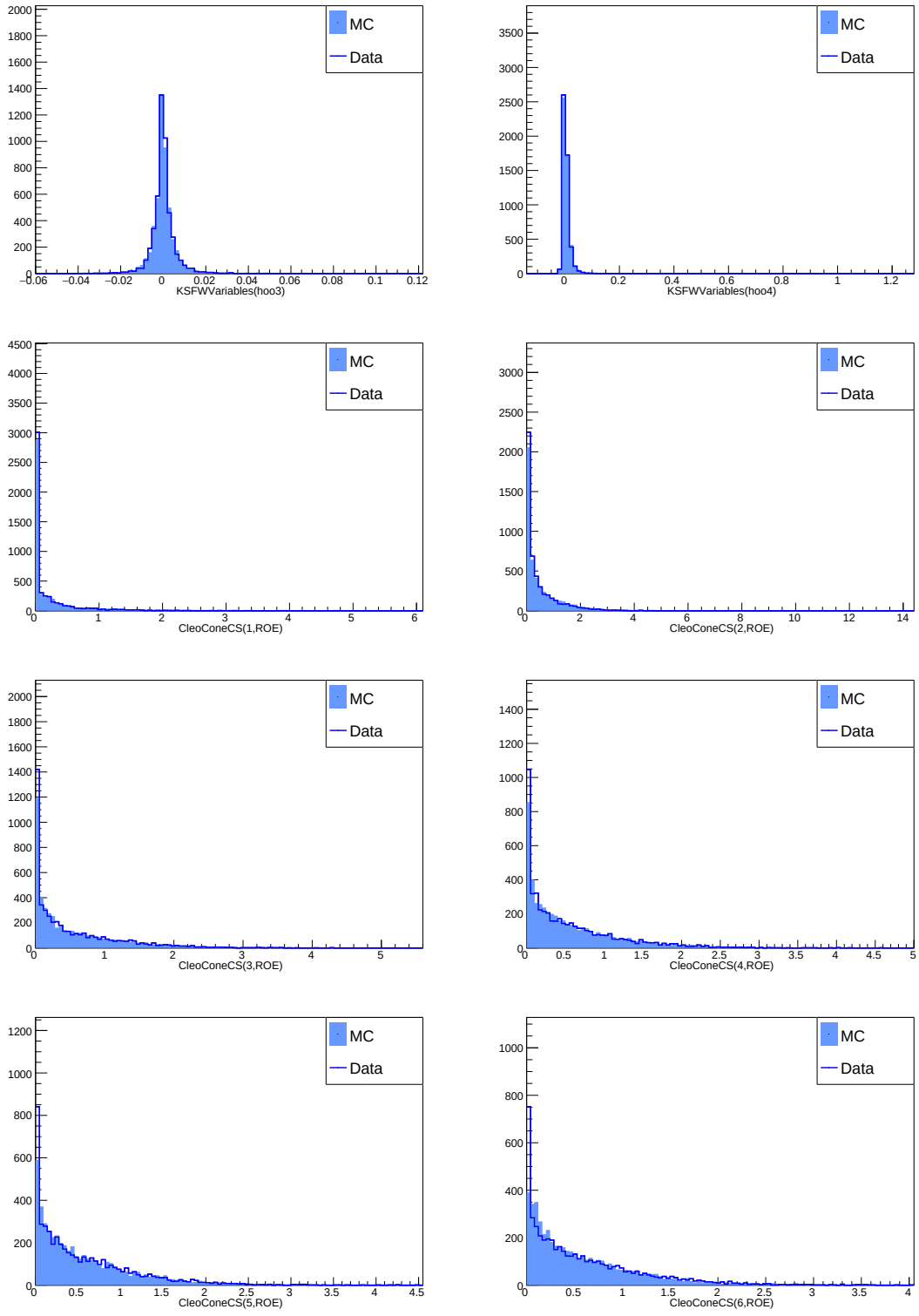


Figure A.5: Comparison between data and simulation for the BDT input variables in the $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$ control channel.

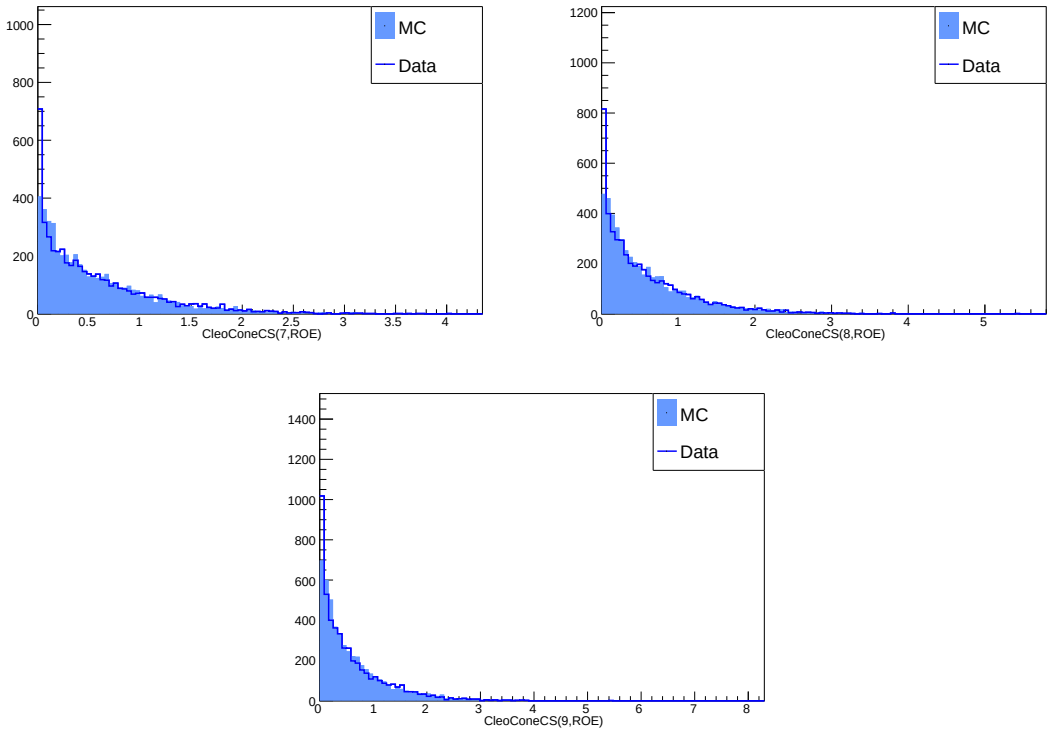


Figure A.6: Comparison between data and simulation for the BDT input variables in the $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$ control channel.

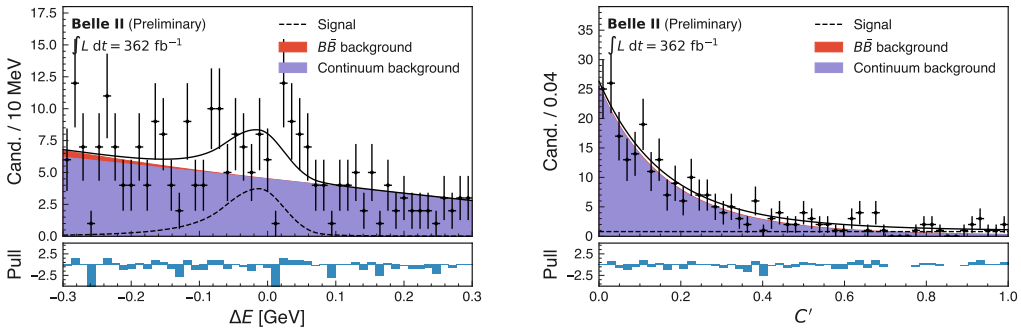
A.4 $B\overline{B}$ Background Components

A total of 200 $B\overline{B}$ background events remains for the fitting stage. There are 21 $B\overline{B}$ events where the reconstructed B mesons are $\Upsilon(4S)$ according MC truth information, 38 $B^0\overline{B}^0$ events and 136 B^+B^- events in $1\,\text{ab}^{-1}$ `MC15ri` generic background. Failure to match the B event type occurred in 5 $B\overline{B}$ events. Failure to match the B decay occurred in 10 $B^0\overline{B}^0$ events and 3 B^+B^- events. The decay information is determined using `basf2` variables `BplusMode`, `BminusMode`, `B0Mode`, `BbarOMode`, and `mcPDG`.

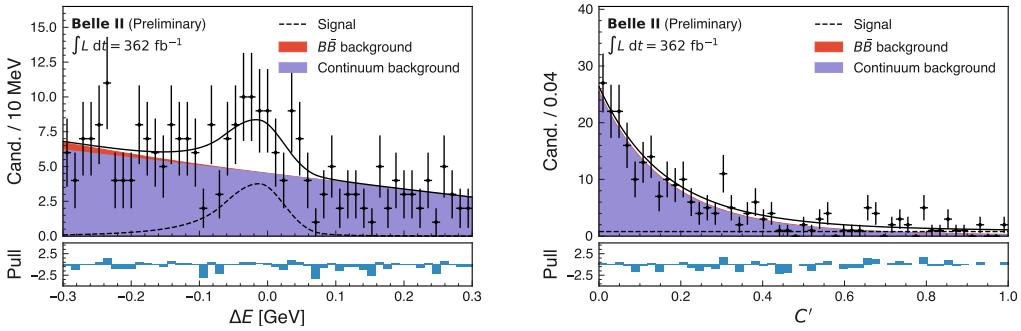
Table A.3: B decays contributing to the $B\overline{B}$ background in $1\,\text{ab}^{-1}$ background simulation.

	Decay mode	Yield in $1\,\text{ab}^{-1}$		Decay mode	Yield in $1\,\text{ab}^{-1}$
B^+	$\rho^+\pi^0$	45	B^0	$K^{*0}\pi^0$	12
	ρ^+K^0	31		$K_0^{*0}\pi^0$	2
	$K^{*+}\pi^0$	28		$D^{*-}a_1^+$	2
	$K_0^{*+}\pi^0$	16		$K_S^0K_S^0$	1
	$\bar{D}^{*0}\rho^+$	2		$K^0\overline{K}^0$	1
	ρ^+D^{*0}	1		$K''^{*+}\pi^-$	1
	$\pi^+K^0\pi^0$	1		$K^0\pi^0\pi^0$	1
	$J/\Psi K_2^{*+}$	1		f_2K^0	1
	$Xsu\gamma$	1		ρ^+D^{*-}	1
	$D_s^{*+}D^-\pi^+$	1		$\overline{D}^0\pi^0$	1
	$\bar{D}^0\pi^+$	1		$D_{s1}^{'+}D^{*-}$	1
	$\bar{D}^{*0}D^+K^{*0}$	1		$D_2^{*-}\mu^+\nu_\mu$	1
	$D^{*-}\pi^0\pi^+\pi^+$	1		$D_1^-\tau^+\nu_\tau$	1
	$\bar{D}^0D_{s1}^{'+}$	1		$D_2^{*-}\tau^+\nu_\tau$	1
	$\bar{D}^{*0}e^+\nu_e$	1		$\phi K_0^{*+}\pi^-$	1
	$\bar{D}_1^0\mu^+\nu_\mu$	1	$\Upsilon(4S)$	–	21

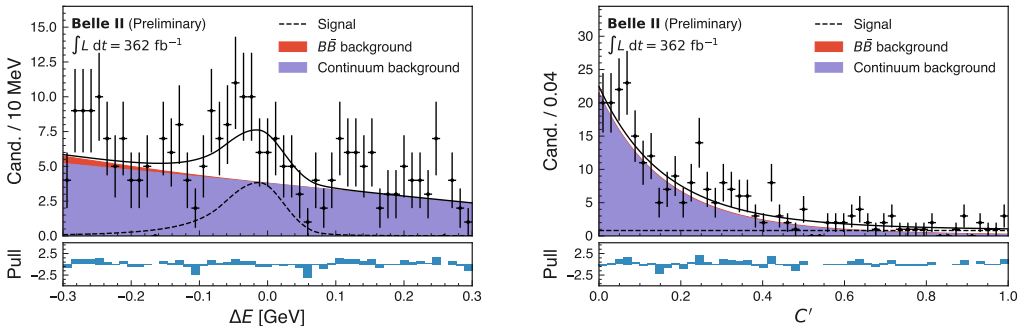
A.5 Flavour-Specific Projections of C' to r-Bins



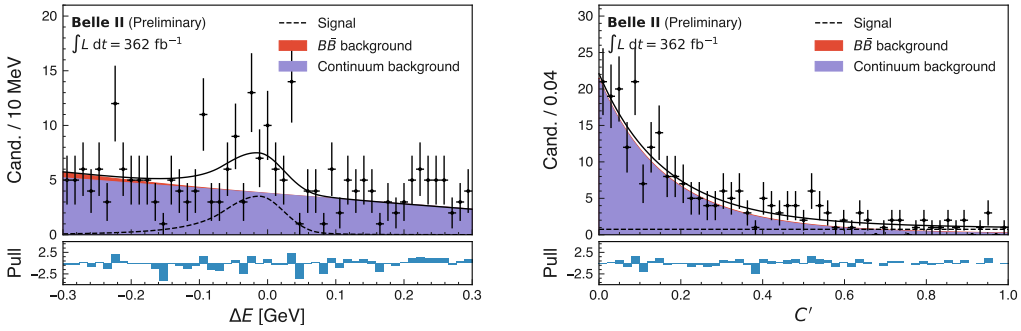
(a) 0th r-bin, $q = -1$



(b) 0th r-bin, $q = 1$

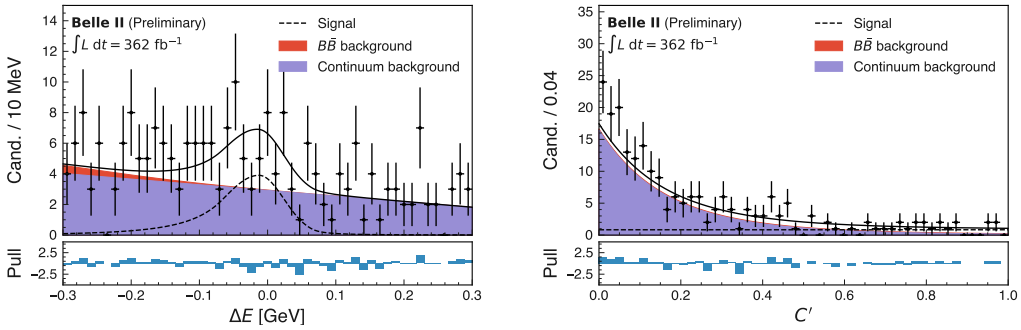


(c) 1st r-bin, $q = -1$

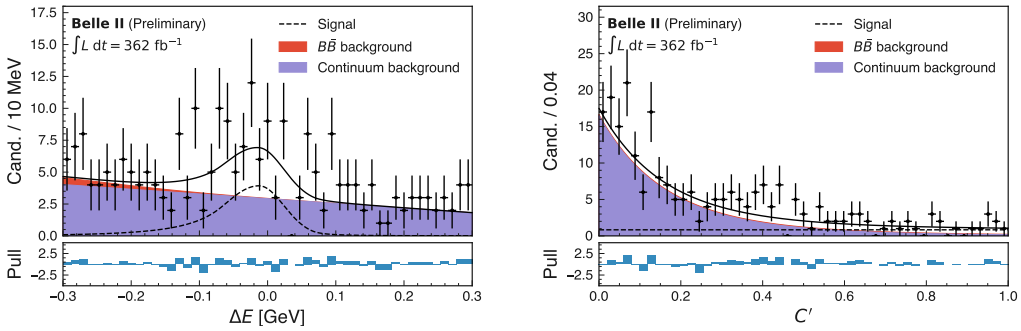


(d) 1st r-bin, $q = 1$

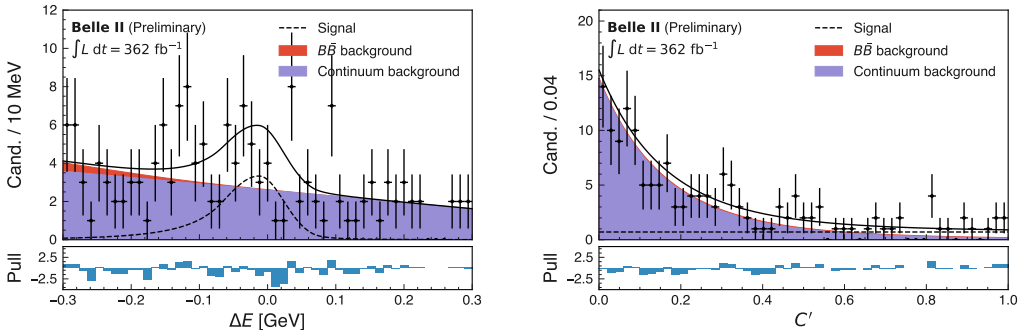
Figure A.7: Flavour-specific C' distributions of $362\ \mathrm{fb}^{-1}$ data projected to each r-bin.



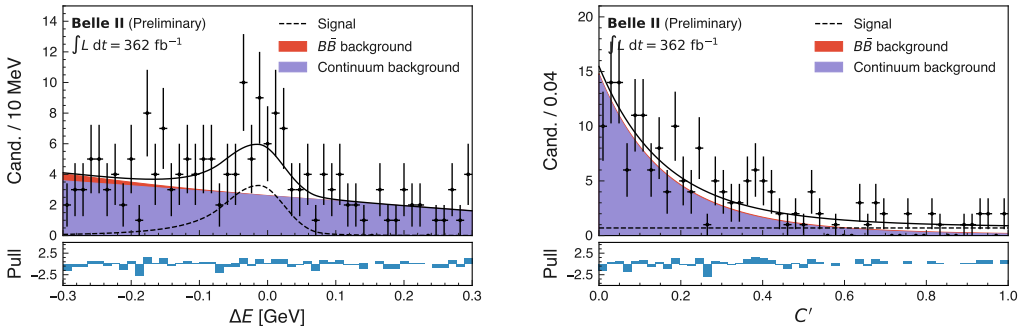
(a) 2nd r-bin, $q = -1$



(b) 2nd r-bin, $q = 1$

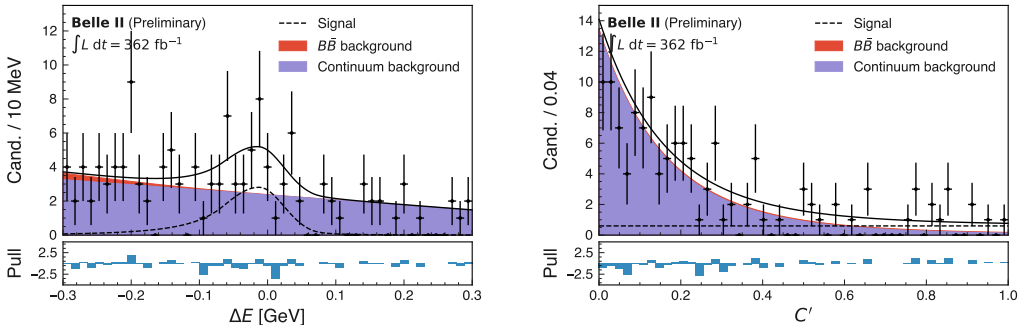


(c) 3rd r-bin, $q = -1$

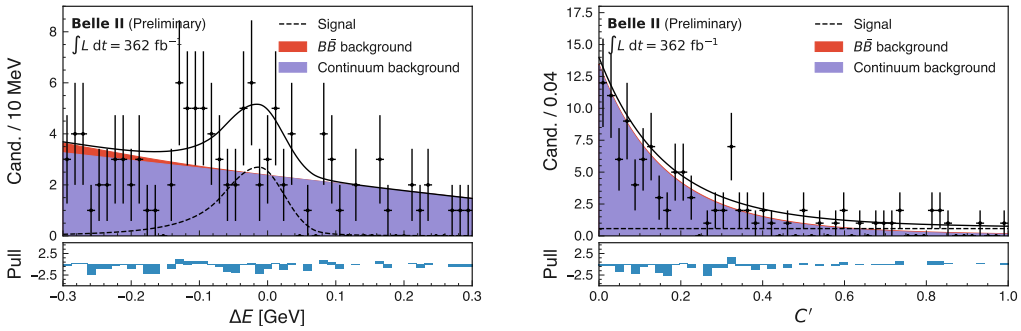


(d) 3rd r-bin, $q = 1$

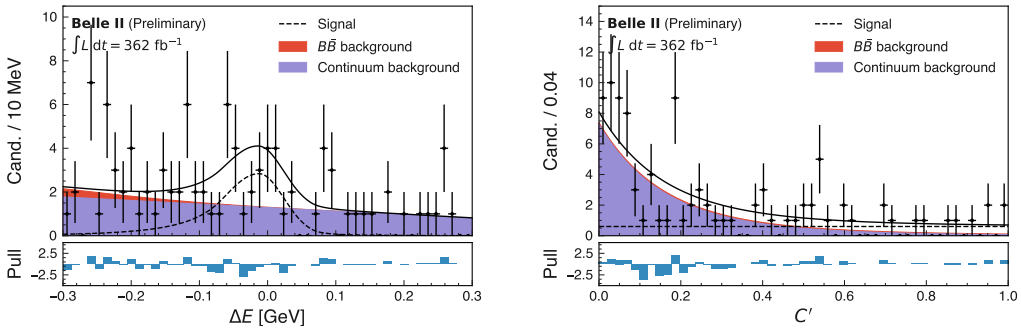
Figure A.8: Flavour-specific C' distributions of $362\,\text{fb}^{-1}$ data projected to each r-bin.



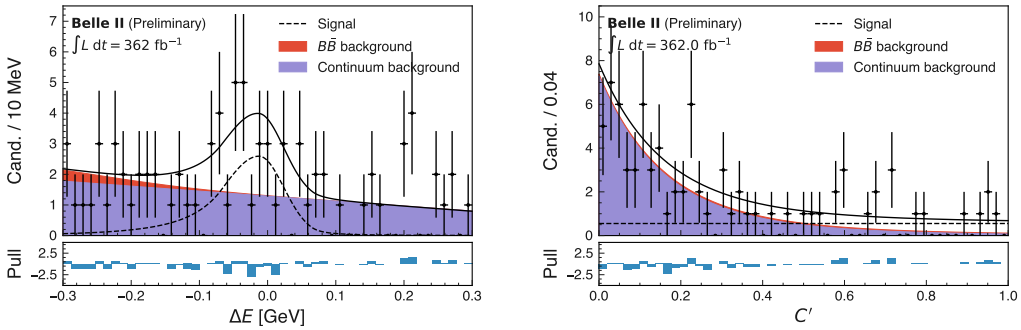
(a) 4th r-bin, $q = -1$



(b) 4th r-bin, $q = 1$

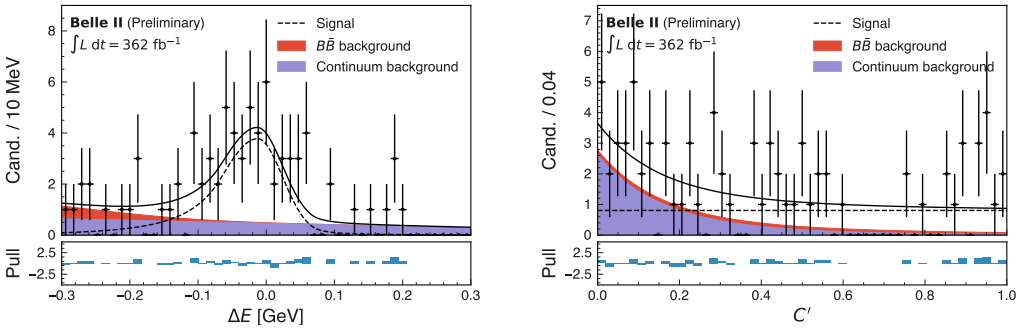


(c) 5th r-bin, $q = -1$

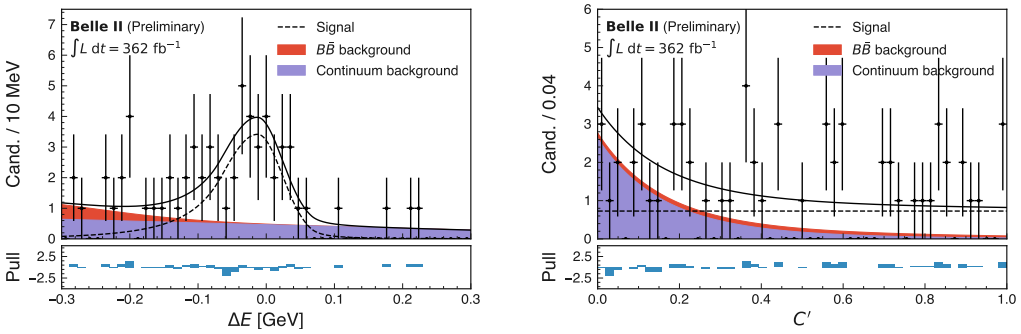


(d) 5th r-bin, $q = 1$

Figure A.9: Flavour-specific C' distributions of $362\,\text{fb}^{-1}$ data projected to each r-bin.



(a) 6th r-bin, $q = -1$



(b) 6th r-bin, $q = 1$

Figure A.10: Flavour-specific C' distributions of $362\,\text{fb}^{-1}$ data projected to each r-bin.

A.6 Correlation Matrices of Fit Parameters

Table A.4: Correlations between fitting parameters of the fits to $1\,\text{ab}^{-1}$ simulation, $62.8\,\text{fb}^{-1}$ partial unblinding, and $362\,\text{fb}^{-1}$ full unblinding samples.

$1\,\text{ab}^{-1}$ simulation

	$\mathcal{A}_{K^0\pi^0}$	$\mathcal{B}_{K^0\pi^0}$	$N_{B\bar{B}}$	$N_{q\bar{q}}$	$q\bar{q}\ C'\ a$	$q\bar{q}\ C'\ b$	$q\bar{q}\ \Delta E\ a_1$
$\mathcal{A}_{K^0\pi^0}$	1.00	-0.01	-0.01	0.01	0.00	0.01	-0.01
$\mathcal{B}_{K^0\pi^0}$		1.00	0.05	-0.32	-0.05	-0.10	0.11
$N_{B\bar{B}}$			1.00	-0.46	-0.08	-0.14	0.42
$N_{q\bar{q}}$				1.00	0.07	0.12	-0.26
$q\bar{q}\ C'\ a$					1.00	-0.89	-0.04
$q\bar{q}\ C'\ b$						1.00	-0.08
$q\bar{q}\ \Delta E\ a_1$							1.00

$62.8\,\text{fb}^{-1}$ data

	$\mathcal{A}_{K^0\pi^0}$	$\mathcal{B}_{K^0\pi^0}$	$N_{B\bar{B}}$	$N_{q\bar{q}}$	$q\bar{q}\ C'\ a$	$q\bar{q}\ C'\ b$	$q\bar{q}\ \Delta E\ a_1$
$\mathcal{A}_{K^0\pi^0}$	1.00	0.14	0.01	-0.07	0.03	-0.09	0.02
$\mathcal{B}_{K^0\pi^0}$		1.00	0.19	-0.38	-0.02	-0.17	0.17
$N_{B\bar{B}}$			1.00	-0.63	-0.20	-0.09	0.54
$N_{q\bar{q}}$				1.00	0.12	0.13	-0.39
$q\bar{q}\ C'\ a$					1.00	-0.88	-0.11
$q\bar{q}\ C'\ b$						1.00	-0.06
$q\bar{q}\ \Delta E\ a_1$							1.00

$362\,\text{fb}^{-1}$ data

	$\mathcal{A}_{K^0\pi^0}$	$\mathcal{B}_{K^0\pi^0}$	$N_{B\bar{B}}$	$N_{q\bar{q}}$	$q\bar{q}\ C'\ a$	$q\bar{q}\ C'\ b$	$q\bar{q}\ \Delta E\ a_1$
$\mathcal{A}_{K^0\pi^0}$	1.00	0.01	-0.02	0.00	-0.01	0.01	-0.01
$\mathcal{B}_{K^0\pi^0}$		1.00	0.07	-0.33	-0.02	-0.13	0.12
$N_{B\bar{B}}$			1.00	-0.49	-0.08	-0.13	0.43
$N_{q\bar{q}}$				1.00	0.05	0.14	-0.28
$q\bar{q}\ C'\ a$					1.00	-0.89	-0.04
$q\bar{q}\ C'\ b$						1.00	-0.08
$q\bar{q}\ \Delta E\ a_1$							1.00

Denotation

Abbreviations

ARICH	Aerogel ring-imaging Cherenkov
`basf2`	Belle II analysis software
CDC	Central drift chamber
CKM	Cabibbo-Kobayashi-Maskawa
CS	Continuum Suppression
DEPFET	Depleted field effect transistor
ECL	Electromagnetic Calorimeter
FastBDT	Fast boosted decision tree
HLT	High-level trigger
IP	Interaction point
KEK	High Energy Accelerator Research Organization (高エネルギー加速器研究機構)
KLM	K_L^0 and muon detector
MC	Monte-Carlo

MVA	Multivariate analysis
NP	New Physics
PDG	Particle Data Group
PDF	Probability density function
PXD	Pixel detector
ROC	Receiver operating characteristic
ROE	Rest-of-event
RPC	Resistive plate chambers
SCF	Self-cross-feed
SM	Standard model
SVD	Silicon vertex detector
TD	Time-dependent
TDCPV	Time-dependent *CP* Violation
TI	Time-integrated
TOP	Time-of-propagation

Variables

$\mathcal{A}$	Direct *CP* asymmetry
$\mathcal{B}$	Branching fraction
C	Charge conjugation

C'	Transformed continuum suppression output
`CLEOConeCS(ROE)`	CLEO cones for rest-of-event particles
`clusterNHits`	Sum of weights in all crystals of an electromagnetic calorimeter cluster
`clusterTiming`	Time of the electromagnetic calorimeter cluster
`CMScosTheta`	Cosine of the polar angle of the signal B in the centre-of-mass frame
`cosHelicityAngleMomentum`	Cosine of the angle between the line defined by the momentum difference of the two daughters in the frame of the given particle (mother) and the momentum of the given particle in the lab frame
`cosTBTO`	Cosine of the angle between the signal B thrust axis and the rest-of-event thrust axis
`cosTBz`	Cosine of the angle between the signal B thrust axis and the z-axis
`daughterAngle`	Angle between any pair of particles belonging to the same decay tree
`daughterDiffofPhi`	Difference in the azimuthal angles between the two given daughters
`dcosTheta`	Cosine of the polar angle with respect to the interaction point
ΔE	Energy difference
Δm	Mass difference between two B^0 mass eigenstates
Δt	Decay time difference between the signal- and tag-side B^0
Δw_r	Difference of wrong-tag fractions between B^0 and $\overline{B}^0$
`DeltaZ`	Spatial difference between signal and tag-side B decay vertices in the beam direction.

`DeltaZErr`	Error of `DeltaZ`
`dphi`	Azimuthal angle with respect to the interaction point
dr	Transverse distance with respect to the interaction point
dz	z distance with respect to the interaction point
E	Energy
ε_r	Partial efficiency
`flightDistance`	Flight distance of particle
`isSignal`	1.0 if the particle is correctly reconstructed, 0.0 if not, and NaN if no related MC particle could be found.
`KSFWVriables`	Kakuno-Super-Fox-Wolfram moments
`LambdaVeto`	MVA to discriminate K_S^0 from Λ
M_{bc}	Beam-constrained mass
$m(\gamma\gamma)$	Invariant mass of π^0
$m(\pi\pi)$	Invariant mass of K_S^0
μ_r	Difference of partial efficiencies between B^0 and $\overline{B}^0$
`nCDCHits`	Number of CDC hits associated to the track
`nSVDHits`	number of SVD hits associated to the track
P	Parity
p	Momentum
`photonMVA`	MVA trained to discriminate non-signal photons

q	Flavour of tag-side B^0
r	Dilution factor
`R2`	Normalised second Fox-Wolfram moment
`RightCategory`	Output probability of the track being the target particle of the flavour tagger category
$\mathcal{S}$	Mixing-induced *CP* asymmetry
`thetaInCDCAcceptance`	True if particle is within CDC angular acceptance, false otherwise
`thrustAxisCosTheta`	Cosine of the polar angle component of the thrust axis
`thrustOm`	Magnitude of the rest-of-event thrust axis
`V0Selector`	MVA to discriminate K_S^0 from misreconstructed K_S^0 coming from the interaction point
w_r	Wrong-tag fraction
χ_d	Time-integrated B^0 mixing parameter