Technische Universität München
TUM School of Natural Sciences

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

# Measurement of Branching Fractions and Direct $CP$ Asymmetries for $B^+ \to K^0\pi^+$, $B^+ \to K^+\pi^0$, and $B^+ \to \pi^+\pi^0$ Decays at Belle II

Dissertation by

**Justin Skorupa**

München, October 2025

# Abstract


We present measurements of branching fractions and direct CP asymmetries for $B^+ \to K^0\pi^+$, $B^+ \to K^+\pi^0$, and $B^+ \to \pi^+\pi^0$ decays. The analysis uses a data sample containing 387 million bottom-antibottom meson pairs produced in $e^+e^-$ collisions at the $\Upsilon(4S)$ resonance by the SuperKEKB collider and recorded with the Belle II detector. The measured branching fractions are:

$B^+ \to K^0\pi^+$ : $(24.37 \pm 0.71 \pm 0.86)\times 10^{-6}$,
$B^+ \to K^+\pi^0$ : $(13.93 \pm 0.38 \pm 0.71)\times 10^{-6}$, and
$B^+ \to \pi^+\pi^0$ : $(5.10 \pm 0.29 \pm 0.27)\times 10^{-6}$.

The corresponding direct CP asymmetries are:

$B^+ \to K^0\pi^+$ : $0.046 \pm 0.029 \pm 0.007$,
$B^+ \to K^+\pi^0$ : $0.013 \pm 0.027 \pm 0.005$, and
$B^+ \to \pi^+\pi^0$ : $-0.081 \pm 0.054 \pm 0.008$,

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

To test the standard model, we combine our results with Belle II measurements for $B^0 \to K^+\pi^-$ and $B^0 \to K^0\pi^0$ conducted on the same data set to evaluate an isospin-based sum rule. We find

$$-0.03 \pm 0.13 \pm 0.04$$

for the sum rule, consistent with the standard model expectation of zero.

# Zusammenfassung


Wir präsentieren Messungen des Verzweigungsverhältnisses und des CP-verletzenden Parameters für $B^+ \to K^0\pi^+$, $B^+ \to K^+\pi^0$ und $B^+ \to \pi^+\pi^0$ Zerfälle. Die Analyse nutzt einen Datensatz, der 387 Millionen Bottom-Antibottom-Mesonenpaare enthält, die in $e^+e^-$-Kollisionen an der $\Upsilon(4S)$-Resonanz vom SuperKEKB Beschleuniger erzeugt und vom Belle II Detektor aufgezeichnet wurden. Die gemessenen Verzweigungsverhältnisse sind:

$$\begin{aligned}
B^+ \to K^0\pi^+ &: (24.37 \pm 0.71 \pm 0.86)\times 10^{-6}, \\
B^+ \to K^+\pi^0 &: (13.93 \pm 0.38 \pm 0.71)\times 10^{-6}, \text{ und} \\
B^+ \to \pi^+\pi^0 &: \ (5.10 \pm 0.29 \pm 0.27)\times 10^{-6}.
\end{aligned}$$

Die entsprechenden CP-verletzenden Parameter sind:

$$\begin{aligned}
B^+ \to K^0\pi^+ &: \ 0.046 \pm 0.029 \pm 0.007, \\
B^+ \to K^+\pi^0 &: \ 0.013 \pm 0.027 \pm 0.005, \text{ und} \\
B^+ \to \pi^+\pi^0 &: \text{-}0.081 \pm 0.054 \pm 0.008,
\end{aligned}$$

wobei die erste Unsicherheit die statistische und die zweite die systematische Unsicherheit angibt.

Um das Standardmodel zu testen, berechnen wir eine auf Isospin-Symmetrie basierende Summenregel mit unseren Ergebnissen und den Resultaten weitere Belle II Messungen für $B^0 \to K^+\pi^-$ und $B^0 \to K^0\pi^0$ Zerfällen, die mit dem gleichen Datensatz durchgeführt wurden. Wir erhalten

$$-0.03 \pm 0.13 \pm 0.04$$

für die Summenregel. Das Ergebnis stimmt dem dem Erwartungswert des Standardmodell von Null überein.

# Contents

# Chapter 1

# Introduction

The standard model of particle physics (SM) was completed in 2012 with the discovery of the Higgs boson by the ATLAS and CMS experiments. It is a remarkably successful theory. In its minimal form, the SM describes the fundamental particles and their interactions using only 19 free parameter. Its predictions have been tested in thousands of measurements with precision reaching better than one part in a trillion.

Despite these successes, the SM is known to be incomplete. For example, it does not incorporate the gravitational interaction, nor does it provide a viable candidate for dark matter. Furthermore, it offers no explanation for the matter-antimatter asymmetry observed in the universe. These shortcomings motivate searches for physics beyond the SM. A especially fruitful approach are indirect searches in the flavor sector.

The $B$-factory experiments *BABAR* at SLAC in the United States and Belle at KEK in Japan were constructed to test the Cabibbo-Kobayashi-Maskawa (CKM) structure of the universe. Their measurement of time-dependent CP violation in $B^0 \to J/\psi\, K^0_S$ decays established the validity of the CKM description. Kobayashi and Maskawa earned the Nobel prize for their prediction in 2008. Yet, many open questions remain. The amount of CP violation predicted by the SM is insufficient to explain the mattter-antimatter asymmetry observed in the universe. Moreover, several anomalies in measurements of $B$ decays have been observed over the past two decades, which may hint at contributions from physics beyond the SM.

In this context, the Belle II experiment at KEK in Japan is the successor of the original $B$-factories. Designed to integrate a data sample 50 times larger than that of Belle, it will perform precision measurements of rare $B$ decays. Belle II will either reveal deviations from the SM or place some of the most stringent constraints to date on possible physics beyond the SM scenarios.

In this thesis, the branching fraction and direct CP asymmetries for $B^+ \to \pi^+\pi^0$, $B^+ \to K^+\pi^0$, and $B^+ \to K^0\pi^+$ decays will be measured. Unless otherwise stated, charge conjugation is implied throughout this thesis. In these decays, tree-level amplitudes are suppressed in the SM. Hence, they are sensitive to the CKM structure in interference between tree- and loop-level amplitudes as well as contributions from physics beyond the SM entering as virtual heavy particles in the loop. Together with measurement for other $B \to \pi\pi$ decays, the measurement for $B^+ \to \pi^+\pi^0$ allows the extraction of the CKM angle $\phi_2$ ($\alpha$). The measurement for $B^+ \to K^+\pi^0$ and $B^+ \to K^0\pi^+$ allow to test the SM with an isospin based sum rule when combined with measurements for other $B \to K\pi$ decays.

This thesis is structured as follows: Chapter 2 gives a brief description of the SM with an emphasis on flavor physics. Chapter 3 discusses the experimental setup of the Belle II detector and the SuperKEKB accelerator. Chapter 4 outlines the reconstruction and analysis techniques to measure the branching fractions and direct CP asymmetries. Chapter 5 summarizes the findings and provides an outlook for future studies.

# Chapter 2

# Charmless Hadronic $B$ Decays in the Standard Model

## 2.1 The Standard Model of Particle Physics

The standard model of particle physics (SM) is an effective quantum field theory describing the known particles and interactions (bar the gravitational interaction) [1, 2, 3]. Historically, the SM emerged from efforts to incorporate special relativity into quantum theory, which led to the framework of quantum field theory. It has been experimentally tested to be valid up to an energy scale of $\mathcal{O}(1\,\text{TeV})$. It is not expected to remain valid at higher energies, where quantum gravity effects become important, which are currently not incorporated in the SM. Hence, the SM is an *effective* quantum field theory.

In quantum field theory, fields assign values to every point in spacetime. Particles arise as excitations of these fields, i.e., pertubations around the ground state. Due to the quantized nature of the fields, only excitations fulfilling energy and momentum conditions are allowed. Hence, it is not possible to create an excitation corresponding to, for example, half a particle.

The SM describes the interactions and dynamics of these quantum fields. Mathematically, this is achieved using the Lagrangian formalism, constrained by local gauge symmetry under the group $SU(3)_C \times SU(2)_L \times U(1)_Y$. Here, $SU(3)_C$ is the unitary group describing the strong interaction and $C$ is the color charge. The product of groups $SU(2)_L \times U(1)_Y$ describes the unification of weak and electromagnetic interaction, $L$ stands for left, and $Y$ for the weak hypercharge.

The mediators of the strong interaction are eight massless gauge bosons, called gluons, which correspond to the generators of $SU(3)_C$. The strong interaction occurs between particles carrying color charge. Gluons themselves carry color charge, hence they are able to directly interact with each other. The gauge bosons associated to the generators of $SU(2)_L$ are $W^1$, $W^2$, and $W^3$, while the gauge bosons corresponding to the generator of $U(1)_Y$ is $B$. They mix to form the physical mediators of the weak and electromagnetic interactions. The mediators of the weak interaction are two massive charged bosons, $W^{\pm}$, and a neutral boson, $Z^0$. The mediator of the electromagnetic interaction is the massless photon, $\gamma$. The electromagnetic interaction acts on charged particles. The gauge bosons and the physical mediators are related via the following equations:

$$W^{\pm} = \frac{1}{\sqrt{2}}(W^1 \mp W^2) \text{ and } \begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos(\theta_W) & \sin(\theta_W) \\ -\sin(\theta_W) & \cos(\theta_W) \end{pmatrix} \begin{pmatrix} B \\ W^3 \end{pmatrix},$$

where $\theta_W$ is the so called Weinberg angle. The masses of the mediators are related via

$$m_Z = \frac{m_W}{\cos\theta_w},$$

where $m_Z$ is the mass of the $Z^0$ and $m_W$ is the mass of the $W^{\pm}$. Particles acquire mass through their interaction with the Higgs field, whose excitation is the spin-0 Higgs boson.

Matter is composed of spin-half fermions. For each fermion, there is an associated anti-particle with the same mass but opposite quantum numbers. Fermions are subdivided into two groups:

- **Quarks**: There are six types of quarks, grouped into three generations, each forming a weak isospin doublet:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix}.$$

The up-type quarks ($u$, $c$, $t$) carry an electric charge of $+2/3$ and the down-type quarks ($d$, $s$, $b$) have an electric charge of $-1/3$. Each quark carries both a flavor and a baryon quantum number. The flavor quantum number is conserved in electromagnetic and strong interactions but can be violated by the weak interaction, while the baryon quantum number is conserved in the SM. In addition, quarks also carry color charge, which causes them to participate in the strong interaction. Due to color confinement [4], free quarks are not observable. They can only be observed in color-neutral bound states. The primary type of such bound states are mesons, consisting of a quark-antiquark pair, and baryons, composed of three quarks.

- **Leptons**: Similar to quarks, leptons are grouped into three generations of weak isospin doublets:

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}, \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}, \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}.$$

  The massive leptons ($e$, $\mu$, $\tau$) carry an electric charge of $-1$. The neutrinos ($\nu_e$, $\nu_\mu$, $\nu_\tau$) are electrically neutral. Observed neutrino oscillations indicate at least two of the three neutrinos have nonzero mass [5]. However, the SM does not include a mechanism to generate neutrino masses. Each lepton has a lepton-family quantum number corresponding to its generation. This quantum numbers are conserved in the SM but known to be broken by neutrino oscillations. The total lepton number is conserved.

The particle content of the SM is summarized in Fig. 2.1.

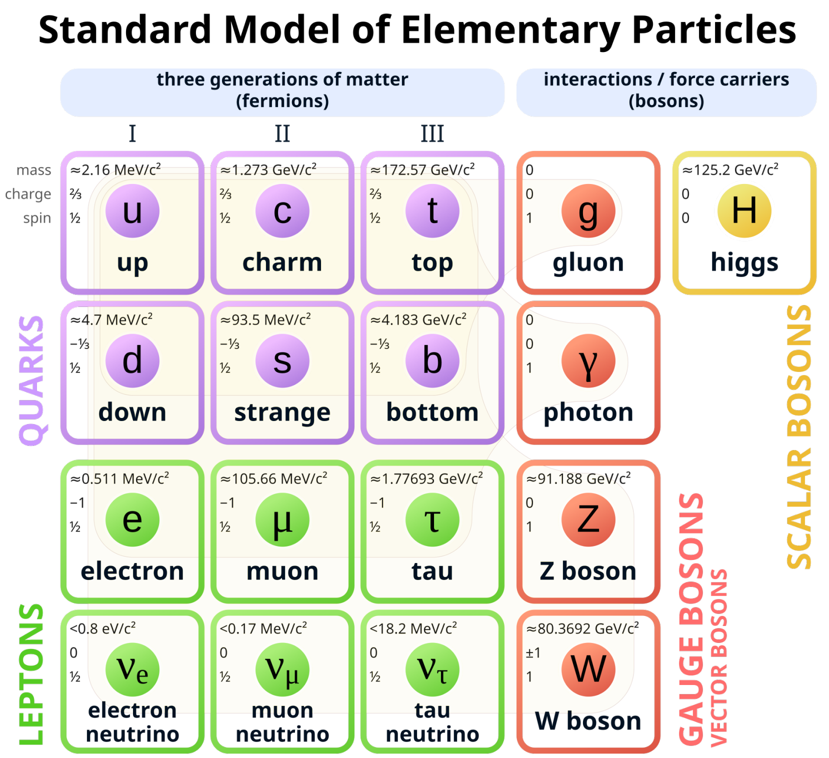


Figure 2.1: Particle content of the standard model of particle physics [6].

## 2.2 Open Questions in the Standard Model

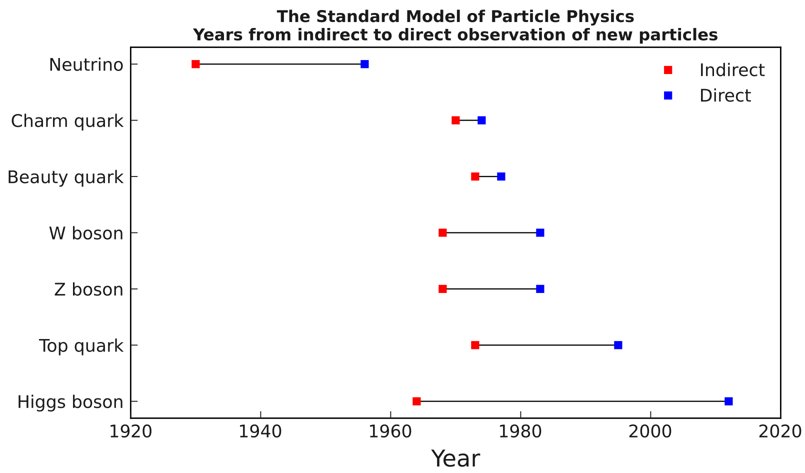


Figure 2.2: Inference of the presence of particles from the effects they have on other observables (indirect observation, red bars) and direct (blue bar) observation of particles.

The final fundamental particle of the SM, the Higgs boson, was discovered by the ATLAS and CMS collaborations in 2012 [7, 8]. The SM is remarkably successful in predicting a wide range of subatomic processes. These predictions have been tested in thousands of experimental measurements, some of which reach a relative precision better than one part per trillion [9, 10].

Despite this success, we know mainly from cosmological observation the SM is incomplete. For example, the outskirts of galaxies rotate faster than expected from their observed mass [11]. In the standard model of cosmology (the $\Lambda$CDM model), this observation is explained by the introduction of so-called dark matter, which does not interact electromagnetically [12, 13]. The standard model of particle physics does not contain any viable dark matter candidate. In addition, the SM can not explain the observed matter-antimatter asymmetry observed in the universe [14]. Currently, the SM also does not incorporate general relativity. It also contains no mechanism for neutrinos to aquire mass nor does it explain why the weak interaction is much stronger than gravity.

These unexplained phenomena motivate to further probe the SM and searches for physics beyond it. To this end, two complementary approaches are pursued: direct and indirect searches.

Direct searches aim to produce particles not included in the SM in high-energy collisions and to detect their decay products. These searches are limited by the maximum center-of-mass energy available at particle colliders.

Indirect searches, in contrast, search for deviations between precise experimental measurements and SM predictions. The conceptual idea is that heavy particles beyond the SM can contribute virtually to certain processes, modifying observable quantities. Indirect searches often have a higher mass reach than direct searches. Historically, several fundamental particles were observed indirectly before being detected directly (see Fig. 2.2).

## 2.3 Flavor Physics

Indirect searches are a central focus of flavor physics, which investigates processes involving quark and lepton flavor transitions. Here, only quark flavor transitions are discussed, and among them, only those phenomena relevant to the measurements presented in this thesis.

Historically, the concept of flavor physics originates from the observation that kaons exhibit a large difference in production and decay rate. To explain this phenomenon, Gell-Mann introduced the strangeness quantum number [15]. Particles containing strangeness are produced in pairs in the strong interaction, which conserves strangeness, but decay via the weak interaction, which allows strangeness to change. Hence explaining the difference in production and decay rate. Nowadays, strangeness is recognized as a flavor quantum number. To explain the violation of strangeness in the weak interactions, Cabbibo introduced the concept of destinct mass and weak eigenstates. Cabibbo postulated the down ($d$) and strange ($s$) quark mix through a rotation by a mixing angle, now known as the Cabibbo angle [16]. This theory successfully described the strangeness-changing weak decays. However, the theory also predicted a rate of $K^0 \to \mu^+\mu^-$ decays incompatible with experimental observations. In 1970, Glashow, Iliopoulos, and Maiani postulated a fourth yet undiscovered quark, the charm quark ($c$). Contributions of the charm quark suppress flavor changing neutral currents, bringing theoretical predictions into agreement with experimental data [17]. The subsequent discovery of the $J/\psi$, a bound $c\bar{c}$ state, confirmed these predictions [18, 19]. The successful resolution of this inconsistency through flavor physics shows the power of indirect searches.

At this time, it was already discovered that the weak interaction violates C, P, and T symmetry. The C transformation replaces particles with their anti-particles, the P transformation flips the sign of spatial coordinates, and the T transformation reverses the direction of time [20]. In addition, it was also known that the weak interaction violates CP symmetry [21]. To explain CP violation Kobayashi and Maskawa postulated a third, yet undiscovered, generation of quarks and introduced the unitary Cabibbo-Kobayashi-Maskawa matrix (CKM matrix) [22]. The CKM matrix relates the weak $(d', s', b')$ and mass $(d, s, b)$ quark eigenstates. The matrix is parameterized as follows:

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

where $|V_{\mathrm{ij}}|^2$ is the probability of a quark of flavor j transitioning into a quark of flavor i. The CKM matrix has $(N-1)^2$ free parameter, where $N$ is the number of quark generations. For two quark generations, there is only one free parameter, the Cabibbo angle. For three quark generations, there are four free parameter, three Euler angle and a complex phase which is responsible for CP-violation phenomena. Hence the introduction of new quark generation was needed to explain CP violation.

The CKM is hierarchical, therefor it is convenient to expand it in the small parameter $\lambda = \sin(V_{\mathrm{us}}) \approx 0.23$. This parameterization is called the Wolfenstein parametrization [23], the resulting CKM matrix is:

$$\begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4),$$

where

$$\lambda = \frac{V_{\mathrm{us}}}{\sqrt{V_{\mathrm{ud}}^2 + V_{\mathrm{us}}^2}}, A\lambda^2 = \lambda \left|\frac{V_{\mathrm{cb}}}{V_{\mathrm{us}}}\right|, \text{ and } A\lambda^3(\rho + i\eta) = V_{\mathrm{ub}}^*.$$

The unitarity of the CKM matrix imposes the condition

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

It is possible to depict this condition graphically as a triangle, where the area of the triangle is related to the amount of CP violation. It is convenient to normalize the sides of the triangle by $V_{\mathrm{cd}}V_{\mathrm{cb}}^*$ such that two of its vertices are aligned at (0,0) and (1,0). The final vertex is situated at $(\overline{\rho}, \overline{\eta})$, where $\overline{\rho} = \rho(1 - \lambda^2/2)$ and $\overline{\eta} = \eta(1 - \lambda^2/2)$ are related to the Wolfenstein parameter $\rho$ and $\eta$. Fig. 2.3 shows the resulting triangle, commonly referred to as the unitarity triangle. The sizes and angles of this triangle are measurable in $B$ decays, determining them is one of the main aims of $B$-factories.

In principle, the unitarity of the CKM matrix imposes five additional conditions, each of which can also be represented as a triangle. However, these triangles are usually squashed, since opposed to the triangle associated to $B$ decays, their sides have different order in the Wolfenstein parameter $\lambda$. An exception is the triangle involving top quark transitions, but experimental access to the top quark is limited due to its short lifetime and production challenges.

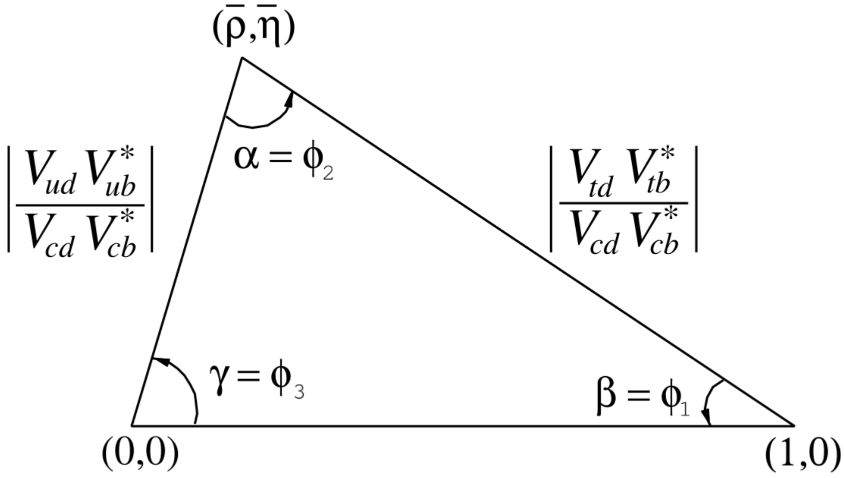


Figure 2.3: The unitary triangle [9].

## 2.4 Charmless Hadronic *B* Decays

Hadronic $B$ decays into a final state without charm quarks (charmless hadronic $B$ decays) provide an interesting playground for indirect searches for physics beyond the SM. In these decays, the tree-level amplitude is suppressed due to the smallness of $V_{\mathrm{ub}} \approx \lambda^3 \approx 0.23^3$ describing the $b \to u$ transition. Hence, loop-level amplitudes, which are susceptible to contributions from virtual heavy particles not included in the SM, contribute significantly. In Fig. 2.4, Feynman diagrams of the amplitudes expected to dominate are shown.

Quantitative predictions of charmless hadronic $B$ decays require the calculation of hadronic matrix elements. These calculations are difficult, since they involve a wide range of energy scales. In particular, computations involving the strong interaction are challenging, as it cannot be treated perturbatively at the hadronic scale ($\approx 0.5$ GeV). Two complementary approaches have been developed to address this issue. The first is based on factorization, and separates the hadronic matrix elements into perturbatively calculable and non-pertubative elements. The second approach relates groups of decays through approximate symmetries in the strong interaction. This approach cannot predict individual branching fractions or direct CP asymmetries but predictions for groups of decays only require a few reduced matrix elements. Since both approaches are complementary, they can be

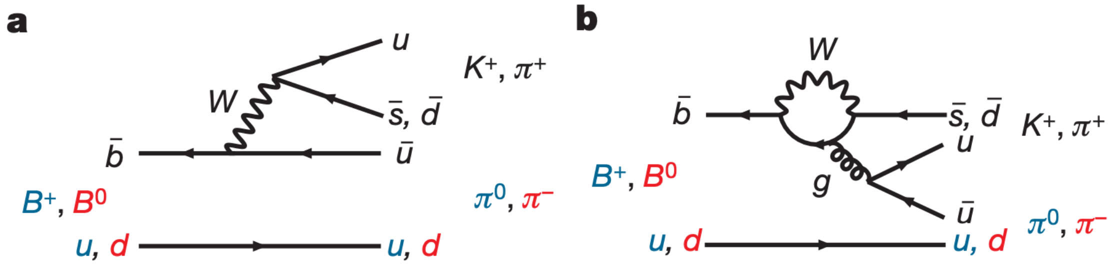


Figure 2.4: Feynman diagram for $B \to \pi\pi$ and $B \to K\pi$ decays [24].

used to crosscheck each other.

### 2.4.1 The $B \to K\pi$ Decays

The $B \to K\pi$ decays receive contributions of similar magnitude from tree-level and loop-level amplitudes. Since these amplitudes carry a different weak phase, they enable the measurement of CP violation. Historically, direct CP violation in $B$ decays was first observed in $B^0 \to K^+\pi^-$ decays [25, 26]. Soon after, it was realized the degree of direct CP violation between $B^+ \to K^+\pi^0$ and $B^0 \to K^+\pi^-$ decays ($\delta(K\pi)$) differs. The significance of this difference eventually exceeded $5\sigma$ [9]. Both decays differ only by the present spectator quark (see Fig. 2.4). Hence, from SU(3) symmetry, it was expected for the direct CP violation in both decays to be similar. This observation is often dubbed the "$K\pi$ puzzle" in literature.

Many theoretical models have been brought forward to explain the $K\pi$ puzzle. Usually, enhancing the contribution of color-suppressed tree-level amplitudes [27, 28, 29] or electroweak loop-level amplitudes [30, 31] to $B^+ \to K^+\pi^0$. While these models are able to explain the $K\pi$ puzzle they are associated with large uncertainties and leave room for contributions from physics beyond the SM [30, 31, 29]. Recent calculation based on SU(3) symmetry but including SU(3) breaking effects for $\delta(K\pi)$ predict $(11.0^{+1.2}_{-1.3})\%$ [32] which is in agreement with experimental measurements [9]. On the other hand, NNLO QCD factorization predicts $(2.10^{+1.45}_{-2.89})\%$ [33]. The unclear understanding of $B \to K\pi$ decays motivates further measurements.

A more complete examination of $K\pi$ decays via SU(3) symmetry is the following sum rule [34]

Table 2.1: Branching fractions ($\times 10^{-6}$) and direct CP asymmetries (%) for $B \to K\pi$ decays measured in this thesis. The first uncertainty is the statistical uncertainty and the second is the systematic uncertainty.

| Observable | LHCb [36, 37] | BABAR [38, 39] | Belle [35] |
|---|---|---|---|
| $\mathcal{B}(K^+\pi^0)$ | | $13.6 \pm 0.6 \pm 0.7$ | $12.62 \pm 0.38 \pm 0.71$ |
| $\mathcal{A}(K^+\pi^0)$ | $2.5 \pm 1.5 \pm 0.7$ | $3.0 \pm 3.9 \pm 1.0$ | $4.3 \pm 2.4 \pm 0.2$ |
| $\mathcal{B}(K^0\pi^+)$ | | $29.9 \pm 1.1 \pm 1.0$ | $23.07 \pm 0.53 \pm 0.71$ |
| $\mathcal{A}(K^0\pi^+)$ | $-2.2 \pm 2.5 \pm 1.0$ | $-2.9 \pm 3.9 \pm 1.0$ | $-1.1 \pm 2.1 \pm 0.6$ |

$$\begin{aligned}\Delta(K\pi) = \mathcal{A}(K^+\pi^-) &+ \mathcal{A}(K^0\pi^-)\frac{\mathcal{B}(K^0\pi^-)}{\mathcal{B}(K^+\pi^-)}\frac{\tau^0}{\tau^+} \\ &- 2\mathcal{A}(K^+\pi^0)\frac{\mathcal{B}(K^+\pi^0)}{\mathcal{B}(K^+\pi^-)}\frac{\tau^0}{\tau^+} - 2\mathcal{A}(K^0\pi^0)\frac{\mathcal{B}(K^0\pi^0)}{\mathcal{B}(K^+\pi^-)},\end{aligned}$$

where $\mathcal{A}(X)$ and $\mathcal{B}(X)$ denote the direct CP violation and branching fraction for the specified decay, $\tau^0$ is the $B^0$ lifetime, and $\tau^+$ is the $B^+$ lifetime. In the SM, $\Delta(K\pi)$ is predicted to be 0 with an uncertainty of the order of 1% [33, 34]. The Belle collaboration measured the sum rule to be $-0.270 \pm 0.132 \pm 0.060$ [35], where the first uncertainty is the statistical uncertainty and the second is the systematic uncertainty. A determination using current world averages for the input parameter yields a value of $-0.13 \pm 0.11$ [9], where correlations between measurements are neglected. In Table 2.1, measurements of the branching fractions and direct CP violation for $B \to K\pi$ decays studied in this thesis are summarized. Since the determination of the sum rule itself as well as the direct CP asymmetry for the individual decay channels are limited by the statistical uncertainty, further measurements of these channels are required.

### 2.4.2 The $B \to \pi\pi$ Decays

The determination of the sides and angles of the CKM triangle is one of the cornerstones of the Belle II physics program. Among the angles, $\phi_2$ and $\phi_3$ are considerably less well measured than $\phi_1$ (see Fig. 2.5). For the decay $B^0 \to \pi^+\pi^-$, the mixing-induced CP violation parameter would, in the absence of contributions from loop-level amplitudes, be equal to $\sin 2\phi_2$ [40]. However, loop-level amplitudes contribute considerably. Hence the measured angle $\phi_2^{\rm meas}$ is shifted from the physical angle $\phi_2$ via

$$\phi_2^{\rm meas} = \phi_2 + \Delta\phi_2,$$

where $\Delta\phi_2$ is the shift induced by the loop-level amplitudes. To disentangle this effect, several $B \to \pi\pi$ decays are measured and isospin symmetry is employed [41].

The decay amplitude for a given $B \to \pi\pi$ decay can be written as

$$A_{ij} = |T_{ij}|e^{i\phi_3} + |P_{ij}|e^{i\delta},$$

where $ij$ denotes the charge configuration of the final state ($+-$, $+0$, or $00$), $|T|$ is the magnitude of the tree-level amplitude with weak phase $\phi_3$, and $|P|$ the magnitude of the loop-level amplitude with strong phase $\delta$.

Since the decaying $B$ is spinless, the two-pion system is produced predominately without angular momentum. Hence, the total isospin $I$ of the two-pion system is even (zero or two). For odd isospin, the wave function would be anti-symmetric under particle exchange which is forbidden by Bose-Symmetry. The $\pi^+\pi^0$ state is a pure $I = 2$ state, while the $\pi^+\pi^-$ and $\pi^0\pi^0$ states are both $I = 0$ and $I = 2$.

The loop-level amplitudes mediate only isospin $1/2$ transitions, while the tree-level amplitude mediates isospin $1/2$ as well as $3/2$ transitions. Hence, for the $B^+ \to \pi^+\pi^0$ decay, only the tree level amplitude contribute while for the other two decays both tree- and loop-level amplitudes contribute.

Expanding the decay amplitudes in terms of the total isospin pieces then yields the relationship:

$$(1/\sqrt{2})A^{+-} + A^{00} = A^{+0}$$

and for the CP conjugated amplitudes

$$(1/\sqrt{2})\overline{A}^{+-} + \overline{A}^{00} = \overline{A}^{-0},$$

where $\overline{A}^{ij}$ are obtained from $A^{ij}$ by changing the sign of the CKM phases.

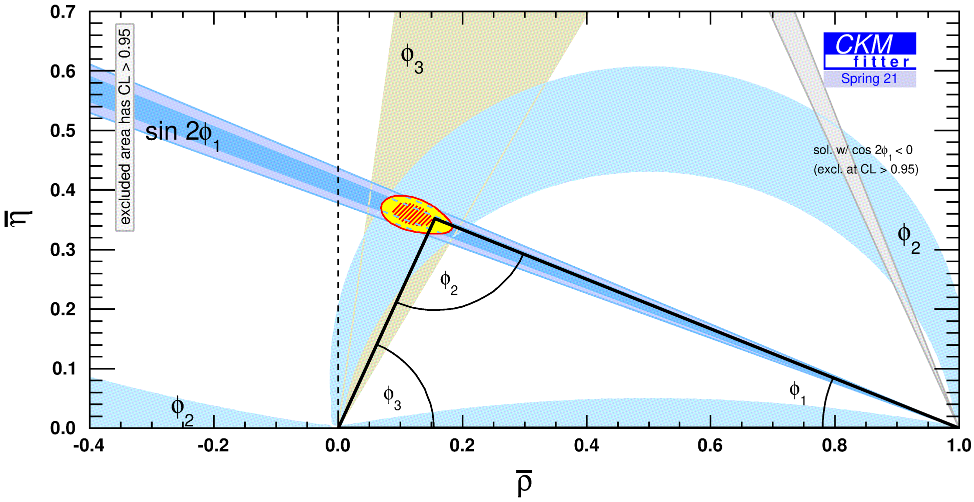


Figure 2.5: Constraints on the CKM triangle from measurements of the angles only [42].

Both of these relations can be displayed as a triangle in the complex plane. Since the $A^{+0}$ amplitude for $B^+ \to \pi^+\pi^0$ receives no loop-level amplitude contribution, it differs from its CP conjugate $\overline{A}^{-0}$ only by a weak phase. Explicitly,

$$\overline{A}^{-0} = e^{2i\phi_3} A^{+0}.$$

By rotating the $\overline{A}^{ij}$ triangle by a weak phase $e^{2i\phi_3}$, the bases of the two triangles can be brought to overlap. The rotated amplitudes are donated as $\tilde{A}^{XX}$. The angle between the side of the triangle associated with $B^0 \to \pi^+\pi^-$ ($A^{+-}/\sqrt{2}$ and $\tilde{A}^{+-}/\sqrt{2}$) is then equal to $\Delta\phi_2$.

The sides and angels of the triangles can be measured my measuring the branching fractions and direct CP asymmetries for the different $B \to \pi\pi$ decays. In Table 2.2, the current status of measurements of $B^+ \to \pi^+\pi^0$ which is the decay measured in this thesis is summarized. At present, the precision of the measurements for $B^+ \to \pi^+\pi^0$ and the isospin analysis are limited by the statistical uncertainty. Further measurements with the larger Belle II dataset are essential to reduce these uncertainties and improve the determination of $\phi_2$.

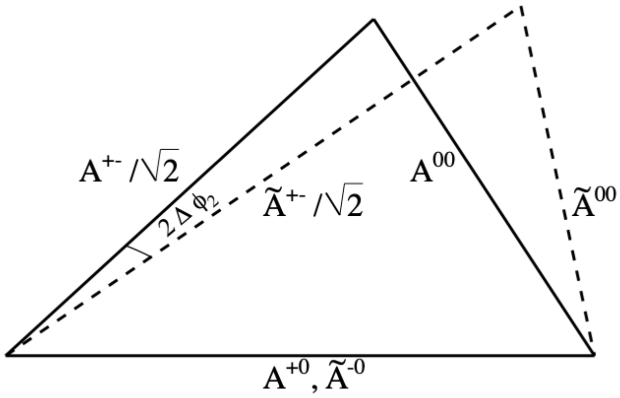


Figure 2.6: Triangles constructed via the isospin relation between the $B \to \pi\pi$ decay amplitudes. The triangles have been rotated to that the $A^{+-}/\overline{A}^{+-}$ are aligned. Hence, the triangle belonging to $\overline{A}^{+-}$ is indicated with a tilde [43].

Table 2.2: Branching fractions ($\times 10^{-6}$) and direct CP asymmetries (%) for $B^+ \to \pi^+\pi^0$. The first uncertainty is the statistical uncertainty and the second is the systematic uncertainty.

| Observable | *BABAR* [38] | Belle [35] |
|---|---|---|
| $\mathcal{B}(\pi^+\pi^0)$ | $5.02 \pm 0.46 \pm 0.26$ | $5.86 \pm 0.26 \pm 0.38$ |
| $\mathcal{A}(\pi^+\pi^0)$ | $0.03 \pm 0.08 \pm 0.01$ | $0.025 \pm 0.043 \pm 0.007$ |

# Chapter 3

# The Belle II Experiment at the SuperKEKB Collider

The $B$-factories, such as Belle II, are collider based particle physics experiments. They were originally designed to measure $B$ decays, but nowadays have a much broader physics program including measurements in the charm, tau, quarkonium, and dark sectors. Currently, two types of $B$-factories are in operation, distinguished by the type of collider they use:

- The LHCb experiment [44] at the large hadron collider at CERN near Geneva, Switzerland.
- The Belle II experiment [45] at the SuperKEKB electron-positron collider.

The remainder of this section discusses briefly advantages of $B$-factories situated at electron-positron colliders compared to those at hadron colliders.

At electron-positron accelerators based $B$-factories, beams are collided with a center-of-mass energy of 10.58 GeV. This energy is at the mass of the $\Upsilon(4S)$ resonance, which almost exclusively decays into $B\overline{B}$ pairs [9], enhancing the $B\overline{B}$ production cross section. In addition, the $B$ mesons are produced almost at rest, simplifying their reconstruction. Due to the low background environment in electron-positron collisions, reconstruction of decays is possible with high efficiency and low trigger biases. This is especially important for the reconstruction of final states with neutral particles such as $\pi^0$ or $K_S^0$, which are comparatively harder to reconstruct at hadron collider due to ubiquitous background. Moreover, the knowledge of the initial-state conditions and almost $4\pi$ hermiticity of the detector allows to infer missing energy and momenta from, e.g., neutrinos. This is especially important for analyses involving semileptonic $B$ decays. Finally, since the absolute number of produced $B$ mesons can be measured at electron-positron colliders, they enable the measurement of absolute branching fractions.

In the following sections, the SuperKEKB collider, the Belle II detector, and the software packages used are discussed in more detail.

## 3.1 SuperKEKB Collider

SuperKEKB [45, 46] is a double-ring positron-electron collider. Positrons are accelerated to an energy of 4 GeV before they are stored in the low-energy ring. Electrons are stored in the high-energy ring with an energy of 7 GeV. Fig. 3.1 shows a schematic of the collider.

Positrons and electrons collide at a center-of-mass energy of 10.58 GeV, near the mass of the $\Upsilon(4S)$ resonance. At this energy, the production of $B\overline{B}$ pairs is enhanced. Table 3.1 summarizes the cross section for dominant processes in $e^+e^-$ collisions at this energy.

Due to the asymmetric beam energies, produced $B\overline{B}$ pairs are boosted in the direction of the electron beam. The Lorentz boost is $\beta\gamma = 0.28$ and increases the average flight length for $B$ to 130 µm, enabling the determination of their decay time difference via measurements of the decay vertex positions. Compared to its predecessor KEKB, the Lorentz boost is reduced (Lorentz boost at KEKB $\beta\gamma = 0.425$), necessitating upgrades to the tracking system to recover the precision required for analyses of the decay-time evolution. These upgrades are discussed in Section 3.2.2. The beam boost was chosen to reduce the emittance and increase the beam lifetime.

A measure for the performance of a collider is the luminosity $\mathcal{L}$. The luminosity is proportional to

$$\mathcal{L} \propto \frac{I_\pm \xi_\pm}{\beta_y^*}, \tag{3.1}$$

where $I$ is the beam current, $\xi$ is the vertical beam-beam parameter, and $\beta_y^*$ is

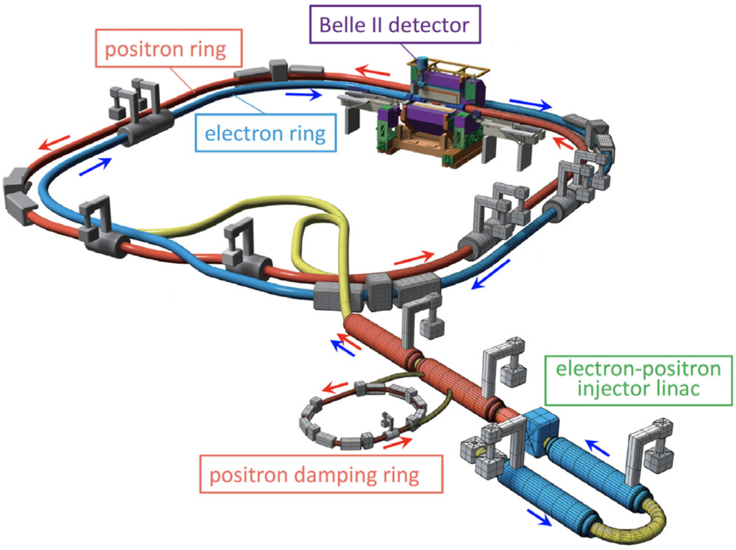


Figure 3.1: Schematic of the SuperKEKB collider. Modified from [46].

the beta function at the interaction point. The suffix + (−) denotes the positron (electron) beam parameter. SuperKEKB aims to reach a luminosity 30 times higher than that of KEKB. To achieve this goal, the aim is to double the beam currents and reduce $\beta_y^*$ by a factor of 20 compared to KEKB. The vertical beam-beam parameter is kept at a value already achieved at KEKB. The reduction of $\beta_y^*$ is achieved by adopting the nanobeam scheme [47].

SuperKEKB and the Belle II detector are in operation since 2018 and scheduled to operate until 2043. Since 2018, SuperKEKB achieved a world record luminosity of $5.1 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ (2.4 times that of KEKB) and Belle II integrated a data set of $575\,\mathrm{fb}^{-1}$. Current projections for the integration of further data consider two scenarios: A scenario where the focusing magnet in front of the interaction point is upgraded and a second scenario where it is not. In the first scenario, SuperKEKB reaches a luminosity 30 times higher than that of its predecessor and Belle II collects $50\,\mathrm{ab}^{-1}$ of data. In the second scenario, SuperKEKB achieves a luminosity 15 times higher than its predecessor and Belle II integrates a data set of $35\,\mathrm{ab}^{-1}$. The luminosity projections are shown in Fig. 3.2.

Table 3.1: The $e^+e^-$ cross section at 10.58 GeV [48].

| Process | Cross Section [nb] |
|---|---|
| $e^+e^- \to b\bar{b}$ | 1.05 |
| $e^+e^- \to c\bar{c}$ | 1.30 |
| $e^+e^- \to s\bar{s}$ | 0.35 |
| $e^+e^- \to u\bar{u}$ | 1.39 |
| $e^+e^- \to d\bar{d}$ | 0.35 |
| $e^+e^- \to \tau^+\tau^-$ | 0.94 |
| $e^+e^- \to \mu^+\mu^-$ | 1.16 |
| $e^+e^- \to e^+e^-$ | 40 |

.

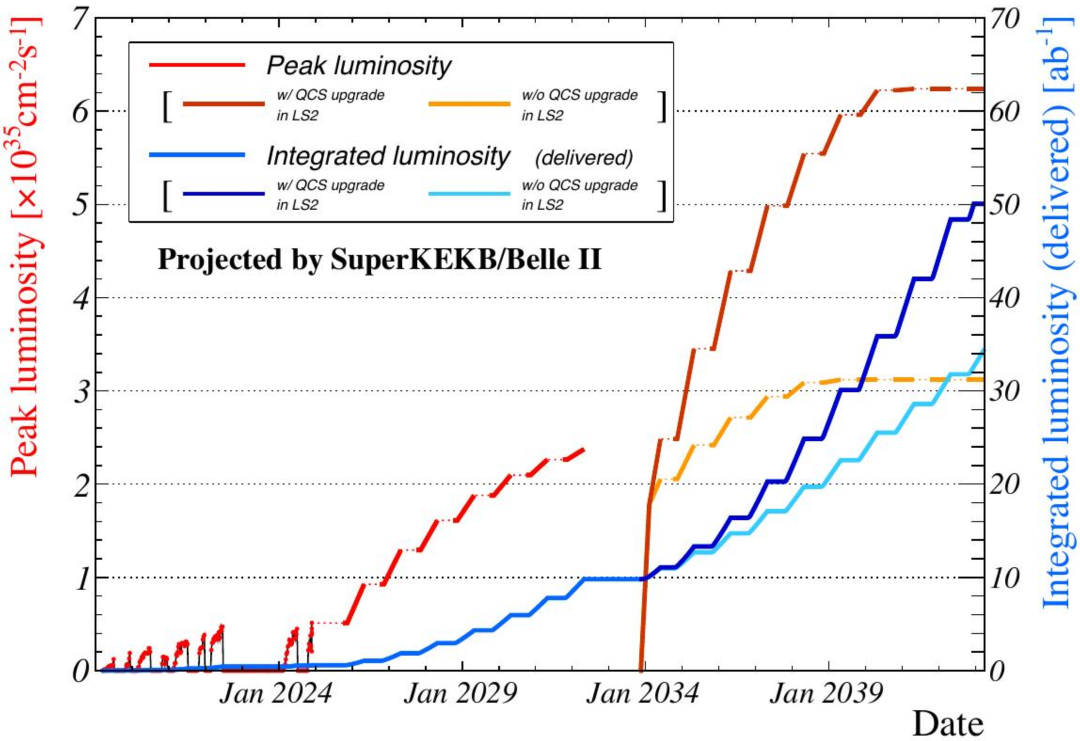


Figure 3.2: Luminosity (red colors) and integrated luminosity (blue colors) projection for SuperKEKB and Belle II under two scenarios [49]. One where the final focusing magnet is upgraded and one where it is not.

## 3.2 Belle II Detector

The Belle II detector (Fig. 3.3) is a general-purpose, large-solid-angle detector surrounding SuperKEKB's interaction point. It measures the energy as well as the momentum of a broad range of neutral and charged particles produced in SuperKEKB's collision and infers their species. The detector is approximately cylindrical, with a width and height of about 8 m and a weight of around 1400 t.

From the interaction point outward, the detector comprises several subdetectors: two-layer silicon-pixel sensors (PXD), four-layer silicon-strip sensors (SVD), a wire chamber (CDC), photomultiplier tubes with quartz-radiator bars (TOP), photosensors with an aerogel radiator (TOP), CsI(Tl) crystals (ECL), and multiple layers of resistive plate counters (KLM). A comprehensive description of the detector can be found in [45, 50], upon which the following detailed sections are based.

### 3.2.1 Coordinate System

The Belle II experiment uses a right-handed coordinate system, with the origin at the interaction point. The $z$-axis, defined as the bisector of the two beams, aligns positively with the electron-beam direction. The $y$-axis points vertically upwards, while the $x$-axis is perpendicular to both the $y$- and $z$-axes.

In the coordinate system, the azimuth angle $\phi$ is measured counterclockwise around the $z$-axis, beginning from the $x$-axis at the 0 position. The zenith angle $\theta$ is determined by the angle between the $z$-axis and the radial line. In Fig. 3.4, the coordinate system is superimposed on the detector.

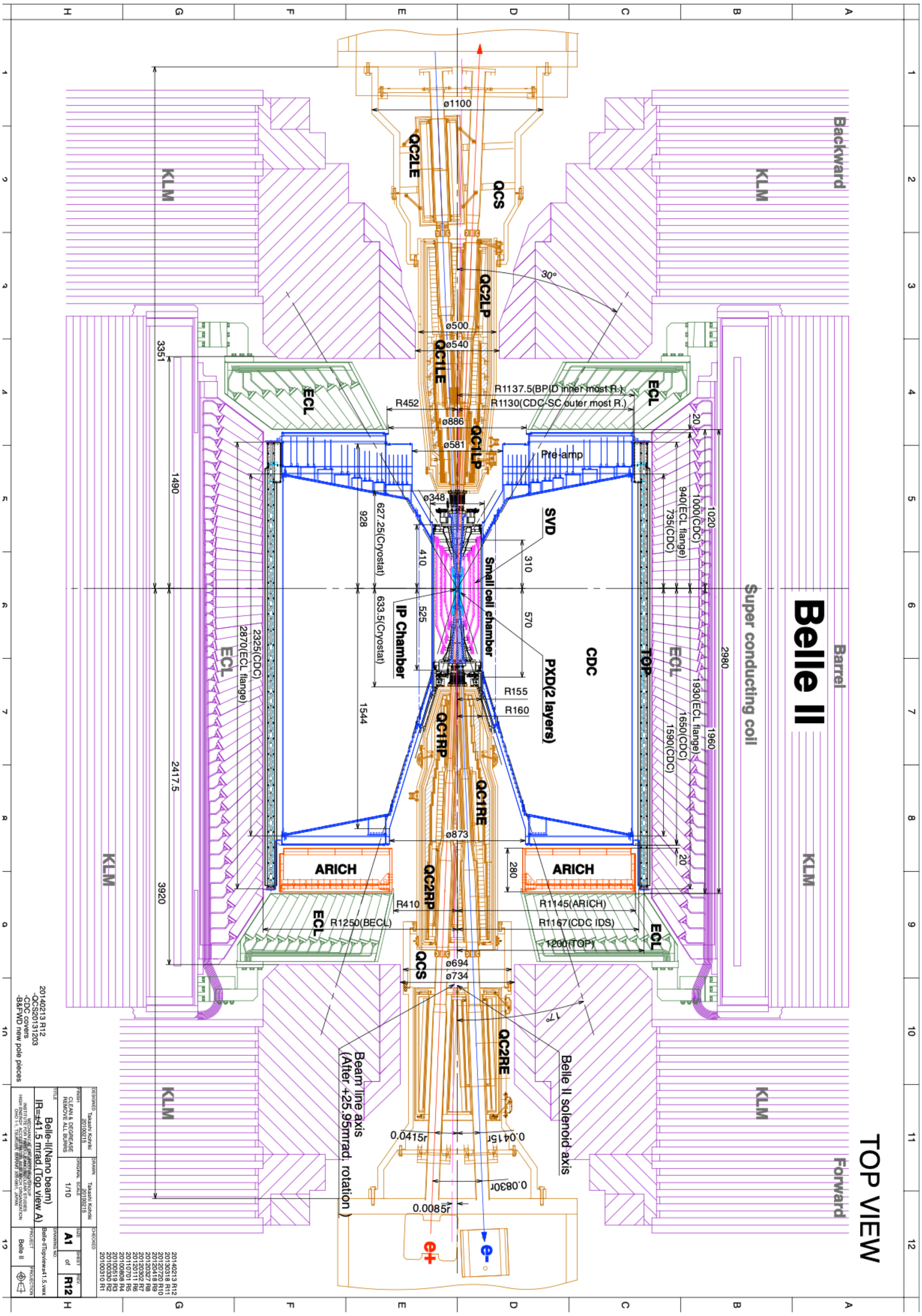


Figure 3.3: Top view of the Belle II detector [50].

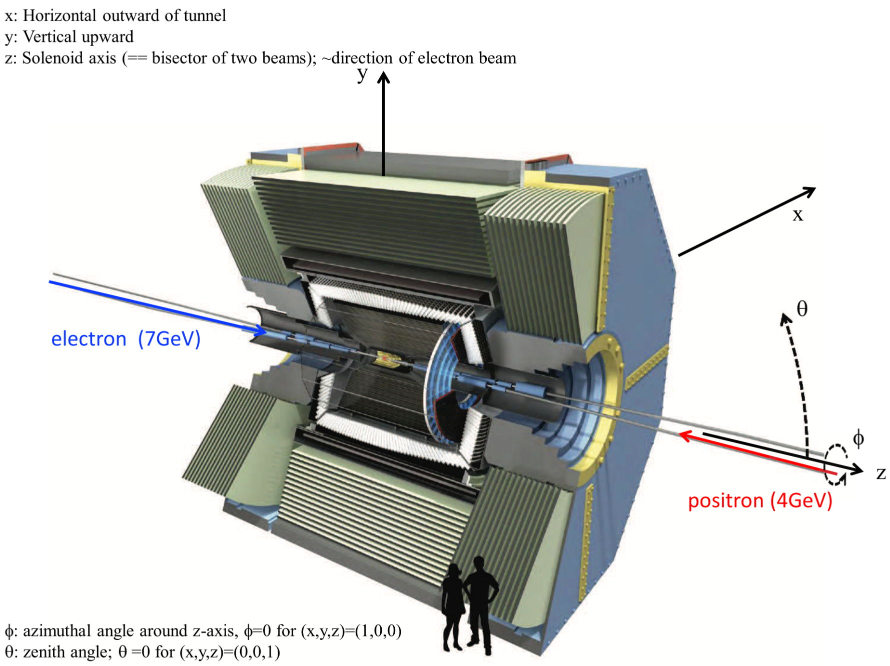


Figure 3.4: The Belle II coordinate system [51].

### 3.2.2 Tracking System

Track reconstruction is essential in Belle II, as the majority of decays of interest involve charged particles. It plays a particularly important role in identifying the decay vertices of short-lived particles such as $B$, $D$, and $\tau$. Accurate vertex reconstruction enables time-dependent analyses and precise lifetime measurements.

The Belle II tracking system, like ancient Gaul, is divided into three parts: the Pixel Detector (PXD), the Silicon Vertex Detector (SVD), and the Central Drift Chamber (CDC). The entire system is submerged in a 1.5 T magnetic field. Due to the magnetic field, trajectories of charged particle traversing the tracking system are bent by the Lorentz force. Measuring the curvature allows their momentum to be determined.

The **PXD** is the innermost subdetector. The design of the PXD has been driven by considerations of the vertex resolution, beam background level, and material budget.

Due to the lower beam boost of SuperKEKB compared to KEKB ($\beta\gamma = 0.28$ vs. 0.43), the separation between the $B$ meson decay vertices is reduced. Consequently, improved vertex resolution is required to match or surpass the resolution of time-dependent analyses achieved at Belle. Thanks to the nano-beam scheme, the beam pipe at Belle II has a radius of 10 mm only two thirds of that of Belle. This allows the innermost layer of the tracking system to be positioned closer to the interaction point than at Belle, improving vertex resolution. However, since beam background scales with the inverse square of the radius, the first detector layers are subject to higher background rates. In addition, beam backgrounds are generally expected to be higher at Belle II compared to Belle due to the higher instantaneous luminosity.

To address this challenge, the PXD consists of silicon pixel sensors instead of silicon strip sensors. Pixel sensors provide a finer granularity due to a much higher number of read-out channels. Using strip sensors would result in a high fraction of read-out channels being triggered in the same event, increasing the likelihood of ambiguities in track reconstruction.

Another design considerations is the low material budget to reduce multiple scattering. The PXD consists of pixelated DEPFET sensors [52], which integrate signal detection and amplification into a thin sensor. These sensors allow read-out electronics to be placed outside the acceptance region hence them not contributing to the material budget. This design fulfills the low material budget requirement: the PXD contributes 0.21% radiation length per layer to the budget [53], while for comparison the beam pipe contributes 0.7%.

The PXD consists of two layers located at radii of 14 mm and 22 mm. These layers are composed of 20 ladders in total, arranged in a windmill geometry. The detector covers a polar angle acceptance region from 17° to 155° in zenith angle. During taking of the data used in this thesis, only the first layer as well as one sixth of the second layer was instrumented, see Fig. 3.5. In 2023, the PXD was replaced by a new PXD which is fully instrumented.

The **SVD** surrounds the PXD and consists of four layers of silicon strip detectors located at radii of 38 mm, 80 mm, 115 mm, and 140 mm. These layers cover a zenith angle range from 17° to 150°. To maximize the acceptance and reduce the amount of material needed, the sensors are slated in direction of the electron beam (in beam boost direction). A schematic of the SVD is shown in Fig. 3.5.

As with the PXD, one of the main design considerations is the beam background. To keep the the fraction of triggered strips in a signal event below 10%, the layer closest to the interaction point is farther away than for Belle's SVD2 and the number of read-out chips has been more than doubled. At the same time, the outermost layer extends farther than in SVD2, as placing the drift chamber at that position would result in a too high background level for a wire chamber. Furthermore, SVD provides timing information. This enables the suppression of hits out of event time caused by the beam background.

Together, PXD and SVD provide a momentum resolution of the order of tens of MeV/c. This resolution is essential to reconstruct low-momentum pions from $D^* \to D\pi$ decays which often do not reach the CDC. High efficiency in the reconstruction of these decays is needed, as they provide information about the flavor of the decaying $B$. In addition, CDC and SVD allow to reconstruct $K_S^0$ which typically decay outside the PXD and leave no hits in it. This capability is especially important for the $B^+ \to K^0\pi^+$ decays studied in this thesis.

The **CDC** is the outermost subdetector of Belle II's tracking system. It is a wire chamber composed of nine superlayers, each consisting of six layers of sense wires. The wire orientations alternate with respect to the beam axis, allowing

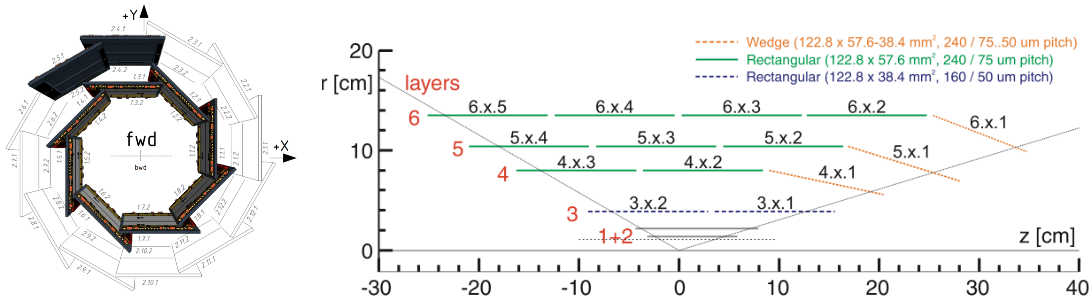


Figure 3.5: Left: PXD instrumentation during data taking relevant for this thesis [56]. Right: Schematic layout of the SVD (layer 3-6). Individual sensors and their properties are shown in blue and green. Slanted sensors and their properties in orange [57].

three-dimensional track reconstruction. The superlayer closest to the interaction point is more compact due to the high beam background level and includes two additional guard wires for the same reason. Compared to the Belle CDC, the Belle II CDC begins at a larger radius to avoid the beam background level close to the interaction point. The Belle II CDC also extends to a larger outer radius, due to the more compact design of the particle identification system. The wire configurations of the Belle and Belle II CDCs are compared in Fig. 3.6.

The CDC also contributes to particle identification, since charged particles loose energy in the CDC due to ionization. The energy loss per unit length depends on its velocity or alternatively $p/m$, where $p$ is the particles momentum and $m$ its mass. Hence measuring the energy loss and momentum of a particle allows to determine its species. In Fig. 3.6, the energy loss as function of momentum for different species of particles in the CDC is shown. Especially for momenta below 1 GeV/c, the energy loss provides strong discriminating power.

Performance studies of the tracking system show a transverse impact parameter resolution of $(13.64 \pm 0.08)\,\mu\text{m}$ and a longitudinal impact parameter resolution of $(14.92 \pm 0.07)\,\mu\text{m}$ [54]. Studies of the $D^0$ and $D^+$ lifetime show a decay-time resolution of 70 fs and 60 fs [55]. Compared to Belle, these parameters are improved by a factor of two.

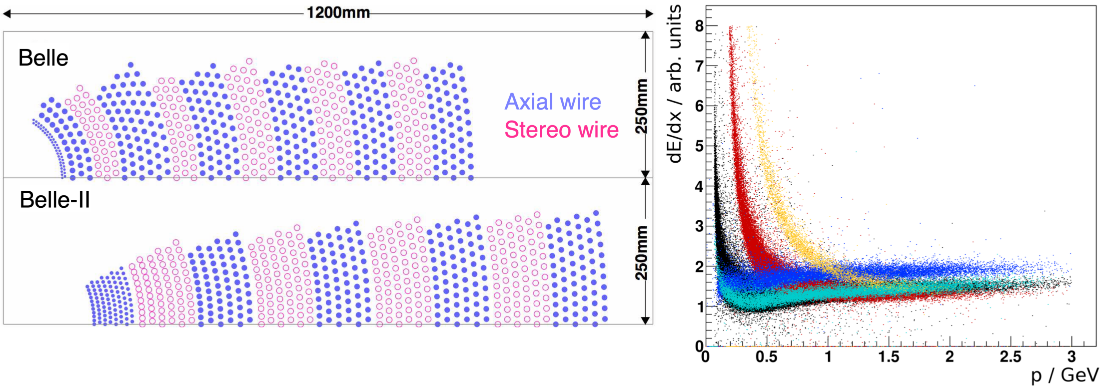


Figure 3.6: Left: Wire configuration of the Belle CDC (upper) and Belle II CDC (lower) [58]. Right: Scatter plot of the energy loss per unit length as a function of the particles momentum in the CDC for pions (black), kaons (red), protons (gold), electrons (blue), muons (cyan) [59].

### 3.2.3 Particle Identification System

The particle identification system consists of two complementary subdetectors. The TOP covers the central part of the detector, corresponding to a zenith angle range from 31° to 128°. The ARICH is located in the direction of the beam boost (electron beam), covering a zenith angle range from 14° to 30°. Both detectors are designed to separate pions and kaons over a wide momentum range, but also provide discrimination for other particle species. TOP and ARICH follow a similar design principle (Fig. 3.7): Cherenkov photons are generated in a radiator (quartz in case of TOP, aerogel in case of ARICH) and detected by photon detectors. In the TOP, photons are internally reflected within the quartz bars until they reach the photon detectors. The impact position as well as the arrival time are used for particle identification. In the ARICH, the photon detectors are positioned 20 cm away from the radiator, allowing Cherenkov rings to form. The measured radius of the rings enables particle identification.

### 3.2.4 Electromagnetic Calorimeter

The ECL measures the energy of particles and is essential for detecting photons and other neutral particles such as the $\pi^0$. Neutral particles make up about one third of all particles produced in $B$ decays. They do not leave a signal in the

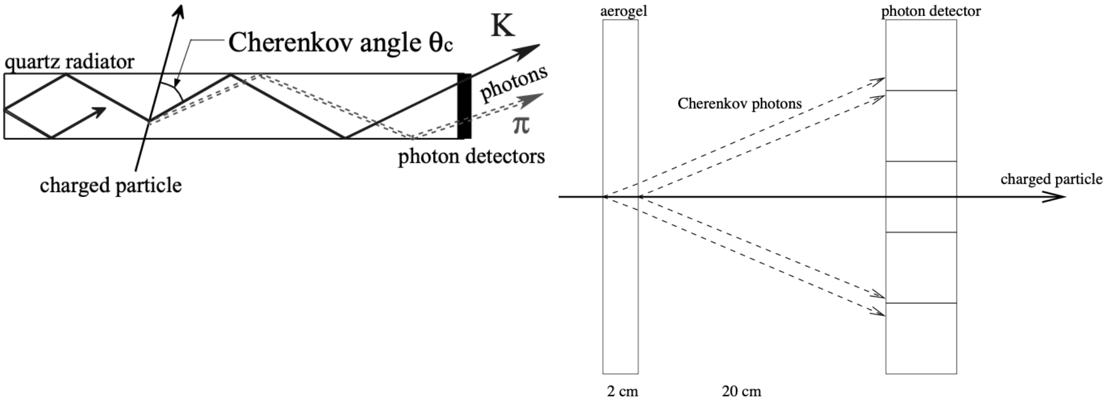


Figure 3.7: Schematic of the TOP (left) and ARICH (right) detectors [45]. Both detectors follow a similar design principle: Cherenkov photons are generated in a radiator (quartz in case of TOP, aerogel in case of ARICH) and detected by photon detectors. The arrival time of the photons (only TOP) and their impact position are used for particle identification.

tracking signal hence it is crucial to observe them with high efficiency in the ECL. Belle II reuses the CsI(Tl) crystals and photon detectors of the Belle ECL, while read-out electronics are upgraded to provide timing information which allows to suppress background. The ECL is subdivided into three distinct regions, covering the zenith ranges [12.5°, 31.4°], [32.2°, 128.7°], and [130.7°, 155.1°].

### 3.2.5 $K_L^0$ and $\mu$ Detector

The KLM is the outermost subdetector of Belle II. Its primary function is the detection of $K_L^0$ and $\mu$ that deposit no or only a part of their energy in the ECL. To achieve this, the KLM provides 3.9 interaction lengths of material, compared to the 0.8 interaction lengths of the ECL. It consists of glass-electrode resistive plate chambers as detectors sandwiched between iron plates as absorbers. The iron plates also serve as flux return for the magnetic field. The KLM covers a zenith angle range of [25°, 155°].

## 3.3 Trigger System

The data rate and absolute data size produced by the Belle II detecor is too large for all the data to be stored. A trigger system is used to suppress beam background events as well as select physics events of interest, and consequently reduce the data rate and size. For hadronic events, the trigger logic is inclusive, meaning no hadronic events are discarded. The Belle II trigger system consists of two layers: A hardware-based trigger (L1) followed by a software-based trigger (HLT) [60, 61].

The L1 is implemented on field-programmable gate arrays to allow configuration of the trigger logic. A global decision logic receives subtrigger information from the CDC, TOP, ECL, and KLM and decides whether to keep an event. The L1 trigger reduces the data rate to 30 kHz with a latency of 5 µs.

The HLT receives events passing the L1 and further reduces the data rate to 10 kHz. The HLT performs a full event reconstruction (excluding the PXD) similar to offline analysis and selects physics events of interest. In addition to reducing the event rate, the HLT also decreases the data size by discarding PXD hit information that is unlikely to be associated with a reconstructed track. This results in a factor of five reduction in the amount of data saved.

## 3.4 The Belle II Analysis Software Framework

The Belle II analysis software framework is the central data processing software used at Belle II [62, 63]. The framework is written in `C++` and interfaced with `python`. It is responsible for generation of simulated data as well as processing of experimental and simulated data. During data processing, it handles low-level task such as track reconstruction and clustering, but also high-level tasks like reconstruction of specific decays, application of selection criteria, and vertex fitting.

To process data, user define a sequence of modules and set their parameters. These modules are executed along a so-called path. Data is stored into a common exchange container and shared between modules. This setup allows subtasks (e.g., track or decay reconstruction) to be split into different independent and interchangeable modules.

After reconstruction, variables of interest for events passing the selection criteria are stored in a `ROOT` ntuple [64, 65] and made available for further analysis.

## 3.5 Simulation

In this thesis, we use simulated data to establish the analysis strategy prior to looking at experimental data. The simulated data is produced using the Monte Carlo method. We utilize `EvtGen` to simulate known $B$ meson decays and their time evolution [66], `KKMC` to generate $e^+e^- \to q\overline{q}$, where $q = u, d, c, s$, events [67], `PYTHIA8` to simulate hadronization [68], `PHOTOS` to simulate final state radiation [69], and `GEANT4` to model the detector response [70]. Additionally, we overlay beam-induced background on the simulated events [71].

# Chapter 4

# Measurements

## 4.1 Analysis Strategy

We aim to measure the branching fraction and direct CP asymmetries for $B^+ \to K^0\pi^+$, $B^+ \to K^+\pi^0$, and $B^+ \to \pi^+\pi^0$ decays with $K^0 \to K^0_S \to \pi^+\pi^-$ and $\pi^0 \to \gamma\gamma$. Unless stated otherwise, charge conjugation is implied during throughout this thesis. For these measurements, we use a data set corresponding to an integrated luminosity of $362\,\mathrm{fb}^{-1}$ taken by the Belle II detector on-resonance from March 2019 until July 2022. The analysis presented herein is a blind analysis, the analysis is developed on simulated data and validated using control channels. The experimental data containing the signal region is only analyzed after the analysis procedure has been finalized and validated. These steps are undertaken to reduce the chance of accidentally biasing the result.

To perform the analysis, we reconstruct the decays from tracks and energy clusters measured in the detector. During the reconstruction, selection requirements are applied to separate the signal from background. A key challenge of this analysis is the treatment of background from $e^+e^- \to q\bar{q}$, where $q = u, d, c, s$, (continuum background) which is strongly predominant compared to signal. To suppress continuum background, we identify 48 discriminating observables and train a multivariate algorithm that combines them non-linearly.

After the selection, the reconstructed samples still contain contributions from background processes. We perform a multidimensional fit to determine the sample composition and direct CP asymmetry. Another key challenge in the fit is the correct treatment of peaking background, where the charged track has been misidentified as a pion or kaon. This type of background peaks at the same or close to the same position as signal, hence it is challenging for the fit to disentangle these components. We fit the $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$ samples simultaneously to properly account for events where a charged kaon or pion as been misidentified as the other. We perform several test of our estimator to test its robustness.

As the analysis is developed blind, it relies heavily on the accurateness of the simulation. We study several control channels to assess differences between the experimental data and simulation and correct for them. Finally, we evaluate the systematic uncertainties.

After the whole analysis procedure has been developed and validated, we fit the experimental data and extract the physics parameter.

## 4.2 Data Sample

In this analysis, we use experimental and simulated data. The experimental data has been collected from March 2019 until July 2022 by the Belle II detector. The data contains 387 million bottom-antibottom meson pairs produced in $e^+e^-$ collisions at the $\Upsilon(4S)$ resonance by the SuperKEKB collider, corresponding to an integrated luminosity of $(362 \pm 2)\,\mathrm{fb}^{-1}$.

The simulated sample corresponds to $1000\,\mathrm{fb}^{-1}$ and contains $e^+e^- \to q\bar{q}$, where $q = u, d, c, s$, events in appropriate proportions. Furthermore, dedicated simulation samples are generated for the signal modes and control channels. These samples contain over a million events for each respective decay, exceeding the amount expected in data by more than three orders of magnitude.

## 4.3 Reconstruction and Selection

We reconstruct $B^+ \to K^0\pi^+$, $B^+ \to K^+\pi^0$, and $B^+ \to \pi^+\pi^0$ decays with $K^0 \to K^0_S \to \pi^+\pi^-$ and $\pi^0 \to \gamma\gamma$. The reconstruction follows a bottom-up approach where $K^+$, $\pi^+$ and $\gamma$ are first reconstructed from tracks as well as energy clusters in the calorimeter and then combined in kinematic fits to reconstruct the decay chains.

### 4.3.1 Signal $B$ Candidate Reconstruction and Selection

To suppress tracks from beam background not originating from the interaction point, tracks must have a distance of closest approach to the interaction point of less than 2.0 cm along the $z$ axis and less than 0.5 cm in the transverse plane. In addition, tracks are required to have a polar angle within the drift-chamber acceptance and be associated with atleast 20 measurement points therein. The later two requirements are applied as they improve the reliability of the particle identification calibration. A track is identified as kaon (pion) candidate if the binary particle identification likelihood $\mathcal{L}_{K/\pi} = \mathcal{L}_K/(\mathcal{L}_K + \mathcal{L}_\pi)$, where $\mathcal{L}_K$ and $\mathcal{L}_\pi$ are the particle identification likelihoods for a kaon and pion, is greater (lower) than 0.4. This requirement correctly identifies 86 % of kaons and pions.

We reconstruct $K_S^0$ through the decay channel $K_S^0 \to \pi^+\pi^-$. Neutral kaon candidates are retained if the dipion mass satisfies $450 < m(\pi\pi) < 550\,\mathrm{MeV}/c^2$. The mass requirement effectively suppresses tracks originating from charged kaons, and $K_S^0$ fly on average 10 cm. Hence, for tracks used in the $K_S^0$ reconstruction, we do not impose a particle identification requirement, nor do we require tracks originate from the interaction point. A vertex fit using the `TreeFit` algorithm [72] is performed and only candidates with a converged fit are retained. Afterwards a fit using the `kFit` algorithm [73] is performed, where the mass of the $K_S^0$ candidate is constrained to its nominal value. Within `basf2` two boosted decision tree classifiers [74] are available that distinguish $K_S^0$ from background, the `V0Selector` and the `LambdaVeto`. Both provide an output between 0 (unambiguous identification of background) and 1 (unambiguous identification of a $K_S^0$). The `V0Selector` is trained to distinguish $K_S^0$ from random combination of tracks. The classifier takes 15 variables as input that describe the kinematics of the $K_S^0$ candidate and the associated pion pair. Real $K_S^0$ have a well-reconstructed vertex separated from interaction point due to the considerable $K_S^0$ flight length while random combination of tracks have a reconstructed vertex of poorer quality closer to the interaction point. In addition, since the $K_S^0$ is moving directly away from the interaction point, the pointing angle between the vector connecting the interaction point and the $K_S^0$ flight direction is small for real $K_S^0$ but tends to larger values for random combination of tracks. Hence the significance of the $K_S^0$ flight distance and the pointing angle are the two most discriminating variables in the `V0Selector`. The `LambdaVeto` suppresses background from $\Lambda^0 \to p^+\pi^-$ decays, where the proton is reconstructed as a pion. The `LambdaVeto` uses eight input variables: the particle identification likelihood of the associated pions to be a proton, kinematic information of the reconstructed pions, and the invariant mass distribution of the pion pair if either particle is assumed to be a proton. With the latter being the most discriminating variable in the classifier. To optimize the requirements on the invariant mass, `V0Selector` output, and the `LambdaVeto` output a random cut optimization is performed. For all three variables, a million sets of cut values are drawn simultaneously from a uniform distribution and the significance defined as $S/\sqrt{S+B}$, where $S$ is the number of signal and $B$ the number of continuum background events ($e^+e^- \to q\overline{q}$, where $q = u, d, c, s$) in the signal region defined as $-0.13 < \Delta E < 0.10\,\mathrm{GeV}$, where $\Delta E$ is difference of the energy of the reconstructed $B$ candidate and the beam center of mass energy, is calculated. The requirements on the variables are tightened to the subset of 100 requirements with highest significance. The final selection is $480 < m(\pi\pi) < 510\,\mathrm{MeV}/c^2$, `V0Selector` output $> 0.2$, and `LambdaVeto` output $> 0.06$. This selection increase the significance for $B^+ \to K_S^0\pi^+$ decays from 3.5 to 7.9.

We form $\pi^0$ candidates via $\pi^0 \to \gamma\gamma$, reconstructing photons from ECL clusters not associated to a track. To suppress beam-background, we require clusters to involve more than one crystal and to have an energy deposit greater than $30\,\mathrm{MeV}/c^2$. An additional background source arises from events in which a large amount of energy is deposited in the ECL. Due to the decay time of ECL signals, a high-energy signal may persist into the next event, mimicking a photon. We require the cluster time to be within 200 ns of the estimated event time to reduce this background. We use a boosted decision-tree (BDT) originally trained for $B^0 \to \pi^0\pi^0$ decays at Belle II to suppress background from misreconstructed photons. This BDT, which employs 10 cluster shape variables, is applied to our data with a threshold requirement of output greater than 0.2. The angle between the photon momenta must be smaller than 0.4 radians, and the diphoton mass must be between 115 and 150 MeV/$c^2$. This range corresponds to roughly 2.5 times the diphoton mass resolution.

We form signal $B$ candidates by combining a $K_S^0$ or $\pi^0$ with a $K^+$ or $\pi^+$, depending on the reconstructed decay chain. We perform a vertex fit using `TreeFit` of each $B$ candidate, constraining the $B$ to originate from the interaction point and the $K_S^0$ and $\pi^0$ to their known masses. To select signal $B$ candidates, we define two variables. The beam-constrained mass $M_{\rm bc} \equiv \sqrt{(E_{\rm beam}/c^2)^2 - (|\vec{p}|/c)^2}$ and the energy difference $\Delta E \equiv E - E_{\rm beam}$, where the beam energy $E_{\rm beam}$, the $B$ momentum $\vec{p}$ and its energy $E$ are in the center-of-mass frame. The $M_{\rm bc}$ distribution of correctly reconstructed signal events peaks at the $B$ mass and their $\Delta E$ distribution at zero. For misreconstructed $B$ events, where one of the charged pions is identified as a kaon or vice versa, the distribution peaks at the $B$ mass in $M_{\rm bc}$ but is shifted by 30–50 MeV in $\Delta E$. For signal $B$ candidates, $M_{\rm bc}$ must be in the range $[5.272, 5.288]$ GeV/$c^2$ and $\Delta E$ in the range $[-0.3, 0.3]$ GeV.

### 4.3.2 Tag *B* Candidate Reconstruction and Selection

Next to information from the signal candidate $B$, the continuum suppression explained in Section 4.3.3 uses also information from the other $B$ in the event. For this purpose, it is necessary to reconstruct also the latter. In this reconstruction, all tracks not used for the signal $B$ candidate reconstruction are combined. These tracks must have a distance of closest approach to the interaction point of less than 2.0 cm along the $z$ axis, less than 0.5 cm in the transverse plane, at least 1 measurement point in the CDC and the SVD and a momentum above 50 MeV/$c$. After the reconstruction, the $B$ vertex is fitted using the `kFit` algorithm. In the fit, the tag $B$ vertex is constrained to lie within a tube defined by the overlap of the interaction point and the signal $B$ vertex projected towards the tag $B$ flight direction. The flavor tagger is used to determine the flavor of the $B$.

### 4.3.3 Continuum Suppression

After reconstructing and selecting signal, the samples are predominantly composed of background from $e^+e^- \to q\bar{q}$, where $q = u, d, c, s$, known as continuum background. In $e^+e^-$ collisions at the center-of-mass energy of the $\Upsilon(4S)$ resonance, less than 25% of the hadronic cross section is due to the production of the resonance itself. The remaining 75% or more is attributed to continuum background as shown in Fig. 4.1. In these events, random combinations of tracks and energy clusters can mimic signal. Given the high production cross section and the low expected branching fraction for the signal, this background outnumbers signal by several orders of magnitude. To suppress continuum background, we identify several discriminating observables and train a boosted decision tree (BDT) using the *lightGBM* framework [75].

For training, we use observables related to the event shape, flavor tagger, vertex fits, and $M_{bc}$. Below, we briefly explain why these observables are able to discriminate between signal and background. Event shape observables describe the topological signature of events. In the $e^+e^-$ collisions, the center-of-mass energy is just above the $B\overline{B}$ production threshold. As a result, $B\overline{B}$ pairs are produced almost at rest in the $\Upsilon(4S)$ frame, leading to isotropically distributed decay products. Conversely, quark pairs in continuum background are produced with high momentum and hadronize in two back-to-back jets. Fig. 4.2 shows the different topological signature of a $B\overline{B}$ and a continuum background event captured by event shape observables. Belle II's flavor tagging algorithms infers the flavor of the bottom quark in a $B$ by exploiting information about its decay products [76]. This is possible since many neutral $B$ decays provide flavor-specific signatures. For example, in semileptonic $B$ decays to a final state $X^-\ell^+\nu$, where $X^-$ is a $D^-$ or $D^{*-}$ and $\ell^+$ a charged lepton, the positive charge of the lepton determines the decayed $B$ was a $\overline{B}$. Similarly, a negatively charged lepton tags a $B$. The algorithm's output also allows to distinguish continuum background, since it is much less likely such background contains leptons with the same kinematic properties as those from $B$ decays. Observables related to the vertex fits can help to discriminate signal from continuum background, as reconstructed $B$s have a well-defined vertex and a specific flight length. In contrast, continuum background often results in a poorly fitted vertex as it arises from a random combination of tracks and has no distinct vertex. Finally, $M_{bc}$ is suitable to differentiate signal and continuum background as signal events peak at the $B$ mass, whereas continuum background events exhibit a flat distribution.

We reconstruct the $B^+ \to \overline{D}(\to K^+\pi^-\pi^0)\pi^+$ decay channel to check the data-simulation agreement for each observable. In addition, we check the correlation between each observable in $\Delta E$ and exclude observables which exhibit a high correlation to circumvent a correlation between the BDT output and $\Delta E$. Both are fit variables in the fit of sample composition and a correlation would make the fit modeling much more complicated. In total, we converge on 48 observables for BDT training. The variables used in training are:

- **foxWolframR1** – ratio of the first to the zeroth fox wolfram moment [77] (1 variable)
- **KSFWVariablesXX** – modified Super Fox Wolfram moments also called Kakuno's super Fox Wolfram moments [78] (11 variables)
- **CleoConeCSXX** – the first to ninth Cleo cone [79] calculated only from tracks and clusters used for the tag $B$ reconstruction (9 variables)
- **thrustAxisCosTheta** – cosine of the polar angle of the thrust axis [80] (1 variable)
- **cosTBz** – cosine of the angle between the thrust axis of the signal and tag $B$ (1 variable)
- **cosTBTO** – cosine of the angle between the thrust axis of the signal $B$ and the $z$-axis (1 variable)
- **CMS_cosTheta** – cosine of the signal $B$ polar angle in the center-of-mass frame (1 variable)
- **FBDT_qrCombined** – product of the flavor of the tag $B$ times probability of the flavor being correctly assigned as estimated by the flavor tagger (1 variable)
- **qpSlowPion** – product of the charge of the track with highest probability of being a slow pion and the corresponding probability as estimated by the flavor tagger (1 variable)

- **qpKinLepton** – product of the charge of the track with highest probability of being a primary lepton and the corresponding probability as estimated by the flavor tagger (1 variable)
- **TagDXX** – distance between the tag $B$ vertex, its uncertainty and its significance of distance from the interaction point in $x$, $y$, and $z$ direction (9 variables)
- **DeltaXX** – the distance between the signal and tag $B$ vertex and the significance of distance in $x$, $y$, and $z$ direction as well as the uncertainty in $z$ direction (7 variables)
- **nTracks** – number of tracks in the event (1 variable)
- **TagVpVal** – the p-value of the tag $B$ vertex fit (1 variable)
- **chiProb** – the $\chi^2$ of the signal $B$ vertex fit (1 variable)
- $\boldsymbol{M_{\mathrm{bc}}}$– beam-constrained mass (1 variable)

We train two separate discriminators. One for the $B^+ \to K_S^0\pi^+$ channel and one jointly for the $B^+ \to \pi^+\pi^0$, and $B^+ \to K^+\pi^0$ channels. We use 250000 (1000000) signal and 250000 (1000000) continuum background events to train the algorithm for $B^+ \to K_S^0\pi^+$ ($B^+ \to \pi^+\pi^0$ and $B^+ \to K^+\pi^0$). Half of the signal and continuum background events are used for training and half of the events are used for testing. In Appendix A.1, the data-simulation agreement for the input variables in a control channel is shown. After training, the distribution of the BDT output in the training and test set is compared to check for overtraining. In addition, it is checked whether applying the BDT sculpts the $\Delta E$ distribution. The result of both is shown in Fig. 4.3. No significant overtraining and sculpting is observed. The feature importance is shown in Fig. 4.4. In Fig. 4.5 and Fig. 4.6, the BDT output and the transformed BDT output between data and simulation for the $B^+ \to \overline{D}(\to K^+\pi^-\pi^0)\pi^+$ control mode are compared.

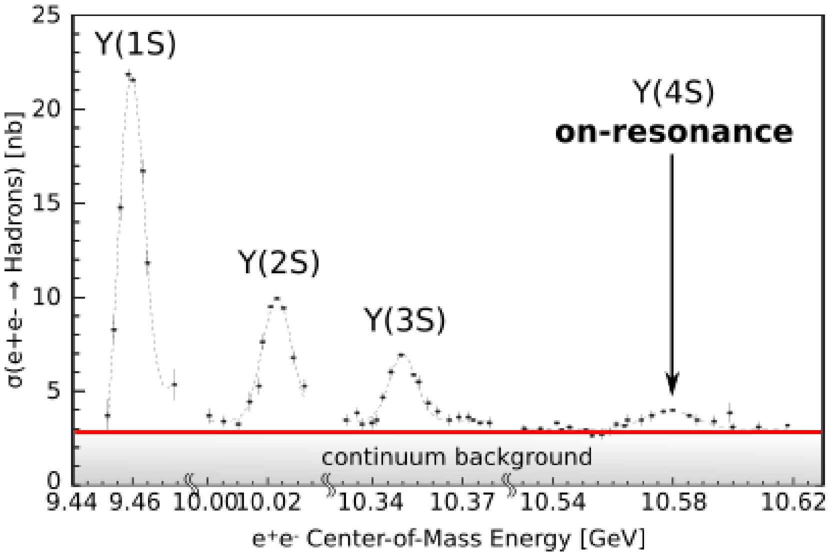


Figure 4.1: Hadron production cross section in $e^+e^-$ collisions as a function of center-of-mass energy. The $\Upsilon$ resonances are visible as prominent peaks above flat background, attributed to lighter quark production (continuum background). The continuum background level is indicated by a red horizontal line.

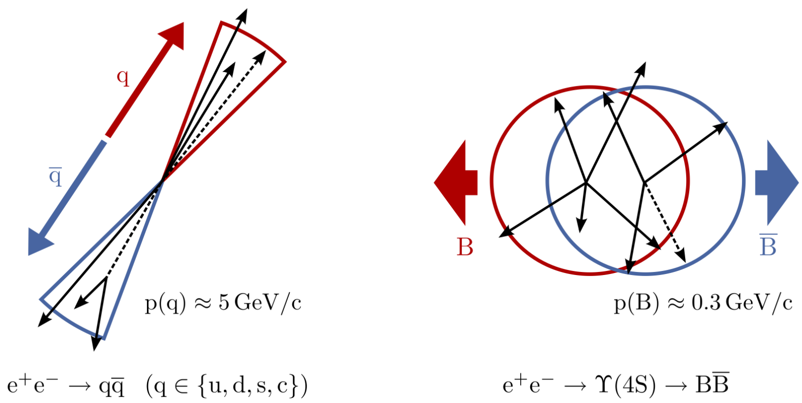


Figure 4.2: Topological signature of a continuum background (left) and a $B\overline{B}$ (right) event. The quark pair in continuum background are produced with high momentum, resulting in two back-to-back jets after hadronization. The $B$ are produced almost at rest and decay isotropically.

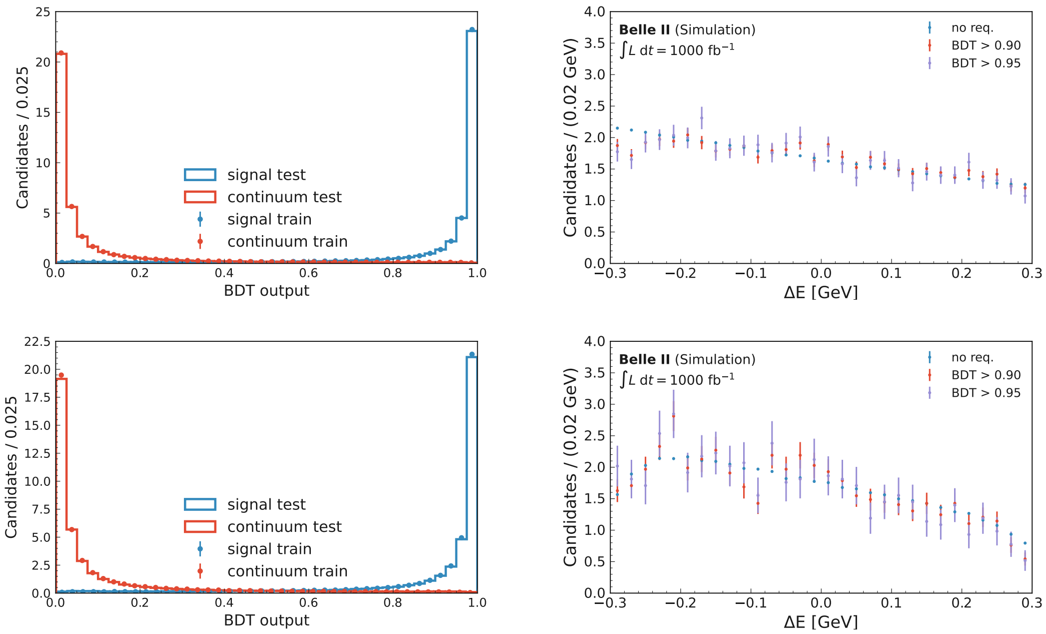


Figure 4.3: Left: Overtraining check for $B^+ \to \pi^+\pi^0$ and $B^+ \to K^+\pi^0$ (upper) and $B^+ \to K_S^0\pi^+$ (lower). Right: The continuum background $\Delta E$ distribution before and after placing a requirement on the continuum suppression output for $B^+ \to \pi^+\pi^0$ and $B^+ \to K^+\pi^0$ (upper) and $B^+ \to K_S^0\pi^+$ (lower).

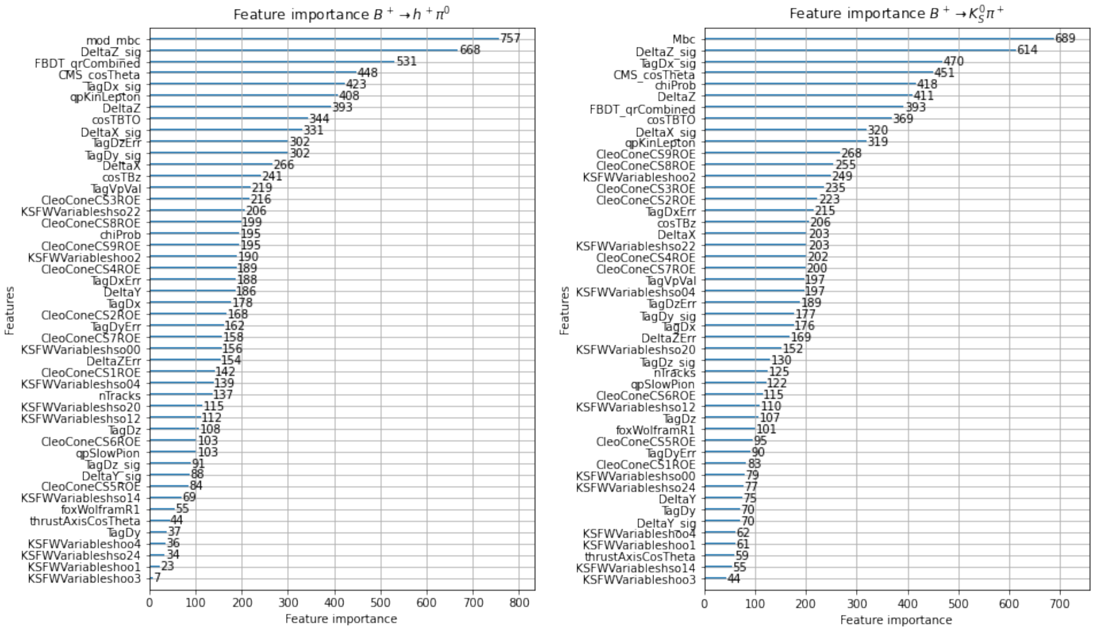


Figure 4.4: The feature importance for the continuum suppression trained for $B^+ \to \pi^+\pi^0$ and $B^+ \to K^+\pi^0$ (left) and $B^+ \to K^0_S\pi^+$ (right). The feature important has been calculated as the number of times a feature is used to split a decision tree.

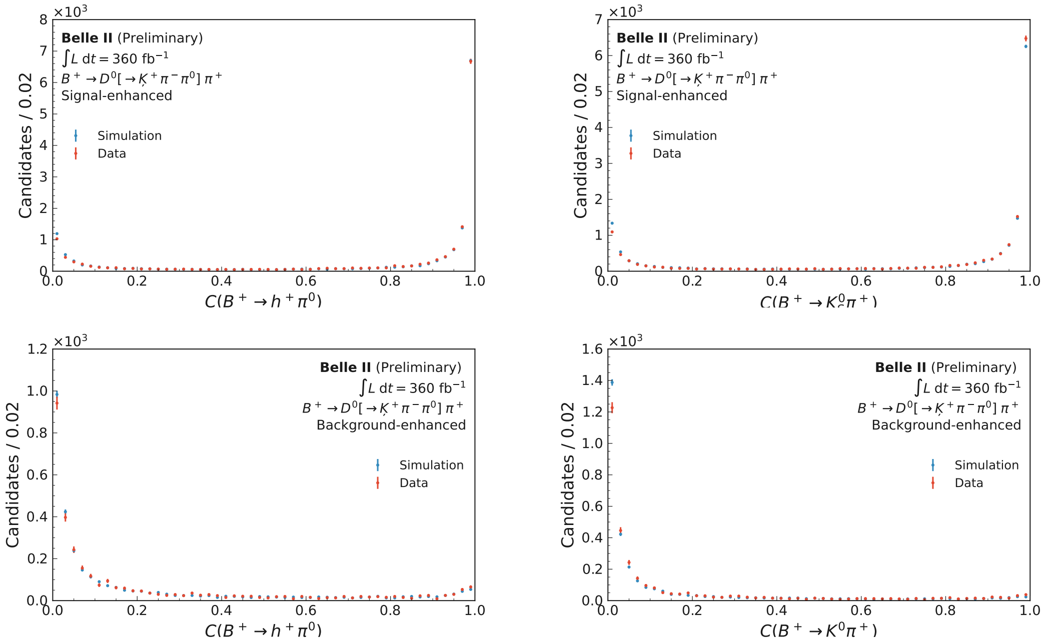


Figure 4.5: Top: Comparison of the continuum suppression output between data and simulation in the $B^+ \to \overline{D}(\to K^+\pi^-\pi^0)\pi^+$ control channel for the continuum suppression trained for $B^+ \to \pi^+\pi^0$ and $B^+ \to K^+\pi^0$ (left) and the continuum suppression trained for $B^+ \to K_S^0\pi^+$ (right). The distributions are shown in the signal-enhanced region defined as $-0.1 < \Delta E < 0.1$. Bottom: Same as before but in the background-enhanced region defined as $\Delta E > 0.1\,\mathrm{GeV}/c^2$.

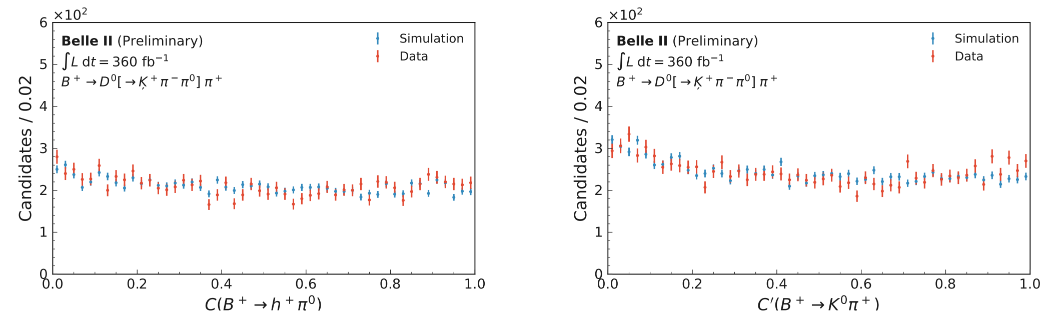


Figure 4.6: Comparison of the data and simulation distribution of the transformed continuum suppression output in the $B^+ \to \overline{D}(\to K^+\pi^-\pi^0)\pi^+$ control channel for the continuum suppression trained for $B^+ \to \pi^+\pi^0$ and $B^+ \to K^+\pi^0$ (left) and the continuum suppression trained for $B^+ \to K_S^0\pi^+$ (right).

### 4.3.4 Final Sample Composition

The $\Delta E$ and transformed continuum suppression output distribution after the full selection are shown in Fig. 4.7.

In all samples four components are visible:

- **Signal** from the decay we do want to measure. This components peaks at 0 in $\Delta E$ and by construction follows a uniform distribution in the transformed continuum suppression output.
- **Peaking background** from decays where a pion has been misidentified as a kaon or vice versa and otherwise the decay has been correctly reconstructed. This component peaks in $\Delta E$, hence the name, but is shifted from 0 by 30–50 MeV due to the wrong mass hypothesis used in the reconstruction. In the transformed continuum suppression output this component follows a linear function.
- $\boldsymbol{B\overline{B}}$ **background** from $B$ decays which are not signal or peaking background. This component clusters at lower $\Delta E$ values since usually the $B$ decay is not fully reconstructed and a track or cluster is missed. In the transformed continuum suppression output this component follows a linear function.
- **Continuum background** from $e^+e^- \to q\overline{q}$, where $q = u, d, c, s$. The $\Delta E$ distribution of this component follows a linear function and the transformed continuum suppression output distribution an exponential.

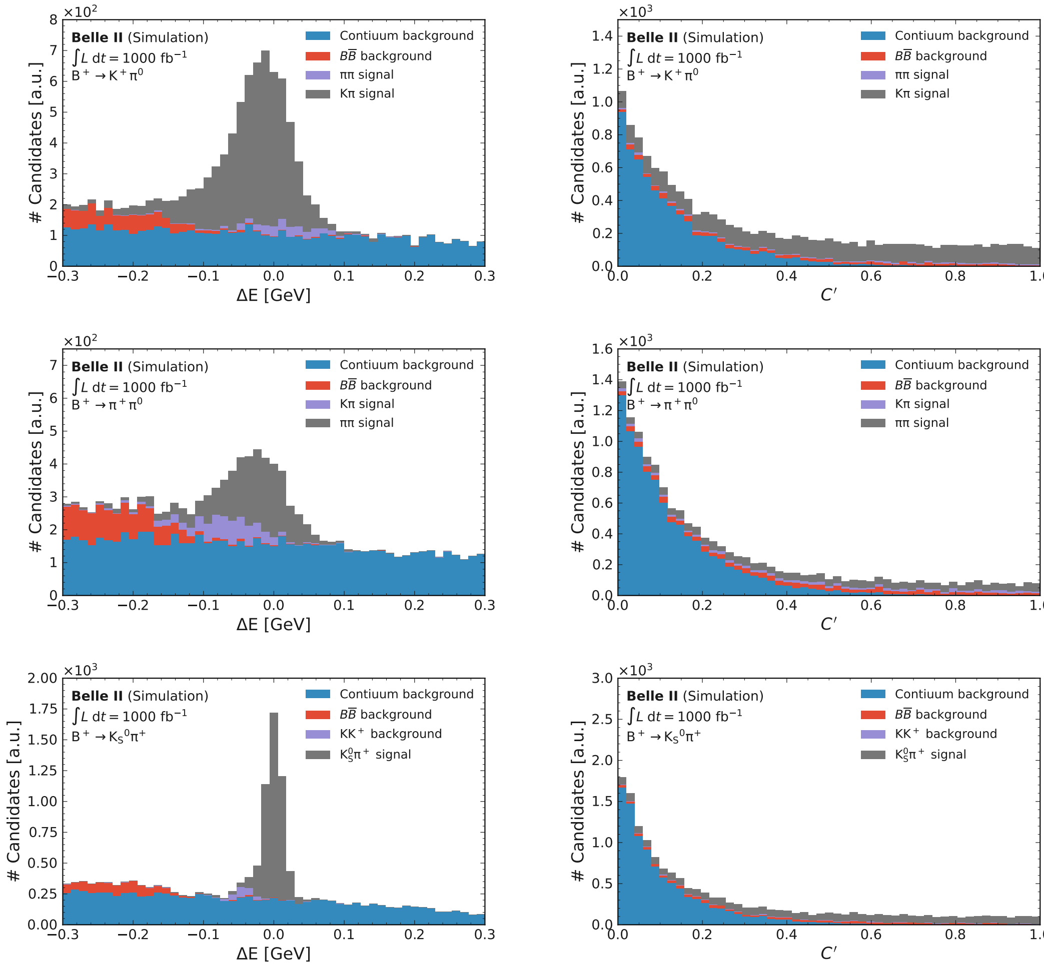


Figure 4.7: The $\Delta E$ (left) and transformed continuum suppression output $C'$ (right) distribution for for the final $B^+ \to K^+\pi^0$ (top), $B^+ \to \pi^+\pi^0$ (center), and $B^+ \to K_S^0\pi^+$ (bottom) sample.

## 4.4 Fit of Sample Composition

After the selection described in Section 4.3, the reconstructed samples are expected to still contain background events. To determine the physics parameter and background contribution, we perform fits of the sample composition using the maximum likelihood method. We use `zfit` [81] to construct the likelihood function and `Minuit` [82] for its minimization.

We perform two extended fits to the unbinned energy-difference $\Delta E$ and transformed CS classifier $C'$ distributions: One simultaneous to the samples of $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$ decays and one to the sample of $B^+ \to K_S^0\pi^+$ decays. To measure $\mathcal{A}_{\mathrm{CP}}$, all samples are split according to the charge of the $B$ candidate. In these fits, we consider three components: signal, continuum background, peaking background, and background from other $B$ decays. For more details on each component, see Section 4.3.

We obtain the shape for each component empirically based on pure simulated samples of these components. We obtain the two dimensional shapes describing the $\Delta E$ and $C'$ distributions as product of the one dimensional shapes, i.e., it is assumed all fit components can be fully factorized. This assumption is valid if correlations between $\Delta E$ and $C'$ can be neglected. The correlations are shown in Table 4.1. We model the $\Delta E$ distribution of the signal component using the sum a double-sided Crystal ball and a Gaussian, allowing the fraction between both components to vary. The Crystal ball function consists of a Gaussian core component and a power-law tail on the left side [83]. The double-sided Crystal ball has an additional power-law tail on the right side. The $C'$ shape of the signal is described using a linear function. The peaking background component is described using the same shapes as the signal with the the exception of the $\Delta E$ distribution for $B^+ \to K_S^0\pi^+$ which is modeled using the sum of a Crystal ball function and a Gaussian. The $B\overline{B}$ background is modeled using a shape obtained via kernel density estimation [84] to describe $\Delta E$ and a linear function for $C'$. The continuum background is modeled using a linear function to describe $\Delta E$ and an exponential function defined as $f(x) = \exp(ax + bx^2)$, where $a$ and $b$ are free parameter, for $C'$. In Table 4.2, all shapes are summarized and in Figs. 4.8 to 4.11, a comparison between the shapes and simulated events are shown. Differences are observed for the $\Delta E$ distribution for some of the signal and peaking background components. These differences are deemed acceptable, as the simulated samples for those components are three orders of magnitude larger than what is expected to be observed in data.

We fit directly for the physics parameter, i.e., the branching fractions and direct CP asymmetries. The respective signal yields can be computed from the physics parameter in the following way,

$$N^{\pm} = 2N_{B^{+}B^{-}}\, \epsilon\, \mathcal{B}_{\rm sub}\, \mathcal{B}_{\rm sig}\, \frac{1 \mp \mathcal{A}}{2}, \tag{4.1}$$

where $N^{\pm}$ is the respective signal yield for positively/negatively charged signal $B$ candidates, $N_{B^{+}B^{-}}$ is the number of produced $B^{+}B^{-}$ pairs (centrally provided), $\epsilon$ is the reconstruction efficiency (estimated from simulation), $\mathcal{B}_{\rm sub}$ are the sub branching fractions of decays happening in the reconstruction chain (taken from PDG), in the case of $B^{+} \to K^{+}\pi^{0}$ and $B^{+} \to \pi^{+}\pi^{0}$ this is $\pi^{0} \to \gamma\gamma$ and for $B^{+} \to K^{0}\pi^{+}$ it it includes the fraction of $K_{S}^{0}$ in $K^{0}$ and the branching fraction of $K_{S}^{0} \to \pi^{+}\pi^{-}$, $\mathcal{B}_{\rm sig}$ is the signal branching fractions (fit parameter) and $\mathcal{A}$ is the direct CP asymmetry (fit parameter). In the simultaneous fit to the samples of $B^{+} \to K^{+}\pi^{0}$ and $B^{+} \to \pi^{+}\pi^{0}$ decays, we also calculate the yield of the peaking background using this formula. In the $B^{+} \to K^{+}\pi^{0}$ sample, the peaking background consists of $B^{+} \to \pi^{+}\pi^{0}$ decays, where the $\pi^{+}$ was misidentified as a Kaon, and vice versa for the $B^{+} \to \pi^{+}\pi^{0}$ sample. By also calculating the peaking background this way, we constrain the fraction of peaking background and signal and determine the amount of peaking background in a data-driven way. For other components, we let the yields free floating separate for each charge. In addition, we let the continuum background shape parameter free to float in the fit, as the distributions are known to be not well reproduced in simulation.

For $B^{+} \to K^{+}\pi^{0}$ and $B^{+} \to \pi^{+}\pi^{0}$, we end up with 18 fit parameter. They are:

- The branching fractions of $B^{+} \to K^{+}\pi^{0}$ and $B^{+} \to \pi^{+}\pi^{0}$ (2 parameters)
- The direct CP asymmetries of $B^{+} \to K^{+}\pi^{0}$ and $B^{+} \to \pi^{+}\pi^{0}$ (2 parameters)
- The continuum background, $\Delta E$ and $C'$ shape parameter (6 parameter)

- The continuum background yields separate for each sample and charge (4 parameters)
- The $BB$ yields separate for each sample and charge (4 parameters)

For $B^+ \to K_S^0\pi^+$, we end up with 11 fit parameter. They are:

- The branching fractions of $B^+ \to K_S^0\pi^+$ (1 parameter)
- The direct CP asymmetries of $B^+ \to K_S^0\pi^+$ (1 parameter)
- The continuum background $\Delta E$ and $C'$ shape parameter (3 parameter)
- The peaking background yields separate for each charge (2 parameter)
- The continuum background yields separate for each charge (2 parameter)
- The $BB$ yields separate for each charge (2 parameter)

Table 4.1: Correlations coefficient between $\Delta E$ and the transformed continuum suppression output for the different samples and fit components.

| | signal | peaking bkg | $B\overline{B}$ bkg | cont. bkg |
|---|---|---|---|---|
| $B^+ \to \pi^+\pi^0$ | 0.03 | 0.03 | 0.06 | 0.01 |
| $B^+ \to K^+\pi^0$ | 0.03 | 0.03 | 0.07 | 0.01 |
| $B^+ \to K_S^0\pi^+$ | 0.02 | 0.01 | 0.08 | 0.02 |

Table 4.2: Summary of fit models for $B^+ \to h^+\pi^0$ and $B^+ \to K_S^0\pi^+$. CB denotes a Crystal Ball function, DCB a double-sided Crystall Ball function, and KDE a kernel density estimator.

Fit shapes for $B^+ \to \pi^+\pi^0$ and $B^+ \to K^+\pi^0$.

| | $\Delta E$ | $C'$ |
|---|---|---|
| Signal | DCB + Gaussian | 1st degree polynomial |
| peaking bkg | DCB + Gaussian | 1st degree polynomial |
| $B\overline{B}$ bkg | KDE | 1st degree polynomial |
| cont. bkg | 1st degree polynomial | $\exp(ax + bx^2)$ |

Fit shapes for $B^+ \to K_S^0\pi^+$.

| | $\Delta E$ | $C'$ |
|---|---|---|
| $B^+ \to K_S^0\pi^+$ signal | DCB + Gaussian | 1st degree polynomial |
| peaking bkg | CB + Gaussian | 1st degree polynomial |
| $B\overline{B}$ bkg | KDE | 1st degree polynomial |
| cont. bkg | 1st degree polynomial | $\exp(ax + bx^2)$ |

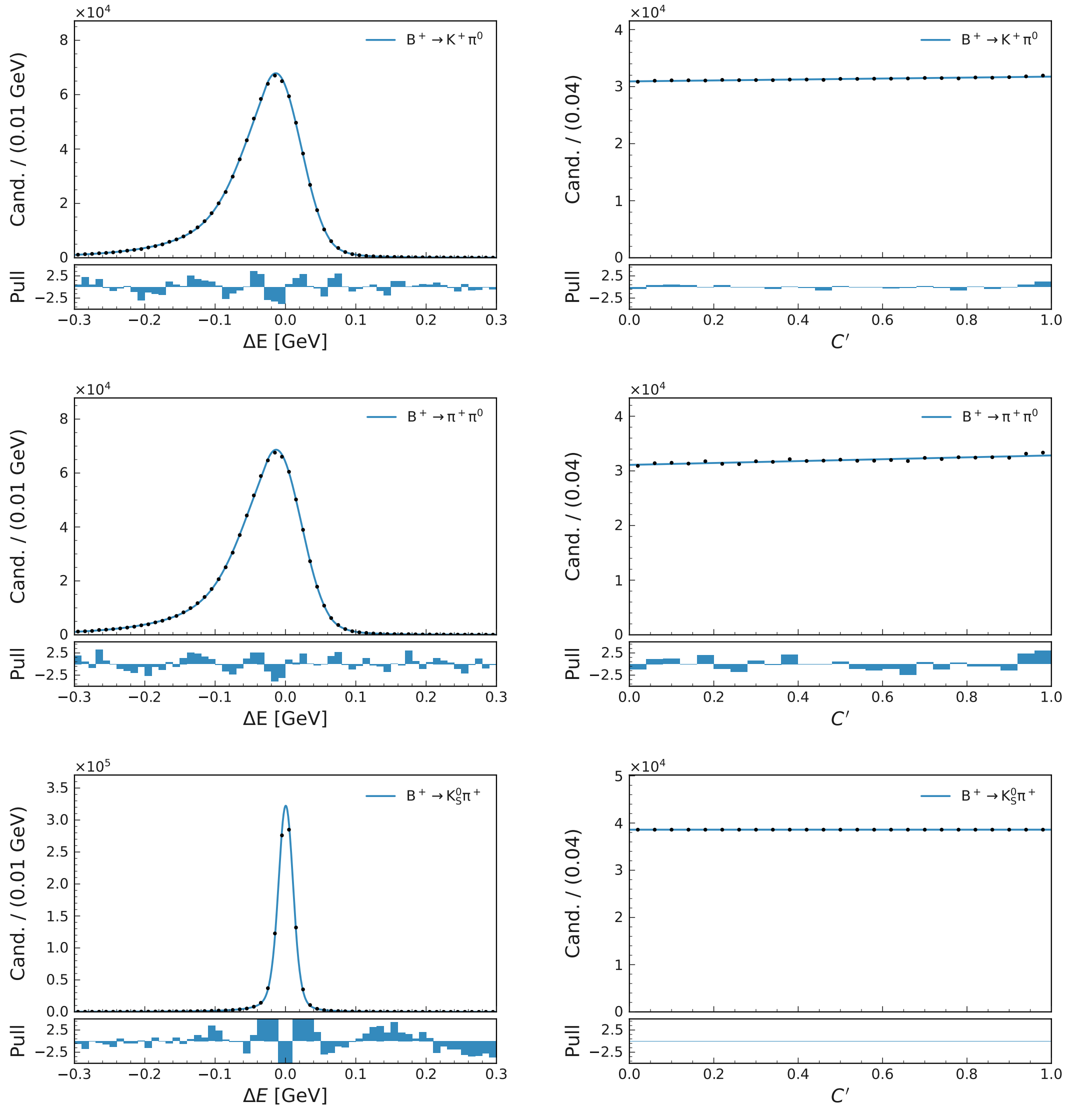


Figure 4.8: Left: The $\Delta E$ distribution (black dots with errorbars) for signal in the $B^+ \to K^+\pi^0$ (top), $B^+ \to \pi^+\pi^0$ (center), and $B^+ \to K^0\pi^+$ (bottom) sample. The result of a fit is shown as a solid blue line. Below differences between simulated data and the fit results, normalized by fit uncertainties, (pulls) are shown. Right: Same but for the transformed continuum suppression output $C'$.

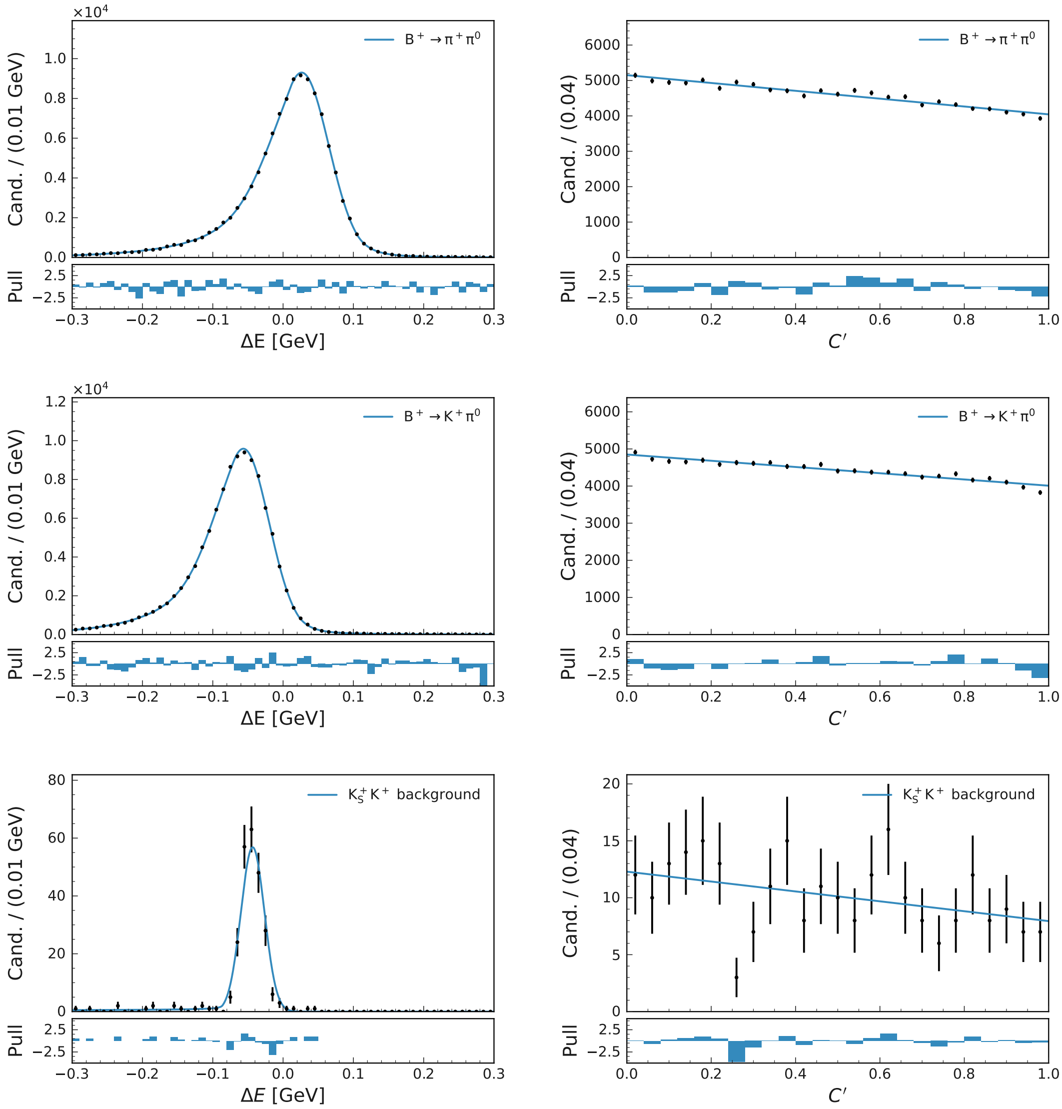


Figure 4.9: Left: The $\Delta E$ distribution (black dots with errorbars) for peaking background in the $B^+ \to K^+\pi^0$ (top), $B^+ \to \pi^+\pi^0$ (center), and $B^+ \to K^0\pi^+$ (bottom) sample. The result of a fit is shown as a solid blue line. Below differences between simulated data and the fit results, normalized by fit uncertainties, (pulls) are shown. Right: Same but for the transformed continuum suppression output $C'$.

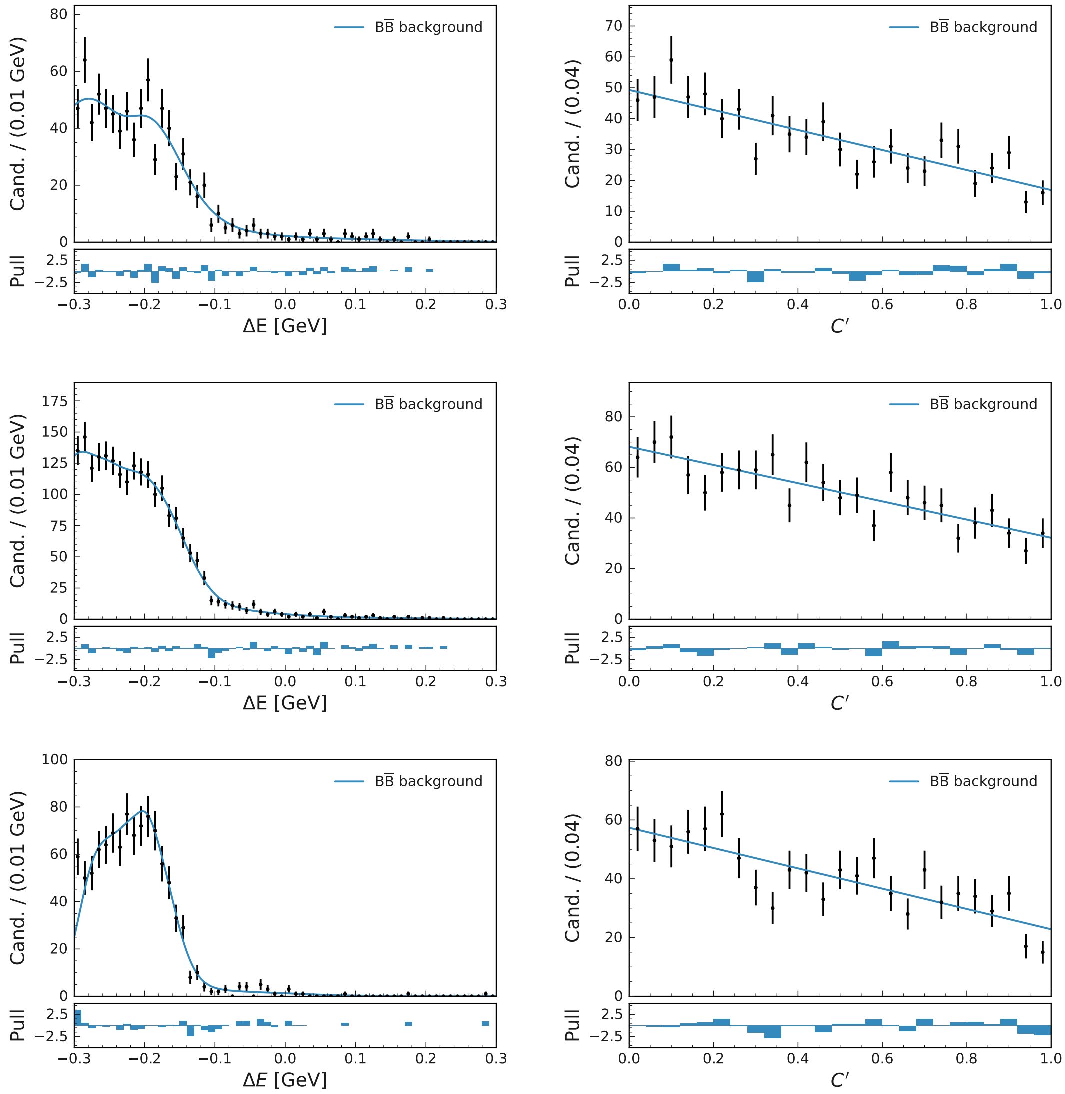


Figure 4.10: Left: The $\Delta E$ distribution (black dots with errorbars) for $B\overline{B}$ background in the $B^+ \to K^+\pi^0$ (top), $B^+ \to \pi^+\pi^0$ (center), and $B^+ \to K^0\pi^+$ (bottom) sample. The result of a fit is shown as a solid blue line. Below differences between simulated data and the fit results, normalized by fit uncertainties, (pulls) are shown. Right: Same but for the transformed continuum suppression output $C'$.

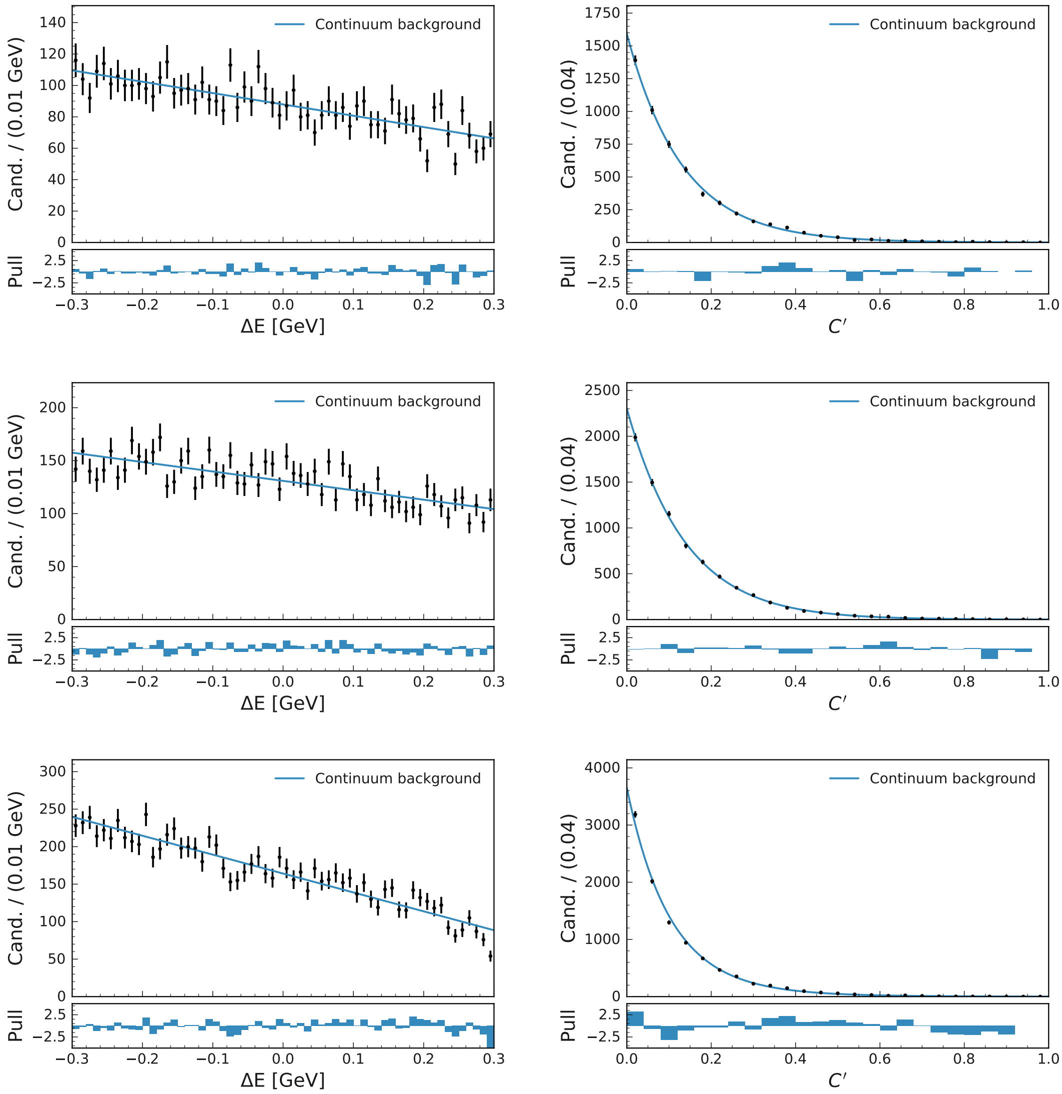


Figure 4.11: Left: The $\Delta E$ distribution (black dots with errorbars) for continuum background in the $B^+ \to K^+\pi^0$ (top), $B^+ \to \pi^+\pi^0$ (center), and $B^+ \to K^0\pi^+$ (bottom) sample. The result of a fit is shown as a solid blue line. Below differences between simulated data and the fit results, normalized by fit uncertainties, (pulls) are shown. Right: Same but for the transformed continuum suppression output $C'$.

### 4.4.1 Estimator Properties

Due to several reasons, for example, over simplification of the fit model the maximum likelihood fits might exhibit biases or non Gaussian uncertainties. In this section, we study the properties of the estimators to identify possible biases and to estimate the statistical precision expected on real data.

#### Estimation Using Large Simulated Sample

We perform a fit to a large sample of simulated data corresponding to $1000\,\mathrm{fb}^{-1}$, roughly three times size of the data sample. Due to the larger size of the sample effects due to mismodeling of the fit shapes, for example due to neglect of correlations between $\Delta E$ and $C'$ are enhanced compared to data. The simulated sample contains a signal and background composition similar to what we expect to observe in data. We run the fit on the simulation sample as it is run on data.

We show the result of a fit to the simulation samples for $B^+ \to \pi^+\pi^0$ and $B^+ \to K^+\pi^0$ in Fig. 4.12. In addition, we show signal- and background-enhanced fit projections in Fig. 4.13 and Fig. 4.14. Similar fit projections for a fit to the $B^+ \to K_S^0\pi^+$ sample are shown in Fig. 4.15. The agreement between the fit projections and the data is good in the nominal plot as well as in the signal- and background-enhanced plots, reassuring that the overall fit modeling as well as the modeling of individual components is working well. The fitted parameters agree also well with the simulation expectation as shown in Tables 4.3 and 4.4. In Tables A.1 and A.2, we show the correlation between fit parameters.

#### Estimation Using Ensemble Tests

We use ensemble tests to determine our estimators properties. A commonly used method for estimation is analyzing the standardized residual, also known as pull distribution. For a set of ensembles generated with parameter values that vary according to a Poisson distribution, the pull distribution is defined as the difference between the observed fit result and the parameter value used in the generation process, divided by the estimated uncertainty. For an unbiased estimator with Gaussian uncertainties, the pull distribution follows a Gaussian with a mean of

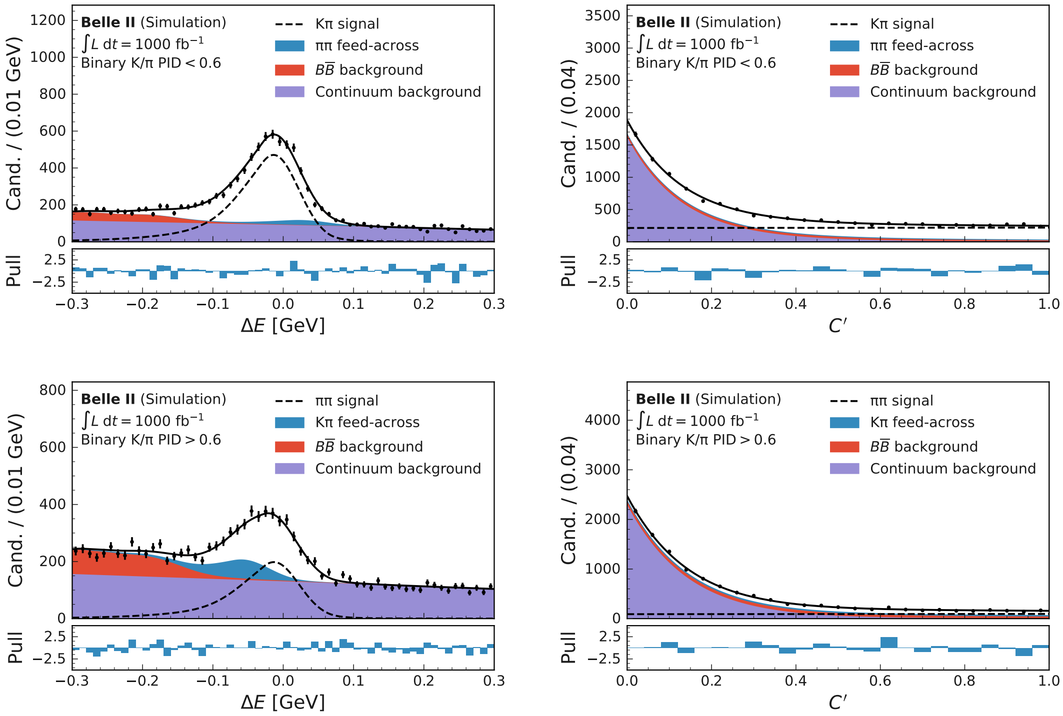


Figure 4.12: Top: The $\Delta E$ distribution (left) and transformed continuum suppression output (right) in the simulation sample corresponding to $1000\,\mathrm{fb}^{-1}$ of data for $B^+ \to K^+\pi^0$. The result of a fit to the sample is shown as a solid black line. Individual fit components are shown as black dashed line (signal), blue shaded area (peaking background), red shaded area ($B\overline{B}$ background) and purple shaded area (continuum background). Below differences between simulated data and the fit results, normalized by fit uncertainties, (pulls) are shown. Bottom: Same but for $B^+ \to \pi^+\pi^0$.

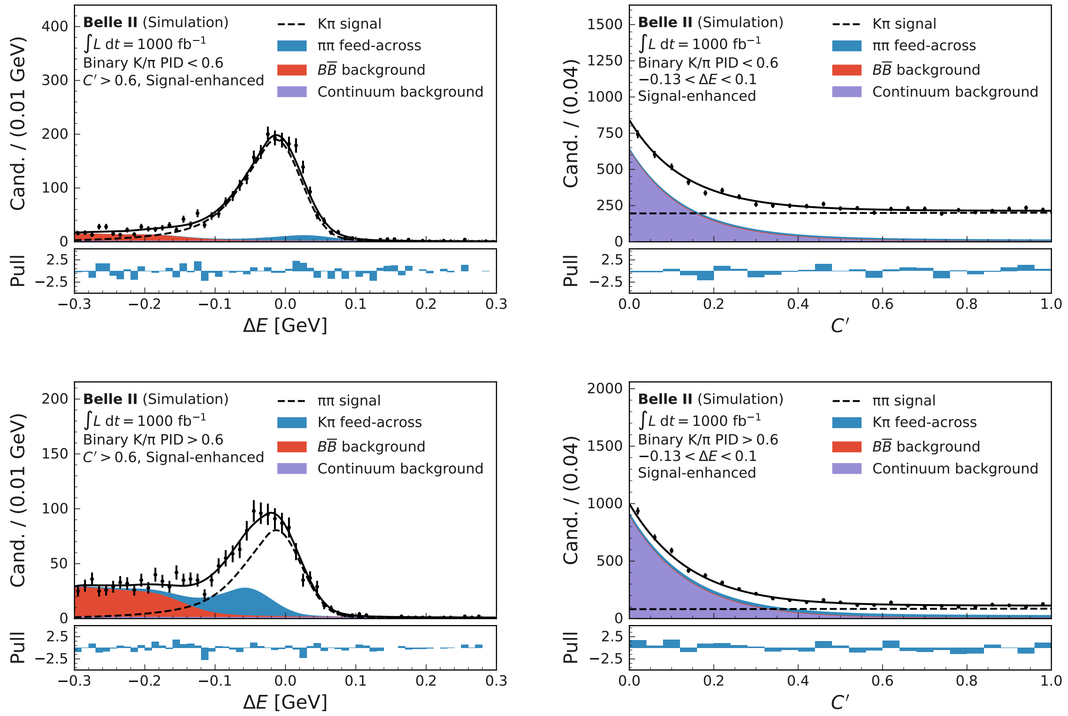


Figure 4.13: Top: The $\Delta E$ distribution (left) and transformed continuum suppression output (right) in the simulation sample corresponding to $1000\,\mathrm{fb}^{-1}$ of data for $B^+ \to K^+\pi^0$. The result of a fit to the sample is shown as a solid black line. Individual fit components are shown as black dashed line (signal), blue shaded area (peaking background), red shaded area ($B\overline{B}$ background) and purple shaded area (continuum background). Below differences between simulated data and the fit results, normalized by fit uncertainties, (pulls) are shown. Bottom: Same but for $B^+ \to \pi^+\pi^0$. The distributions are shown in signal-enhanced projection. The signal-enhanced region is defined as $-0.13 < \Delta E < 0.10\,\mathrm{GeV}/c^2$ and $C' > 0.6$ if the respective variable is not plotted.

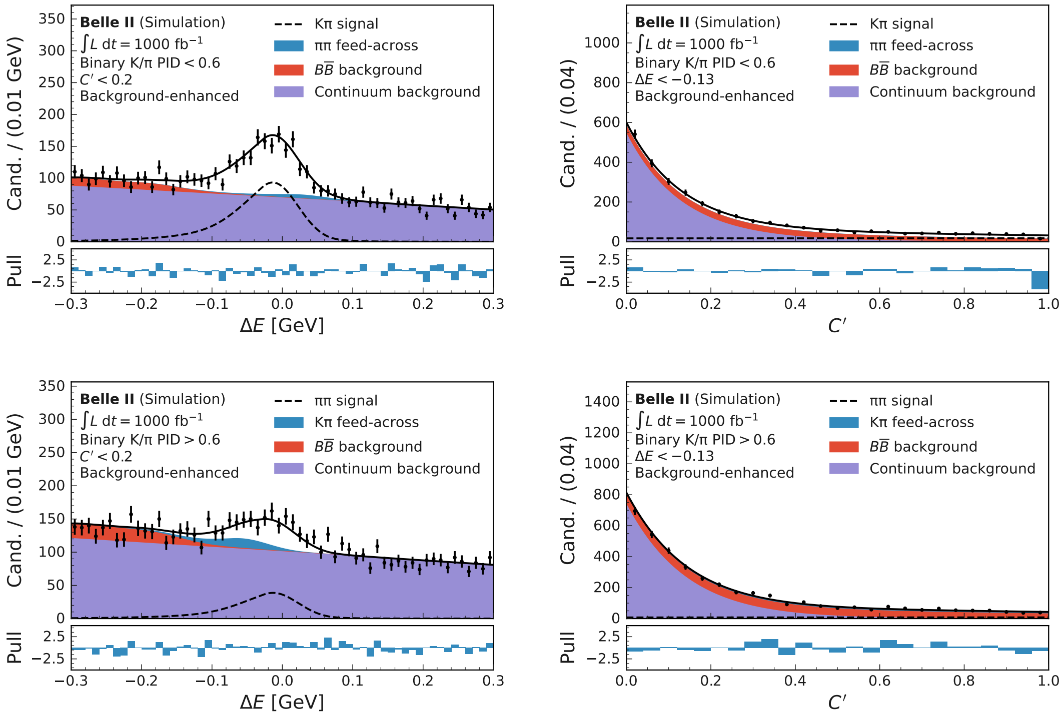


Figure 4.14: Top: The $\Delta E$ distribution (left) and transformed continuum suppression output (right) in the simulation sample corresponding to $1000\,\mathrm{fb}^{-1}$ of data for $B^+ \to K^+\pi^0$. The result of a fit to the sample is shown as a solid black line. Individual fit components are shown as black dashed line (signal), blue shaded area (peaking background), red shaded area ($B\bar{B}$ background) and purple shaded area (continuum background). Below differences between simulated data and the fit results, normalized by fit uncertainties, (pulls) are shown. Bottom: Same but for $B^+ \to \pi^+\pi^0$. The distributions are shown in background-enhanced projection. The background-enhanced region is defined as $\Delta E < -0.13\,\mathrm{GeV}/c^2$ and $C' < 0.2$ if the respective variable is not plotted.

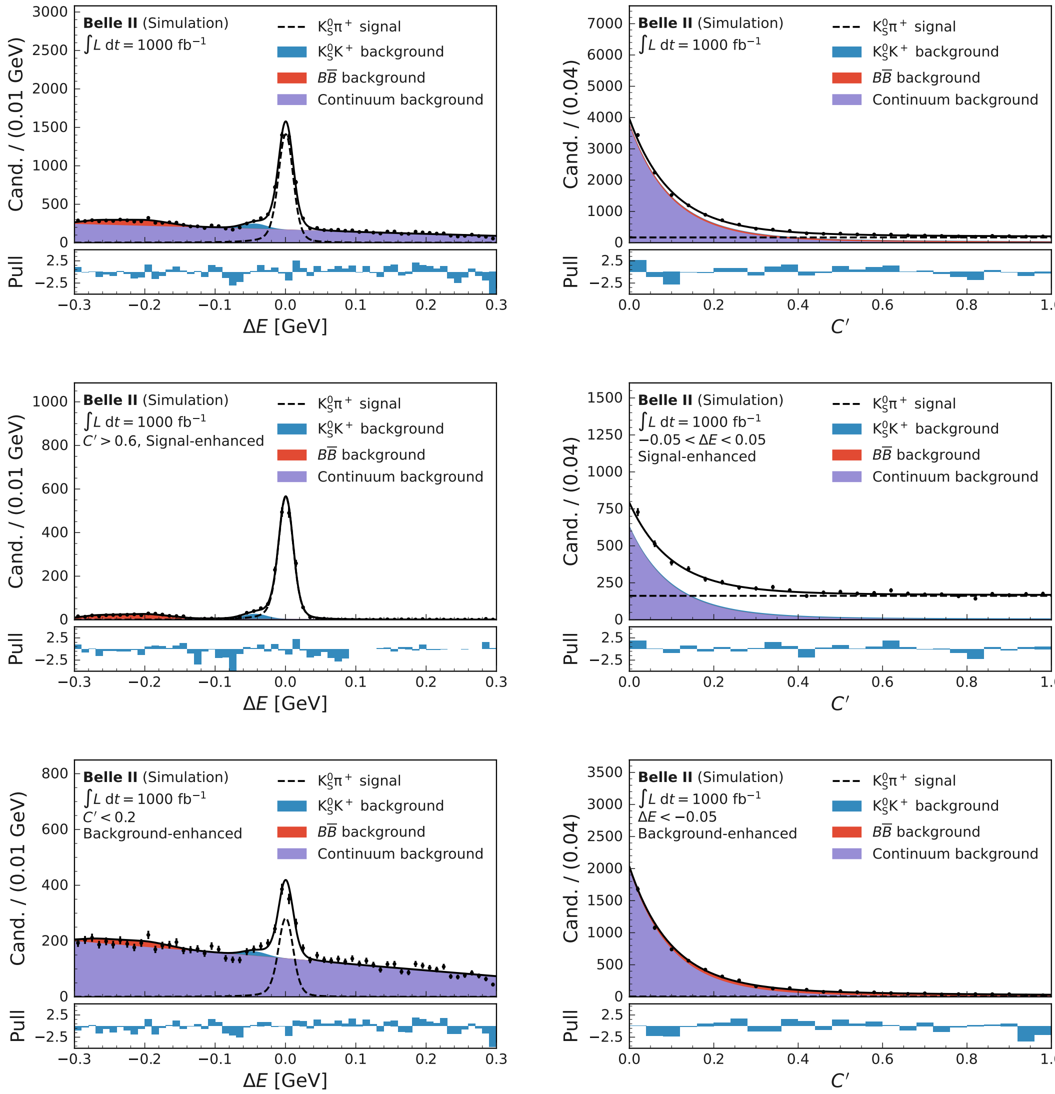


Figure 4.15: Top: The $\Delta E$ distribution (left) and transformed continuum suppression output (right) in the simulation sample corresponding to $1000\,\text{fb}^{-1}$ of data for $B^+ \to K_S^0\pi^+$. The result of a fit to the sample is shown as a solid black line. The fit components are shown as black dashed line (signal), blue shaded area (peaking background), red shaded area ($B\overline{B}$ background) and purple shaded area (continuum background). Below differences between simulated data and the fit results, normalized by fit uncertainties, (pulls) are shown. Middle: Same but in signal-enhanced projection. The signal-enhanced region is defined as $-0.05 < \Delta E < 0.05\,\text{GeV}/c^2$ and $C' > 0.6$ if the respective variable is not plotted. Bottom: Same but in background-enhanced projection. The background-enhanced region is defined as $\Delta E < 0.05\,\text{GeV}/c^2$ and $C' < 0.6$ if the respective variable is not plotted.

Table 4.3: Fit result for the fit to the $1000\,\mathrm{fb}^{-1}$ simulation sample for $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$.

| Parameter | Fit Result | Simulation Truth |
|---|---|---|
| $\mathcal{B}_{K^+\pi^0}$ | $(12.85 \pm 0.22) \times 10^{-6}$ | $12.71 \times 10^{-6}$ |
| $\mathcal{B}_{\pi^+\pi^0}$ | $(5.34 \pm 0.17) \times 10^{-6}$ | $5.43 \times 10^{-6}$ |
| $\mathcal{A}_{K^+\pi^0}$ | -0.020 $\pm$ 0.016 | -0.009 |
| $\mathcal{A}_{\pi^+\pi^0}$ | 0.016 $\pm$ 0.030 | 0.006 |
| $B\overline{B}^+_{K^+\pi^0}$ yield | 383 $\pm$ 41 | 409 |
| $B\overline{B}^-_{K^+\pi^0}$ yield | 369 $\pm$ 41 | 418 |
| $B\overline{B}^+_{\pi^+\pi^0}$ yield | 680 $\pm$ 43 | 620 |
| $B\overline{B}^-_{\pi^+\pi^0}$ yield | 653 $\pm$ 43 | 634 |
| cont. $\mathrm{bkg}^+_{K^+\pi^0}$ yield | 2652 $\pm$ 69 | 2666 |
| cont. $\mathrm{bkg}^-_{K^+\pi^0}$ yield | 2702 $\pm$ 70 | 2609 |
| cont. $\mathrm{bkg}^+_{\pi^+\pi^0}$ yield | 3884 $\pm$ 76 | 3937 |
| cont. $\mathrm{bkg}^-_{\pi^+\pi^0}$ yield | 3865 $\pm$ 76 | 3918 |
| cont. $\mathrm{bkg}_{K^+\pi^0}$ $\Delta E$ 1st param | -0.2683 $\pm$ 0.028 | - |
| cont. $\mathrm{bkg}_{K^+\pi^0}$ $C$ a | -7.65 $\pm$ 0.36 | - |
| cont. $\mathrm{bkg}_{K^+\pi^0}$ $C$ b | 0.52 $\pm$ 0.80 | - |
| cont. $\mathrm{bkg}_{\pi^+\pi^0}$ $\Delta E$ 1st param | -0.1976 $\pm$ 0.023 | - |
| cont. $\mathrm{bkg}_{\pi^+\pi^0}$ $C$ a | -7.03 $\pm$ 0.29 | - |
| cont. $\mathrm{bkg}_{\pi^+\pi^0}$ $C$ b | -1.20 $\pm$ 0.65 | - |

Table 4.4: Fit result for the fit to the $1000\,\mathrm{fb}^{-1}$ simulation sample for $B^+ \to K^0_S\pi^+$.

| Parameter | Fit Result | Simulation Truth |
|---|---|---|
| $\mathcal{B}_{K^0_S\pi^+}$ | $(2.343 \pm 0.040) \times 10^{-5}$ | $2.352 \times 10^{-5}$ |
| $\mathcal{A}_{K^0_S\pi^+}$ | 0.021 $\pm$ 0.017 | 0.006 |
| peaking $\mathrm{bkg}^-$ yield | 128 $\pm$ 23 | 115 |
| peaking $\mathrm{bkg}^+$ yield | 171 $\pm$ 24 | 138 |
| $B\overline{B}^+$ yield | 520 $\pm$ 37 | 494 |
| $B\overline{B}^-$ yield | 532 $\pm$ 36 | 508 |
| cont. $\mathrm{bkg}^+$ yield | 4948 $\pm$ 81 | 4946 |
| cont. $\mathrm{bkg}^-$ yield | 4862 $\pm$ 79 | 4903 |
| cont. bkg $C'$ a | -10.07 $\pm$ 0.23 | - |
| cont. bkg $C'$ b | 2.64 $\pm$ 0.49 | - |
| cont. bkg $\Delta E$ 1st param | -0.453 $\pm$ 0.018 | - |

zero and a standard deviation of one. A deviation from a mean of zero indicates a biased estimator. Similarly, a deviation from a standard deviation of one indicates an over- or underestimation of the uncertainties. The mean value of the estimated uncertainties is an estimate of the statistical uncertainty.

In addition to pull distributions, we also use linearity tests to assess the properties of our estimators. In the linearity tests, we generate several sets of ensembles by deviating the parameter values used in the generation process from their nominal values. We then allow them to fluctuate according to a Poisson distribution around these deviated values. For each ensemble, we plot the average of the fitted parameter values with their standard deviation as uncertainty against the parameter value used in the generation process. For an unbiased fitter, a linear fit through the set of points has a slope of one and an offset of zero. Deviations from this behavior indicate a bias [85]. In our linearity tests, we vary the branching fractions from 60% to 140% of their nominal value in steps of 20% and the direct CP asymmetries from -0.150 to 0.150 in steps of 0.005. These variations correspond to more than $3\sigma$ around the current world averages.

We generate ensembles using two different methods:

1. Toy ensembles: These are created by generating samples from the shape of each fit component. They serve as preliminary tests, for example to check whether the estimators can distinguish between different fit components.

2. Secondly bootstrapped ensembles: These are generated by bootstrapping [86], i.e., sampling with replacement, from the large simulation sample. The ensembles created by bootstrapping more closely resemble the data. They are used to assess the ability of the fit models to generalize to new data with features that are missing in the fit models themselves, for example correlations.

In Fig. 4.16, Gaussian fits to the pull distributions of toy ensembles are shown. All pull distributions follow a Gaussian shape and have a mean compatible with zero and a standard deviation compatible with one, indicating no bias. This is supported by the results of linearity tests using toy ensembles (see Fig. 4.17), which also shows no bias. The statistical uncertainty on the branching fractions estimated from these toy ensembles is 2.8% for $B^+ \to K^+\pi^0$, 5.1% for $B^+ \to \pi^+\pi^0$, and 2.9% for $B^+ \to K^0_S\pi^+$. The estimated statistical uncertainty on the direct CP asymmetry is 0.0274 for $B^+ \to K^+\pi^0$, 0.0491 for $B^+ \to \pi^+\pi^0$, and 0.0287 for $B^+ \to K^0_S\pi^+$.

The results of linearity tests using bootstrapped ensembles are shown in Fig. 4.18. Small deviations from an intercept of zero for the direct CP asymmetries of $B^+ \to \pi^+\pi^0$ and $B^+ \to K^0_S\pi^+$ are visible. Since these deviations are small compared to the estimated statistical uncertainty, we assign them as a systematic uncertainty. For the branching fraction of $B^+ \to \pi^+\pi^0$ a deviation of 2.3% from an intercept of zero is visible. This value is not negligible compared to the statistical uncertainty but small compared to the systematic uncertainties, see Section 4.6. The deviation disappears when the continuum background component is sampled from the shape, while all other components remain bootstrapped (see Fig. 4.19). This indicates the continuum background shape insufficiently describing features present in the data, is the cause of this bias. We correct this bias by scaling the fit result by its value.

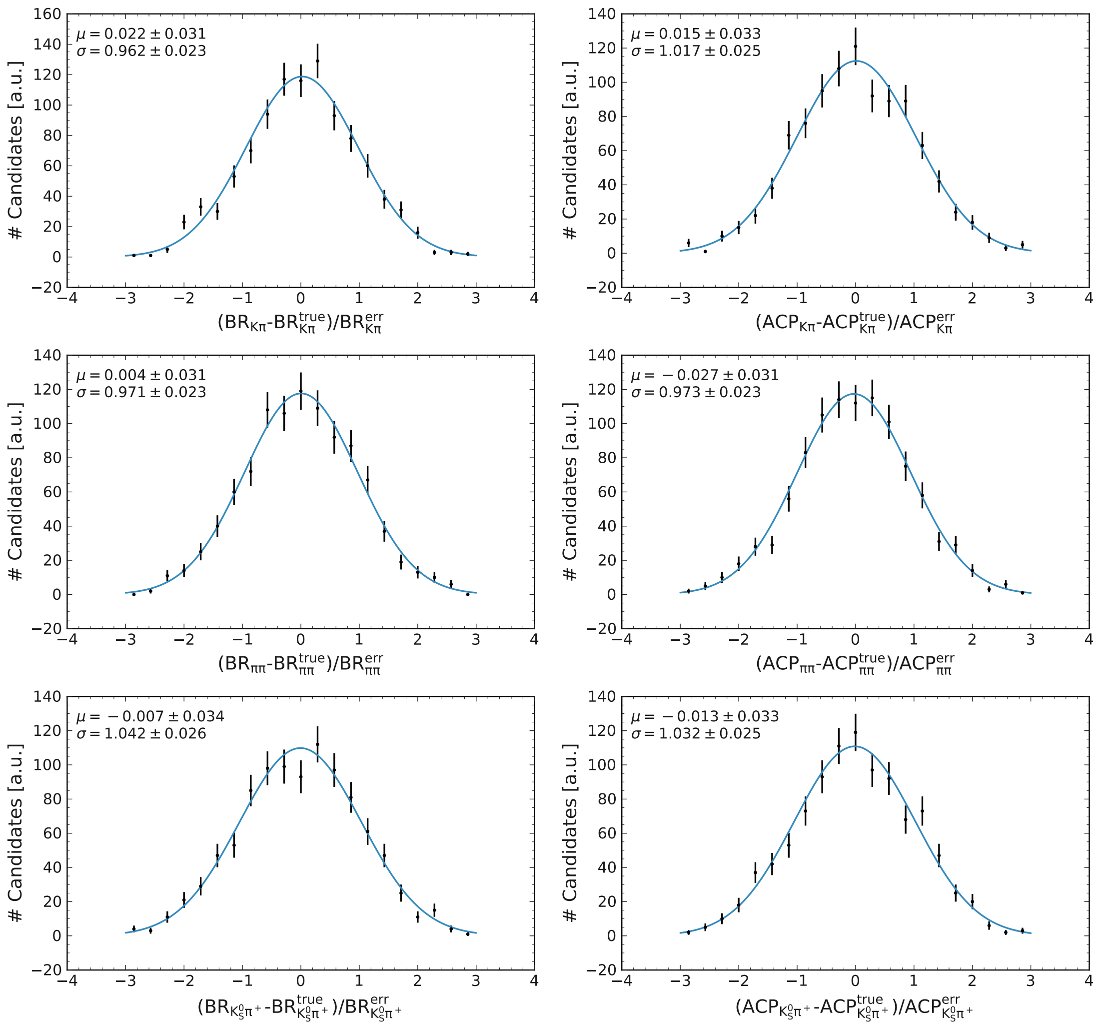


Figure 4.16: Distribution of differences between the fit result and simulated data, normalized by fit uncertainties, (pull distribution) for the branching fraction (left) and direct CP asymmetry (right) for $B^+ \to K^+\pi^0$. Middle (Bottom): Same but for $B^+ \to \pi^+\pi^0$ ($B^+ \to K_S^0\pi^+$).

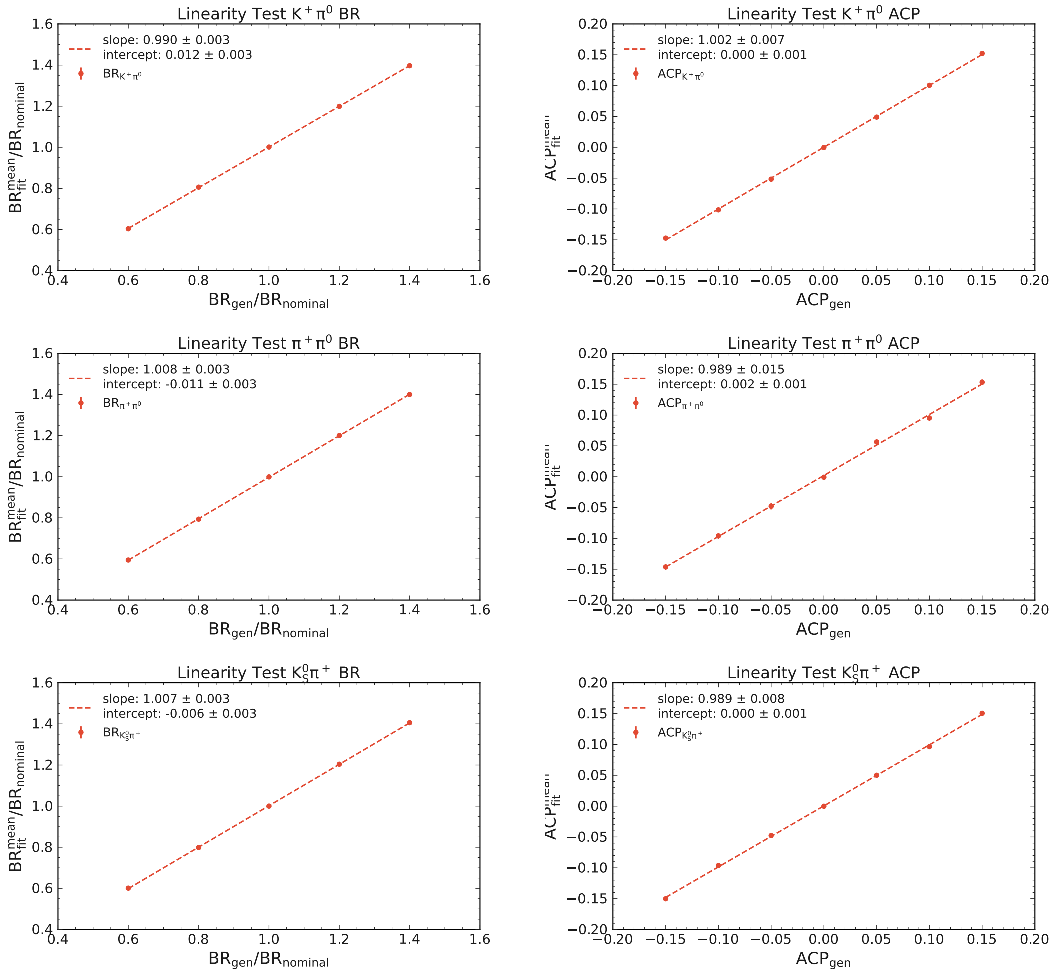


Figure 4.17: Upper: Result of the linearity test using toy ensembles for the branching fraction (left) and direct CP asymmetry (right) of $B^+ \to K^+\pi^0$. Middle (Bottom): Same but for $B^+ \to \pi^+\pi^0$ ($B^+ \to K_S^0\pi^+$).

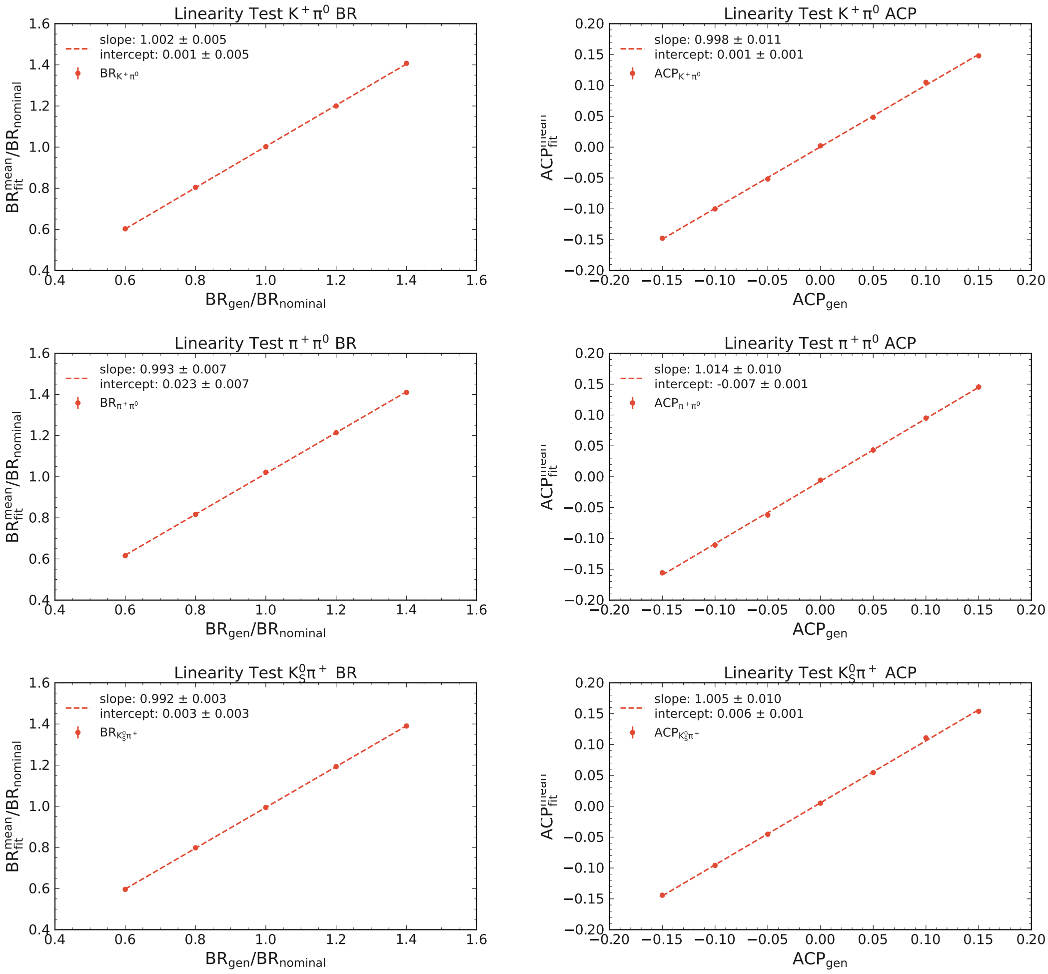


Figure 4.18: Upper: Result of the linearity test using bootstrapped ensembles for the branching fraction (left) and direct CP asymmetry (right) of $B^+ \to K^+\pi^0$. Middle (Bottom): Same but for $B^+ \to \pi^+\pi^0$ ($B^+ \to K^0_S\pi^+$).

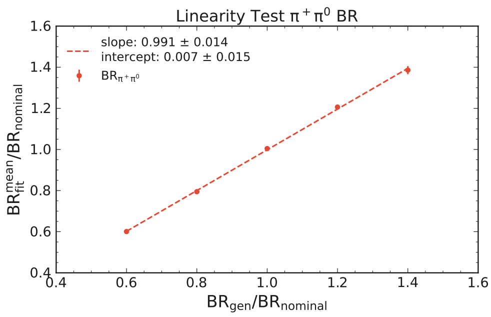


Figure 4.19: Linerarity test for the branching fraction of $B^+ \to K_S^0\pi^+$ when the continuum-background component is sampled from the fit model while all other components are bootstrapped.

## 4.5 Simulation Correction

For the computation of the reconstruction efficiencies introduced in Section 4.3.4 and the fit modeling introduced in Section 4.4, we rely solely on simulated data. However, the simulation shows discrepancies to the experimental data. This is due to difficulties in simulating the physics processes and imperfect detector calibration. If not accounted for, these discrepancies impair the reliable extraction of the physics parameters. We use control modes with a topology similar to the signal modes and a higher branching fraction to correct these discrepancies.

We adjust the reconstruction efficiencies for data-simulation differences in several areas: The efficiency of the track reconstruction, continuum suppression and particle identification requirements, the $K_S^0$ reconstruction, and the $\pi^0$ reconstruction. We also account for discrepancies in the instrumental asymmetries, i.e., discrepancies related to the differences in reconstruction efficiency between particles and antiparticles. To correct the signal and peaking background fit model, we introduce shift and scaling (fudge) factors. The shift factor shifts the mean of the Gaussian component of the models and the scale factor scales their widths.

In the remainder of this section, we describe the extraction of the different correction factors.

### 4.5.1 Tracking Efficiency Correction

The performance of Belle II's track finding algorithm may differ between data and simulation, e.g., due to poorly modeled beam background. A correction for such effects is centrally provided by the Belle II Collaboration. The correction is obtained from $e^+e^- \to \tau^+\tau^-$ events, where one $\tau$ decays into an electron or a muon and corresponding neutrinos, and the other $\tau$ decays into three charged pions, a $\tau$ neutrino, and any number of neutral pions. The electron/muon as well as two of the three charged pions are reconstructed to tag $\tau$-pair events. The efficiency of the track finding algorithm is then calculated by dividing the number of events in which the fourth track is found by the total number of events. The ratio of track finding efficiency between data and simulation is $(99.94 \pm 0.24)\%$, where the uncertainty includes the statistical and systematic uncertainty. This ratio is consistent with 100%, i.e., no correction, hence we do not scale with this ratio and assign the uncertainty as systematic uncertainty per track.

### 4.5.2 Continuum Suppression Correction

The efficiency of the continuum suppression requirement differs between data and simulation due to small differences in the input variables. We use $B^+ \to \overline{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays to assess the difference in continuum suppression response and extract correction factors. We use this decay because its branching fraction is more than an order of magnitude greater than that of the signal decays [9] and also contains a neutral particle.

The reconstruction follows the same approach as for the signal decays (see Section 4.3), where appropriate we use the same selection criteria. Compared to the signal decays, the amount of background is greatly reduced since we place a mass window requirement around the nominal $D^0$ mass. The width of this window corresponds to three times the detector mass resolution. We split the reconstructed sample into two disjoint samples: one that passes and one that fails the continuum suppression criterion. We then perform a simultaneous fit the $\Delta E$ distribution of both. Using the fit result, the continuum suppression correction factor is calculated from the following double ratio:

$$\eta = \left(\frac{N_{\text{pass}}}{N_{\text{pass}} + N_{\text{fail}}}\right)_{\text{data}} / \left(\frac{N_{\text{pass}}}{N_{\text{pass}} + N_{\text{fail}}}\right)_{\text{simulation}} , \tag{4.2}$$

where $N_{\text{pass}}$ is the number of events passing the continuum suppression requirement and $N_{\text{fail}}$ is the number of events failing it. The subscript "data" and "simulation" refers to the ratio being computed using real or simulated data. The double ratio approach has the advantage that systematic uncertainties, such as the uncertainty on the number of produced $B$ pairs, cancel.

The fit procedure follows the same approach as for the signal decays. We extract shapes for each component (signal, $B\overline{B}$ background and continuum background) by fitting dedicated simulation samples. In the fit to data, the signal as well as the $B\overline{B}$ shape parameter are fixed and only the continuum shape parameter is allowed to vary. Additional fit parameter are the yield of each component and

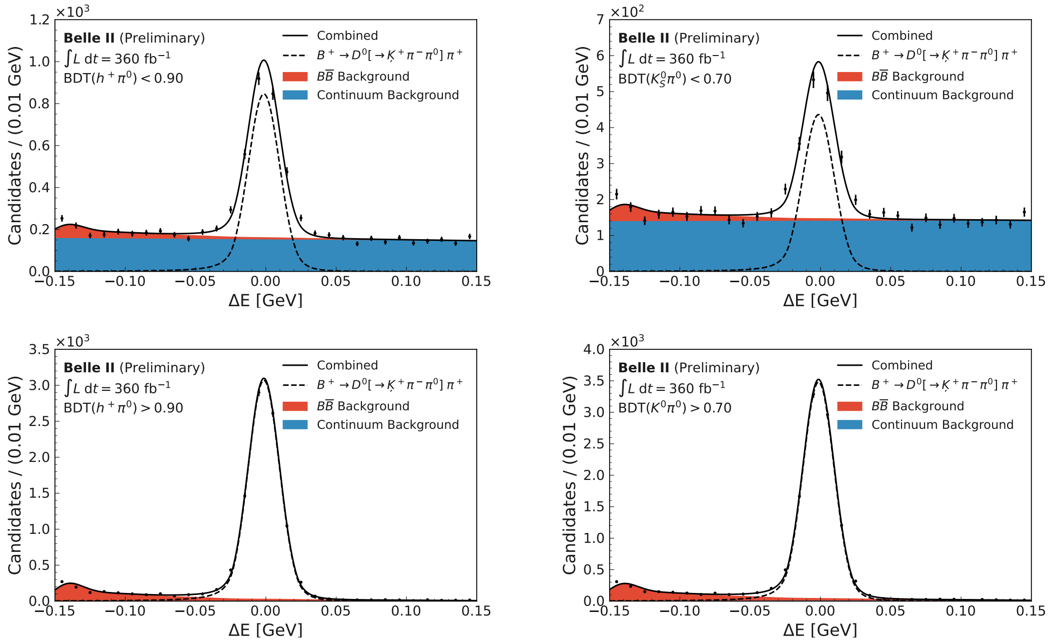


Figure 4.20: Result of the fit to the $\Delta E$ distribution of $B^+ \to \overline{D}(\to K^+\pi^-\pi^0)\pi^+$ for the extraction of the efficiency of the continuum suppression cut for the $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$ (left) and $B^+ \to K_S^0\pi^+$ (right) continuum suppression. Above (Below) the fit to the sample failing (passing) the continuum suppression requirement is shown. The data is shown as black dots with errorbars. The result of the fit is overlaid as a solid black line, the signal component as black dashed line, the $B\overline{B}$ background as red shaded area and the continuum background as blue shaded are.

fudge factors for the signal model.

In Fig. 4.20, fit projections on the simulation and data sample are shown for the continuum suppression of the $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$ samples, as well for the continuum suppression of the $K^0\pi^+$ sample. For the continuum suppression of $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$, we obtain a correction factor of $(96.81 \pm 0.66)\%$ and for the continuum suppression of $B^+ \to K_S^0\pi^+$ we obtain a correction factor of $(99.10 \pm 0.51)\%$. The uncertainties represent statistical uncertainties, which we assign as systematic uncertainty.

### 4.5.3 Particle Identification Correction

We identify a track as kaon or pion candidate via a requirement on the binary kaon–pion particle identification likelihood. The efficiency of this requirement differs between data and simulation. We derive a correction factor for the simulation-based efficiency estimate by measuring the efficiency in data and simulation. For these measurements, we use $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$ and $K^0_S \to \pi^+\pi^-$ decays, where the species of the tracks is inferred from kinematic information.

During reconstruction, we apply the same requirements on the tracks as for the signal decays bar the particle identification requirement. After reconstruction, we perform a fit to the mass distribution of the $D^0$ and the $K^0_S$. Using the fit result and the sWeight technique [87], we obtain the particle identification likelihood, momentum, and polar angle distribution of signal events. We calculate the efficiency of the particle identification criterion by summing the signal sWeights of events meeting the criterion and dividing by the total sum of all signal sWeights. This procedure is repeated for both data and simulation.

The correction factor is determined by dividing the efficiency obtained from data and simulation. This analysis is conducted across bins of track momentum and polar angle, as the correction factor depends on these variables. The entire procedure is automated within the `Systematic Corrections Framework` [88] provided by the Belle II Collaboration.

We apply the correction factors to our signal samples and determine the necessary scaling of the simulation-based efficiency estimate to match the data efficiency. The efficiency estimate for correctly identified $B^+ \to K^+\pi^0$ events needs to be scaled by $(97.34 \pm 0.10)\%$, and for correctly identified $B^+ \to \pi^+\pi^0$ events by $(97.22 \pm 0.13)\%$. For mis-identified $B^+ \to K^+\pi^0$ events, the scale factor is $(120.40 \pm 0.72)\%$ and for mis-identified $B^+ \to \pi^+\pi^0$ events the scale factor is $(116.70 \pm 0.70)\%$. The scaling factors for $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$ events show similar behavior. We are overestimating, the number of correctly identified events and are underestimating the number of mis-identified events. In addition, the corrections for mis-identified events are larger than for correctly identified events. This behavior is expected.

### 4.5.4 Efficiency Correction of $K_S^0$ Reconstruction

In this section, we investigate differences in the $K_S^0$ reconstruction efficiency between data and simulation. First, we study discrepancies due to the displaced-vertex reconstruction, i.e., finding two tracks from a $K_S^0$ decay displaced from the interaction point and reconstructing the corresponding vertex. Second, we examine whether the $K_S^0$ selection causes a difference in reconstruction efficiency.

#### Displaced Vertex $K_S^0$ Efficiency Correction

The $K_S^0$ flies on average 10 cm before it decays, hence its decay vertex is far displaced from the interaction point. As a consequence, most pions generated in the $K_S^0$ decay traverse only parts of the tracking system. This makes the $K_S^0$ reconstruction especially susceptible to beam background effects. These effects are known to be badly modeled in the simulation, necessitating a correction of the $K_S^0$ reconstruction efficiency.

We use $D^{*+} \to D^0(\to K_S^0(\to \pi^+\pi^-)\pi^+\pi^-)\pi^+$ decays to derive a correction factor for the $K_S^0$ reconstruction efficiency. For the reconstruction, we use the same approach as for the signal decays (see Section 4.3). Due to requirements on the $D^0$ mass and the mass difference between $D^0$ and $D^{*+}$ background is greatly reduced. After reconstruction and selection, we compare the momentum distribution of $K_S^0$ in the $D^{*+} \to D^0(\to K_S^0(\to \pi^+\pi^-)\pi^+\pi^-)\pi^+$ and $B^+ \to K^0\pi^+$ sample. This comparison is shown in Fig. 4.21. Since both distributions differ and the $K_S^0$ reconstruction efficiency depends on the $K_S^0$ momentum, we require the $K_S^0$ momentum to be greater than two GeV/$c^2$.

To derive the reconstruction efficiency correction, we obtain the yield for reconstructed $K_S^0$ candidates in simulation and data by fitting the $K_S^0$ mass distribution. In the fits, two components are considered: signal, containing correctly reconstructed $K_S^0$, and background, containing all other events. We model the signal using a double-sided Crystal Ball function and the background using a first-order polynomial. In Fig. 4.22, the fit result of the signal and background model to dedicated simulation samples are shown. In the fits to extract the yields, the tail parameter of the double-sided Crystal Ball are fixed to those derived by fitting the signal simulation sample. All other signal and background shape parameter are left free.

To account for the dependence of the reconstruction efficiency correction on the $K_S^0$ kinematics, fits are performed in bins of flight distance and flight direction (polar angle). The bins in flight direction are chosen in such a way to have the same number of signal candidates in each bin. The bins in flight distance have been inherited from a previous internal Belle II study of the $K_S^0$ reconstruction efficiency correction. The binning scheme is summarized in Section 4.5.4. We extract independent fit models for each bin.

In each bin, we calculate the ratio between the fitted $K_S^0$ candidate yield in simulation and data. We normalize the ratio in the first flight distance bin to one and scale the ratio in all other bins accordingly, assuming the reconstruction efficiency correction in the first flight distance bin is given by the tracking efficiency (see Section 4.5.1). In addition, we assume there is no dependence of the correction in flight direction for the *first* flight distance bin, hence each *first* flight distance bin is independently normalized. In Fig. 4.23, the ratios between data and simulation for the first flight-distance bin as a function of flight direction are shown, supporting our assumption.

We consider two sources of systematic uncertainty: The first is due to normalizing the first bin of flight distance; the second is due to mismodeling of tails of the $K_S^0$ mass distribution in simulation. To account for the first source, we assign the statistical uncertainty of the ratio of yields in the first bin of flight distance as systematic uncertainty. There are three such bins (one per bin in flight direction). We compute the total uncertainty as the weighted sum of them, using following equation:

$$\sigma^2_{\text{norm}} = \sum_{i=1}^{3} f_i^2 \sigma^2_{1,i} \tag{4.3}$$

where $\sigma_{1,i}$ is the uncertainty of the $K_S^0$ yield ratio in the first bin of flight distance of the $i$-th bin of flight direction, and $f_i^2$ is the fraction of total $K_S^0$ candidates in the $i$-th bin of flight direction integrated over flight distance. This uncertainty is considered as an overall systematic uncertainty.

We account for the second source by repeating the study using the sum of two Gaussians as signal model. In Fig. 4.22, the alternative model alongside the

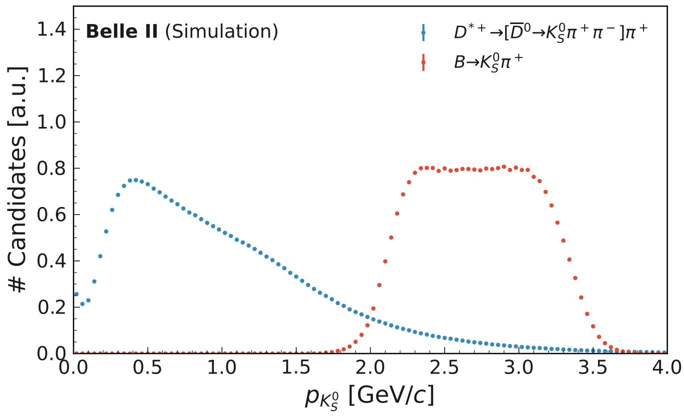


Figure 4.21: Comparison of the momentum of signal $K_S^0$ for the $D^{*+} \to [\overline{D}^0 \to K_S^0\pi^+\pi^-]\pi^+$ channel (blue dots) and one of the signal channels (red dots).

nominal model is shown. The alternative model does not describe the tails of the $K_S^0$ mass distribution. We take the difference in result to our nominal model as systematic uncertainty. This uncertainty is conservative, as we compare our model to a model that does not describe the tails of the $K_S^0$ mass distribution.

The $K_S^0$ reconstruction efficiency factors as a function of flight direction and flight distance are shown in Fig. 4.24. As expected, the correction factors vary depending on both. We apply the correction factors to our signal sample and find we need to scale the $K_S^0$ reconstruction efficiency resulting from the simulation by $(95.2 \pm 0.3(\text{stat}) \pm 1.2(\text{norm}) \pm 1.5(\text{fit}))\%$ to match the efficiency in data. The first uncertainty is the statistical uncertainty, the second is due to the normalization uncertainty and the third due to the fit model uncertainty. The sum in quadrature of all uncertainties is 1.9%.

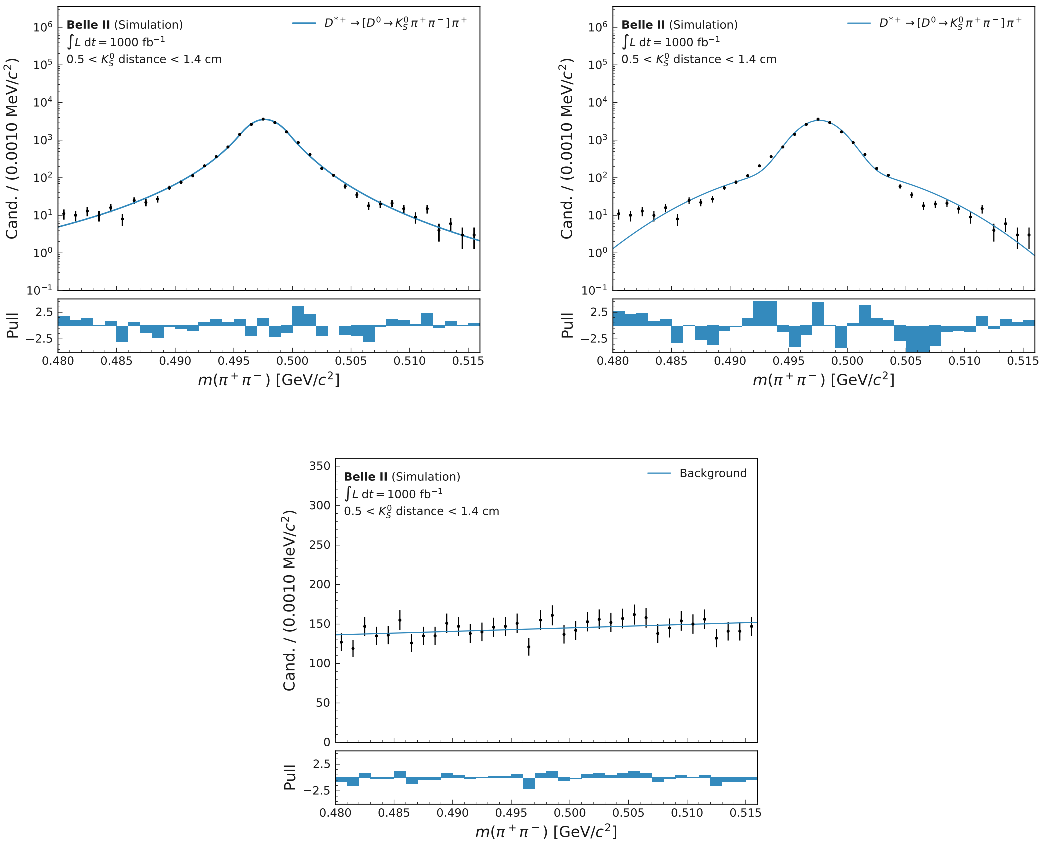


Figure 4.22: Upper: Fit to the signal $K_S^0$ mass distribution in the first flight distance bin for $\cos(\theta) < 0.3$ using a double-sided Crystal Ball (left) and two Gaussians (right). Below differences between simulated data and the fit results, normalized by fit uncertainties, (pulls) are shown. Lower: Fit to the background distribution in the same bin using a first-order Chebyshev polynomial.

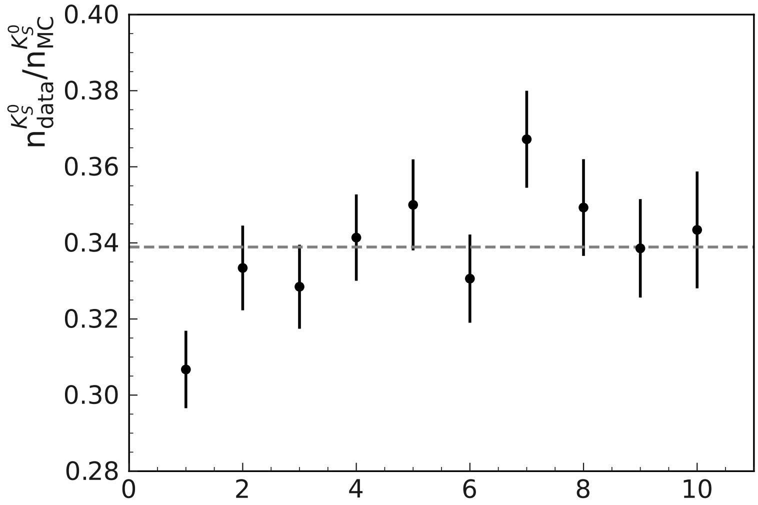


Figure 4.23: Ratio of $K_S^0$ reconstruction efficiency between data and simulation for the first flight-distance bin, divided into ten $\cos(\theta)$ bins with equal number of candidates. The displayed uncertainty represents the statistical uncertainty only. The grey dashed line shows the averages of all bins.

Table 4.5: Binning scheme for the displaced vertex $K_S^0$ efficiency correction study. We remind the cut $p(K_S^0) > 2\,\mathrm{GeV}/c^2$ that is also applied.

| 3D flight distance [ cm ] | $\cos(\theta)$ |
|---|---|
| (0.5,1.4) | (-1, 0.3) |
| (1.4,2.4) | (0.3, 0.7) |
| (2.4,3.6) | (0.7, 1) |
| (3.6,5.2) | |
| (5.2,7.0) | |
| (7.0,9.0) | |
| (9.0,13.) | |
| (13.,22.) | |
| (22.,35.) | |
| (35.,55.) | |

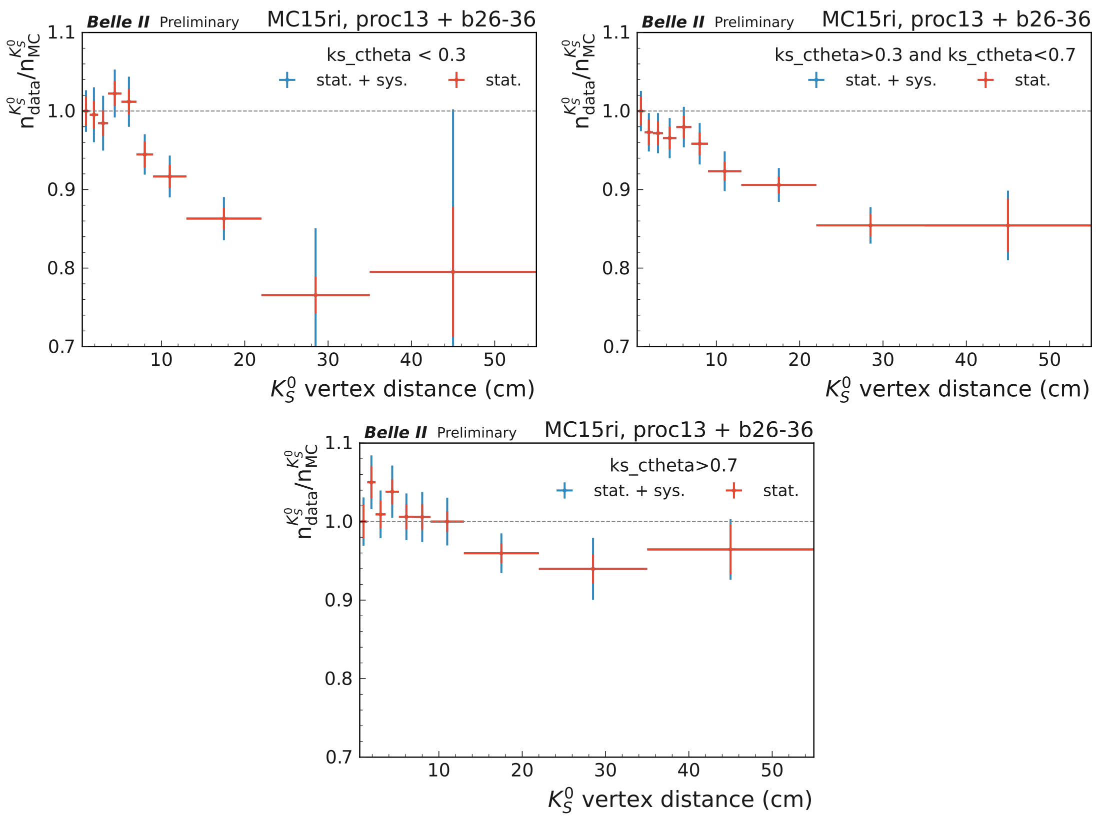


Figure 4.24: Ratio of $K_S^0$ reconstruction efficiency between data and simulation for $\cos(\theta) < 0.3$ (upper left), $0.3 < \cos(\theta) < 0.7$ (upper right), and $\cos(\theta) > 0.7$ (lower). Vertical error bars in red show the statistical uncertainty and in blue show the sum in quadrature of statistical and systematic uncertainty.

### Selection $K_S^0$ Efficiency Correction

During the $K_S^0$ reconstruction, we apply requirements on the $K_S^0$ mass distribution and the `V0Selector` as well as `LambdaVeto` boosted decision trees. Due to mismodeling of the $K_S^0$ mass distribution or input variables of the boosted decision trees, the efficiency of these requirements can differ between data and simulation. We use $D^{*+} \to \overline{D}^0(\to K_S^0\pi^0)\pi^+$ decays to assess these differences.

To evaluate the correction factor, we use the double-ratio approached introduced for the continuum suppression efficiency correction (see Section 4.5.2). During reconstruction, we apply a loose selection on the $K_S^0$ to obtain an almost unbiased sample. Afterwards, we apply the `V0Selector` and `LambdaVeto` requirements as for the signal channels to define pass and fail samples. The $K_S^0$ yield in both samples is obtained by fitting the $K_S^0$ mass distribution. After the fit, the signal shape is integrated across a mass window that corresponds to that of the signal channels. This integration is carried out instead of directly applying the mass window requirement as the fit is otherwise unstable. Using the result of the integration, we evaluate the double ratio.

Apart from the $K_S^0$ selection, we place similar selection requirements as in our signal modes during reconstruction. The amount of background is greatly reduced due to requirements on the $\overline{D}^0$ mass and the mass difference between $\overline{D}^0$ and $D^{*+}$. To suppress $B\overline{B}$ events, the $D^{*+}$ momentum must be greater than 2.5 GeV/$c^2$. We require the $K_S^0$ momentum to be in the range [2.0,3.5] GeV/$c^2$, to restrict it to a region populated by the signal channel (see Fig. 4.25).

Two components are taken into account in the fit: signal, containing correctly reconstructed $K_S^0$, and background, containing all other events. The signal is modeled with a double-sided Crystal Ball function. We use separate shapes for the signal passing and failing the $K_S^0$ selection. For background failing the $K_S^0$ selection we use a second-order polynomial, and for background passing the $K_S^0$ selection the sum of a second-order polynomial and a Gaussian. All shape parameter are determined using dedicated simulation samples. The Gaussian of the background peaks at the same position as the signal which could lead to background events being misidentified as signal. Such an effect is neglected since the signal is three orders of magnitude more abundant than the portion of the background contained

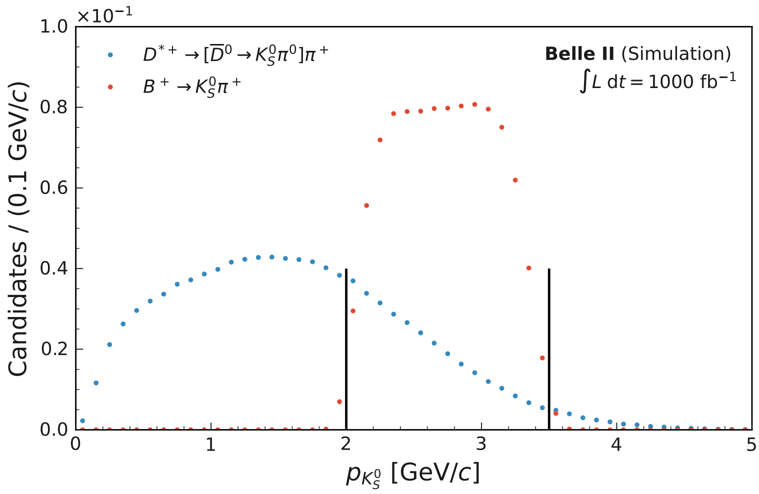


Figure 4.25: Comparison of the momentum of signal $K_S^0$ for the $D^{*+} \to [\overline{D}^0 \to K_S^0\pi^0]\pi^+$ channel (blue dots) and one of the signal channels (red dots). The vertical solid black lines indicate the cuts applied in this analysis.

in the Gaussian. In the fits to extract the yields, we introduce parameter shifting and smearing the mean and width of the Gaussian core of the signal shape, which are shared between the signal shape for the events failing and passing the $K_S^0$ selection.

The fit is carried out in bins of the $K_S^0$ flight distance and direction (polar angle), due to a possible dependence of the correction on both. In Fig. 4.26, an example for a fit in one of the bins is shown. Applying the correction factors to the $B^+ \to K^0\pi^+$ signal sample, yields a scaling factor for the branching fraction of $(99.56 \pm 0.25)\%$, which is in agreement with 1, i.e., no correction within two standard deviations. Hence we do not scale using this factor and assign the uncertainty on it as a systematic uncertainty.

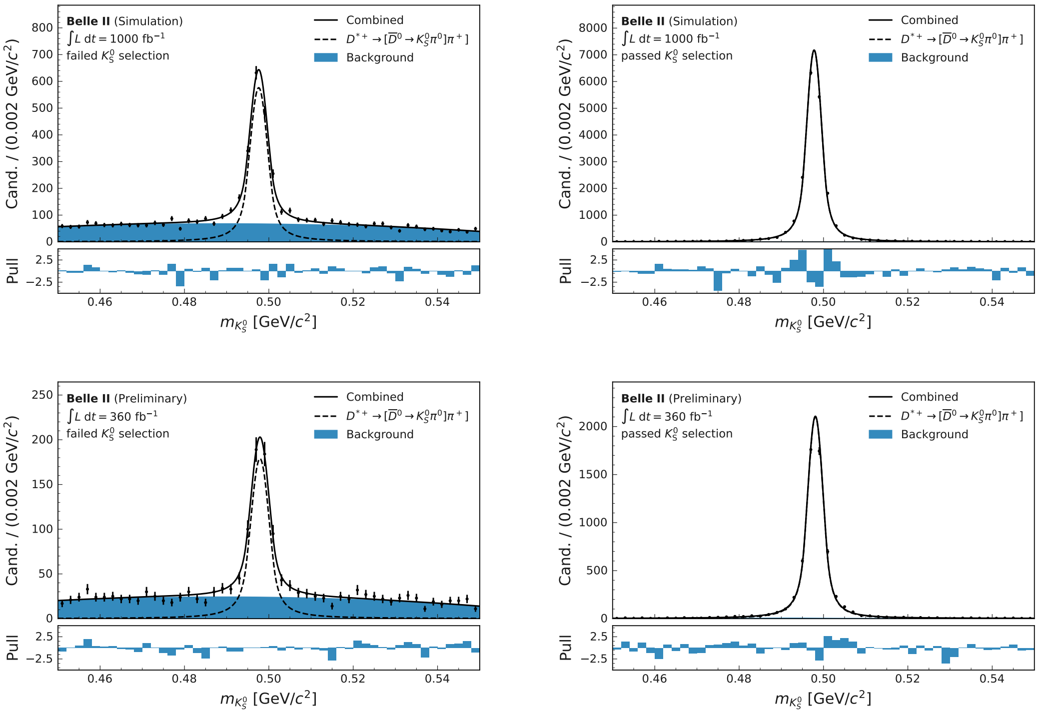


Figure 4.26: Upper: Fit to the invariant $K_S^0$ mass distribution (black dots) to the simulation sample failing (left) and passing (right) the $K_S^0$ selection. The signal component is shown as a black dashed line, the background component as a blue shaded area and the combination of the fit shapes as a solid black line. Below differences between data and the fit results, normalized by fit uncertainties, (pulls) are shown. Lower: Same but for real data.

### 4.5.5 Efficiency Correction of $\pi^0$ Reconstruction

Due to mismodeling of cluster variables used in the $\pi^0$ selection, the $\pi^0$ reconstruction efficiency differs between data and simulation. The Belle II Collaboration provides centrally corrections for the $\pi^0$ reconstruction efficiency. For the calculation of the correction, $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^0$ decays are reconstructed and fits to the invariant mass distribution are performed to estimate their yields in data and simulation. The correction factor is determined by comparing the ratio of yields in data and simulation for both decays and taking the known values of their branching fractions as input [9]. The correction factors are determined as a function of the $\pi^0$ momentum and polar angle as it depends on both. We apply the correction factors to our simulation and obtain a scaling factor of $(99.2 \pm 3.8)\%$ for $K^+\pi^0$ and $(99.3 \pm 3.8)\%$ for $\pi^+\pi^0$, where the uncertainties include the statistical and systematic uncertainty. The biggest contribution to the uncertainty stems from the uncertainty of the $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^0$ branching fraction.

### 4.5.6 Corrections for Instrumental Charge-Asymmetries

The measured asymmetry for $B^+$ and $B^-$ decays is the raw asymmetry $\mathcal{A}_{\text{raw}}$, it is the sum of the direct CP asymmetry $\mathcal{A}_{\text{CP}}$, we want to measure, and the instrumental asymmetry $\mathcal{A}_{\text{det}}$. The instrumental asymmetry is caused by differences in interaction and reconstruction probabilities between particles and antiparticles. One known cause is that positively charged tracks have on average one associated hit less in the central drift chamber in data due to the non-uniform electric field around the sense wires. Due to drift-time mismodeling this effect is opposite in simulation compared to data. To measure the direct CP asymmetry, we need to subtract the instrumental asymmetry from the measured raw asymmetry. Estimations of the instrumental asymmetry are provided by the Belle II Collaboration.

For the estimation of the charged pion instrumental asymmetry $D^+ \to K_S^0\pi^+$ decays are reconstructed and the $D^+$ mass distribution is fitted. The known value of the direct CP asymmetry is subtracted from the measured value [9]. Due to $\gamma^* - Z^0$ interference in $e^+e^- \to c\bar{c}$ are produced antisymmetrically in forward-backward direction. To account for this effect, the instrumental asymmetry is measured in bins of the polar angle in the center-of-mass system and averaged over bins with opposite sign, since the asymmetry is anti-symmetric in this variable. The charged kaon instrumental asymmetry is measured by reconstructing $D^0 \to K^-\pi^+$ events. Here, the same strategy is followed to remove the effect from the production asymmetry. The direct CP asymmetry is measured to be below 0.1% and hence neglected. The charged kaon asymmetry can than be obtained from the measured value of the $D^0 \to K^-\pi^+$ asymmetry by subtracting the measured asymmetry for charged pions.

Since the instrumental asymmetry depends on the flight direction, momentum, and number of hits in the drift chamber, we reweight these distributions to match those of our signal channels. We find as value of instrumental asymmetry $\mathcal{A}_{\rm det}(\pi^+)$(for $B^+ \to K^0_S\pi^+$) $= (0.78 \pm 0.42)\%$, $\mathcal{A}_{\rm det}(\pi^+)$(for $B^+ \to \pi^+\pi^0$) $= (0.78 \pm 0.42)\%$, and $\mathcal{A}_{\rm det}(K^+)$(for $B^+ \to K^+\pi^0$) $= (1.41 \pm 0.46)\%$.

### 4.5.7 Correction for $\Delta E$ Fit Model

Due to an imperfect energy calibration, the $\Delta E$ distribution differs between data and simulation. This leads to a shift and smearing of the peaks in the $\Delta E$ distribution. To correct the signal and peaking background fit model, we introduce shift and scaling (fudge) factors. The shift factor shifts the mean of the Gaussian component of the models and the scale factor scales their widths. We analyze $B^0 \to \overline{D}(\to K^+\pi^-)\pi^0$ decays to extract fudge factors for the $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$ channel. To extract fudge factors for the $B^+ \to K_S^0\pi^+$ channel, we use $B^+ \to \overline{D}(\to K_S^0\pi^+\pi^-)\pi^+$ and $B^0 \to D^-(\to K_S^0\pi^-)\pi^+$ decays.

During reconstruction, we apply selection requirements as for the signal channels. Compared to the signal channels, the background is greatly reduced due to a mass window requirement on the $D^0$ mass. For the $B^+ \to \overline{D}(\to K_S^0\pi^+\pi^-)\pi^+$ channel, we require the $K_S^0$ momentum to be greater than $0.5\,\mathrm{GeV}/c$. This requirement suppresses background from other $B\overline{B}$ decays.

After reconstruction, we perform a two dimensional fit of the $\Delta E$ and $C'$ distribution. The fit procedure follows that for the signal channels. First, we extract models by fitting dedicated simulation samples for each of the fit components (signal, $B\overline{B}$ background and continuum background). Afterwards, we introduce fudge factors to the signal shape. In the fit to data, the yields, the fudge factors, and the continuum background shape parameter are free to float.

Since the fudge factor might depend on the $\pi^0$ or $K_S^0$ momentum, we compare them between the control channels and the signal channels (see Fig. 4.27). The momentum distribution shows adequate agreement for the $\pi^0$. For the $K_S^0$, the control channel contains more low momentum events than the signal channel. We assess the momentum dependence by dividing the sample into two bins of $K_S^0$ momentum (One above and one below $1.5\,\mathrm{GeV}/c$) and perform the fit in both bins. The fit results for the fudge factors agree within uncertainty, therefore we do not divide the sample into different momentum bins for the nominal fit to extract the shift and scale parameter.

In Figs. 4.28 to 4.30, fit projections to the fit to data for the various channels are shown. All show good agreement between the fit result and data. For $B^0 \to \overline{D}(\to K^+\pi^-)\pi^0$, the extracted shift parameter is $(-0.0005 \pm 0.0017)\,\mathrm{GeV}$

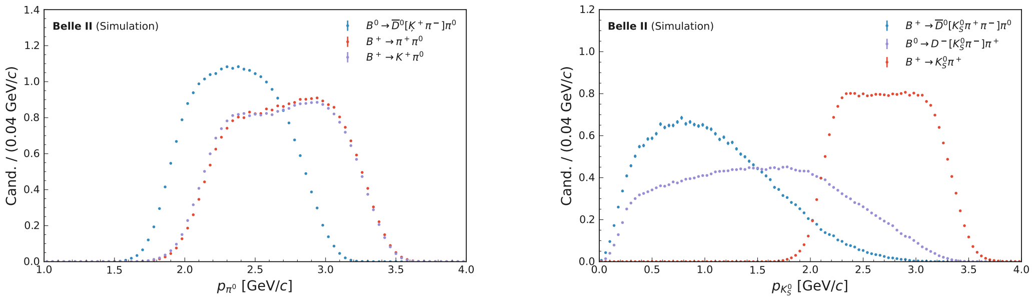


Figure 4.27: The $\pi^0$ momentum distribution for simulated $B^0 \to \overline{D}(\to K^+\pi^-)\pi^0$ (blue dots with errorbars), $B^+ \to \pi^+\pi^0$ (red dots with errorbars) and $B^+ \to K^+\pi^0$ (purple dots with errobars) signal events. The $K_S^0$ momentum distribution for simulated $B^+ \to \overline{D}(\to K_S^0\pi^+\pi^-)\pi^+$ (blue dots with errorbars), $B^0 \to D^-(\to K_S^0\pi^-)\pi^+$ (purple dots with errorbars), and $B^+ \to K_S^0\pi^+$ (red dots with errorbars) signal events.

and the scaling parameter is $1.034 \pm 0.046$. For $B^+ \to \overline{D}(\to K_S^0\pi^+\pi^-)\pi^+$ ($B^0 \to D^-(\to K_S^0\pi^-)\pi^+$), the extracted shift parameter is $(-0.96 \pm 0.18)$ MeV/$c$ ($(-0.71 \pm 0.22)$ MeV/$c$) and the scaling parameter is $1.120 \pm 0.017$ ($1.086 \pm 0.020$). By combining the measurements we obtain a shift of $(-0.84 \pm 0.14)$ MeV/$c$ and a scale of $1.10 \pm 0.013$.

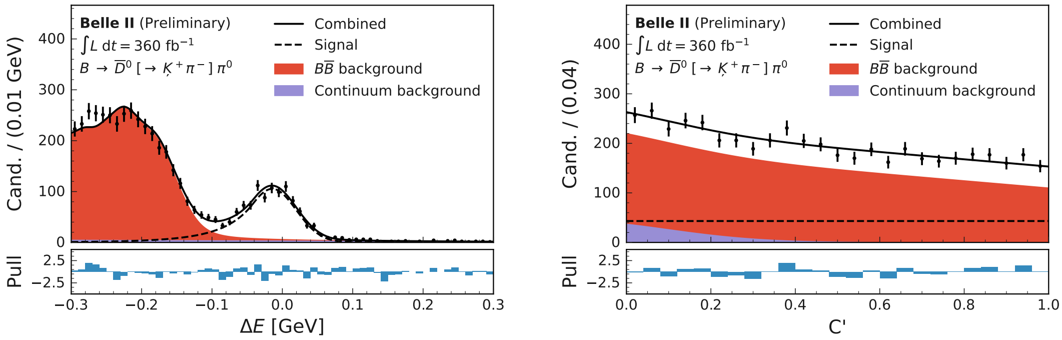


Figure 4.28: Left: $\Delta E$ distribution for $B^0 \to \overline{D}(\to K^+\pi^-)\pi^0$ shown as black dots with errorbars. Right: Transformed continuum suppression output distribution for $B^0 \to \overline{D}(\to K^+\pi^-)\pi^0$ shown as black dots with errorbars. The result of fits to the distributions 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 (continuum background). Below differences between data and the fit results, normalized by fit uncertainties, (pulls) are shown.

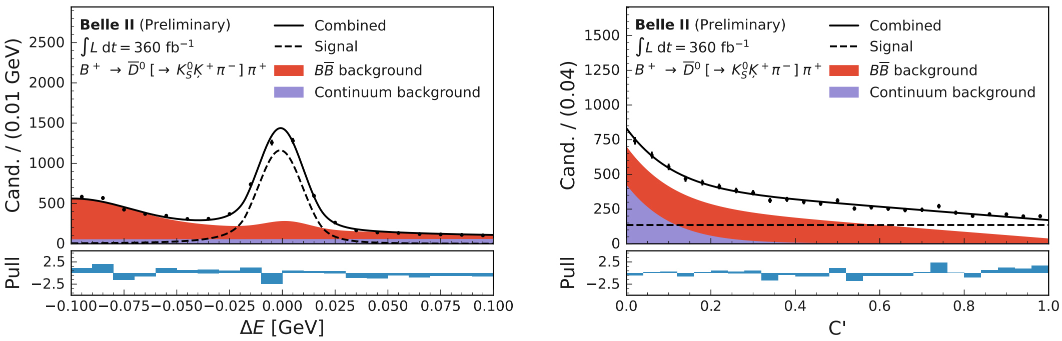


Figure 4.29: Left: $\Delta E$ distribution for $B^+ \to \overline{D}(\to K_S^0\pi^+\pi^-)\pi^+$ shown as black dots with errorbars. Right: Transformed continuum suppression output distribution for $B^+ \to \overline{D}(\to K_S^0\pi^+\pi^-)\pi^+$ shown as black dots with errorbars. The result of fits to the distributions 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 (continuum background). Below differences between data and the fit results, normalized by fit uncertainties, (pulls) are shown.

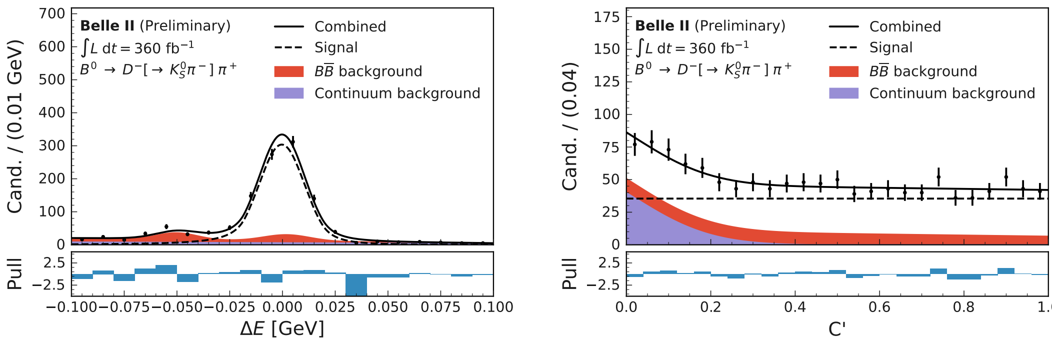


Figure 4.30: Left: $\Delta E$ distribution for $B^0 \to D^-(\to K_S^0\pi^-)\pi^+$ shown as black dots with errorbars. Right: Transformed continuum suppression output distribution for $B^0 \to D^-(\to K_S^0\pi^-)\pi^+$ shown as black dots with errorbars. The result of fits to the distributions is shown 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 (continuum background). Below differences between data and the fit results, normalized by fit uncertainties, (pulls) are shown.

### 4.5.8 Correction for $C'$ Fit Model

Due to differences between data and simulation in the continuum suppression input variables, the transformed continuum suppression output $C'$ can differ between data and simulation. This leads to a non zero slope for the signal $C'$ shape, where the slope in simulation is zero by construction. Using $B^+ \to \overline{D}^0 [\to K^+\pi^-\pi^0] \pi^+$ decays, we extract a parameter that adjust the slope of the signal $C'$ shape and evaluate its impact on the fit result with a Pearson's chi-squared test.

We reconstruct the control channel with selection requirements similar to that of the signal modes and subsequently fit their $\Delta E$ distribution. For the fit to data, we first extract fit models for our fit components (signal, $B\overline{B}$ background, and continuum background) using dedicated simulation samples and fix them. Afterwards, we introduce fudge factors to the signal model. In the fit to data, the yields, continuum background shape, and the fudge factors are free to float.

After fitting, we use the sWeight technique [87] to obtain the $C'$ distribution of signal events. We perform a fit using a first order polynomial to the background free $C'$ distribution to obtain the correction factor for the $C'$ signal shape. Before fitting the real data, we perform the analysis procedure with simulated data corresponding to $1000\,\mathrm{fb}^{-1}$. The analysis to the simulation sample, finds a correction factor of $0.013 \pm 0.011$ for the continuum suppression of $B^+ \to K^0_S\pi^+$ and a correction factor of $0.020 \pm 0.011$ for the continuum suppression of $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$. Both parameter are in agreement with the expectation of zero within two standard deviations. Subsequently, we carry out the analysis using the real data. Comparisons of the $C'$ distribution from simulation truth information and the sWeighted distribution from simulation and data are shown in Fig. 4.31. We extract a correction parameter of $0.031 \pm 0.018$ ($0.048 \pm 0.017$) for the continuum suppression of $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$ ($B^+ \to K^0_S\pi^+$). The $\Delta\chi^2$ between the sWeightes distribution and a shape with a slope of zero and a slope adjusted using the correction factor is 0.10 (0.24) for the continuum suppression of $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$ ($B^+ \to K^0_S\pi^+$). Since the $\Delta\chi^2$ in both cases is small, we do not shift the nominal fit shape.

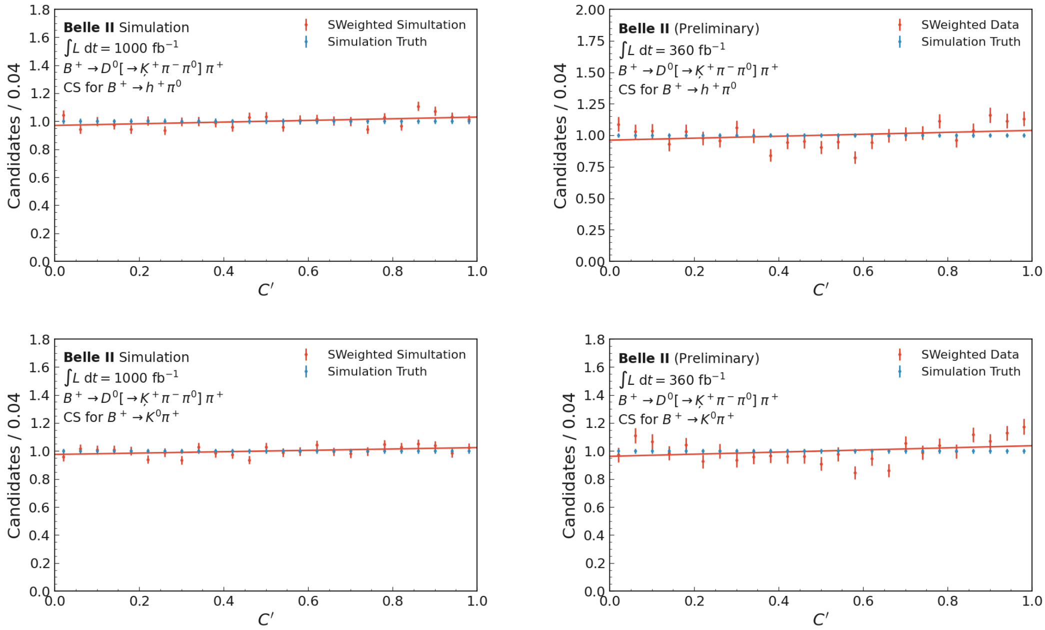


Figure 4.31: Upper left: The signal simulation transformed continuum suppression distribution $C'$ for the continuum suppression trained for $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$ (blue dots) and the same distribution after fitting and sWeigthing the simulation sample (red dots). Upper right: Signal simulation $C'$ distribution (red dots) and $C'$ distribution obtained by fitting and sWeighting the real data. The result of a fit to the $C'$ distribution of the real data is shown as solid blue line. Lower: Same as upper but for the continuum suppression trained for $B^+ \to K_S^0\pi^+$.

## 4.6 Systematic Uncertainties

In this chapter, we discuss the systematic uncertainties. They are summarized in Table 4.6 for the branching fractions and in Table 4.7 for the direct CP asymmetries.

For the branching fractions, the systematic uncertainties are dominating compared to the statistical uncertainty.

The largest contribution for $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$ originates from the uncertainty associated to the reconstruction efficiency of $\pi^0$. This uncertainty itself if dominated from the uncertainty of the ratio of branching fractions for the control channels used in the derivation of the correction factor. Hence, it is not expected to shrink with an increasing data sample size. Reducing the uncertainty will either require measuring the branching fractions of the decays or extending the $\pi^0$ efficiency correction estimation to additional channels.

The largest uncertainty for the branching fractions of $B^+ \to K^0\pi^+$ is dominated by the uncertainty of the $K_S^0$ reconstruction efficiency correction and the uncertainty of the branching fraction of $\Upsilon(4S) \to B^+B^-$. While the former is expected to shrink with increasing data sample size, the latter requires a measurement of the $\Upsilon(4S)$ branching fractions.

Table 4.6 and Table 4.7 summarizes the systematic uncertainties for the measured branching fractions and direct CP asymmetry. Each contribution is detailed in the following subsections.

### 4.6.1 Number of Produced $B^+B^-$ Pairs

In the fit to data, we fit directly for the branching fractions and the direct CP asymmetries. For the purpose of fitting, we need to calculate from those the signal yields using Eq. (4.1). The equation uses the number of produced $B^+B^-$ pairs as input. We calculate the number of produced $B^+B^-$ pairs by multiplying the total number of produced $B\overline{B}$ pairs ($N_{B\overline{B}}$) with the fraction of $\Upsilon(4S) \to B^+B^-$ ($f^{+-}$).

The number of produced B mesons in centrally provided by the Belle II Collaboration. It is estimated by subtracting the number of reconstructed hadronic events in off-resonance data from the number of hadronic events in on-resonance

Table 4.6: Summary of the fractional uncertainties on the branching fractions [%].

| Source | $B^+ \to K^+\pi^0$ | $B^+ \to \pi^+\pi^0$ | $B^+ \to K^0\pi^+$ |
|---|---|---|---|
| $N_{B\bar{B}}$ | 1.5 | 1.5 | 1.5 |
| $f^{+-}$ | 2.4 | 2.4 | 2.4 |
| Tracking | 0.2 | 0.2 | 0.7 |
| $\pi^0$ efficiency | 3.8 | 3.8 | - |
| $K_S^0$ efficiency | - | - | 2.0 |
| CS efficiency | 0.7 | 0.7 | 0.5 |
| PID correction | 0.1 | 0.2 | - |
| Fudge Factors | 1.2 | 2.0 | 0.3 |
| $K\pi$ signal model | 0.1 | <0.1 | <0.1 |
| $\pi\pi$ signal model | <0.1 | <0.1 | - |
| $K\pi$ CF model | <0.1 | 0.1 | - |
| $\pi\pi$ CF model | <0.1 | 0.1 | - |
| $K_S^0 K^+$ model | - | - | 0.1 |
| $B\overline{B}$ model | 0.3 | 0.5 | <0.1 |
| Multiple candidates | 1.0 | 0.3 | 0.1 |
| Total | 5.1 | 5.2 | 3.6 |
| Stat uncert. | 2.8 | 5.1 | 2.9 |

Table 4.7: Summary of the absolute uncertainties on the direct CP asymmetry.

| Source | $B^0 \to K^+\pi^0$ | $B^+ \to \pi^+\pi^0$ | $B^+ \to K^0\pi^+$ |
|---|---|---|---|
| $\Delta E$ shift and scale | 0.001 | 0.002 | 0.001 |
| $K_S^0 K^+$ model | - | - | 0.001 |
| Fitting bias | - | 0.007 | 0.006 |
| Instrumental asymmetry | 0.007 | 0.005 | 0.004 |
| Total | 0.005 | 0.008 | 0.007 |
| Stat uncert. | 0.0274 | 0.0491 | 0.0287 |

data. During the calculation, variations of the cross section and efficiency as a function of energy as well as the difference in luminosity between the on- and off-resonance sample are taken into account. The number of produced $B\overline{B}$ events is $(387 \pm 6) \times 10^6$.

We calculate $f^{+-}$ using the result of [89] and assuming the $\Upsilon(4S)$ decays only to $B\overline{B}$ pairs. It is $0.516 \pm 0.052$.

We assign the uncertainty on $N_{B\overline{B}}$ and $f^{+-}$ a systematic uncertainty on the branching fractions.

### 4.6.2 Efficiency Correction

We assign the uncertainty on the various corrections factors extracted in Section 4.5 as systematic uncertainty on the branching fractions. For the tracking uncertainty, we assign an uncertainty of 0.24% per track on the signal side. For the $\pi^0$ efficiency correction, the uncertainty is 3.8% and for the $K_S^0$ efficiency correction it is 2.0%. A systematic uncertainty of 0.7% is added for the continuum suppression efficiency correction (CS efficiency) trained for $B^0 \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$, and an uncertainty of 0.5% for the continuum suppression efficiency correction trained for $B^+ \to K_S^0\pi^+$.

To evaluate the systematic uncertainty associated to the particle identification (PID) correction, we perform a toy ensemble study. We generate toy ensembles by sampling from the fit shape of each fit component, where the yield of each component varies randomly according to a Poisson distribution. We fit each ensemble twice: Once with nominal PID corrections and once with varied PID corrections. We draw the varied PID corrections from a Gaussian distribution centered at the nominal PID correction and a width corresponding to their uncertainty. Afterwards, the difference between the fit result for the physics parameters using nominal and varied PID corrections is calculated. We assign the standard deviation of the difference distribution as systematic uncertainty.

### 4.6.3 Fudge Factors

We generate toy ensembles to calculate a systematic uncertainty for the fudge factors extracted in Section 4.5.7. We generate toy ensembles by sampling from the fit shape of each fit component, where the yield of each component varies randomly according to a Poisson distribution. Each ensemble is fitted twice: Once with nominal fudge factors and once with varied fudge factors. We draw the varied fudge factors from a two dimensional Gaussian distribution which is centered at the nominal fudge factor and takes the correlation between them into account. We calculate the difference of the fit result for the physics parameter using nominal and varied fudge factors. We assign the standard deviation of the difference distribution as systematic uncertainty.

### 4.6.4 Signal and Peaking Background Modelling

To assess a systematic uncertainty associated to the modeling of the signal and peaking background component, we generate toy ensembles. We generate toy ensembles by sampling from the nominal fit shape of each fit component, where the yield of each component varies randomly according to a Poisson distribution. Afterwards, we fit the generated ensembles twice: Once with the nominal fit shape and once with an alternative fit shape. The alternative fit shape is determined by varying the shape parameter according to a Gaussian centered at the nominal value and width corresponding to the respective uncertainty as obtained by a fit to a dedicated simulation sample. We calculate the difference of the fit result for the physics parameter using nominal and the alternative fit shape. We assign the standard deviation of the difference distribution as systematic uncertainty.

### 4.6.5 Uncertainty of $B\overline{B}$ Modelling

We assess a systematic uncertainty associated to the $B\overline{B}$ shape modeling by replacing the nominal $\Delta E$ shape (a kenerl density estimator) with a Gaussian. In Fig. 4.32, the alternative fit shape is imposed over a dedicated simulation sample. In addition, we vary the $C'$ shape within its uncertainties. Afterwards, we generate toy ensembles by sampling from his alternative shape with yields of each component fluctuated according to a Poisson distribution. The toy ensembles are fitted twice, once with the nominal and once with the alternative fit shape. The difference of the fit results is computed. The mean of the resulting distribution is assigned as a systematic uncertainty.

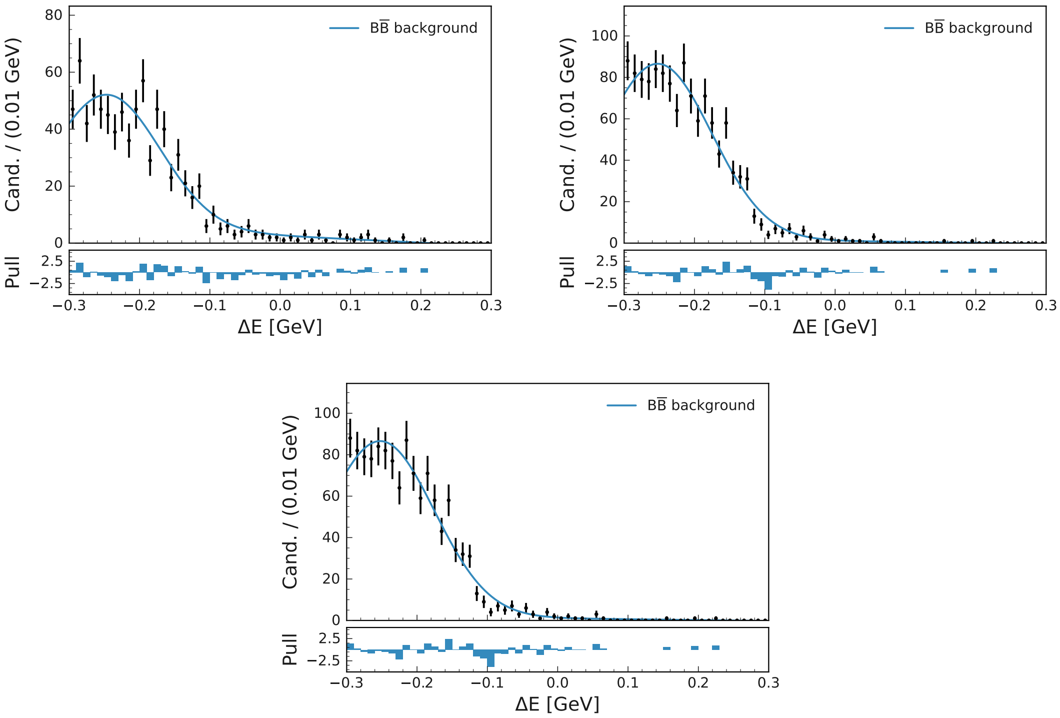


Figure 4.32: Alternative $B\overline{B}$ $\Delta E$ model consisting of a Gaussian + 1st degree polynomial for $B^+ \to K^+\pi^0$ (top left), $B^+ \to \pi^+\pi^0$ (top right), and $B^+ \to K_S^0\pi^+$ (bottom). Below differences between data and the fit results, normalized by fit uncertainties, (pulls) are shown.

### 4.6.6 Multiple Candidates

We do not perform a best candidate selection. To assess a systematic uncertainty associated to this choice, we fit the data with a random best candidate selection. The difference to the nominal fit result is assigned as systematic uncertainty.

### 4.6.7 Instrumental Asymmetries

The uncertainty on the correction of the instrumental asymmetry (see Section 4.5.6) is assigned as a systematic uncertainty on the measurement of the direct CP asymmetry.

## 4.7 Fit on Data

After developing and validating the analysis, we perform a fit of sample composition to the real data sample. The fit contains four components, signal, peaking background, background from $B\overline{B}$ events and continuum background. The fit to data is two-dimensional, we fit the energy difference $\Delta E$ and the transformed continuum suppression output $C'$ distributions. We perform two fits, one simultaneously to the $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$ samples and one to the $B^+ \to K^0\pi^+$ sample. The former fit has 18 free parameter and the latter 11. Free parameter are branching fractions and direct CP asymmetries of the signal decays, background yields and the continuum background shape parameter.

In Fig. 4.33, the result of the fits to data are shown. The data samples contain 5658 ($B^+ \to K^0\pi^+$), 4148 ($B^+ \to K^+\pi^0$), and 4135 ($B^+ \to \pi^+\pi^0$) events. We obtain branching fractions of

$$\begin{aligned} B^+ &\to K^0\pi^+ : (24.37 \pm 0.71 \pm 0.86)\times10^{-6}, \\ B^+ &\to K^+\pi^0 : (13.93 \pm 0.38 \pm 0.71)\times10^{-6}, \text{ and} \\ B^+ &\to \pi^+\pi^0 \ : \ (5.10 \pm 0.29 \pm 0.27)\times10^{-6}, \end{aligned}$$

and direct CP asymmetries of

$$\begin{aligned} B^+ &\to K^0\pi^+ : \ 0.046 \pm 0.029 \pm 0.007, \\ B^+ &\to K^+\pi^0 : \ 0.013 \pm 0.027 \pm 0.005, \text{ and} \\ B^+ &\to \pi^+\pi^0 \ : \text{-}0.081 \pm 0.054 \pm 0.008, \end{aligned}$$

where the first uncertainties are statistical and the second are systematic. These branching fraction correspond to a signal yield of $2052\pm57$ for $B^+ \to K^0\pi^+$, $1547 \pm 45$ for $B^+ \to K^+\pi^0$, and $785 \pm 44$ for $B^+ \to \pi^+\pi^0$.

We combine our results with measurements of $B^0 \to K^+\pi^-$ and $B^0 \to K^0\pi^0$ conducted on the same data set [90, 91] to calculate an isospin-based sum rule to test the standard model. In the calculation, we take (anti-)correlations between common systematic uncertainties such as the tracking uncertainty and the number of produced $B$ mesons into account. We obtain a result of

$$\Delta(K\pi) = -0.03 \pm 0.13 \pm 0.04,$$

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

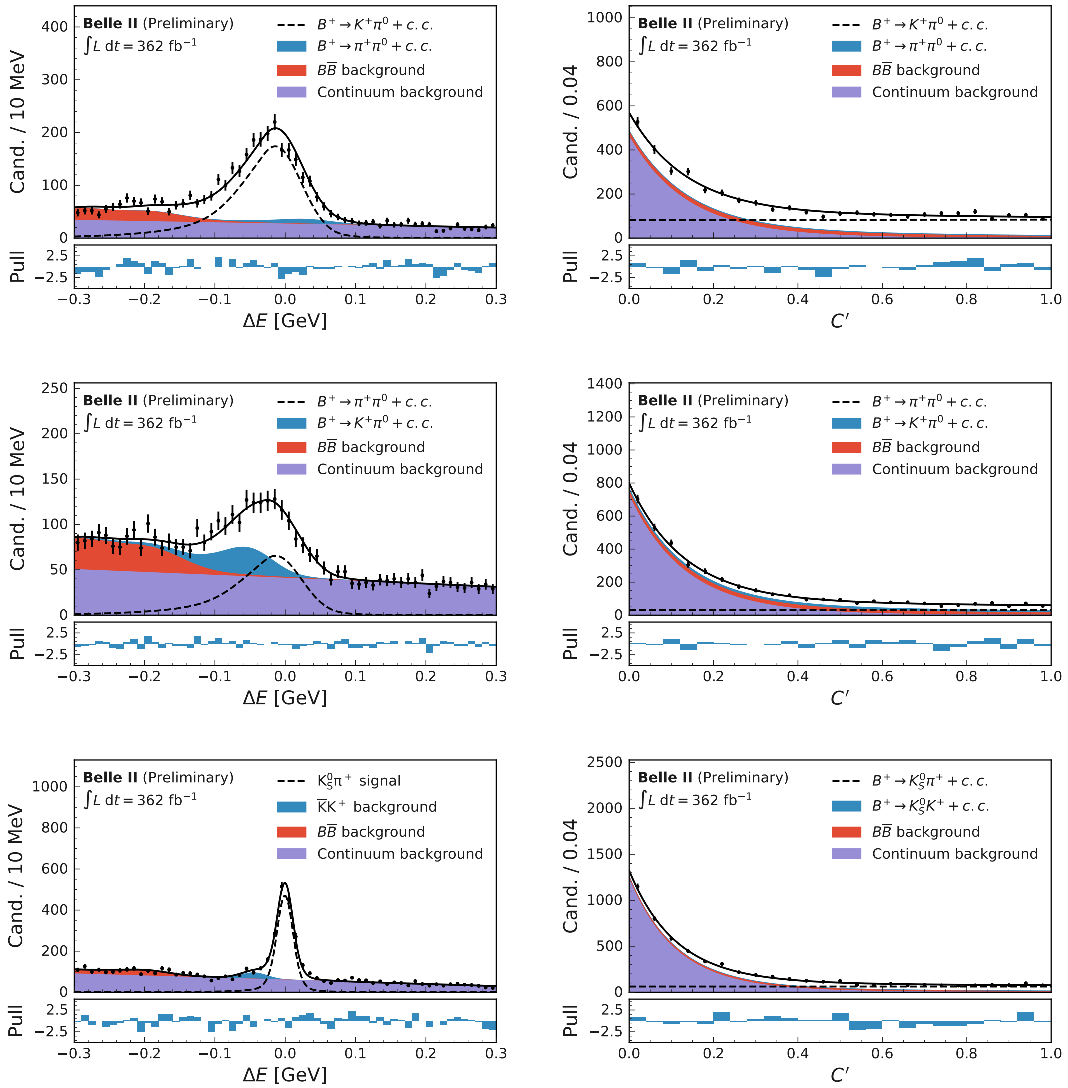


Figure 4.33: The $\Delta E$ (left) and transformed continuum suppression output $C'$ (right) distribution (black dots with error bars) for $B^+ \to K^+\pi^0$ (Top), $B^+ \to \pi^+\pi^0$ (center), and $B^+ \to K^0\pi^+$ (bottom). The result of a fit to the sample is shown as a solid black line. Individual fit components are shown as black dashed line (signal), blue shaded area (peaking background), red shaded area ($B\overline{B}$ background), and purple shaded area (continuum background). Below differences between data and the fit results, normalized by fit uncertainties, (pulls) are shown.

# Chapter 5

# Conclusion

Due to the smallness of the CKM matrix element $|V_{\mathrm{ub}}|$, tree-level amplitudes for $B \to K\pi$ and $B \to \pi\pi$ decays are suppressed in the standard model of particle physics (SM). Hence, these decays receive non-negligible contribution from loop-level amplitudes and are therefor sensitive to contributions from non-SM physics. Measurements of these decays are currently limited by the statistical precision.

The Belle II experiment started data taking in 2019 and collected a data set containing 387 million bottom-antibottom meson pairs before the 2022 shutdown. This data set is approximately half of that collected by its predecessor Belle. Using this data set, we measure the branching fraction and direct CP asymmetry for $B^+ \to K^0\pi^+$, $B^+ \to K^+\pi^0$, and $B^+ \to \pi^+\pi^0$ decays. A key challenge is the treatment of background from $e^+e^- \to q\overline{q}(q = u, d, s, c)$, which is the dominant background source. To address this, discriminating observables are identified and fed into a multivariate algorithm. Physics parameter are determined by a two-dimensional maximum-likelihood fit to the unbinned transformed output of the multivariate algorithm and the energy difference between the center-of-mass beam energy and the energy of the reconstructed signal $B$ candidate.

We obtain branching fractions of

$$
\begin{aligned}
B^+ \to K^0\pi^+ &: (24.37 \pm 0.71 \pm 0.86)\times 10^{-6}, \\
B^+ \to K^+\pi^0 &: (13.93 \pm 0.38 \pm 0.71)\times 10^{-6}, \text{ and} \\
B^+ \to \pi^+\pi^0 &: \ (5.10 \pm 0.29 \pm 0.27)\times 10^{-6}.
\end{aligned}
$$

and direct CP asymmetries of

$B^+ \to K^0\pi^+$ : 0.046 ± 0.029 ± 0.007,
$B^+ \to K^+\pi^0$ : 0.013 ± 0.027 ± 0.005, and
$B^+ \to \pi^+\pi^0$ : -0.081 ± 0.054 ± 0.008,

where the first uncertainties are statistical and the second are systematic. To test the standard model, we combine our results with Belle II measurements of $B^0 \to K^+\pi^-$ and $B^0 \to K^0\pi^0$ conducted on the same data set to evaluate an isospin-based sum rule. We find

$$\Delta(K\pi) = -0.03 \pm 0.13 \pm 0.04$$

for the sum rule, consistent with the standard model expectation of zero.

The determination of the isospin-based sum rule and the individual input measurements are competitive with the best current determinations. The determination of the sum rule and the direct CP asymmetries are still statistically limited. The branching fraction measurements are dominated by systematic uncertainties but the dominant contributions are expected to decrease in the future with increasing data set size. The Belle II experiment is still in its startup phase. Following the shutdown from 2022-2024, data taking has resumed, and the experiment is increasing its data set. In the future, further measurements of $B \to \pi\pi$ and $B \to K\pi$ decays by the Belle II experiment will be crucial to test the SM. This is especially true for decay channels involving neutral particles, which are difficult or impossible to access at other experiments.

# Bibliography


[1] D. Griffiths, *Introduction to Elementary Particles.* New York, USA: John Wiley & Sons, 1987.

[2] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory.* Reading, USA: Addison-Wesley, 1995.

[3] M. Thomson, *Modern Particle Physics.* Cambridge University Press, 2013.

[4] K. G. Wilson, "Confinement of quarks," *Phys. Rev. D*, vol. 10, pp. 2445–2459, Oct 1974.

[5] S. Bilenky, "Neutrino oscillations: From a historical perspective to the present status," *Nuclear Physics B*, vol. 908, pp. 2–13, 2016.

[6] Wikimedia Foundation, "Standard model of elementary particles." `https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg`, 2019. Accessed: 2025-06-02.

[7] G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Physics Letters B*, vol. 716, no. 1, pp. 1–29, 2012.

[8] S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Physics Letters B*, vol. 716, no. 1, pp. 30–61, 2012.

[9] S. Navas *et al.*, "Review of particle physics," *Phys. Rev. D*, vol. 110, no. 3, p. 030001, 2024.

[10] X. Fan, T. G. Myers, B. A. D. Sukra, and G. Gabrielse, "Measurement of the Electron Magnetic Moment," *Phys. Rev. Lett.*, vol. 130, no. 7, p. 071801, 2023.

[11] V. C. Rubin and W. K. Ford, Jr, "Rotation of the Andromeda nebula from a spectroscopic survey of emission regions," *Astrophys. J. 159: 379-403(Feb 1970).*, 12 1969.

[12] S. Weinberg, *Cosmology.* Oxford University Press, 2008.

[13] N. Aghanim *et al.*, "Planck2018 results: VI. Cosmological parameters," *Astronomy & Astrophysics*, vol. 641, p. A6, Sept. 2020.

[14] M. Dine and A. Kusenko, "Origin of the matter-antimatter asymmetry," *Rev. Mod. Phys.*, vol. 76, pp. 1–30, Dec 2003.

[15] M. Gell-Mann, "Isotopic spin and new unstable particles," *Phys. Rev.*, vol. 92, pp. 833–834, Nov 1953.

[16] N. Cabibbo, "Unitary symmetry and leptonic decays," *Phys. Rev. Lett.*, vol. 10, pp. 531–533, Jun 1963.

[17] S. L. Glashow, J. Iliopoulos, and L. Maiani, "Weak interactions with lepton-hadron symmetry," *Phys. Rev. D*, vol. 2, pp. 1285–1292, Oct 1970.

[18] J. E. Augustin *et al.*, "Discovery of a narrow resonance in $e^+e^-$ annihilation," *Phys. Rev. Lett.*, vol. 33, pp. 1406–1408, Dec 1974.

[19] J. J. Aubert *et al.*, "Experimental observation of a heavy particle $J$," *Phys. Rev. Lett.*, vol. 33, pp. 1404–1406, Dec 1974.

[20] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, "Experimental test of parity conservation in beta decay," *Phys. Rev.*, vol. 105, pp. 1413–1415, Feb 1957.

[21] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the $2\pi$ decay of the $K_2^0$ meson," *Phys. Rev. Lett.*, vol. 13, pp. 138–140, Jul 1964.

[22] M. Kobayashi and T. Maskawa, "CP-Violation in the renormalizable theory of weak interaction," *Progress of Theoretical Physics*, vol. 49, pp. 652–657, 02 1973.

[23] L. Wolfenstein, "Parametrization of the Kobayashi-Maskawa Matrix," *Physical Review Letters*, vol. 51, pp. 1945–1947, Nov. 1983.

[24] S. W. Lin *et al.*, "Difference in direct charge-parity violation between charged and neutral $B$ meson decays," *Nature*, vol. 452, pp. 332–335, 2008.

[25] B. Aubert *et al.*, "Direct $CP$ violating asymmetry in $B^0 \to K^+\pi^-$ decays," *Phys. Rev. Lett.*, vol. 93, p. 131801, Sep 2004.

[26] Y. Chao *et al.*, "Evidence for direct $CP$ violation in $B^0 \to K^+\pi^-$decays," *Phys. Rev. Lett.*, vol. 93, p. 191802, Nov 2004.

[27] Y.-Y. Charng and H.-N. Li, "Determining weak phases from the $B \to \pi\pi$, $K\pi$ decays," *Phys. Rev. D*, vol. 71, p. 014036, Jan 2005.

[28] H.-n. Li, S. Mishima, and A. I. Sanda, "Resolution to the $B \to \pi K$ puzzle," *Phys. Rev. D*, vol. 72, p. 114005, Dec 2005.

[29] N. B. Beaudry, A. Datta, D. London, A. Rashed, and J.-S. Roux, "The $B \to K\pi$ puzzle revisited," *Journal of High Energy Physics*, vol. 2018, no. 1, p. 74, 2018.

[30] S. Mishima and T. Yoshikawa, "Large electroweak penguin contribution in $B \to K\pi$ and $\pi\pi$ decay modes," *Phys. Rev. D*, vol. 70, p. 094024, Nov 2004.

[31] V. Barger, C.-W. Chiang, P. Langacker, and H.-S. Lee, "Solution to the $B \to K\pi$ puzzle in a flavor-changing Z' model," *Physics Letters B*, vol. 598, no. 3, pp. 218–226, 2004.

[32] M. Burgos Marcos, M. Reboud, and K. K. Vos, "Detailed $SU(3)$ Flavour Symmetry Analysis of Charmless Two-Body $B$-Meson Decays Including Factorizable Corrections," 4 2025.

[33] G. Bell, M. Beneke, T. Huber, and X.-Q. Li, “Two-loop current–current operator contribution to the non-leptonic QCD penguin amplitude,” *Physics Letters B*, vol. 750, pp. 348–355, 2015.

[34] M. Gronau and J. L. Rosner, “Combining CP asymmetries in $B \to K\pi$ decays,” *Phys. Rev. D*, vol. 59, p. 113002, Apr 1999.

[35] Y. T. Duh *et al.*, “Measurements of branching fractions and direct CP asymmetries for B→K$\pi$, B→$\pi\pi$ and B→KK decays,” *Phys. Rev. D*, vol. 87, no. 3, p. 031103, 2013.

[36] R. Aaij *et al.*, “Measurement of CP Violation in the Decay $B^+ \to K^+\pi^0$,” *Phys. Rev. Lett.*, vol. 126, no. 9, p. 091802, 2021.

[37] R. Aaij *et al.*, “Branching fraction and CP asymmetry of the decays $B^+ \to K^0_{\mathrm{S}}\pi^+$ and $B^+ \to K^0_{\mathrm{S}}K^+$,” *Phys. Lett. B*, vol. 726, pp. 646–655, 2013.

[38] B. Aubert *et al.*, “Study of $B^0 \to \pi^0\pi^0$, $B^\pm \to \pi^\pm\pi^0$, and $B^\pm \to K^\pm\pi^0$ Decays, and Isospin Analysis of $B \to \pi\pi$ Decays,” *Phys. Rev. D*, vol. 76, p. 091102, 2007.

[39] B. Aubert *et al.*, “Observation of $B^+ \to \bar{K}^0K^+$ and $B^0 \to K^0\bar{K}^0$,” *Phys. Rev. Lett.*, vol. 97, p. 171805, 2006.

[40] I. I. Y. Bigi, V. A. Khoze, N. G. Uraltsev, and A. I. Sanda, “The Question of CP Noninvariance - as Seen Through the Eyes of Neutral Beauty,” *Adv. Ser. Direct. High Energy Phys.*, vol. 3, pp. 175–248, 1989.

[41] M. Gronau and D. London, “Isospin analysis of CP asymmetries in B decays,” *Phys. Rev. Lett.*, vol. 65, pp. 3381–3384, Dec 1990.

[42] J. Charles, A. Höcker, H. Lacker, S. Laplace, F. R. Le Diberder, J. Malclés, J. Ocariz, M. Pivk, and L. Roos, “CP violation and the CKM matrix: assessing the impact of the asymmetric B factories,” *The European Physical Journal C*, vol. 41, p. 1–131, May 2005.

[43] A. J. Bevan *et al.*, “The physics of the B factories,” *The European Physical Journal C*, vol. 74, Nov. 2014.

[44] S. Amato *et al.*, "LHCb technical proposal: A Large Hadron Collider Beauty Experiment for Precision Measurements of CP Violation and Rare Decays," *CERN-LHCC-98-04*, Feb 1998.

[45] T. Abe *et al.*, "Belle II technical design report," *arXiv:1011.0352*, 2010.

[46] K. Akai, K. Furukawa, and H. Koiso, "SuperKEKB collider," *Nucl. Instrum. Meth. A*, vol. 907, pp. 188–199, 2018. Advances in Instrumentation and Experimental Methods (Special Issue in Honour of Kai Siegbahn).

[47] P. Raimondi, "New developments in Super B-factories," *Conf. Proc. C*, vol. 070625, p. 32, 2007.

[48] A. Olchevski, ed., *High-energy physics. Proceedings, European School, ESHEP'99, Casta-Papiernicka, Slovak Republic, August 22-September 4, 1999*, CERN Yellow Reports: School Proceedings, 6 2000.

[49] The Belle II collaboration, "Belle II Luminosity webpage." `https://www.belle2.org/research/luminosity/`, 2011. Accessed: 2025-06-02.

[50] E. Kou *et al.*, "The Belle II Physics Book," *Progress of Theoretical and Experimental Physics*, vol. 2019, Dec. 2019.

[51] T. Hara, T. Kuhr, and Y. Ushiroda, "Belle II coordinate system and guideline of Belle II numbering scheme." `https://indico.mpp.mpg.de/event/2308/contributions/4092/attachments/3414/3799/Belle2NumberingScheme.pdf`, 2011. Accessed: 2025-06-02.

[52] J. Kemmer and G. Lutz, "New detector concepts," *Nucl. Instrum. Meth. A*, vol. 253, no. 3, pp. 365–377, 1987.

[53] G. Giakoustidis *et al.*, "Status of the BELLE II Pixel Detector," *PoS*, vol. Pixel2022, p. 005, 2023.

[54] Q. Liu *et al.*, "Operational Experience and Performance of the Belle II Pixel Detector," *JPS Conf. Proc.*, vol. 34, p. 010002, 2021.

[55] F. Abudinén *et al.*, “Precise measurement of the $D^0$ and $D^+$ lifetimes at Belle II,” *Phys. Rev. Lett.*, vol. 127, p. 211801, Nov 2021.

[56] B. Wang *et al.*, “Operational experience of the Belle II pixel detector,” *Nucl. Instrum. Meth. A*, vol. 1032, p. 166631, 2022.

[57] K. Adamczyk *et al.*, “The design, construction, operation and performance of the Belle II silicon vertex detector,” *Journal of Instrumentation*, vol. 17, p. P11042, Nov. 2022.

[58] N. Taniguchi, “Central Drift Chamber for Belle II,” *Journal of Instrumentation*, vol. 12, p. C06014, jun 2017.

[59] C. Pulvermacher, “dE/dx particle identification and pixel detector data reduction for the Belle II experiment.” IEKP-KA/2012-09, 2012. Karlsruhe Institute of Technology (KIT).

[60] Y. Iwasaki, B. Cheon, E. Won, X. Gao, L. Macchiarulo, K. Nishimura, and G. Varner, “Level 1 trigger system for the Belle II experiment,” *IEEE Transactions on Nuclear Science*, vol. 58, no. 4, pp. 1807–1815, 2011.

[61] S. Lee, R. Itoh, T. Higuchi, M. Nakao, S. Y. Suzuki, and E. Won, “Belle II high level trigger at SuperKEKB,” *Journal of Physics: Conference Series*, vol. 396, p. 012029, dec 2012.

[62] T. Kuhr, C. Pulvermacher, M. Ritter, T. Hauth, and N. Braun, “The Belle II core software: Belle II framework software group,” *Computing and Software for Big Science*, vol. 3, Nov. 2018.

[63] The Belle II Collaboration, “Belle II analysis software framework (basf2).” Zenodo. DOI:10.5281/zenodo.14710811, Jan. 2025.

[64] R. Brun and F. Rademakers, “ROOT: An object oriented data analysis framework,” *Nucl. Instrum. Meth. A*, vol. 389, pp. 81–86, 1997.

[65] R. Brun *et al.*, “root-project/root: v6.18/02.” Zenodo. DOI:10.5281/zenodo.3895860, Aug. 2019.

[66] D. J. Lange, "The EvtGen particle decay simulation package," *Nucl. Instrum. Meth. A*, vol. 462, no. 1, pp. 152–155, 2001. BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.

[67] S. Jadach, B. Ward, and Z. Was, "The precision Monte Carlo event generator KK for two-fermion final states in $e^+e^-$ collisions," *Computer Physics Communications*, vol. 130, no. 3, pp. 260–325, 2000.

[68] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, "An introduction to PYTHIA 8.2," *Computer Physics Communications*, vol. 191, pp. 159–177, 2015.

[69] "Photos — a universal Monte Carlo for QED radiative corrections in decays," *Computer Physics Communications*, vol. 66, no. 1, pp. 115–128, 1991.

[70] S. Agostinelli *et al.*, "Geant4—a simulation toolkit," *Nucl. Instrum. Meth. A*, vol. 506, no. 3, pp. 250–303, 2003.

[71] Z. Liptak *et al.*, "Measurements of beam backgrounds in SuperKEKB phase 2," *Nucl. Instrum. Meth. A*, vol. 1040, p. 167168, 2022.

[72] J.-F. Krohn, F. Tenchini, P. Urquijo, F. Abudinén, S. Cunliffe, T. Ferber, M. Gelb, J. Gemmler, P. Goldenzweig, T. Keck, I. Komarov, T. Kuhr, L. Ligioi, M. Lubej, F. Meier, F. Metzner, C. Pulvermacher, M. Ritter, U. Tamponi, and A. Zupanc, "Global decay chain vertex fitting at Belle II," *Nucl. Instrum. Meth. A*, vol. 976, p. 164269, 2020.

[73] P. Avery, "Applied fitting theory III: non-optimal least squares fitting and multiple scattering," *CLEO Note CBX91-74, CLEO*, 1991.

[74] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.

[75] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree," in *Advances*

*in Neural Information Processing Systems* (I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, eds.), vol. 30, Curran Associates, Inc., 2017.

[76] F. Abudinén *et al.*, "B-flavor tagging at Belle II," *The European Physical Journal C*, vol. 82, Apr. 2022.

[77] G. C. Fox and S. Wolfram, "Observables for the analysis of event shapes in $e^+e^-$ annihilation and other processes," *Phys. Rev. Lett.*, vol. 41, pp. 1581–1585, Dec 1978.

[78] K. Abe *et al.*, "Measurement of branching fractions for $B \to \pi\pi$, $K\pi$, and $KK$ decays," *Phys. Rev. Lett.*, vol. 87, p. 101801, Aug 2001.

[79] D. M. Asner *et al.*, "Search for exclusive charmless hadronic B decays," *Phys. Rev. D*, vol. 53, pp. 1039–1050, Feb 1996.

[80] E. Farhi, "Quantum chromodynamics test for jets," *Phys. Rev. Lett.*, vol. 39, pp. 1587–1588, Dec 1977.

[81] J. Eschle, A. Puig Navarro, R. Silva Coutinho, and N. Serra, "zfit: Scalable pythonic fitting," *SoftwareX*, vol. 11, p. 100508, Jan. 2020.

[82] F. James and M. Roos, "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations," *Comput. Phys. Commun.*, vol. 10, pp. 343–367, 1975.

[83] M. Oreglia *et al.*, "Study of the reaction $\psi' \to \gamma\gamma\frac{J}{\psi}$," *Phys. Rev. D*, vol. 25, pp. 2259–2277, May 1982.

[84] K. Cranmer, "Kernel estimation in high-energy physics.," *Computer Physics Communications*, vol. 136, no. 3, pp. 198–207, 2001.

[85] O. Behnke, K. Kröninger, T. Schörner-Sadenius, and G. Schott, eds., *Data Analysis in High Energy Physics.* John Wiley & Sons, Ltd, 2013.

[86] B. Efron, "Bootstrap Methods: Another Look at the Jackknife," *The Annals of Statistics*, vol. 7, no. 1, pp. 1 – 26, 1979.

[87] M. Pivk and F. L. Diberder, “sPlot: A statistical tool to unfold data distributions,” *Nucl. Instrum. Meth. A*, vol. 555, no. 1-2, pp. 356–369, 2005.

[88] Belle II Collaboration, “Systematic corrections framework.” `https://syscorrfw.readthedocs.io/en/latest/`, 2024. Accessed: 2025-04-11.

[89] S. Choudhury *et al.*, “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*, vol. 107, p. L031102, Feb 2023.

[90] I. Adachi *et al.*, “Measurement of branching fractions and direct $CP$ asymmetries for $B \to K\pi$ and $B \to \pi\pi$ decays at Belle II,” *Phys. Rev. D*, vol. 109, p. 012001, Jan 2024.

[91] I. Adachi *et al.*, “Measurement of $CP$ Violation in $B^0 \to K^0_S\pi^0$ Decays at Belle II,” *Phys. Rev. Lett.*, vol. 131, p. 111803, Sep 2023.

# List of Figures

# List of Tables

# Appendix A

# Appendix

## A.1 Comparison of Continuum-Suppression Training Variable Distributions between Data and Simulation

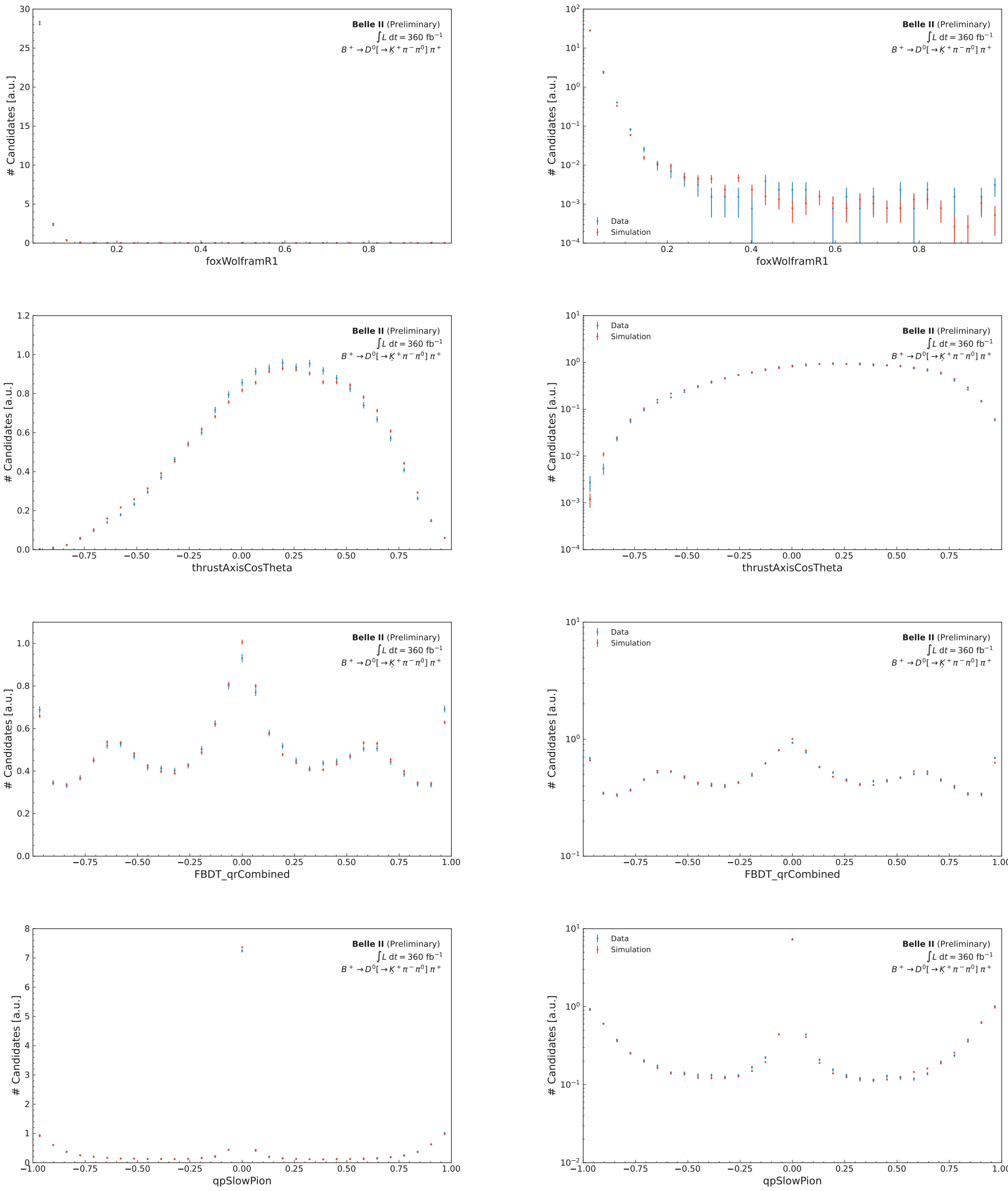


Figure A.1: Comparison between data (blue dots) and simulation (red dots) for the BDT input variables in the $B^+ \to D^0[\to K^+\pi^-\pi^0]\,\pi^+$ control channel.

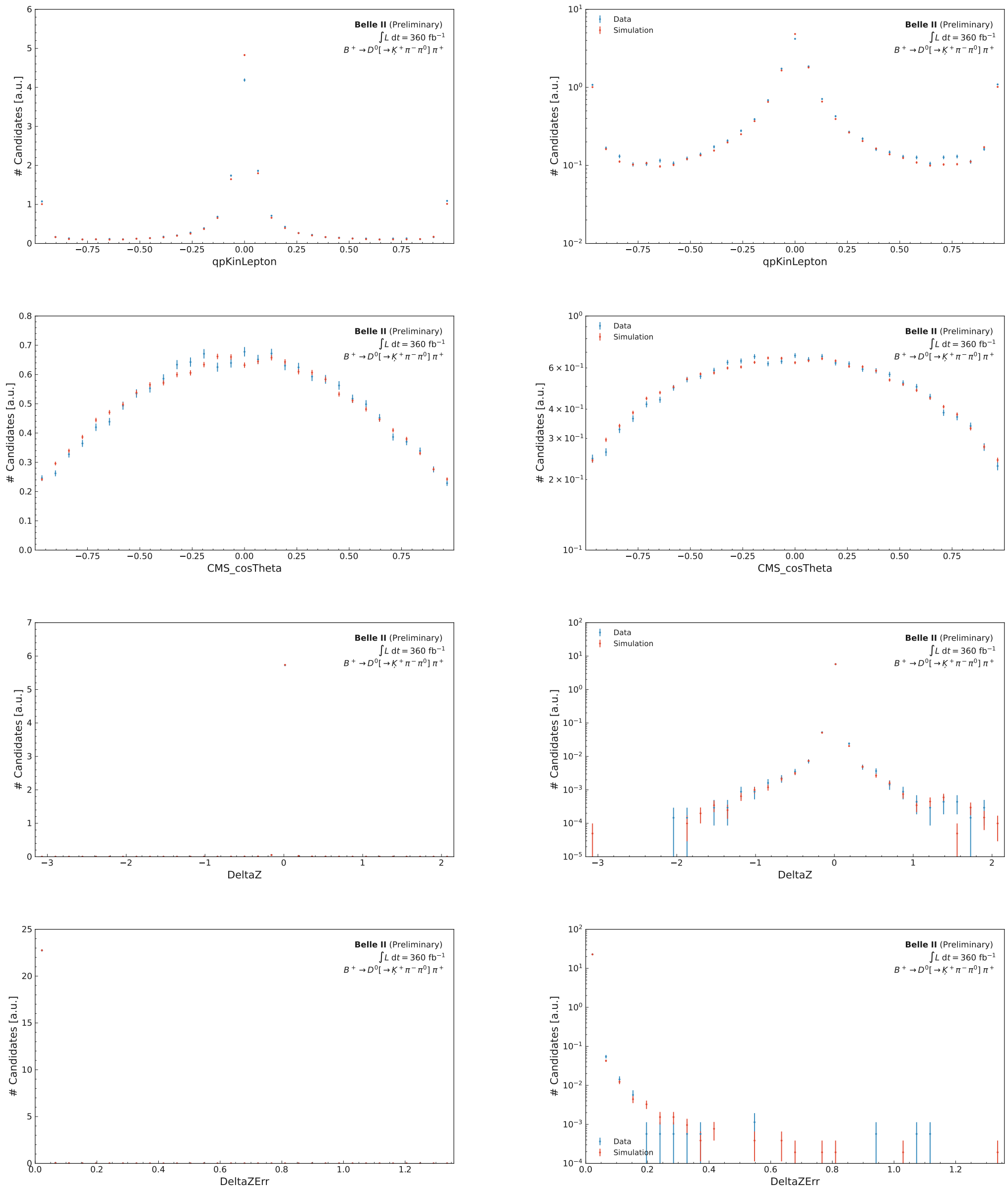


Figure A.2: Comparison between data (blue dots) and simulation (red dots) for the continuum suppression input variables in the $B^+ \to D^0[\to K^+\pi^-\pi^0]\ \pi^+$ control channel.

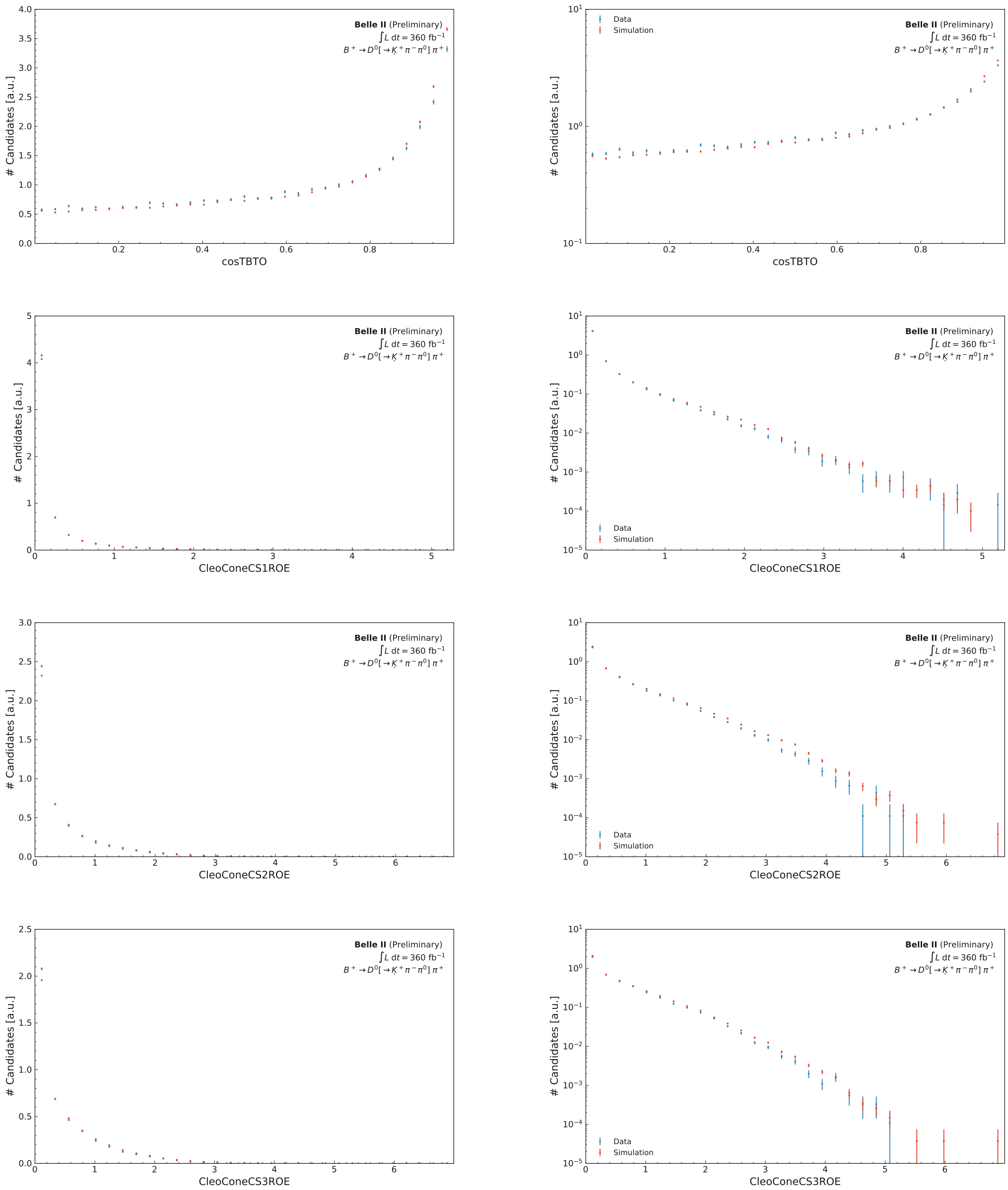


Figure A.3: Comparison between data (blue dots) and simulation (red dots) for the continuum suppression input variables in the $B^+ \to D^0[\to K^+\pi^-\pi^0]\ \pi^+$ control channel.

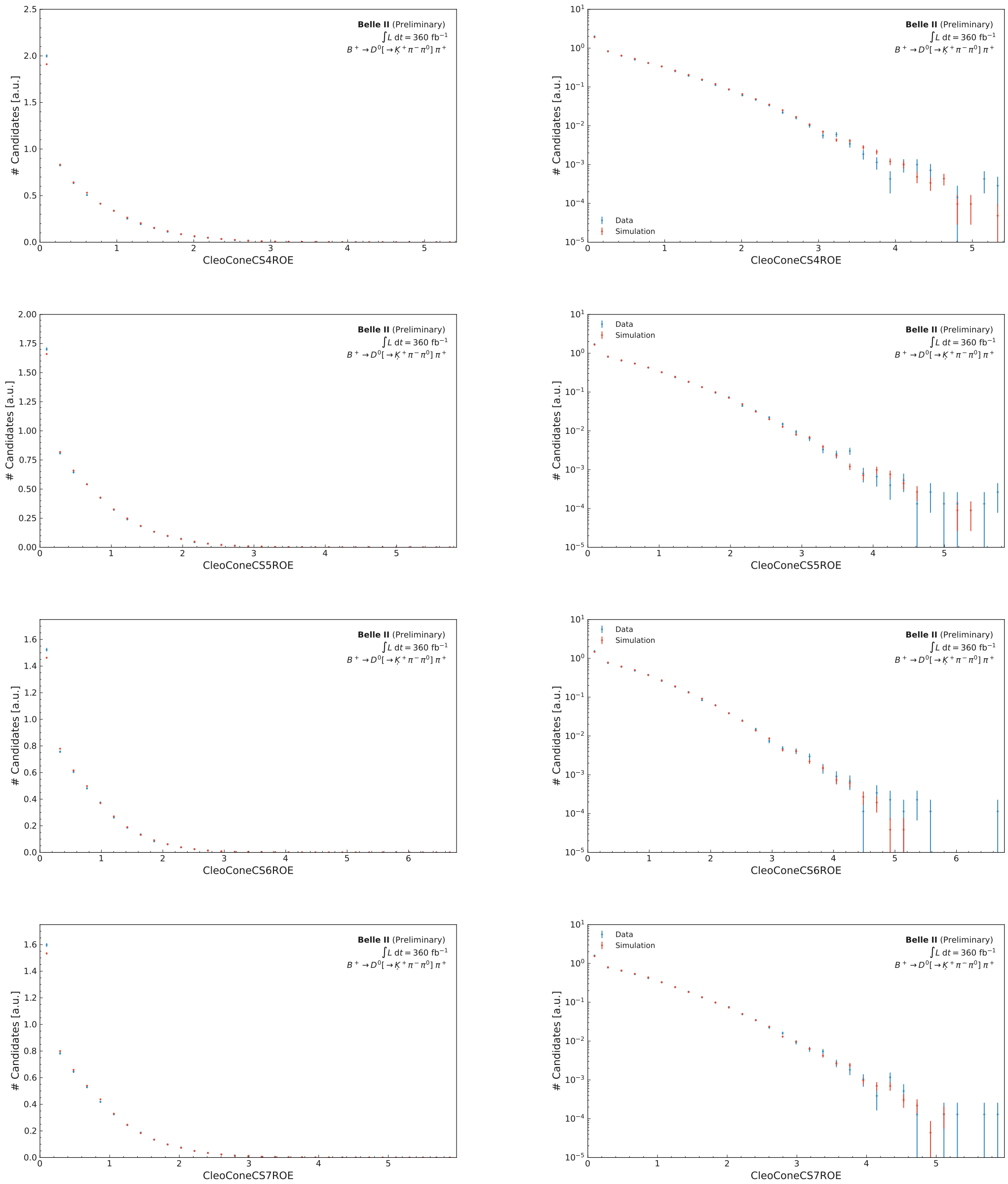


Figure A.4: Comparison between data (blue dots) and simulation (red dots) for the continuum suppression input variables in the $B^+ \to D^0[\to K^+\pi^-\pi^0]\ \pi^+$ control channel.

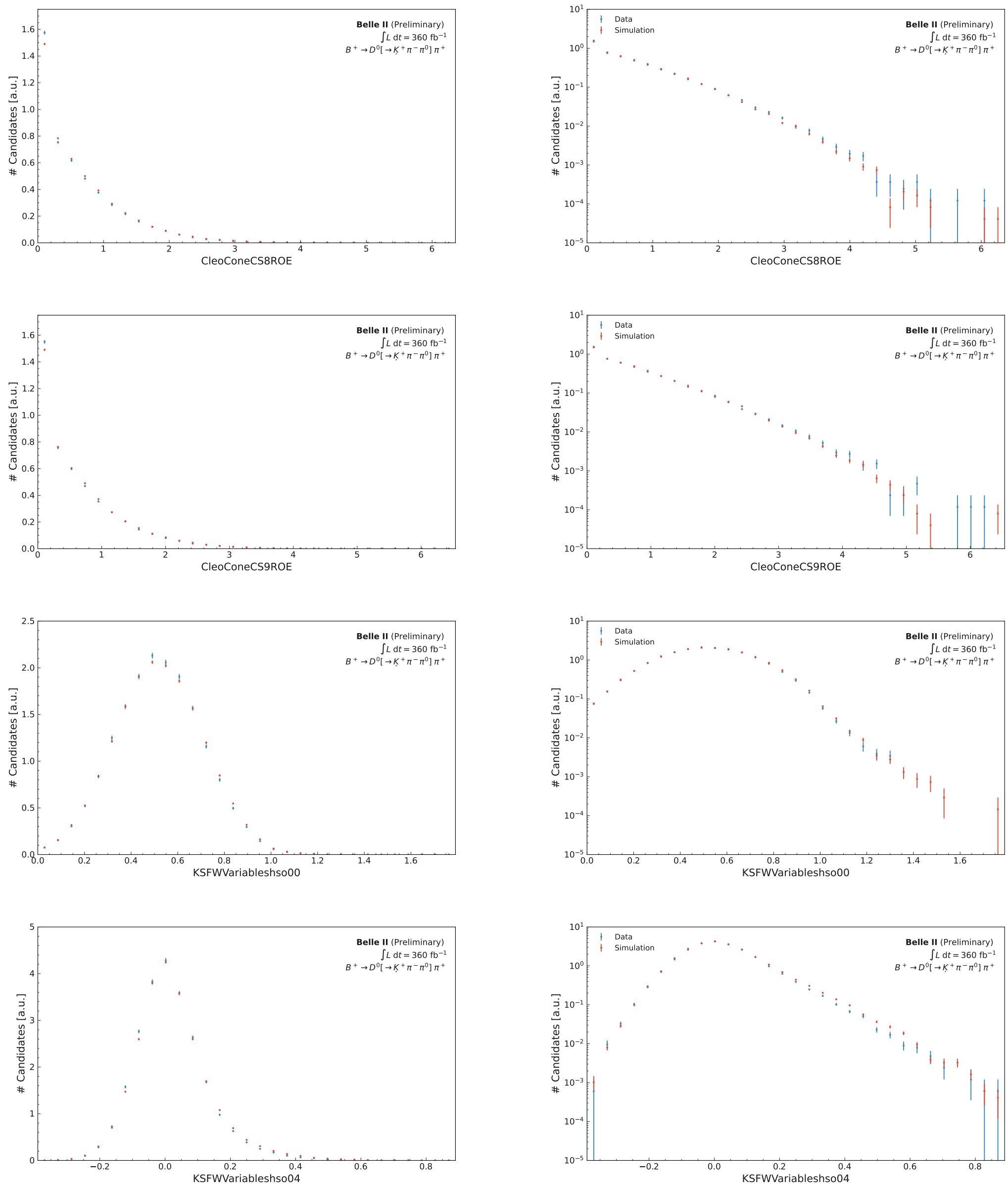


Figure A.5: Comparison between data (blue dots) and simulation (red dots) for the continuum suppression input variables in the $B^+ \to D^0[\to K^+\pi^-\pi^0]\;\pi^+$ control channel.

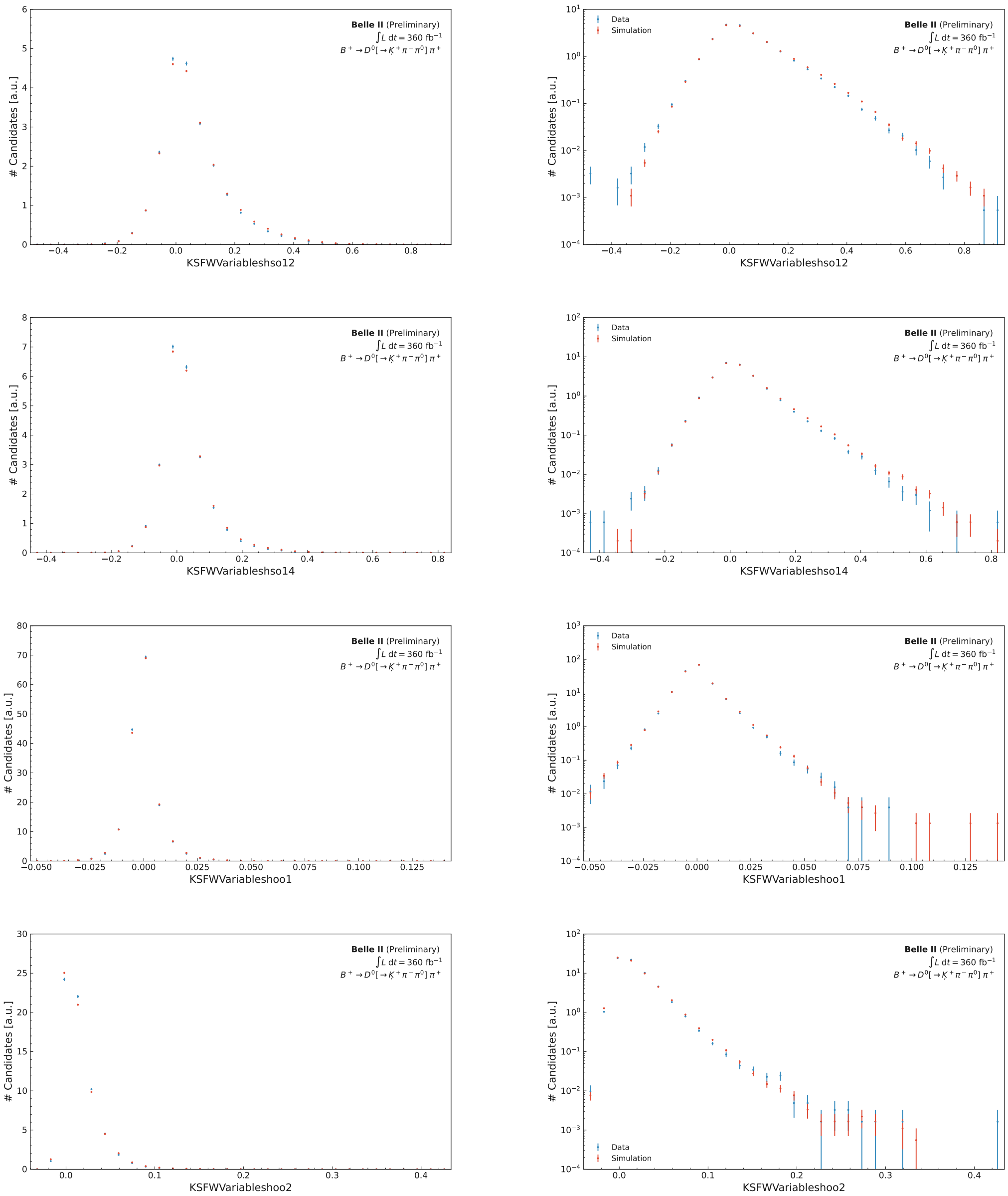


Figure A.6: Comparison between data (blue dots) and simulation (red dots) for the continuum suppression input variables in the $B^+ \to D^0[\to K^+\pi^-\pi^0]\ \pi^+$ control channel.

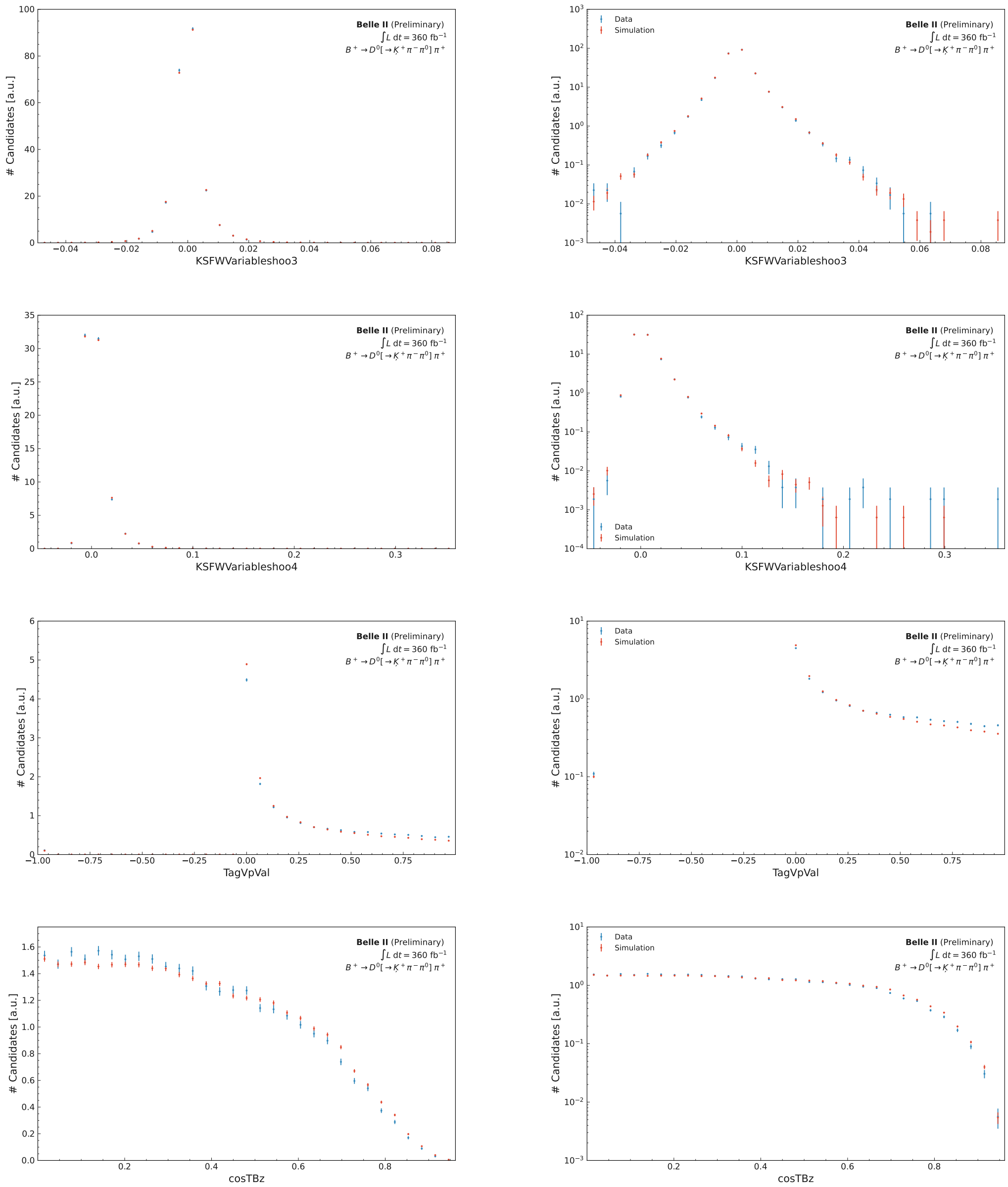


Figure A.7: Comparison between data (blue dots) and simulation (red dots) for the continuum suppression input variables in the $B^+ \to D^0[\to K^+\pi^-\pi^0]\ \pi^+$ control channel.

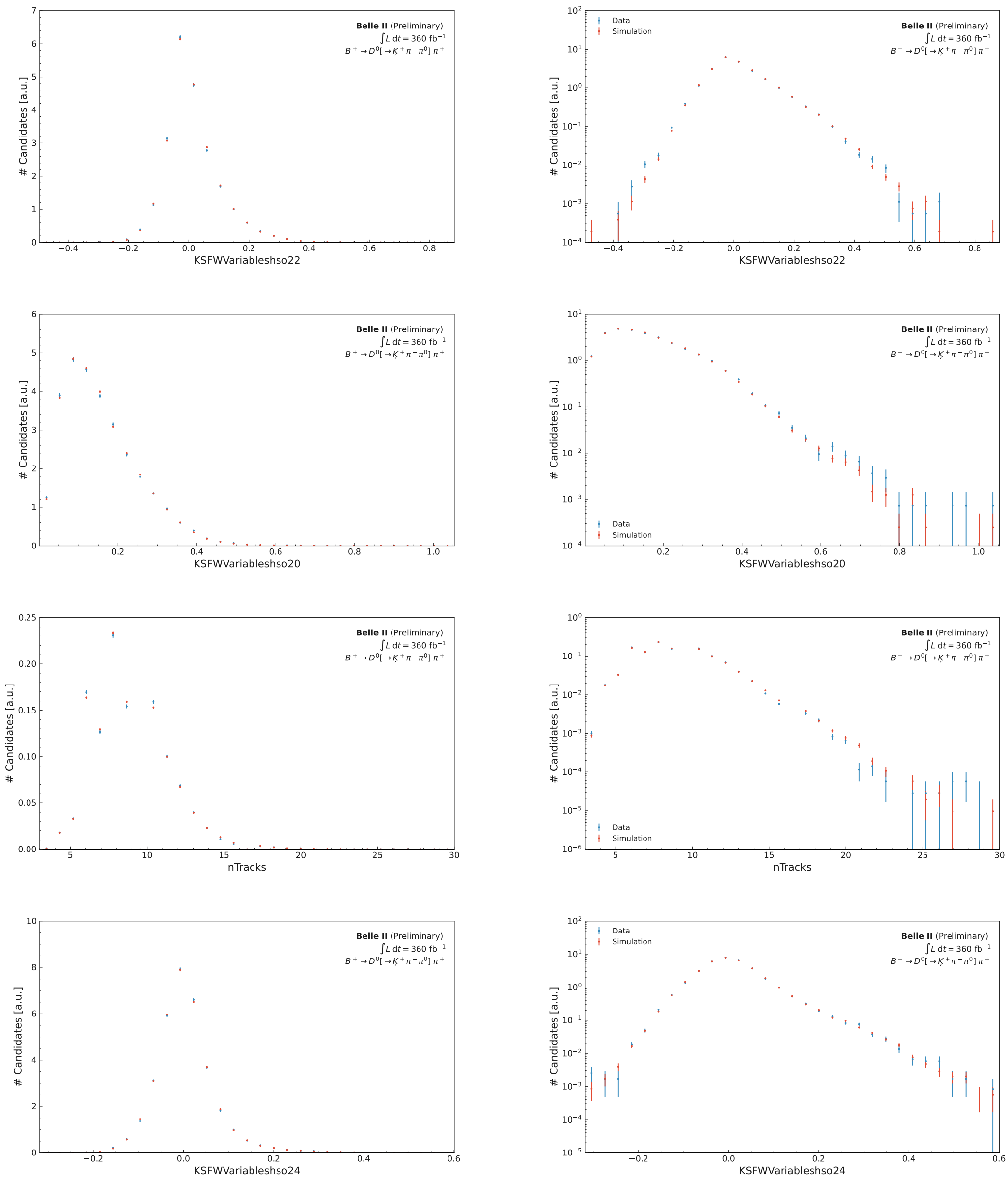


Figure A.8: Comparison between data (blue dots) and simulation (red dots) for the continuum suppression input variables in the $B^+ \to D^0[\to K^+\pi^-\pi^0]\ \pi^+$ control channel.

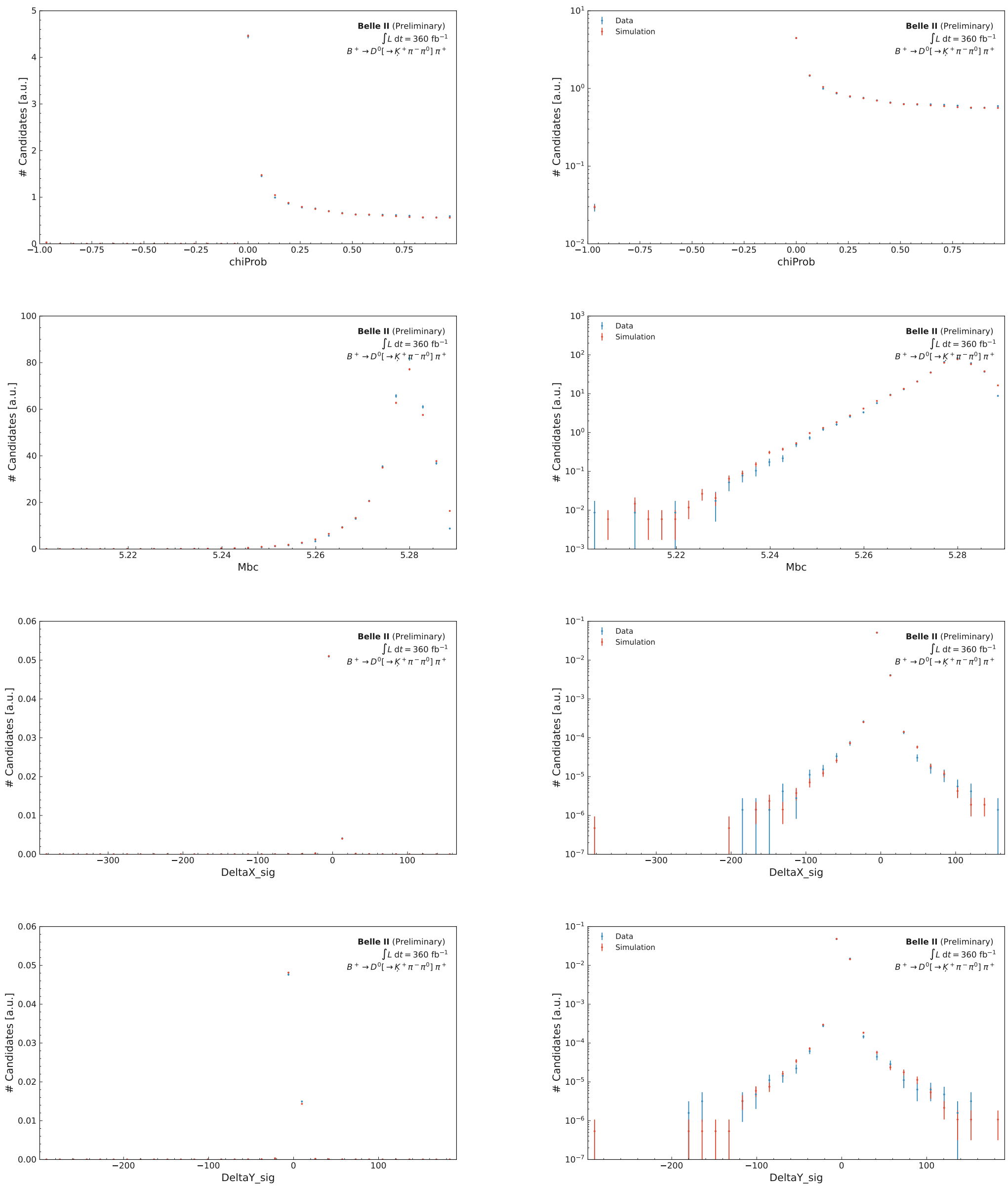


Figure A.9: Comparison between data (blue dots) and simulation (red dots) for the continuum suppression input variables in the $B^+ \to D^0[\to K^+\pi^-\pi^0]\ \pi^+$ control channel.

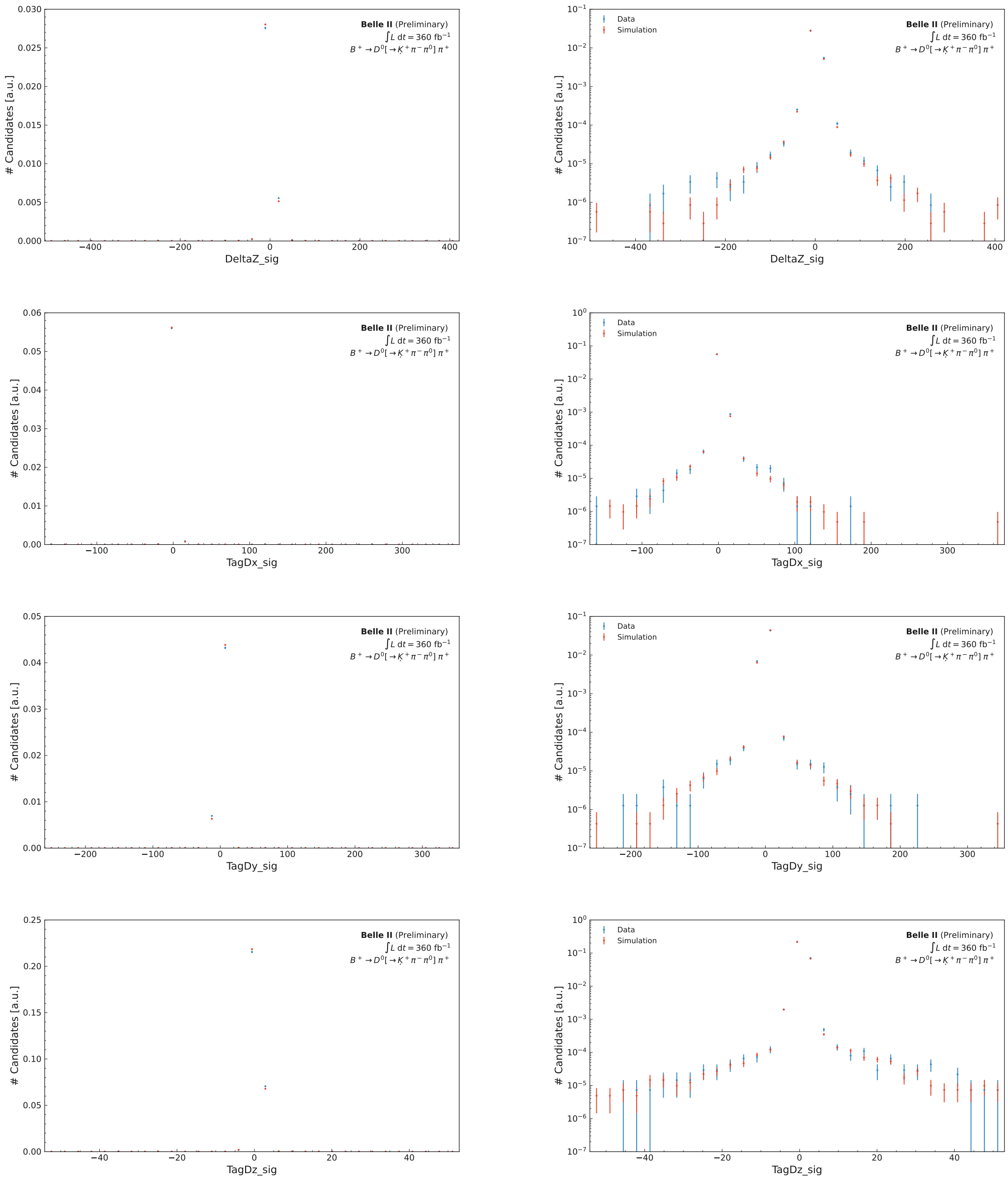


Figure A.10: Comparison between data (blue dots) and simulation (red dots) for the continuum suppression input variables in the $B^+ \to D^0[\to K^+\pi^-\pi^0]\,\pi^+$ control channel.

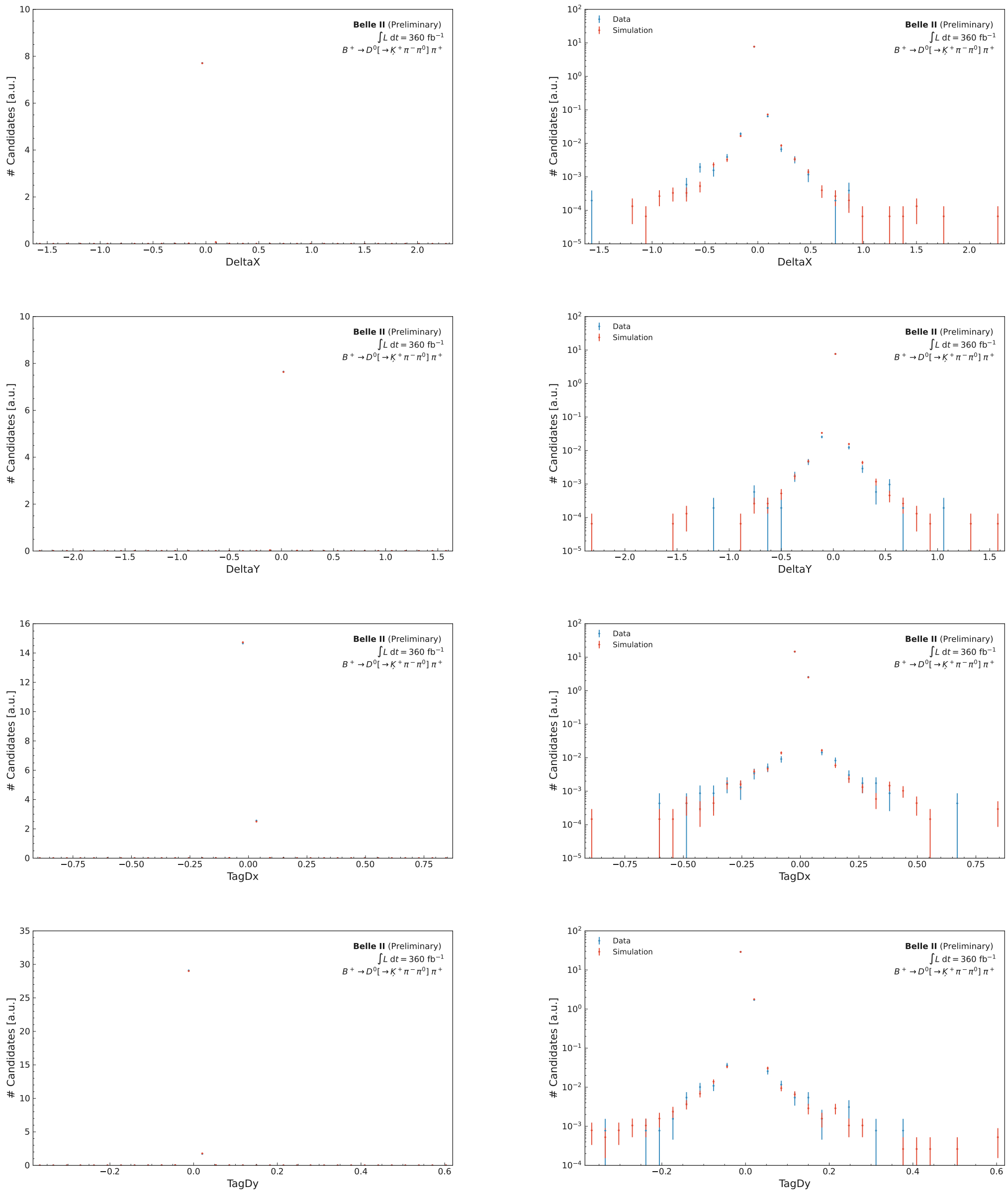


Figure A.11: Comparison between data (blue dots) and simulation (red dots) for the continuum suppression input variables in the $B^+ \to D^0[\to K^+\pi^-\pi^0]\,\pi^+$ control channel.

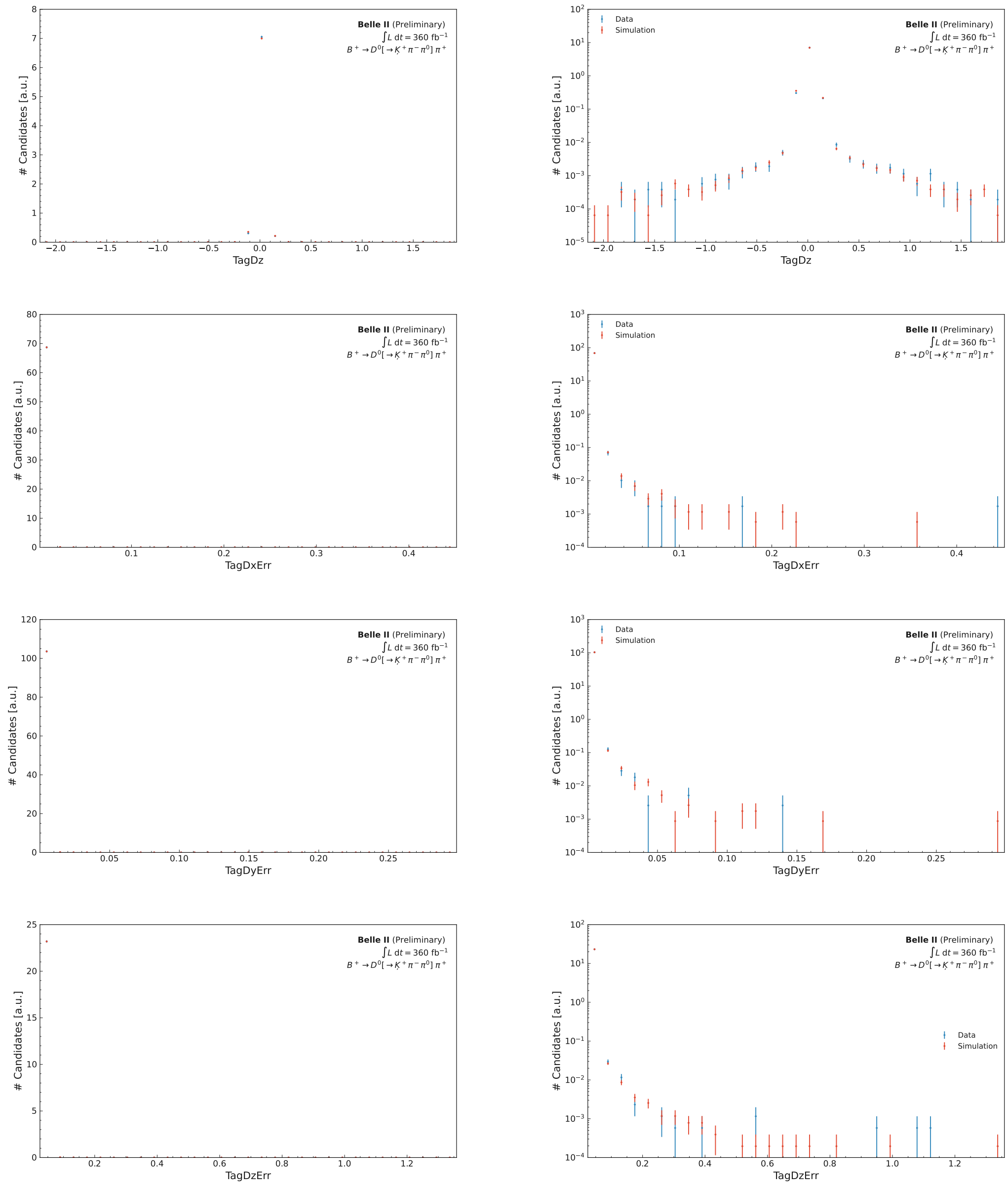


Figure A.12: Comparison between data (blue dots) and simulation (red dots) for the continuum suppression input variables in the $B^+ \to D^0[\to K^+\pi^-\pi^0]\,\pi^+$ control channel.

## A.2 Correlation between Fit Parameter

Table A.1: Correlations between fit parameter for a fit to 1000 $\mathrm{fb}^{-1}$ of simulated data for $B^+ \to K^0\pi^+$.

| | $\mathcal{B}_{K^0_S\pi^+}$ | $\mathcal{A}_{K^0_S\pi^+}$ | $B\overline{B}^+$ yield | $K^0_S\pi^+$ yield | $q\overline{q}^+$ yield | $q\overline{q}$ $C'$ a |
|---|---|---|---|---|---|---|
| $\mathcal{B}_{K^0_S\pi^+}$ | 1.00 | -0.01 | 0.03 | -0.10 | -0.08 | -0.01 |
| $\mathcal{A}_{K^0_S\pi^+}$ | | 1.00 | -0.01 | 0.11 | 0.07 | -0.00 |
| $B\overline{B}^+$ yield | | | 1.00 | 0.02 | -0.30 | -0.03 |
| $K^0_S\pi^+$ yield | | | | 1.00 | -0.12 | -0.02 |
| $q\overline{q}^+$ yield | | | | | 1.00 | 0.03 |
| $q\overline{q}$ $C'$ a | | | | | | 1.00 |

| | $q\overline{q}$ $C'$ b | $q\overline{q}$ $\Delta E$ 1st param | $B\overline{B}^-$ yield | $K^0_S\pi^-$ yield | $q\overline{q}^-$ yield |
|---|---|---|---|---|---|
| $\mathcal{B}_{K^0_S\pi^+}$ | -0.03 | 0.02 | 0.03 | -0.11 | -0.08 |
| $\mathcal{A}_{K^0_S\pi^+}$ | 0.01 | -0.00 | 0.00 | -0.12 | -0.06 |
| $B\overline{B}^+$ yield | -0.13 | 0.21 | 0.12 | 0.04 | -0.08 |
| $K^0_S\pi^+$ yield | -0.04 | 0.05 | 0.04 | 0.01 | -0.03 |
| $q\overline{q}^+$ yield | 0.09 | -0.12 | -0.08 | -0.03 | 0.05 |
| $q\overline{q}$ $C'$ a | -0.86 | -0.02 | -0.03 | -0.02 | 0.03 |
| $q\overline{q}$ $C'$ b | 1.00 | -0.08 | -0.13 | -0.04 | 0.08 |
| qqbar $\Delta E$ 1st param | | 1.00 | 0.20 | 0.05 | -0.11 |
| $B\overline{B}^-$ yield | | | 1.00 | 0.02 | -0.29 |
| $K^0_S\pi^\mp$ yield | | | | 1.00 | -0.12 |
| $q\overline{q}^-$ yield | | | | | 1.00 |

Table A.2: Correlations between fit parameter for a fit to $1000\,\mathrm{fb}^{-1}$ of simulated data for $B^+ \to K^+\pi^0$ and $B^+ \to \pi^+\pi^0$.

| | $\mathcal{B}_{K^+\pi^0}$ | $\mathcal{A}_{K^+\pi^0}$ | $\mathcal{B}_{\pi^+\pi^0}$ | $\mathcal{A}_{\pi^+\pi^0}$ | $B\overline{B}^+_{K^+\pi^0}$ yield | $q\overline{q}^+_{K^+\pi^0}$ yield |
|---|---|---|---|---|---|---|
| $\mathcal{B}_{K^+\pi^0}$ | 1.00 | 0.01 | -0.22 | 0.01 | -0.00 | -0.19 |
| $\mathcal{A}_{K^+\pi^0}$ | | 1.00 | -0.00 | -0.23 | 0.09 | 0.11 |
| $\mathcal{B}_{\pi^+\pi^0}$ | | | 1.00 | -0.02 | 0.01 | 0.01 |
| $\mathcal{A}_{\pi^+\pi^0}$ | | | | 1.00 | -0.02 | -0.01 |
| $B\overline{B}^+_{K^+\pi^0}$ yield | | | | | 1.00 | -0.41 |
| $q\overline{q}^+_{K^+\pi^0}$ yield | | | | | | 1.00 |

| | $q\overline{q}_{K^+\pi^0}$ $C$ a | $q\overline{q}_{K^+\pi^0}$ $C$ b | $q\overline{q}_{K^+\pi^0}$ $\Delta E$ 1st param | $B\overline{B}^-_{K^+\pi^0}$ yield | $q\overline{q}^-_{K^+\pi^0}$ yield | $B\overline{B}^+_{\pi^+\pi^0}$ yield |
|---|---|---|---|---|---|---|
| $\mathcal{B}_{K^+\pi^0}$ | 0.01 | -0.10 | 0.10 | -0.01 | -0.19 | -0.01 |
| $\mathcal{A}_{K^+\pi^0}$ | 0.01 | -0.01 | 0.00 | -0.09 | -0.12 | 0.01 |
| $\mathcal{B}_{\pi^+\pi^0}$ | 0.00 | -0.00 | -0.01 | 0.01 | 0.01 | 0.00 |
| $\mathcal{A}_{\pi^+\pi^0}$ | 0.00 | -0.00 | 0.00 | 0.02 | 0.01 | 0.06 |
| $B\overline{B}^+_{K^+\pi^0}$ yield | -0.00 | -0.16 | 0.32 | 0.17 | -0.16 | 0.00 |
| $q\overline{q}^+_{K^+\pi^0}$ yield | -0.00 | 0.16 | -0.25 | -0.16 | 0.15 | 0.01 |
| $q\overline{q}_{K^+\pi^0}$ $C$ a | 1.00 | -0.89 | -0.00 | 0.02 | -0.02 | 0.00 |
| $q\overline{q}_{K^+\pi^0}$ $C$ b | | 1.00 | -0.13 | -0.18 | 0.18 | 0.00 |
| $q\overline{q}_{K^+\pi^0}$ $\Delta E$ 1st param | | | 1.00 | 0.32 | -0.25 | -0.00 |
| $B\overline{B}^-_{K^+\pi^0}$ yield | | | | 1.00 | -0.40 | -0.00 |
| $q\overline{q}^-_{K^+\pi^0}$ yield | | | | | 1.00 | -0.00 |
| $B\overline{B}^+_{\pi^+\pi^0}$ yield | | | | | | 1.00 |

| | $q\overline{q}^+_{\pi^+\pi^0}$ yield | $q\overline{q}_{\pi^+\pi^0}$ $C$ a | $q\overline{q}_{\pi^+\pi^0}$ $C$ b | $q\overline{q}_{\pi^+\pi^0}$ $\Delta E$ 1st param | $B\overline{B}^-_{\pi^+\pi^0}$ yield | $q\overline{q}^-_{\pi^+\pi^0}$ yield |
|---|---|---|---|---|---|---|
| $\mathcal{B}_{K^+\pi^0}$ | 0.01 | -0.00 | 0.01 | 0.00 | -0.01 | 0.01 |
| $\mathcal{A}_{K^+\pi^0}$ | -0.01 | -0.00 | -0.00 | 0.00 | -0.01 | 0.01 |
| $\mathcal{B}_{\pi^+\pi^0}$ | -0.19 | 0.00 | -0.11 | 0.09 | 0.01 | -0.18 |
| $\mathcal{A}_{\pi^+\pi^0}$ | 0.13 | 0.01 | 0.00 | -0.00 | -0.07 | -0.11 |
| $B\overline{B}^+_{K^+\pi^0}$ yield | -0.00 | -0.00 | -0.00 | 0.00 | -0.00 | -0.00 |
| $q\overline{q}^+_{K^+\pi^0}$ yield | 0.01 | 0.00 | 0.00 | -0.00 | 0.00 | 0.00 |
| $q\overline{q}_{K^+\pi^0}$ $C$ a | -0.00 | 0.00 | -0.00 | 0.00 | -0.00 | -0.00 |
| $q\overline{q}_{K^+\pi^0}$ $C$ b | 0.00 | 0.00 | 0.00 | -0.00 | 0.00 | 0.00 |
| $q\overline{q}_{K^+\pi^0}$ $\Delta E$ 1st param | -0.00 | -0.00 | -0.00 | 0.00 | -0.00 | -0.00 |
| $B\overline{B}^-_{K^+\pi^0}$ yield | -0.00 | 0.00 | -0.00 | 0.00 | 0.00 | -0.00 |
| $q\overline{q}^-_{K^+\pi^0}$ yield | 0.00 | 0.00 | 0.00 | -0.00 | 0.01 | 0.01 |
| $B\overline{B}^+_{\pi^+\pi^0}$ yield | -0.31 | -0.01 | -0.11 | 0.27 | 0.12 | -0.10 |
| $q\overline{q}^+_{\pi^+\pi^0}$ yield | 1.00 | 0.01 | 0.11 | -0.20 | -0.10 | 0.09 |
| $q\overline{q}_{\pi^+\pi^0}$ $C$ a | | 1.00 | -0.89 | -0.01 | -0.00 | -0.00 |
| $q\overline{q}_{\pi^+\pi^0}$ $C$ b | | | 1.00 | -0.10 | -0.12 | 0.12 |
| $q\overline{q}_{\pi^+\pi^0}$ $\Delta E$ 1st param | | | | 1.00 | 0.27 | -0.20 |
| $B\overline{B}^-_{\pi^+\pi^0}$ yield | | | | | 1.00 | -0.32 |
| $q\overline{q}^-_{\pi^+\pi^0}$ yield | | | | | | 1.00 |

## A.3 Normalized Residuals for Background Yields

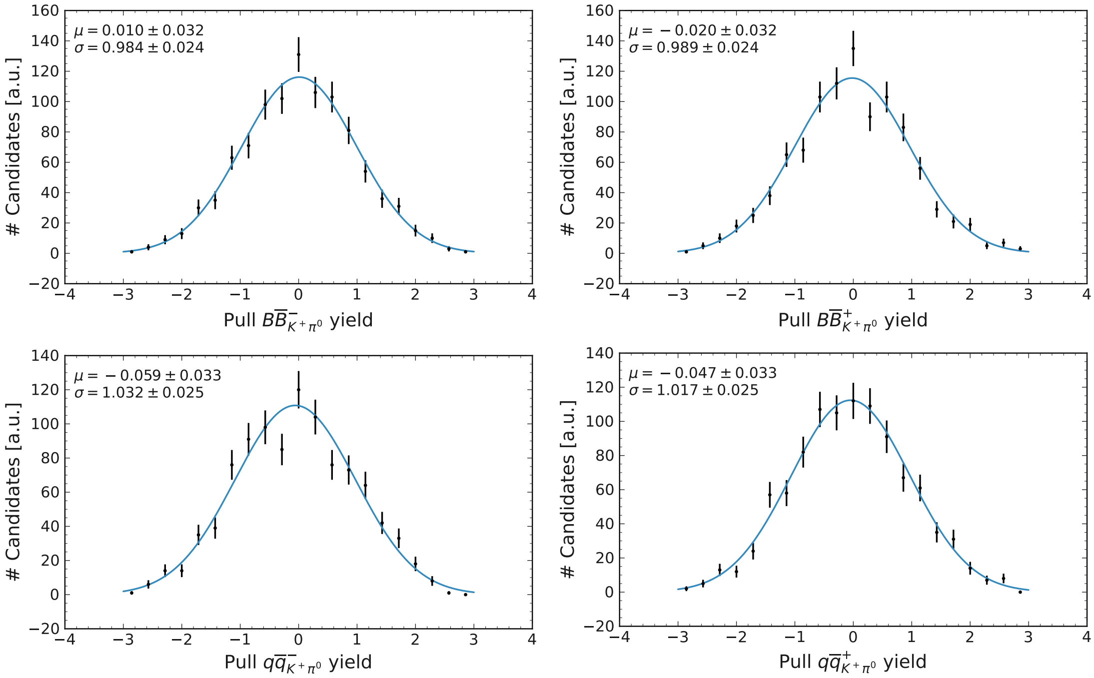


Figure A.13: Top: Distribution of differences between the fit result and simulated data, normalized by fit uncertainties, (pull distribution) for the $B\overline{B}$ background yield for negative (left) and positive (right) charge in the $B^{+} \to K^{+}\pi^{0}$ sample. Bottom: Same but for the continuum background yields.

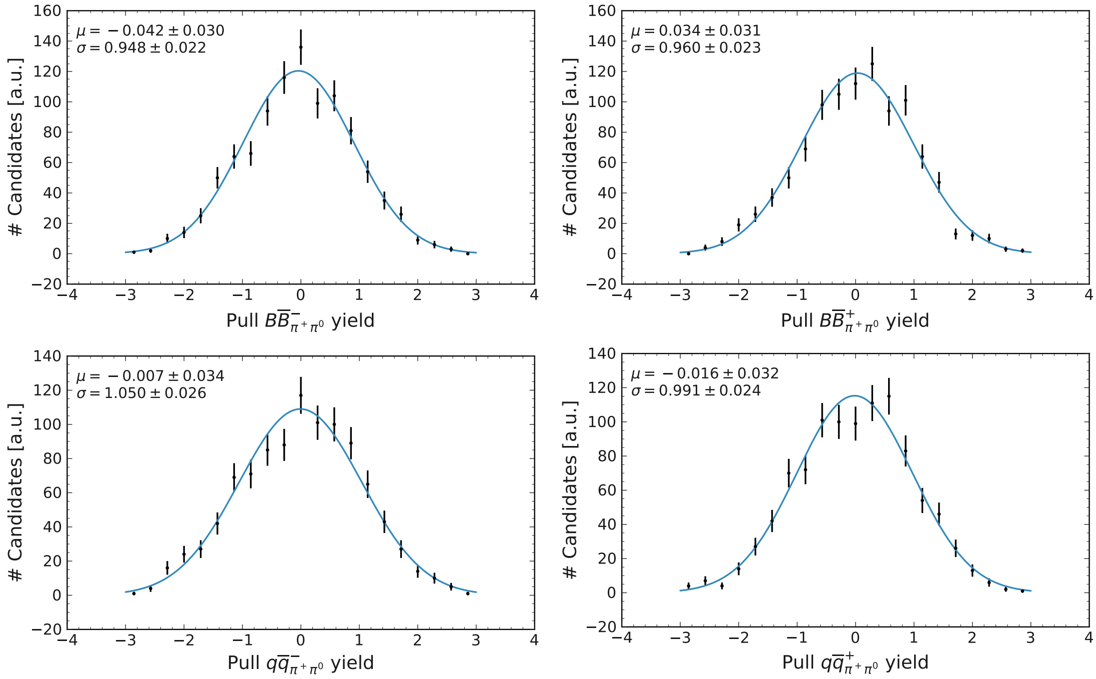


Figure A.14: Top: Distribution of differences between the fit result and simulated data, normalized by fit uncertainties, (pull distribution) for the $B\overline{B}$ background yield for negative (left) and positive (right) charge in the $B^+ \to \pi^+\pi^0$ sample. Bottom: Same but for the continuum background yields.

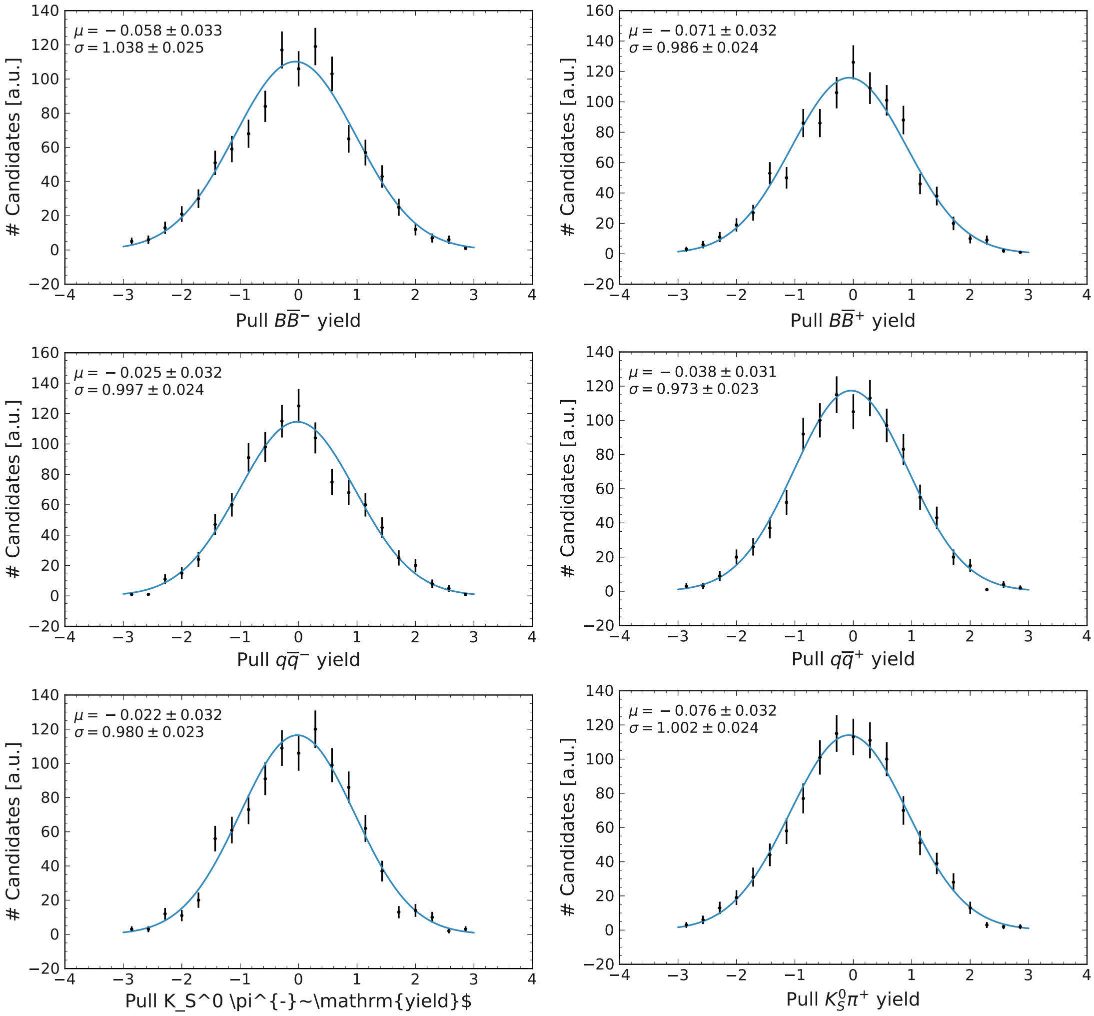


Figure A.15: Top: Distribution of differences between the fit result and simulated data, normalized by fit uncertainties, (pull distribution) for the $B\overline{B}$ background yield for negative (left) and positive (right) charge in the $B^+ \to K_S^0\pi^+$ sample. Bottom: Same but for the continuum background yields.

# Acknowledgments

I would like to thank everyone who made my time in the Belle II collaboration and the Max Planck Institute for Physics great and enjoyable. In particular I would like to thank:

Allen for giving me the opportunity to pursue my PhD at the Max Planck Institute.

Hans-Günther for the day-to-day supervision and continuous guidance throughout my research.

Thibaud and Daniel for the fruitful physics discussions that consistently deepened my understanding and improved my work.

Diego and Mirco for their supervision within Belle II's BtoHadronic group, and for their support in overcoming the pitfalls that can arise when publishing an analysis.

My various office mates, Thomas, Lukas, Martin, and Godo, for making every day at the office enjoyable and productive.

All the people that made the time at KEK and Japan so memorable, Björn, Chia-Ling, Moritz, Philipp, Riccardo, Shu-Ping, and all the others.

Finally, I would like to thank my parents and sister, who never stopped supporting me.