Technische Universität München
TUM School of Natural Sciences

Measurement of Branching Fractions and Direct CP Violation in $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$ Decays at the Belle II Experiment

Benedikt Wach

Vollständiger Abdruck der von der TUM School of Natural Sciences der Technischen Universität München zur Erlangung eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Andreas Weiler

Prüfer der Dissertation:

1. Hon.-Prof. Dr. Allen C. Caldwell
2. Priv.-Doz. Dr. Jan Friedrich

Die Dissertation wurde am 23.10.2023 bei der Technischen Universität München eingereicht und durch die TUM School of Natural Sciences am 08.12.2023 angenommen.

Measurement of Branching Fractions and Direct CP Violation in $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$ Decays at the Belle II Experiment

Benedikt Wach

Technische Universität München

Max-Planck-Institut für Physik

München 2023

Abstract

We present measurements of branching fractions of $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ decays and the direct CP-violation parameter in $B^0 \to K^+\pi^-$ decays. The analyzed data contain 387×10^6 bottom-antibottom meson pairs generated in e^+e^- collisions at the $\Upsilon(4S)$ resonance produced by the SuperKEKB asymmetric-energy collider and are collected with the Belle II detector. We obtain

$$\begin{aligned}
\mathcal{B}(B^0 \to \pi^+\pi^-) &= (5.83 \pm 0.22\,(\text{stat}) \pm 0.17\,(\text{syst})) \times 10^{-6}\ , \\
\mathcal{B}(B^0 \to K^+\pi^-) &= (20.67 \pm 0.37\,(\text{stat}) \pm 0.62\,(\text{syst})) \times 10^{-6}\ , \\
\mathcal{A}_{CP}(B^0 \to K^+\pi^-) &= -0.072 \pm 0.019\,(\text{stat}) \pm 0.007\,(\text{syst})\ .
\end{aligned}$$

The values are consistent with current global averages and demonstrate precision comparable to the best determinations available, despite employing a substantially smaller dataset. The measurement of $\mathcal{B}(B^0 \to \pi^+\pi^-)$ represents the most precise determination to date by a single experiment.

We combine our findings with measurements of related $B \to K\pi$ transitions to test the Standard Model employing a sum rule based on isospin symmetry. We find

$$I_{K\pi} = -0.03 \pm 0.13 \pm 0.04\ ,$$

consistent with the theoretical Standard Model prediction of zero.

Kurzfassung

Wir präsentieren Messungen der Verzweigungsverhältnisse von $B^0 \to \pi^+\pi^-$ und $B^0 \to K^+\pi^-$ Zerfällen sowie der direkten CP-Verletzung in $B^0 \to K^+\pi^-$ Zerfällen. Die ausgewertete Datenmenge enthält 387×10^6 Bottom-Antibottom-Mesonenpaare, die in e^+e^--Kollisionen an der $\Upsilon(4S)$-Resonanz durch den SuperKEKB Collider produziert, und mithilfe des Belle II Detektors aufgezeichnet wurden. Wir messen

$$\begin{aligned}
\mathcal{B}(B^0 \to \pi^+\pi^-) &= (5.83 \pm 0.22\,(\text{stat}) \pm 0.17\,(\text{syst})) \times 10^{-6}\ , \\
\mathcal{B}(B^0 \to K^+\pi^-) &= (20.67 \pm 0.37\,(\text{stat}) \pm 0.62\,(\text{syst})) \times 10^{-6}\ , \\
\mathcal{A}_{CP}(B^0 \to K^+\pi^-) &= -0.072 \pm 0.019\,(\text{stat}) \pm 0.007\,(\text{syst})\ .
\end{aligned}$$

Die Ergebnisse sind konsistent mit den aktuellen globalen Durchschnittswerten. Die Präzision ist trotz der Verwendung einer signifikant geringeren Datenmenge mit jener der derzeit genauesten experimentellen Bestimmungen vergleichbar. Die Messung von $\mathcal{B}(B^0 \to \pi^+\pi^-)$ stellt gegenwärtig die genaueste Einzelmessung weltweit dar.

Wir verwenden die Ergebnisse in Kombination mit Messungen verwandter $B \to K\pi$ Übergänge, um mittels einer auf Isospin-Symmetrie basierenden Summenregel das Standardmodell zu prüfen. Wir erhalten

$$I_{K\pi} = -0.03 \pm 0.13 \pm 0.04\ ,$$

in Übereinstimmung mit dem Erwartungswert des Standardmodells von Null.

Contents

1. Introduction

The Standard Model (SM) of particle physics is a theoretical framework that describes the observed fundamental constituents of nature and three of the four known forces that govern their interactions. It presents our current best understanding of particle dynamics at the subatomic level and has undergone rigorous experimental verification. While the theory has arguably achieved many triumphs, it leaves many fundamental questions of theoretical and phenomenological nature unanswered.

The SM lacks explanations for various phenomena, such as the appearance of precisely three quark-lepton families or the origins of observed hierarchies in their masses and the mixtures between them. In the context of the Big Bang Model, one puzzling observation at the cosmological scale is the abundance of matter in the Universe. In 1967, A. Sakharov postulated the violation of Charge-Parity (CP) symmetry, corresponding to differences between transitions of particles and antiparticles, to be a necessary condition for the evolution of a matter-antimatter asymmetry [1]. Within the Standard Model, CP violation is introduced in weak interactions via the Cabibbo-Kobayashi-Maskawa (CKM) matrix, which describes the mixing between different quark flavors. However, the amount predicted by the Standard Model is many orders of magnitude too small to explain the dominance of matter in the Universe [2].

Regarding its shortcomings, it seems evident, that the SM is merely an effective theory, able to successfully describe the phenomenology at energy scales probed up to now, but incomplete at higher scales. Nowadays, the majority of high-energy physics experiments therefore aims to find New Physics (NP) beyond the SM. To this end, various extension mechanisms have been proposed that typically involve the introduction of new massive particles into the existing framework.

On the so-called *energy frontier*, experiments at the Large-Hadron-Collider (LHC) aim to directly produce new particles in high-energy proton-proton collisions at center-of-mass energies of up to 14 TeV. The interaction between constituent gluons or quarks restricts the mass reach for new particles to few TeV/c^2. So far, direct searches have yielded null results in the pursuit of NP.

Rather than directly producing new particles, indirect searches at the *intensity frontier* aim to measure the decay properties of known particles with high precision to find deviations from SM predictions. The latter could indicate the presence of unobserved massive particles that contribute to transition amplitudes in quantum

loops. Assuming coupling strengths of $\mathcal{O}(1)$, these loop contributions render indirect searches sensitive to masses up to hundreds of TeV/c^2 [3].

The Belle II experiment, located at the next-generation e^+e^- collider SuperKEKB, focuses on precise measurements of quark-flavor dynamics, in particular in B-meson decays. Due to their comparatively long lifetimes and large predicted CP-violating effects, B mesons offer a rich phenomenology to test a wide range of SM extensions.

Specifically, B-meson decays to hadronic two-body final states without charm quarks are heavily suppressed in the SM and exhibit only small branching ratios of $\mathcal{O}(10^{-5})$ or less. Consequently, loop-diagram contributions to these transitions are often of leading order, rendering them particularly sensitive to non-SM physics. However, large uncertainties in the theoretical predictions, due to non-factorizable hadronic amplitudes that hinder the application of perturbation theory, spoil a straightforward interpretation of these measurements. To mitigate this limitation, dynamical symmetries such as isospin symmetry can be exploited to construct linear combinations of branching fractions and CP asymmetries in related decay modes.

This thesis presents the measurements of branching ratios of $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ decays[1] and the direct CP-violation parameter in $B^0 \to K^+\pi^-$ decays in data corresponding to an integrated luminosity of $362\,\text{fb}^{-1}$, recorded at the $\Upsilon(4S)$ resonance with the Belle II detector. The results are combined with those of related $B \to K\pi$ transitions, obtained on the same dataset, in an isospin-based sum rule that serves as a stringent test of the SM.

The outline of this work is as follows. Chapter 2 gives a brief outline of the Standard Model of particle physics, along with the theoretical foundation for CP-violation measurements in B-meson decays. In Chapter 3, the SuperKEKB collider and the Belle II experiment are described. Chapter 4 is devoted to the detailed description of the analysis of $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ decays, covering the techniques employed to measure the branching ratios and CP asymmetry. Chapter 5 is dedicated to the differences between Monte-Carlo simulations and real data, and the methods employed to mitigate them. In Chapter 6, the final results are presented, including a breakdown of all considered systematic uncertainties and a comparison with previous results. Chapter 7 concludes the presented study and provides an outlook for future measurements.

[1] The charge-conjugated transition is implied throughout this document unless stated otherwise.

2. The Standard Model and B-Meson Physics

2.1. The Standard Model of Particle Physics

The Standard Model (SM) of particle physics is a theory that describes the elementary particles and their interactions. It incorporates three of the four fundamental forces, currently believed to exist: the electromagnetic force, the weak nuclear force, and the strong nuclear force. Gravity is not explained by the SM and is instead described by the theory of general relativity. The SM was formulated on the principles of quantum field theory and local gauge symmetries. It has been remarkably successful in explaining a wide range of experimental observations and is currently our best understanding of the microscopic world. In the following, a short introduction to the model is given, starting with the description of its fundamental particles and some of the most important properties of their interactions. As CP violation is uniquely associated with the weak interaction, we put an emphasis on the latter and its unification with the electromagnetic force, the *electroweak unification*. We also give a short description of the Higgs mechanism, by which all massive particles obtain their masses in the SM. This chapter is largely based on derivations found in the standard works [4] and [5], and more specific references such as [6, 7].

2.1.1. Particles of the Standard Model

The SM classifies particles into two main categories, fermions and bosons. Fermions are particles with half-integer spins and follow the rules of Fermi-Dirac statistics. They are the building blocks of all visible matter and include *quarks* and *leptons*. Bosons are the force-carrying particles or *gauge bosons* and have integer-spins, obeying Bose-Einstein statistics. They mediate the fundamental forces of nature described by the SM by coupling to the charges of the corresponding interactions.

Fermions

Quarks come in six flavors, further classified into three generations: up- (u) and down-quarks (d), which make up the first generation and are the constituents of

protons and neutrons, charm- (c) and strange-quarks (s), and top- (t) and bottom-quarks (b), which are the heaviest and least stable. They carry fractional electron charges of $\frac{2}{3}$ or $-\frac{1}{3}$ as well as one of three color charges of the strong interaction.

Leptons, on the other hand, do not carry any color-charge and hence do not participate in the strong interaction. They can also be divided into three generations of progressively larger masses, each consisting of a charged lepton and its associated neutrino. The first generation includes the electron (e), carrying an electric charge of -1, and its corresponding, electrically-neutral neutrino, the electron-neutrino (ν_e). The second and third generations feature the muon (μ) and the muon-neutrino (ν_μ), and the tau lepton (τ) and its corresponding tau-neutrino (ν_τ).

In the SM, for each of the fermions, there exists a corresponding anti-particle with identical mass but opposite charge.

Bosons

The gauge bosons are spin-1 particles and act as force carriers between interacting particles. The photon (γ) is the mediator of the electromagnetic force. The W^+, W^- and Z bosons mediate the weak force, responsible for radioactive decays and neutrino interactions. The gluon (g) mediates the strong force, which binds quarks to hadrons. Finally there is the Higgs boson, the spin-0 boson of the scalar Higgs field, which lends masses to the other particles of the SM.

Figure 2.1 shows all particles of the SM, illustrating the mass hierarchy of the three generations of fermions.

2.1.2. Fundamental Forces and Gauge Symmetries

In the mathematical formulation of quantum field theory (QFT), the Lagrangian (density) is used to describe the dynamics of a system. Following the Hamilton principle of least action, we can derive field equations, like the Dirac equation, the solutions to which represent the quantum states of the physical system described by the field. In order to ensure a consistent description of the laws of physics, the Lagrangian is required to remain invariant under certain transformations. For example, any physical process must remain unaffected by a translation in space-time. These transformations are expressed by global symmetry groups, like the Poincaré group, required for all relativistic quantum field theories.

Following Noether's theorem, for every continuous global symmetry of the Lagrangian, there exists a conserved quantity associated with that symmetry [9]. In this sense, the conservation of quantities, such as electric charge, baryon number, and lepton number, can be understood in terms of the global symmetries of the

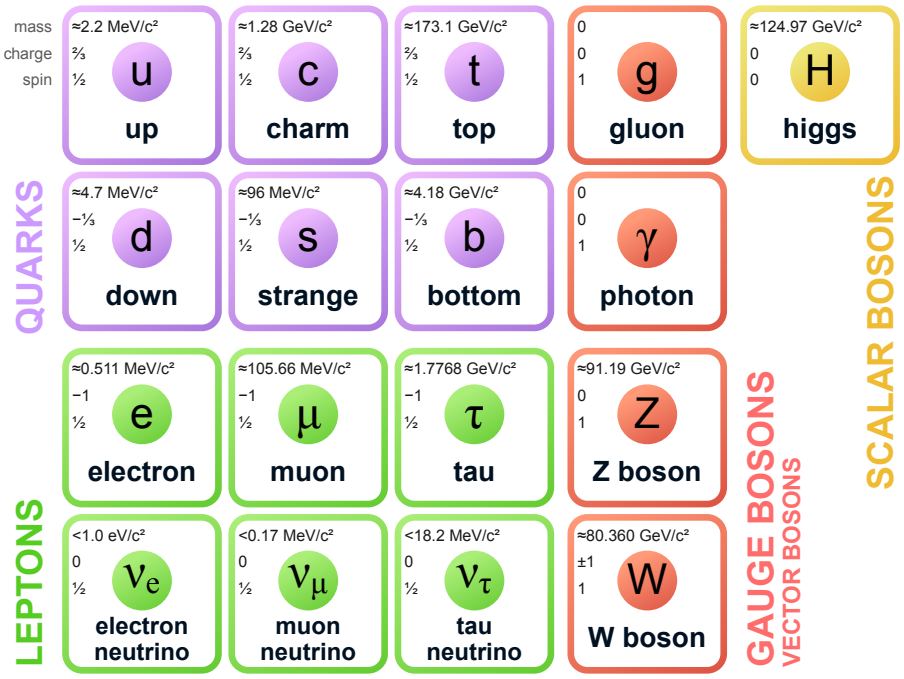


Figure 2.1.: Elementary particles of the Standard Model of particle physics [8]. They are categorized into three generations of quarks and leptons, force-mediating gauge bosons, and the Higgs boson.

Lagrangian.

In addition to *global* symmetries, the Standard Model incorporates *local* symmetries, which are functions of space-time. *Local gauge invariance* is a fundamental principle of the theory, which requires that the Lagrangian must remain unchanged under local phase transformations that depend on space-time coordinates. When promoting a global symmetry to a local one, the gauge invariance is restored via addition of a vector field – the gauge field. The quanta of these gauge fields are called gauge bosons and represent the observed force carrier particles. Mathematically, they correspond to the generators of the respective local symmetry group, which means that they are associated with the transformations that leave the Lagrangian invariant.

The fundamental forces in the Standard Model are described by a a local $SU(3) \otimes SU(2)_L \otimes U(1)_Y$ gauge symmetry, in which each contribution approximately corresponds to one of the three interactions, discussed in the following.

Quantum Electrodynamics (QED)

Quantum Electrodynamics (QED) is the quantum field theory that describes the electromagnetic force, which is responsible for interactions of photons with electrically

charged particles. The gauge symmetry associated with QED is the local $U(1)$ symmetry. The generator of this group corresponds to the massless gauge boson of QED, the photon. The latter does not itself carry the interaction charge and hence does not couple to other photons. Further, the photon has no rest mass, rendering the range of the interaction to be infinite.

Quantum Chromodynamics (QCD)

Quantum Chromodynamics (QCD) is the theory that describes the strong nuclear force, which binds quarks together to form composite particles called hadrons. The gauge symmetry of QCD is based on the $SU(3)$ symmetry group. Similar to the electric charge in QED, quarks carry one of three color charges: red, green, or blue. Antiquarks carry the respective anti-colors.

QCD dictates systems of particles to be color-neutral, a phenomenon called color confinement. Accordingly, individual quarks cannot exist in isolation, leading to the observation of two classes of hadrons, namely mesons and baryons.[1] Mesons are quark-antiquark pairs, carrying combinations of color and the respective anti-color to form color-neutral objects. Baryons on the other hand are composed of three quarks, each carrying a different color charge, summing up to form a "white", color-neutral hadron.

The generators of the $SU(3)$ group produce eight gauge bosons of QCD, called gluons, which mediate the strong interaction between quarks. One crucial aspect of QCD is the fact that gluons themselves carry combinations of color charges, opening the door for self-interactions. One consequence of this is that the interaction strength increases indefinitely with increasing distance between two quarks. As quarks move farther apart, the energy stored in the field between them increases until it is sufficiently large to create additional quark-antiquark pairs, resulting in the formation of new, color-neutral hadrons. This renders the effective range of the strong force to be limited to distances on the order of femtometers. At very short distances (or high energies) on the other hand, the interaction strength weakens, rendering the quarks to behave almost as free particles [10, 11].

Electroweak unification

The weak interaction is the only mechanism that allows for flavor-changing processes and is therefore responsible for most particle decays. In the framework of the Standard Model, the weak interaction is unified with the electromagnetic interaction to form the electroweak interaction [12, 13]. The combined force is based on the

[1] Exotic hadrons, like tetraquarks or pentaquarks, have been created and measured in dedicated experiments. However, they are still an active area of research and not known to occur naturally.

symmetry group $SU(2)_L \otimes U(1)_Y$. Here, the $SU(2)_L$ group describes the symmetry of the weak isospin, I. The latter is the additive quantum charge of the weak interaction. The $U(1)_Y$ group represents the symmetry of the weak hypercharge, $Y = 2(Q_{\mathrm{EM}} - I_3)$, a quantum number relating the electric charge, Q_{EM}, with the third component of the weak isospin.

One important peculiarity of the electroweak interaction is its chiral nature. In the electroweak theory, left-handed fermions carry weak isospin charges of $I = \frac{1}{2}$, forming doublets that include both up-type and down-type particles. On the other hand, right-handed fermions have weak isospin charges of $I = 0$, forming singlets under $SU(2)_L$:

$$\psi_L = \left[\begin{pmatrix} u \\ d \end{pmatrix}_L , \begin{pmatrix} l \\ \nu_l \end{pmatrix}_L \right] , \ \psi_R = [u_R, \ d_R, \ l_R] \ . \tag{2.1}$$

So far, no right-handed neutrinos have been observed, which is why they are not included in the SM. The W-boson triplet couples exclusively to the doublets of left-handed fermions (or right-handed anti-fermions). This chiral nature of the electroweak interaction leads to a maximal symmetry breaking of parity and charge conjugation, which are discussed in more detail in Section 2.2.1.

The fermions transform under the combined local $SU(2)_L \otimes U(1)_Y$ symmetry as [7, 14]:

$$\psi_L \to \psi'_L = \exp\left(-\alpha \cdot \sigma^a/2 - i\beta Y\right) \psi_L \tag{2.2}$$

$$\psi_R \to \psi'_R = \exp\left(-i\beta Y\right) \psi_R \ , \tag{2.3}$$

where α is a three-component vector of gauge transformation parameters associated with the $SU(2)_L$ group, σ^a are the Pauli matrices that describe the three generators of the $SU(2)_L$ group, β is a gauge transformation parameter associated with the $U(1)_Y$ group, and Y is the weak hypercharge.

To ensure that the Lagrangian remains invariant under these local gauge transformations, we introduce the covariant derivative, a generalization of the partial derivative, that incorporates the gauge fields:

$$D_\mu = \partial_\mu - ig\frac{\sigma^a}{2}W^a_\mu - ig'\frac{Y}{2}B_\mu \tag{2.4}$$

Here, ∂_μ is the partial derivative with respect to the space-time coordinates, g and g' are the coupling constants associated with the $SU(2)_L$ and $U(1)_Y$ groups, respectively, W^a_μ is a vector of the three gauge fields associated with the $SU(2)_L$

group, and B_μ is the gauge field associated with the $U(1)_Y$ group.

The kinetic term of the Lagrangian for the electroweak interaction can now be written as [15]:

$$\mathcal{L}_{kin} = -\frac{1}{4} W^a_{\mu\nu} W^{\mu\nu a} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} + \sum_{\text{fermions}} (D^\mu \psi)^\dagger (D_\mu \psi) \tag{2.5}$$

where $W^a_{\mu\nu}$ is the field strength tensor for the $SU(2)$ gauge field,

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g\epsilon^{abc} W^b_\mu W^c_\nu \tag{2.6}$$

and $B_{\mu\nu}$ is the field strength tensor for the $U(1)$ gauge field,

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \ . \tag{2.7}$$

The Lagrangian in Eq. 2.5 describes a system with massless fermions and massless gauge bosons. To obtain the observed physical particles, $W^\pm$, Z^0, and γ, while preserving gauge invariance, the symmetry of $SU(2)_L \otimes U(1)_Y$ is broken spontaneously via the Higgs mechanism.

2.1.3. Higgs Mechanism

The Higgs mechanism provides the means for particles to acquire mass without violating the gauge symmetry of the theory, postulating an additional spin-0 boson, the Higgs boson. The Higgs field, ϕ, is represented by a complex scalar field of a particle with imaginary mass μ. In the SM, the Higgs field is associated to a doublet under $SU(2)$, which can be decomposed into real and imaginary parts,

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \ . \tag{2.8}$$

The scalar part of the Lagrangian of the Higgs field is:

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi) \ , \tag{2.9}$$

where $V(\phi)$ is the Higgs potential, given by:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \ , \tag{2.10}$$

with λ as self-coupling constant of the Higgs field. Taking $\lambda > 0$ and $\mu^2 < 0$, the potential follows a shape illustrated in Figure 2.2. The minimum is obtained for all

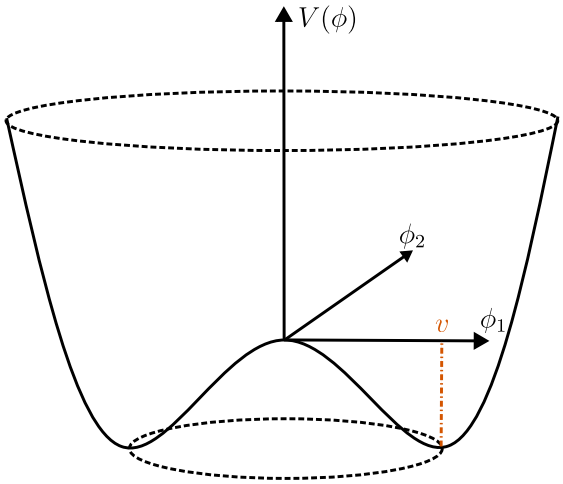


Figure 2.2.: Higgs potential, $V(\phi)$, for $\lambda > 0$ and $\mu^2 < 0$. The minimum is located at points with $|\phi|^2 = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}$.

points fulfilling

$$|\phi|^2 = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2} \text{ with } v \equiv \sqrt{-\frac{\mu^2}{\lambda}} \ . \tag{2.11}$$

We choose the specific point, $\phi_1 = v$, $\phi_2 = \phi_3 = \phi_4 = 0$ which corresponds to the vacuum expectation value (VEV):

$$\langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} v \\ 0 \end{pmatrix} \ . \tag{2.12}$$

We can now perturbate the complex scalar field $\phi(x)$ around this point in the following way[2]:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} v + h(x) \\ 0 \end{pmatrix} \ . \tag{2.13}$$

The Lagrangian rewritten in such a way with respect to the lowest eigenstate has lost its original symmetry — it is *spontaneously broken*. Plugging this into the Lagrangian of Eq. 2.9 with the covariant derivative from Eq. 2.4, we obtain the

[2]When expanding the field around the minimum, we add new fields, yielding additional boson fields, which need to be "rotated away" such that the acquired Goldstone bosons disappear. This is referred to as *unitary gauge* transformation.

contribution

$$\mathcal{L}_{\text{Higgs}} = -\frac{1}{8}v^2 \begin{pmatrix}1\\0\end{pmatrix}^T \begin{pmatrix} gW_\mu^3 - g'B_\mu & g(W_\mu^1 - iW_\mu^2) \\ g(W_\mu^1 + iW_\mu^2) & -gW_\mu^3 - g'B_\mu \end{pmatrix}^2 \begin{pmatrix}1\\0\end{pmatrix} + \dots \tag{2.14}$$

$$= -\frac{1}{8}v^2 g^2 \left((W_\mu^1)^2 + (W_\mu^2)^2\right) + \frac{1}{8}v^2 \left(g'B_\mu - gW_\mu^3\right)^2 + \dots \tag{2.15}$$

We can use W_μ^1 and W_μ^2 to construct the following linear combinations

$$W_\mu^+ = \frac{1}{\sqrt{2}}W_\mu^1 - iW_\mu^2 \ , \ W_\mu^- = \frac{1}{\sqrt{2}}W_\mu^1 + iW_\mu^2 \ , \tag{2.16}$$

which, after comparison to the first term of Eq. 2.15, can be identified as the physical charged gauge bosons of the weak interaction, with mass

$$M_W = \frac{1}{2}vg \ . \tag{2.17}$$

To interpret the second term in Eq. 2.15, we first define the weak mixing angle, or *Weinberg angle*,

$$\theta_W \equiv \tan^{-1}\left(\frac{g'}{g}\right) \ . \tag{2.18}$$

The observed Z^0 boson can then be expressed as linear combination of the neutral components W_μ^3 and B_μ as follows:

$$Z_\mu^0 = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \quad \text{with} \quad M_Z = \frac{1}{2}v\sqrt{g^2 + g'^2} \ . \tag{2.19}$$

Lastly, for the massless photon field, with no coupling to the Higgs field, we obtain

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu \quad \text{with} \quad M_\gamma = 0 \ . \tag{2.20}$$

We have now attained the massive gauge bosons, that we observe in nature, all of which obtain their masses due to a non-zero VEV, breaking the symmetry at energies below the electroweak scale.

The masses of all charged fermions in the Standard Model are generated analogously through Yukawa couplings to the Higgs field. Specifically, the Yukawa interaction terms of the Lagrangian for quarks can be written as:

$$\mathcal{L}_{Yuk} = -\sum_{i,j}^{3} \left(Y_{ij}^u \bar{Q}_{Li}\widetilde{\phi}u'_{Rj} + Y_{ij}^d \bar{Q}_{Li}\phi d'_{Rj}\right) + \text{h.c.} \tag{2.21}$$

where $Y_{ij}^{u,d}$ are the Yukawa coupling matrices for up-type and down-type quarks, Q_L represents the left-handed quark doublet, u'_R and d'_R are the right-handed up-type and down-type quark singlets, respectively, and ϕ is the Higgs doublet. The indices i, j run from 1 to 3, corresponding to the three generations of quarks.

When the Higgs doublet acquires a non-zero vacuum expectation value of $v/\sqrt{2}$, mass terms are generated through the Yukawa couplings:

$$M_u = \frac{v}{\sqrt{2}} Y^u, \qquad M_d = \frac{v}{\sqrt{2}} Y^d \ . \tag{2.22}$$

For each fermion, the respective Yukawa coupling determines the relative strength of the interaction, while the VEV of the Higgs field sets the overall mass scale.

In their immediate form, the mass matrices, M_u and M_d, are not diagonal, which means that the mass eigenstates are not identical to the weak interaction eigenstates. To obtain the physical mass eigenstates, we diagonalize M_u and M_d via transformations of the form

$$M_{diag}^f = \frac{v}{\sqrt{2}} V_L^f Y^f V_R^{f\dagger} \quad \text{with} \quad f = u, d \ , \tag{2.23}$$

where the four $V_{L,R}^{u,d}$ matrices are unitary. We can now express the charged-current interaction term of the Lagrangian, responsible for flavor-changing transitions in the SM, in terms of the quark mass eigenstates:

$$\mathcal{L}_{CC} = -\frac{g}{\sqrt{2}} \sum_{i,j,k,l}^{3} \left(\bar{u}_{Lk} \gamma^\mu W_\mu^+ (V_L^u)_{ki} (V_L^d)_{lj}^\dagger d_{Ll} \right) + \text{h.c.} \tag{2.24}$$

or in a more explicit matrix notation,

$$\mathcal{L}_{CC} = -\frac{g}{\sqrt{2}} (\bar{u}_L, \bar{c}_L, \bar{t}_L) \gamma^\mu W_\mu^+ \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} + \text{h.c.} \ , \tag{2.25}$$

where g is the weak coupling constant and W_μ^+ is the charged weak boson field. We can identify the Cabibbo-Kobayashi-Maskawa (CKM) matrix, V_{CKM}, as the product of the unitary matrices that diagonalize the up-type and down-type quark mass matrices:

$$V_{\text{CKM}} \equiv V_L^u V_L^{d\dagger} \ . \tag{2.26}$$

The CKM matrix is a 3×3 unitary matrix that parametrizes the mixing between quark flavors in the weak charged current interactions and plays a central role in the phenomenon of CP violation.

2.1.4. Physics beyond the SM

Over the past several decades, the Standard Model has undergone rigorous experimental verification, consistently matching the results of countless experiments with extraordinary precision. The last missing piece, the Higgs boson, was experimentally discovered in 2012 by the ATLAS and CMS collaborations at the Large Hadron Collider, completing the SM [16, 17]. Despite its success, the SM of particle physics falls short in explaining several important phenomena observed in nature.

In spite of decades of research, the integration of general relativity, describing the curvature of space-time into the quantum field framework of the Standard Model remains an unsolved problem in theoretical physics [18]. Similarly, there is the hope to combine the strong nuclear force with the electroweak force under a larger, encompassing gauge symmetry, that remains unbroken at higher energy scales [19]. The most prominent approach to such a Grand Unified Theory (GUT) is the super-symmetric (SUSY) extension of the SM, which predicts at minimum one superpartner for each particle in the SM [20, 21]. Such a model could also address the observed hierarchy of the fermion masses, which is, much like the number of generations or the coupling strengths between them, not founded on any particular rationale in the SM. So far, no particles pointing towards SUSY being realized in nature have been observed.

Additionally, there are several observations at the cosmological scale that cannot be understood in the context of the SM. For one, there exist gravitational effects in galaxies and galaxy clusters, which cannot be accounted for by visible matter alone, suggesting the presence of an unseen mass, known as dark matter [22]. In the SM, there exists no particle that could serve as a viable candidate for such type of matter. Moreover, there remains the mystery of the abundance of matter in the observable Universe. According to our most successful theories of the early Universe, the Big Bang should have produced roughly equal amounts of matter and antimatter. Sakharov's conditions, proposed in 1967, address the observed asymmetry [1]. Among these conditions is the requirement of Charge-Parity (CP) violation, which corresponds to differences in the behavior of particles and antiparticles and is explained in more detail in the following section. Cronin and Fitch's discovery of CP violation in the neutral kaon system in 1964 [23] opened up the possibility that similar violations could be responsible for the matter-dominated Universe.

In the Standard Model, quark-flavor transitions in weak decays are governed by the unitary CKM mixing matrix (cf. Section 2.2.2). This matrix includes a complex phase, which leads to differences in the decay properties between matter and antimatter particles. However, the observed baryon asymmetry in the Universe is many orders of magnitude larger than our current theory accounts for [2]. It is

evident that new sources of CP violation are needed.

Nowadays, the majority of elementary particle physics experiments therefore aim to find physics beyond the SM, that could help address these unsolved puzzles. Precision experiments, such as Belle II, focus on overconstraining a set of parameters that describe the flavor dynamics in B-meson decays. By precisely measuring a vast array of observables, such as branching ratios and CP-violation parameters, the unitarity of the CKM matrix is tested. Any significant deviations from the expected unitarity could provide crucial hints pointing towards the existence of new particles or interactions.

2.2. CP Violation

2.2.1. Discrete Symmetries

The fundamental principle of symmetry plays a crucial role in elementary particle physics and our understanding of the workings of the Universe. Continuous symmetries, such as gauge symmetries, form the basis of the Standard Model formulation, providing a framework to describe the fundamental forces and particles. Noether's theorem connects symmetries to conservation laws, stating that every continuous symmetry transformation corresponds to a conserved quantity [9]. For instance, the invariance of physical processes under space and time transformations leads to the conservation of momentum and energy. Next to continuous symmetries, three related, discrete symmetries play a significant role in the SM: the discrete space-time transformations of parity (P) and time reversal (T), as well as charge conjugation (C) which corresponds to inverting the sign of all charges.

Parity

The parity operator, denoted as P, performs a spatial inversion through the origin,

$$P\Psi(\vec{r}) = \Psi(-\vec{r}) \ . \tag{2.27}$$

Alternatively, it can be thought of a reflection, followed by a rotation of 180° around the axis of reflection. Applying the operation twice results in the original state, that is $P^2 = \mathbf{1}$. The eigenvalues are therefore $\eta_P = 1$, for even parity, and $\eta_P = -1$, for odd parity. Fermions have even intrinsic parity, while anti-fermions have odd intrinsic parity.

Parity is a multiplicative quantum number. The parity of a many-particle system is the product of the intrinsic parities of the particles and the parity of the spatial

wavefunction, given by $(-1)^L$, where L corresponds to the angular momentum of the system.

Historically, the discovery of P-symmetry violation in weak interactions came as a surprise in 1956, when C.S. Wu and her collaborators conducted an experiment that involved beta decay of ^{60}Co atoms [24]. They observed a preferential emission of electrons in a specific direction, indicating a violation of the parity symmetry. This discovery shattered the long-held belief that the laws of physics are symmetric under parity reflection. In 1957 T.D. Lee and C.N. Yang were awarded the Nobel Prize in Physics for their theoretical prediction of the violation of parity symmetry one year earlier [25].

Charge conjugation

The charge conjugation operator C transforms a particle into its anti-particle,

$$C\Psi = \bar{\Psi} \ . \tag{2.28}$$

It inverts all additive quantum numbers such as electric charge, flavor, baryon number, lepton number and weak isospin. As for the parity transformation, applying C twice results in the original state and the possible eigenvalues are $\eta_C = \pm 1$. Particles with vanishing additive quantum numbers, like the photon or the neutral pion, are identical to their anti-particles and therefore evidently eigenstates of C.

As eluded to in the previous section, the absence of right-handed neutrino couplings and the presence of right-handed anti-neutrino couplings in the weak interaction violates the C-symmetry maximally. After the observation of parity violation by Wu et al. it was generally assumed that the combined transformation CP remained a perfect symmetry in nature. However, in 1964, J. Cronin and V. Fitch conducted experiments at the Brookhaven National Laboratory, observing asymmetric decay rates of neutral kaons and their antiparticles into specific final states [23]. This groundbreaking discovery provided the first direct evidence of the violation of this combined transformation. As a result, in 1980 they were awarded the Nobel Prize in Physics. Many years later, in 2001 and 2019, the violation of CP-symmetry was also observed in B-meson and D-meson decays [26, 27, 28].

Time reversal

The discrete symmetry operation T reverses the flow of time,

$$T\Psi(t) = \Psi(-t) \ . \tag{2.29}$$

Whereas positions in space and all charges are left unaffected, it inverts the momenta and spins of a system. At the macroscopic scale, the second law of thermodynamics dictates an increase of entropy over time and therefore a violation of T-symmetry. However, at the microscopic scale of individual particle interactions, the symmetry is mostly conserved. As for P and C, the electromagnetic and strong interactions are invariant under time reversal.

It can be shown that the dynamics of a system described by a local Lorentz invariant gauge theory, like the SM, must be invariant under the combined transformation of CPT [29]. This so-called CPT theorem implies that particles must, consequently, have masses and lifetimes identical to their respective antiparticles. So far, all experimental tests have confirmed CPT to be a perfect symmetry in nature.

Another consequence of the CPT theorem is its implication of T-symmetry violation: In order to compensate for the breaking of CP-symmetry, T-symmetry must necessarily be broken. A direct evidence of this was obtained by the CPLEAR collaboration at CERN in 1998 [30]. The experiment involved measurements of semileptonic decays of neutral kaons and anti-kaons. The observed asymmetry in the decay rates provided conclusive evidence for T-symmetry violation in the neutral kaon system.

2.2.2. CKM Mechanism

In the SM, flavor-changing transitions are governed by charged currents of the weak interaction. Mediated by charged $W^{\pm}$ bosons, left-handed up-type quarks can transition to left-handed down-type quarks, and vice versa. The corresponding interaction term of the Lagrangian can be written as:

$$\mathcal{L}_{CC} = -\frac{g}{\sqrt{2}} \bar{u}_{Li} \gamma^{\mu} (V_{\mathrm{CKM}})_{ij} d_{Lj} W_{\mu}^{+} + \mathrm{h.c.} \tag{2.30}$$

where g is the weak coupling constant, $i, j = 1, 2, 3$ refer to the quark generations, $\bar{u}_{Li}$ and d_{Lj} are the left-handed components of the quark fields and V_{CKM} is the quark mixing matrix (cf. Eq. 2.25):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_{weak} = V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{mass} \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{mass} . \tag{2.31}$$

The above Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix is a unitary matrix that characterizes the mixing of quark flavors, relating the weak eigenstates to their corresponding mass eigenstates. As the physical mass eigenstates are not

identical to the eigenstates of the weak interaction, this matrix contains off-diagonal elements, allowing for transitions between quark generations. Its elements govern the probabilities of such transitions and need to be determined experimentally.

In order to explain the groundbreaking observation of CP violation consistently within the quark model of the SM, the Lagrangian of Eq. 2.30 must not be invariant under CP conjugation, and hence, $V_{\mathrm{CKM}} \neq V_{\mathrm{CKM}}^{\dagger}$. Accordingly, the matrix needs to contain at least one irreducible complex phase, which changes sign under CP.

In principle, for a unitary $N \times N$ matrix, there exist N^2 free parameters, $N(N-1)/2$ real parameters and $N(N+1)/2$ complex phases. As the Lagrangian is invariant under simultaneous phase shifts of the physical quark fields, $(2N-1)$ complex phases in the V_{CKM} matrix can be eliminated by a proper phase convention. The remaining $(N-1)(N-2)/2$ independent complex phases accommodate any CP violation, which becomes viable with a minimum of three quark generations. In this case, the matrix is defined by three real parameters and a single complex phase.

At the time of discovery of CP violation in the neutral kaon system, only the three lightest quarks, u, d and s, had been observed experimentally. In 1973, M. Kobayashi and T. Maskawa provided a consistent explanation for CP violation, postulating the existence of an unknown third generation quark doublet [31].[3] In recognition of their revolutionary finding, they were awarded the Nobel Prize in Physics in 2008.

The original parametrization used by the authors was given in terms of three Euler angles, θ_1, θ_2, θ_3, and one complex phase, δ:

$$V_{\mathrm{CKM}} = \begin{pmatrix} c_1 & -s_1 c_3 & -s_1 s_3 \\ s_1 c_2 & c_1 c_2 c_3 - s_2 s_3 e^{i\delta} & c_1 c_2 s_3 + s_2 c_3 e^{i\delta} \\ s_1 c_2 & c_1 s_2 c_3 + c_2 s_3 e^{i\delta} & c_1 s_2 s_3 - c_2 s_3 e^{i\delta} \end{pmatrix} , \tag{2.32}$$

where we denote the sine and cosine of the corresponding angle θ_i as c_i and s_i.

An alternative, widely-used representation is given by the so-called Wolfenstein parametrization [33], which provides a more intuitive visualization of the relative sizes of the individual transition probabilities. It follows from a series expansion of the matrix elements in powers of the parameter $\lambda = |V_{us}| = |V_{cd}| = \sin(\theta_C)$, which corresponds to the sine of the Cabibbo angle, governing transitions between the first two generations. The approximate form of the CKM matrix using this

[3] The idea of a mixing mechanism connecting different quark flavors via the superposition of their weak eigenstates was first proposed by N. Cabibbo for two generations [32]. Its central result, the mixing matrix is, therefore, nowadays referred to as the Cabibbo-Kobayashi-Maskawa matrix.

parametrization, up to $\mathcal{O}(\lambda^3)$, reads:

$$V_{\text{CKM}} = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} , \tag{2.33}$$

where A, ρ and η are real parameters in $(0, 1)$. The only complex contribution is given by $i\eta$, which hence accounts for any CP violation in this representation.

From the unitarity condition of the CKM matrix follows

$$(VV^\dagger)_{ij} = (V^\dagger V)_{ij} = \delta_{ij} , \tag{2.34}$$

which results in twelve relations between matrix elements. Six of those with $i \neq j$ can be represented as triangles in the complex plane with sides $V_{ik}V_{jk}^*$. The latter can be obtained in decay rate measurements of the corresponding transitions. The interior angles are determined in measurements of CP violation.

One important relation relevant for B-meson decays is given by

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 , \tag{2.35}$$

where all terms are of equal order in λ, $\mathcal{O}(\lambda^3)$, which is unique to the B-meson system. As a result, the interior angles are all $\mathcal{O}(60°)$ and CP violation effects are significantly larger than in other systems. For practicality, we can rotate the triangle and normalize it by one of its sides, $V_{cd}V_{cb}^*$, to align it with the real axis. The third corner is then located at $(\bar{\rho}, \bar{\eta})$, with

$$\bar{\rho} \equiv \left(1 - \frac{\lambda^2}{2}\right)\rho, \qquad \bar{\eta} \equiv \left(1 - \frac{\lambda^2}{2}\right)\eta . \tag{2.36}$$

The corresponding angles are given by

$$\phi_1 \equiv \beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) , \tag{2.37}$$

$$\phi_2 \equiv \alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) , \tag{2.38}$$

$$\phi_3 \equiv \gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) , \tag{2.39}$$

where the different notations correspond to varying conventions. The triangle is schematically shown in Figure 2.3.

The collaborative goal of experiments such as Belle II or LHCb is to overconstrain

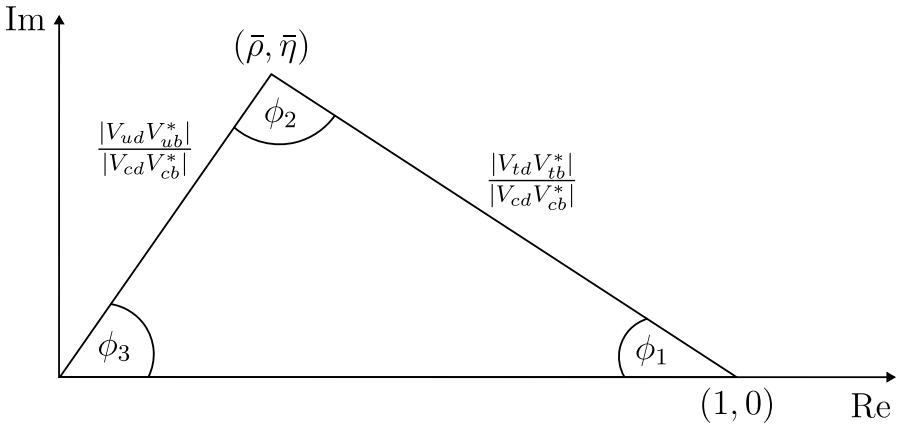


Figure 2.3.: Unitarity triangle for the B-meson system, corresponding to Eq. 2.35. The sides are normalized by $|V_{cd}V_{cb}^*|$.

the parameters of the CKM triangle. By experimentally verifying the closure of the triangle, one can test the unitarity condition and detect any deviations, which would indicate the presence of physics beyond the Standard Model. Additional observables, like the mass difference of neutral B mesons, Δm_d, can provide further constraints. Figure 2.4 shows a combination of constraints from numerous measurements along with the result of a global fit to determine the shape of the triangle. Evidently, the fit result coincides consistently with the 95% CL regions, indicating good agreement at the current level of precision [20].

2.2.3. Time-Evolution of Neutral Mesons

Off-diagonal elements in the CKM matrix introduce flavor-changing transitions as part of the weak interaction in the SM. Accordingly, the eigenstates of the weak interaction are not identical to the physical mass eigenstates. As a consequence, neutral mesons, such as B^0 and $\bar{B}^0$, exhibit time-dependent flavor oscillations, which result in a change of their flavor quantum number by two units. The underlying mechanism driving these oscillations depends on a small mass difference between the mass eigenstates, giving rise to a mixing of the flavor eigenstates as the system evolves. In the following, the dynamics of the neutral meson system will be described, establishing the necessary framework for a quantitative interpretation of CP-violation measurements.[4]

We can express a neutral, flavored meson state as a linear combination of its two

[4]Though we use the notation of B-meson states, B^0 and $\bar{B}^0$, already in this section, the results are generally applicable to other heavy neutral mesons, as well. The specifics of B-meson dynamics are discussed at the end of this section.

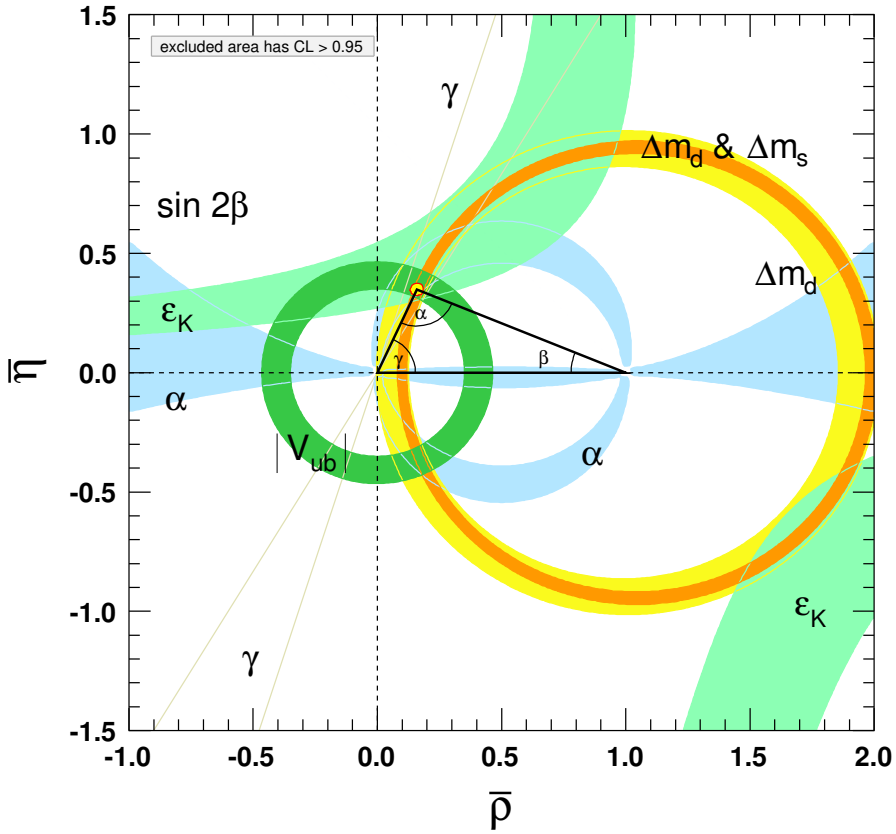


Figure 2.4.: Constraints from numerous measurements on the unitary triangle along with the result of a global fit. Shaded areas indicate the 95% CL [20].

flavor-eigenstates at a given time, $t_0 = 0$,

$$|B(t_0 = 0)\rangle = a(0)|B^0\rangle + b(0)|\overline{B}^0\rangle \ . \tag{2.40}$$

For $t > 0$, this system will evolve into

$$|B(t)\rangle = a(t)|B^0\rangle + b(t)|\overline{B}^0\rangle + \sum_i c_i(t)|f_i\rangle \ , \tag{2.41}$$

where $|f_i\rangle$ denote final states accessible via decay. In the Weisskopf-Wigner approximation, that is, for time scales significantly larger than those of the strong interaction, and neglecting the terms $c_i(t)$, we can approximate the time evolution of the system, using the time-dependent Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\psi(t) = \mathcal{H}_{\text{eff}}\,\psi(t) \quad \text{with} \quad \psi(t) = \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \ , \tag{2.42}$$

where $\mathcal{H}_{\text{eff}}$ is a non-hermitian 2×2 matrix, representing the effective Hamiltonian. The latter can be expressed in terms of two hermitian matrices, $\boldsymbol{M}$ and $\boldsymbol{\Gamma}$:

$$\mathcal{H}_{\text{eff}} = \boldsymbol{M} - \frac{i}{2}\boldsymbol{\Gamma} \ . \tag{2.43}$$

The diagonal elements of $\mathcal{H}_{\text{eff}}$ correspond to flavor-conserving transitions, and assuming CPT invariance, we can infer $M_{11} = M_{22} \equiv m$ and $\Gamma_{11} = \Gamma_{22} \equiv \Gamma$. Additionally, hermiticity of $\boldsymbol{M}$ and $\boldsymbol{\Gamma}$ dictates $M_{12} = M_{21}^*$ and $\Gamma_{12} = \Gamma_{21}^*$, such that we can write

$$\mathcal{H}_{\text{eff}} = \begin{pmatrix} m - \frac{i}{2}\Gamma & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & m - \frac{i}{2}\Gamma \end{pmatrix} \ . \tag{2.44}$$

The physical eigenstates of this matrix can be expressed as linear combinations of the flavor-eigenstates,

$$\begin{aligned} |B_L\rangle &= p|B^0\rangle + q|\overline{B}^0\rangle \\ |B_H\rangle &= p|B^0\rangle - q|\overline{B}^0\rangle \ , \end{aligned} \tag{2.45}$$

where the suffix L and H indicate the light and heavy mass eigenstates, respectively, and normalization requires $|p|^2 + |q|^2 = 1$. With explicit diagonalization of $\mathcal{H}_{\text{eff}}$, it can further be shown that q and p satisfy

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} \quad [34]. \tag{2.46}$$

The physical mass states obey the typical decay law,

$$\begin{aligned} |B_L(t)\rangle &= e^{-i\omega_L t}|B_L\rangle = e^{-i\omega_L t}(p|B^0\rangle + q|\overline{B}^0\rangle) \\ |B_H(t)\rangle &= e^{-i\omega_H t}|B_H\rangle = e^{-i\omega_H t}(p|B^0\rangle - q|\overline{B}^0\rangle) \ , \end{aligned} \tag{2.47}$$

with eigenvalues

$$\omega_{L,H} = m_{L,H} - \frac{i}{2}\Gamma_{L,H} \ . \tag{2.48}$$

We can express the parameters of $\omega_{L,H}$ in terms of the observable average lifetime and mass, and the difference between the light and heavy mass eigenstates:

$$\Delta m \equiv m_H - m_L \ , \quad \Delta\Gamma \equiv \Gamma_L - \Gamma_H \ , \quad m = \frac{m_L + m_H}{2} \ , \quad \Gamma = \frac{\Gamma_L + \Gamma_H}{2} \ . \tag{2.49}$$

To obtain the time-evolution in the basis of the flavor-eigenstates, we solve Eq. 2.45

for $|B^0\rangle$ and $|\bar{B}^0\rangle$, and subsequently plug in the time-dependent terms for $|B_L(t)\rangle$ and $|B_H(t)\rangle$ from Eq. 2.47:

$$\begin{aligned}
|B^0(t)\rangle &= \frac{1}{2p}\left(e^{-i\omega_L t}(p|B^0\rangle + q|\bar{B}^0\rangle) + e^{-i\omega_H t}(p|B^0\rangle - q|\bar{B}^0\rangle)\right) \\
&= \frac{1}{2}\left(e^{-i\omega_L t} + e^{-i\omega_H t}\right)|B^0\rangle + \frac{q}{2p}\left(e^{-i\omega_L t} - e^{-i\omega_H t}\right)|\bar{B}^0\rangle \\
|\bar{B}^0(t)\rangle &= \frac{1}{2q}\left(e^{-i\omega_L t}(p|B^0\rangle + q|\bar{B}^0\rangle) - e^{-i\omega_H t}(p|B^0\rangle - q|\bar{B}^0\rangle)\right) \\
&= \frac{p}{2q}\left(e^{-i\omega_L t} - e^{-i\omega_H t}\right)|B^0\rangle + \frac{1}{2}\left(e^{-i\omega_L t} + e^{-i\omega_H t}\right)|\bar{B}^0\rangle \ .
\end{aligned} \tag{2.50}$$

These expressions can be simplified further by introducing the following parameter:

$$g_\pm = \frac{1}{2}\left(e^{-i\omega_L t} \pm e^{-i\omega_H t}\right) \ , \tag{2.51}$$

which can be written in explicit terms of the observables in Eq. 2.49,

$$\begin{aligned}
g_+ &= e^{-imt}e^{-\Gamma t/2}\left(\cosh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} - i\sinh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2}\right) \\
g_- &= e^{-imt}e^{-\Gamma t/2}\left(-\sinh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} + i\cosh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2}\right) \ .
\end{aligned} \tag{2.52}$$

Using the definitions of $g_\pm$ from Eq. 2.51, we can rewrite the time-evolution in Eq. 2.50,

$$\begin{aligned}
|B^0(t)\rangle &= g_+(t)|B^0\rangle + \frac{q}{p}g_-(t)|\bar{B}^0\rangle \\
|\bar{B}^0(t)\rangle &= \frac{p}{q}g_-(t)|B^0\rangle + g_+(t)|\bar{B}^0\rangle \ .
\end{aligned} \tag{2.53}$$

Assuming a pure flavor initial state, $|B^0\rangle$, we can now readily determine the probability to observe the same or opposite flavor-state after some time t,

$$\begin{aligned}
\left|\langle B^0|B^0(t)\rangle\right|^2 &= |g_+(t)|^2 = \frac{e^{-\Gamma t}}{2}\left(\cosh(\frac{\Delta\Gamma t}{2}) + \cos(\Delta m t)\right) \\
\left|\langle \bar{B}^0|B^0(t)\rangle\right|^2 &= \left|\frac{q}{p}g_-(t)\right|^2 = \left|\frac{q}{p}\right|^2\frac{e^{-\Gamma t}}{2}\left(\cosh(\frac{\Delta\Gamma t}{2}) - \cos(\Delta m t)\right) \ ,
\end{aligned} \tag{2.54}$$

where we have used the orthogonality of $|B^0\rangle$ and $|\bar{B}^0\rangle$ and the explicit expressions for $g_\pm$ in Eq. 2.52. Evidently, non-vanishing $\Delta\Gamma$ and Δm will lead to an oscillation of the two flavor eigenstates.

Finally, we consider the time-dependant decay rates of initially pure flavor states into a common final state f. For this, we denote the amplitudes for transitions of flavored mesons $|B\rangle$ and $|\bar{B}\rangle$ to some final states $|f\rangle$ and $|\bar{f}\rangle$ as

$$\begin{aligned} A_f &= \langle f|\mathcal{H}|B\rangle \quad \bar{A}_f = \langle f|\mathcal{H}|\bar{B}\rangle \\ A_{\bar{f}} &= \langle \bar{f}|\mathcal{H}|B\rangle \quad \bar{A}_{\bar{f}} = \langle \bar{f}|\mathcal{H}|\bar{B}\rangle \ , \end{aligned} \tag{2.55}$$

with the effective Hamiltonian $\mathcal{H}$. Further, we introduce the following CP-violation parameter

$$\lambda_{CP} = \frac{q}{p}\frac{\bar{A}_f}{A_f} = \left|\frac{q}{p}\right|\left|\frac{\bar{A}_f}{A_f}\right| e^{-i(\phi_M+\phi_D)} \ , \tag{2.56}$$

where ϕ_M is the weak mixing phase, and ϕ_D is the decay phase.

Neglecting phase-space normalization, we obtain the decay rates by taking the absolute squares of the transition amplitudes,

$$\Gamma(B^0(t)\to f) = |\langle f|\mathcal{H}|B^0(t)\rangle|^2 \quad \text{and} \quad \Gamma(\bar{B}^0(t)\to f) = |\langle f|\mathcal{H}|\bar{B}^0(t)\rangle|^2 \ . \tag{2.57}$$

Inserting Eq. 2.53 into the above equations, and using the definition of λ_{CP}, we obtain,

$$\begin{aligned} \Gamma(B^0(t)\to f) = \frac{1}{2}|A_f|^2 e^{-\Gamma t}\Bigg[&(1+|\lambda_{CP}|^2)\cosh\frac{\Delta\Gamma t}{2} + (1-|\lambda_{CP}|^2)\cos(\Delta m t) \\ &- 2\,\mathrm{Re}(\lambda_{CP})\sinh\frac{\Delta\Gamma t}{2} - 2\,\mathrm{Im}(\lambda_{CP})\sin(\Delta m t)\Bigg] \end{aligned} \tag{2.58}$$

$$\begin{aligned} \Gamma(\bar{B}^0(t)\to f) = \frac{1}{2}\left|\frac{p}{q}\right|^2 |A_f|^2 e^{-\Gamma t}\Bigg[&(1+|\lambda_{CP}|^2)\cosh\frac{\Delta\Gamma t}{2} - (1-|\lambda_{CP}|^2)\cos(\Delta m t) \\ &- 2\,\mathrm{Re}(\lambda_{CP})\sinh\frac{\Delta\Gamma t}{2} + 2\,\mathrm{Im}(\lambda_{CP})\sin(\Delta m t)\Bigg] \ . \end{aligned} \tag{2.59}$$

Time-evolution of neutral B mesons

The flavor oscillation in neutral B-mesons is primarily mediated by box diagrams, with the dominant contribution arising from virtual t-quark exchange due to the corresponding CKM matrix element $V_{tb} \sim 1$. The contributions from lighter quarks (u, c, and d) can be neglected as a result of the Glashow-Iliopoulos-Maiani (GIM)

mechanism [35], which largely cancels out their amplitudes, minimizing their impact on the $B^0 - \bar{B}^0$ mixing process. The leading-order Feynman diagrams are shown in Figure 2.5, where b- and d-quarks undergo intermediate flavor changes to primarily t-quarks via the exchange of $W^{\pm}$ bosons.

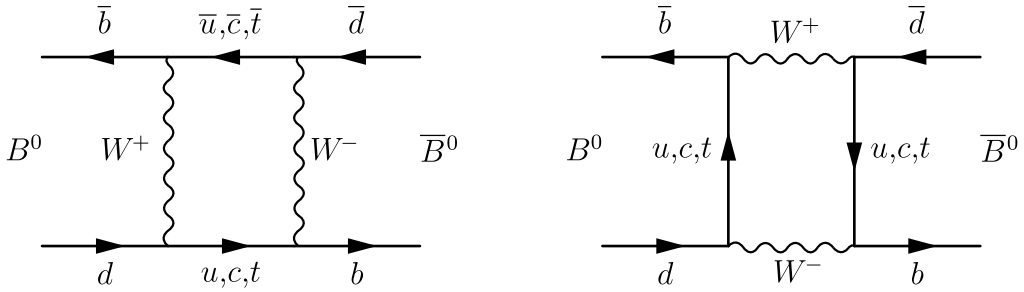


Figure 2.5.: Feynman diagrams of dominant contributions to the flavor-mixing between neutral B-meson eigenstates.

For the neutral B-meson system, we can make two simplifications. Firstly, the lifetime difference between the two mass eigenstates, B_H and B_L, is practically negligible with $\Delta\Gamma/\Gamma \approx \mathcal{O}(10^{-3})$, such that we can write $\Gamma_L = \Gamma_H = \Gamma$.

Secondly, we have $\Gamma_{12}/M_{12} \approx \mathcal{O}(m_b^2/m_t^2) \approx \mathcal{O}(10^{-3})$ and hence $|\Gamma_{12}| \ll |M_{12}|$. We can therefore expand the mixing parameter as follows [36]:

$$\frac{q}{p} = \sqrt{\frac{M_{12}^*}{M_{12}}} + \mathcal{O}\left(\frac{\Gamma_{12}}{M_{12}}\right) \approx \frac{M_{12}^*}{|M_{12}|} \ . \tag{2.60}$$

Consequently, the phase of q/p can be approximated by the phase of the box diagrams governing the $B^0 - \bar{B}^0$ oscillation. As the latter is dominated by virtual t-quark contributions, we can infer

$$\frac{q}{p} \approx \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \ . \tag{2.61}$$

Evidently, in this case, the mass eigenstates B_L and B_H can be considered as flavor eigenstates, as $|q/p| \approx 1$, rendering CP violation in mixing to be very small (see Section 2.2.4).

Using $\Delta\Gamma = 0$, we can simplify the expression for observing the same or opposite

flavor of a pure initial flavor state after some time t (Eq. 2.53) to read

$$\begin{aligned} \left|\langle B^0|B^0(t)\rangle\right|^2 &= \frac{e^{-\Gamma t}}{2}\big(1+\cos(\Delta m t)\big) \\ \left|\langle \bar{B}^0|B^0(t)\rangle\right|^2 &\overset{\left|\frac{q}{p}\right|\approx 1}{=} \frac{e^{-\Gamma t}}{2}\big(1-\cos(\Delta m t)\big)\ . \end{aligned} \tag{2.62}$$

Analogously, the decay rates of neutral B-meson states into a final state f, described in Eq. 2.58 and 2.59, can now be expressed as

$$\Gamma(B^0(t)\to f) = \frac{1}{2}|A_f|^2 e^{-\Gamma t}\Big[1+|\lambda_{CP}|^2+(1-|\lambda_{CP}|^2)\cos(\Delta m t) - 2\,\mathrm{Im}(\lambda_{CP})\sin(\Delta m t)\Big] \tag{2.63}$$

$$\Gamma(\bar{B}^0(t)\to f) = \frac{1}{2}|A_f|^2 e^{-\Gamma t}\Big[1+|\lambda_{CP}|^2-(1-|\lambda_{CP}|^2)\cos(\Delta m t) + 2\,\mathrm{Im}(\lambda_{CP})\sin(\Delta m t)\Big] \tag{2.64}$$

2.2.4. Types of CP Violation

Based on the above derivations, we can classify the mechanisms of CP violation in neutral meson decays into three categories.

CP violation in mixing

CP violation in mixing can occur during the oscillation of neutral mesons. In this case, the mass eigenstates are not pure CP eigenstates, i.e., $|q/p| \neq 1$. This condition becomes evident when explicitly comparing the asymmetric transition probabilities between the flavor eigenstates,

$$\left|\langle B^0|\bar{B}^0(t)\rangle\right|^2 \neq \left|\langle \bar{B}^0|B^0(t)\rangle\right|^2\ , \tag{2.65}$$

which is equivalent to

$$\left|\frac{p}{q}g_-(t)\right|^2 \neq \left|\frac{q}{p}g_-(t)\right|^2 \quad\Rightarrow\quad \left|\frac{q}{p}\right| \neq 1\ . \tag{2.66}$$

CP violation in decay

CP violation in decay (or *direct CP violation*) occurs when the decay amplitudes of a specific transition differ between particle and its anti-particle,

$$A_f \neq \bar{A}_{\bar{f}} \ . \tag{2.67}$$

In general, a decay is governed by a number of possible intermediate states that contribute to the total decay amplitude:

$$A_f = \sum_i |A_i| \, e^{i(\delta_i + \phi_i)} \quad \bar{A}_{\bar{f}} = \sum_j |A_j| \, e^{i(\delta_j + \phi_j)} \ , \tag{2.68}$$

where δ_i and ϕ_i represent strong and weak phases, respectively. Whereas the former remain invariant under CP transformation, the weak phases change signs. We can rewrite the condition for direct CP violation in the following form:

$$|A_f|^2 - \left|\bar{A}_{\bar{f}}\right|^2 = -2 \sum_{i,j} |A_i| \, |A_j| \sin(\delta_i - \delta_j) \sin(\phi_i - \phi_j) \ . \tag{2.69}$$

From the above equation it becomes evident that there needs to be a minimum of two intermediate states governing the decay with differing strong and weak phases for direct CP violation to occur.

Mixing-induced CP violation

Mixing induced CP violation can occur when a final state, f_{CP} is accessed by both flavor states. Figure 2.6 schematically shows the contribution of both states, after undergoing flavor mixing, to a common final state. Even for $\left|A_f/\bar{A}_f\right| = |q/p| = 1$, the interference of the mixing and decay processes can generate a CP asymmetry, if there is a non-vanishing phase difference. In terms of the complex quantity λ_{CP}, the condition for mixing induced CP violation can be expressed as $\mathrm{Im}(\lambda_{CP}) \neq 0$. Experimentally, the determination requires a time-dependant measurement, which is discussed in the following section in detail.

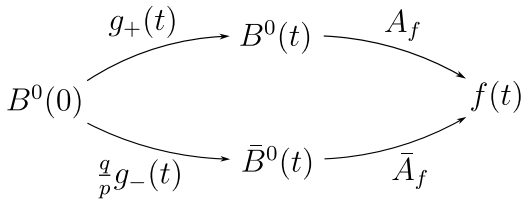


Figure 2.6.: Interference of two decay amplitudes leading to time-dependent CP violation in the decay of a neutral meson to a CP eigenstate.

2.3. B-Meson Sector

At the Belle II experiment, electrons and positrons are collided primarily at a center-of-mass energy of 10.58 GeV, corresponding to the third radially-excited bottomonium ($b\bar{b}$) resonance, denoted as $\Upsilon(4S)$. The rest mass of the $\Upsilon(4S)$ lies just above the production threshold for a pair of B mesons ($m_B = 5.279\,\text{GeV}$). As a result, the resonance predominantly decays at roughly equal rates into a pair of B^+B^- or $B^0\bar{B}^0$ mesons, which are nearly at rest in the center-of-mass frame.

B mesons exhibit comparably long lifetimes of $\tau_{B^0} = 1.52\,\text{ps}$ and $\tau_{B^\pm} = 1.62\,\text{ps}$. These extended lifetimes facilitate the measurement of time-dependent CP violation and can be attributed to the small couplings of the third quark generation to the other two generations, which are CKM-suppressed by factors of $V_{ub} \sim \mathcal{O}(\lambda_C^3)$ and $V_{cb} \sim \mathcal{O}(\lambda_C^2)$.

The combination of long lifetimes and sizeable CP violation in the B-meson system renders the latter an excellent laboratory for studying CP violation and precisely testing the predictions of the Standard Model. In the following, we examine CP-violation measurements in the $B^0\bar{B}^0$ system, followed by a more detailed discussion of the specific decay channels relevant to this thesis, $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$.

2.3.1. CP Violation in the B-Meson Sector

With negligible CP violation in mixing, B mesons can in principle still exhibit CP violation in the decay, directly or mixing-induced, leading to observable differences in the decay rates. We define the time-dependant rate asymmetry,

$$a_{CP}(t) \equiv \frac{\Gamma(\bar{B}^0(t) \to f) - \Gamma(B^0(t) \to f)}{\Gamma(\bar{B}^0(t) \to f) + \Gamma(B^0(t) \to f)}. \tag{2.70}$$

Plugging in the time-dependant decay rates from Eq. 2.63 and 2.64, the above asymmetry reads

$$a_{CP}(t) = \frac{|\lambda_{CP}|^2 - 1}{|\lambda_{CP}|^2 + 1}\cos(\Delta m t) + \frac{2\,\mathrm{Im}(\lambda_{CP})}{|\lambda_{CP}|^2 + 1}\sin(\Delta m t)$$

$$= \mathcal{A}_{CP}\cos(\Delta m t) + \mathcal{S}_{CP}\sin(\Delta m t)\ , \tag{2.71}$$

where, for the last step, we have introduced the CP-violation parameters[5]

$$\mathcal{A}_{CP} \equiv \frac{|\lambda_{CP}|^2 - 1}{|\lambda_{CP}|^2 + 1} \quad \text{and} \quad \mathcal{S}_{CP} \equiv \frac{2\,\mathrm{Im}(\lambda_{CP})}{|\lambda_{CP}|^2 + 1}\ . \tag{2.72}$$

We have established, that for neutral B mesons $q/p \approx 1$ and hence $|\lambda_{CP}| = |\bar{A}/A|$, indicating that $\mathcal{A}_{CP} \neq 1$ corresponds to non-vanishing direct CP violation. $\mathcal{S}_{CP} \neq 0$, on the other hand, implies the presence of mixing-induced CP violation, resulting from an interference of mixing and decay amplitudes. From an experimental point-of-view, the task is to measure the time evolution of $a_{CP}(t)$ in terms of the coefficients of $\cos(\Delta m t)$ and $\sin(\Delta m t)$, which in turn, determine $|\lambda_{CP}|$ and $\mathrm{Im}(\lambda_{CP})$.

Lastly, we note, that the above relations are valid for decays of neutral B mesons into CP eigenstates. In the case of charged mesons or decays to flavor-specific final states, such as $B^0 \to K^+\pi^-$, there is evidently no mixing-induced CP violation. The direct CP-violation parameter can then be readily expressed as difference between time-*in*dependent decay rates,

$$\mathcal{A}_{CP}^{\mathrm{dir}} \equiv \frac{\Gamma(\bar{B} \to \bar{f}) - \Gamma(B \to f)}{\Gamma(\bar{B} \to \bar{f}) + \Gamma(B \to f)}\ . \tag{2.73}$$

2.3.2. Time-Dependent CP-Violation Measurement

The time-evolution described in the preceding section and the corresponding rate asymmetry in Eq. 2.71 rely on prior knowledge of the initial flavor of the decaying B meson. As the $\Upsilon(4S)$ resonance decay produces pairs of B mesons, a neutral decaying particle, in principle, has an equal probability of being created as either a B^0 or a $\bar{B}^0$. At the Belle II experiment, we take advantage of quantum entanglement between the two mesons to address this challenge.

Due to conservation of total angular momentum of the $\Upsilon(4S)$ ($J^{PC} = 1^{--}$), the B mesons ($J^P = 0^-$) must have a relative angular momentum $L = 1$. Following Bose-Einstein statistics, the symmetric states B^0B^0 and $\bar{B}^0\bar{B}^0$ are prohibited.

[5] An alternative notation in use is $\mathcal{C}_{CP} = -\mathcal{A}_{CP}$. We stick to the above convention to remain consistent with references used in Section 2.3.3.

Accordingly, the two mesons form a quantum-mechanically entangled state and oscillate coherently until one of the B mesons decays. The combined state can be expressed as

$$|B_1(t_1), B_2(t_2)\rangle = \frac{1}{\sqrt{2}} \left(|B_1^0(t_1)\rangle |\bar{B}_2^0(t_2)\rangle - |\bar{B}_1^0(t_1)\rangle |B_2^0(t_2)\rangle \right) \ . \tag{2.74}$$

At the time of decay, t_1, the flavor of the decaying meson, B_1, dictates the flavor of the other, B_2, to be opposite. Following the time-evolution derived in the previous sections, the second meson then oscillates until its decay at t_2. Experimentally, this circumstance is used by inferring the flavor of neutral B mesons decaying into CP eigenstates from flavor-specific final state decays, which *tag* the flavor of both mesons at the time of their decay. From the time-dependant decay rates, one can derive the following expression for the probability of a decay to a CP final state f_{CP} for a given flavor tag $q = \pm 1$:

$$\mathcal{P}\left(\Delta t, q\right) = \frac{1}{4\tau_B} e^{-\frac{|\Delta t|}{\tau_B}} \left[\left(1 + q\big(\mathcal{A}_{CP} \cos(\Delta m \Delta t) + \mathcal{S}_{CP} \sin(\Delta m \Delta t)\big)\right] \ , \tag{2.75}$$

where $\Delta t = t_2 - t_1$ is the decay-time difference, and τ_B is the average lifetime for neutral B mesons. From the above expression it becomes clear, that fitting the decay rates as a function of Δt for both flavors allows to determine the CP violation parameters $\mathcal{A}_{CP}$ and $\mathcal{S}_{CP}$. In practice, due to the short lifetime of $\mathcal{O}$(ps), the time-difference is accessed via the spatial difference in decay vertex positions. To this end, the Υ(4S) is produced with a Lorentz boost $\beta\gamma$. With the $B\bar{B}$ pairs almost at rest in the center-of-mass frame, the time difference can be approximated with

$$\Delta t \approx \frac{\Delta z}{\beta\gamma c} \ , \tag{2.76}$$

with Δz corresponding to the spatial decay position difference along the boost axis, z, and c representing the speed of light.

2.3.3. Hadronic Two-Body Decays

Hadronic two-body B-meson decays play a crucial role in exploring flavor dynamics within the Standard Model. Specifically, transitions without charm quarks in the final state are often heavily CKM-suppressed and therefore exhibit only small branching ratios of $\mathcal{O}(10^{-5})$ or less. Despite this challenge, the suppression opens up possibilities to precisely test SM predictions for the interference between different higher-order decay amplitudes in the form of CP violation.

This study presents the analysis of neutral B-meson decays into two charged

pseudo-scalar mesons, $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$. In both cases, the tree diagram contribution to the transition amplitude is heavily CKM-suppressed, rendering the so-called penguin (loop) diagram to be of leading order. The latter refers to one-loop processes in which a quark temporarily changes flavor via the exchange of virtual $W^\pm$ bosons, and engages in a typically strong tree interaction (see Figure 2.7). These flavor-changing neutral current diagrams offer a valuable avenue to probe new physics beyond the Standard Model. Unobserved particles might participate in the virtual loop, introducing an extra amplitude to the decay process. Such effects could be measured in terms of deviations from SM predictions of decay rates or CP asymmetries.

The theoretical understanding of charmless two-body B-meson decays is challenging due to the non-perturbative nature of QCD in the low-energy regime. To tackle this, various approaches have been developed, such as QCD factorization, which aims to factorize the decay amplitude into perturbatively calculable and non-perturbative hadronic matrix elements. Additionally, $SU(3)$ flavor symmetry has been employed to relate different decay modes, providing valuable insight and reducing the number of independent parameters. As these methods make use of the world average values for the CKM matrix elements and the related components of the unitarity triangle, which are obtained only through experiment, precise measurements are of vital importance.

In this section, the specifics of $b \to u\bar{u}d$ and $b \to sud$ transitions are discussed. The focus lies on two methods to combine measurements of isospin-related transitions to significantly reduce theoretical uncertainties and constrain NP contributions.

$b \to u\bar{u}d$ transitions

Transitions of the kind $b \to u\bar{u}d$ are closely related to the angle ϕ_2 of the CKM triangle. Specifically the time-dependent measurement of $B^0 \to \pi^+\pi^-$ decays can be used to extract ϕ_2 in terms of experimentally accessible CP-violation parameters. The CP violation arises due to the interference of the B^0-$\bar{B}^0$ mixing process with the decay amplitude. Assuming a single contribution of the tree diagram to the decay, and hence no direct CP violation, the angle ϕ_2 is directly related to the mixing-induced CP-violation parameter [37],

$$\mathcal{S}_{CP} \equiv \frac{2\,\mathrm{Im}(\lambda_{CP})}{|\lambda_{CP}|^2 + 1} = \sin(2\phi_2)\ . \tag{2.77}$$

We can visualize the topology of the leading contributions to the decay amplitude in terms of their Feynman diagrams. In the case of $B^0 \to \pi^+\pi^-$, these are given by the charged-current tree diagram and the one-loop penguin diagram, depicted in Figure 2.7. The magnitude of both amplitudes can be compared in terms of the

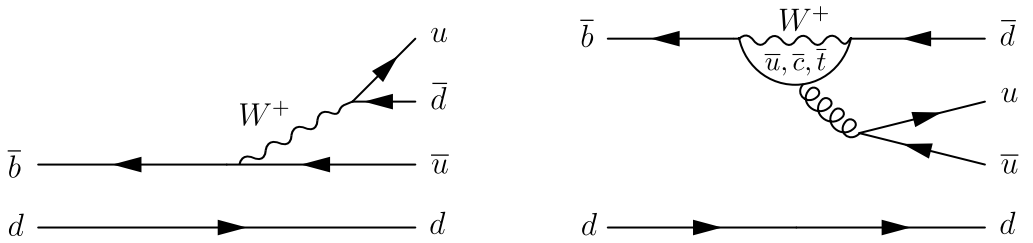


Figure 2.7.: Contributions of different topologies to the decay amplitude of $B^0 \to \pi^+\pi^-$: (left) tree diagram, (right) one-loop QCD penguin diagram.

CKM matrix elements governing the respective transitions. With

$$|V_{ud}| \sim 1, |V_{ub}| \sim \lambda_C^3 \quad \text{and} \quad |V_{tb}| \sim 1, |V_{td}| \sim \lambda_C^3 \ , \tag{2.78}$$

where λ_C represents the sine of the Cabibbo angle, we find a similar suppression of $\mathcal{O}(\lambda_C^3)$ for both transitions. Consequently, the penguin diagram with a different weak phase *pollutes* the extraction of ϕ_2 and leads to sizeable direct CP violation. In 1990, M. Gronau and D. London proposed to combine the measurements of all three isospin-related $B^0 \to \pi^+\pi^-$, $B^0 \to \pi^0\pi^0$ and $B^0 \to \pi^+\pi^0$ decay modes to retrieve the unpolluted angle ϕ_2 [38].

The proposed analysis relies on the approximate conservation of isospin symmetry in the hadronization process of $B \to \pi\pi$ transitions. With pions carrying an isospin of 1, the total isospin of the di-pion final-state, I, can in principle be 0, 1 or 2. The decay of B mesons, being spin-0 particles, dictates a total angular momentum of $J = 0$. As such, a final state with $I = 1$ would be anti-symmetric, following $(-1)^{J+I}$. Hence, adhering to Bose-Einstein statistics, only final states with a total isospin of either $I = 0$ or $I = 2$ are allowed.

Whereas the tree diagram can lead to both of these cases, with the gluon carrying no isospin, the penguin diagram results only in final states with $I = 0$. In the case of $B^0 \to \pi^+\pi^0$ with $I = 2$, the decay is only mediated by the tree diagram. The amplitudes governing the different decays can be decomposed into components of $I = 0$ and $I = 2$. Evaluating the Clebsch-Gordon coefficients, we can write

$$\frac{1}{\sqrt{2}} A^{+-} = A_2 - A_0 \ , \tag{2.79}$$

$$A^{00} = 2A_2 + A_0 \ , \tag{2.80}$$

$$A^{+0} = 3A_2 \ , \tag{2.81}$$

where A^{ij} represent the total amplitude of the respective decay, and A_I denote the components leading to $I = 0$ or 2. Rearranging the above equations yields the

following triangle relation in the complex plane:

$$\frac{1}{\sqrt{2}}A^{+-} + A^{00} = A^{+0} \ , \tag{2.82}$$

and accordingly for the charge-conjugated case:

$$\frac{1}{\sqrt{2}}\bar{A}^{+-} + \bar{A}^{00} = \bar{A}^{-0} \ . \tag{2.83}$$

The magnitudes of the decay amplitudes are obtained experimentally through the respective decay rate measurements, which, in the case of neutral decays, need to account for mixing (cf. Eq. 2.63). Using the definition of the CP-violation parameter λ_{CP}, and defining the following amplitude ratios,

$$z \equiv \frac{A_0}{A_2} \quad \text{and} \quad \bar{z} \equiv \frac{\bar{A}_0}{\bar{A}_2} \ , \tag{2.84}$$

we can write for $B^0 \to \pi^+\pi^-$,

$$\lambda_{CP} = \frac{q}{p}e^{-2i\phi_t}\left(\frac{1-\bar{z}}{1-z}\right) \overset{\left|\frac{q}{p}\right|\approx 1}{=} e^{-2i(\phi_t+\phi_M)}\left(\frac{1-\bar{z}}{1-z}\right) \ , \tag{2.85}$$

with ϕ_t and ϕ_M corresponding to the CKM phase of the tree-level amplitude and the weak mixing phase, respectively. Identifying the CKM elements governing both of these processes, we can relate them with the angles of the CKM triangle, $\phi_t = \phi_3$ and $\phi_M = \phi_1$. Further assuming CPT symmetry, $\phi_2 = \pi - \phi_1 - \phi_3$, we find for the imaginary part,

$$\mathrm{Im}(\lambda_{CP}) = \mathrm{Im}\left(e^{-2i\phi_2}\left[\frac{1-\bar{z}}{1-z}\right]\right) = \mathrm{Im}\left(e^{-2i\phi_2}\left[\frac{1-|\bar{z}|e^{\pm i\bar{\theta}}}{1-|z|e^{\pm i\theta}}\right]\right) \ . \tag{2.86}$$

For the second equality, we have introduced the angles θ and $\bar{\theta}$, to express the phases of z and $\bar{z}$. Following geometrical considerations of the triangle relations visualized in Figure 2.8, we can obtain all four parameters, z, $\bar{z}$, θ and $\bar{\theta}$, from measuring the decay rates, with a twofold ambiguity regarding the sign of the phases. The phase of the term in the square bracket can therefore take on four values, which we denote as $\pm\epsilon_\pm$ and $\pm\eta_\pm$. The angle ϕ_2 is then a solution to one of the four equations,

$$\sin(2\phi_2 \pm \epsilon_\pm) = \mathrm{Im}(\lambda_{CP,\pm})/m_\pm \tag{2.87}$$

$$\sin(2\phi_2 \pm \eta_\pm) = \mathrm{Im}(\lambda_{CP,\pm})/m_\pm \ , \tag{2.88}$$

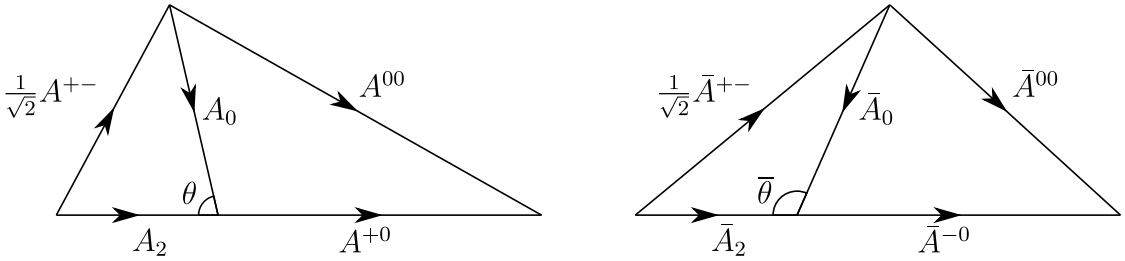


Figure 2.8.: Visual representation of the triangle relations between the three isospin-related decay modes of the $B \to \pi\pi$ system (Eq. 2.82 (left) and Eq. 2.83 (right)) [39]

where $m_\pm$ is the magnitude of the square bracket above. The fourfold ambiguity can be resolved by establishing the analogous expressions for the $B^0 \to \pi^0\pi^0$ system, which yields

$$\sin(2\phi_2 \pm \epsilon_{00}) = \mathrm{Im}(\lambda_{CP,00})/m_{00} \tag{2.89}$$
$$\sin(2\phi_2 \pm \eta_{00}) = \mathrm{Im}(\lambda_{CP,00})/m_{00}\ . \tag{2.90}$$

The time-dependant measurement of $B^0 \to \pi^0\pi^0$ is experimentally very challenging, as it requires the precise reconstruction of the decay vertex from four photons. Additionally, with CP-eigenstate final states, both neutral decay modes require the use of a flavor-tagger algorithm, which further complicates the extraction of the CP-violation parameters. At this stage of the Belle II experiment, a precise measurement of ϕ_2 using the outlined analysis is not feasible with the limited data currently available. A dedicated study estimates the precision of Belle II with an integrated luminosity of $50\,\mathrm{ab}^{-1}$ to $\delta\phi_2 \approx 0.6°$ [40]. This projection includes the related analysis of $B \to \rho\rho$ and would correspond to a five-fold increase compared to current world average precision.

$b \to sud$ transitions

The decay amplitude of $B^0 \to K^+\pi^-$ transitions receives dominantly contributions from the tree diagram and the QCD one-loop penguin diagram, which are visualized in Figure 2.9. Comparing the magnitudes of both topologies in terms of the CKM-matrix elements entering the transitions,

$$|V_{ub}| \sim \lambda_C^3, |V_{us}| \sim \lambda_C \quad \text{and} \quad |V_{tb}| \sim 1, |V_{ts}| \sim \lambda_C^2\ , \tag{2.91}$$

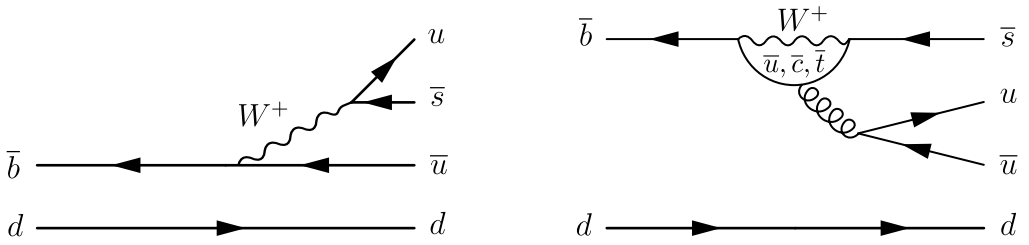


Figure 2.9.: Contributions of different topologies to the decay amplitude of $B^0 \to K^+\pi^-$: (left) tree diagram, (right) one-loop QCD penguin diagram.

we find that the penguin-diagram contribution dominates with a suppression of $\mathcal{O}(\lambda_C^2)$. This feature renders the decay potentially sensitive to heavy new-physics particles entering the loop, which could lead to measurable deviations from SM expectations.

Further, the interference between both diagrams, with different weak and strong phases, can lead to direct CP violation. As the final state is *self-tagging*, the flavor of the decaying B meson is unambiguously determined by the charge of the kaon. As such, the direct CP-violation parameter can be obtained via a time-integrated measurement without the need to infer the initial flavor using the tag-side decay (cf. Eq. 2.73). However, large theoretical uncertainties in predicted strong interaction phases impede the extraction of the weak phase ϕ_3 from the direct CP asymmetry.

Considerable interest in this decay mode was sparked by the observation of a significant discrepancy between CP asymmetries in $B^0 \to K^+\pi^-$ and $B^+ \to K^+\pi^0$ decays. In the limit of $SU(3)$ symmetry and the absence of color-suppressed tree and electroweak penguin contributions, the asymmetries were expected be equal, $\Delta\mathcal{A}_{CP} = \mathcal{A}_{CP}(K^+\pi^-) - \mathcal{A}_{CP}(K^+\pi^0) \approx 0$ [41]. However, independent measurements by Belle, BaBar, LHCb and CDF have consistently confirmed a deviation from this naive expectation with a significance of more than 5σ [20]. Many theoretical conjectures have since been brought forward to try to resolve this $\Delta\mathcal{A}_{CP}$ puzzle (see [42, 43, 44, 45] and references therein). Whereas an enhanced color-suppressed tree contribution is at odds with well-established SM parameters, it would resolve the discrepancy without the need for NP. The second prominent approach is the inclusion of a sizable electroweak penguin contribution, which would indicate physics beyond the SM. With considerable uncertainties in many of the assumptions, it is fair to conclude that the theoretical interpretation of these measurements remains somewhat unclear [46].

To obtain a model-independent test of the SM, in [47] M. Gronau suggests to employ isospin symmetry among all $B \to K\pi$ modes to relate dominant terms in the CP asymmetries. One significant advantage of the proposed sum rule is the

cancellation of common systematic uncertainties in theoretical predictions as well as experimental measurements. In the following, we will briefly outline the derivation, following the author's line of arguments. The application, using the most recent results for the measurements of all $B \to K\pi$ modes at Belle II, which includes the measurement presented in this thesis, is detailed in Section 6.4.

Based on the effective Hamiltonian describing $B \to K\pi$ decays, in the diagrammatic approach the total transition amplitudes can be decomposed into terms of distinct topologies, which correspond to the operators in the Hamiltonian [48, 49]. We use the following six diagrams: the color-favored and color-suppressed tree amplitudes T and C, the gluonic penguin amplitude P, color-favored and color-suppressed electroweak penguin amplitudes P_{EW} and P_{EW}^C, and the annihilation amplitude A. Assuming unitarity of the CKM triangle, the transition amplitudes of the $B \to K\pi$ system can be written as:

$$-A\left(K^+\pi^-\right) = \lambda_u(P_{uc} + T) + \lambda_t(P_{tc} + \frac{2}{3}P_{EW}^C) \ , \tag{2.92}$$

$$-\sqrt{2}A\left(K^+\pi^0\right) = \lambda_u(P_{uc} + T + C + A) + \lambda_t(P_{tc} + P_{EW} + \frac{2}{3}P_{EW}^C) \ , \tag{2.93}$$

$$\sqrt{2}\left(K^0\pi^0\right) = \lambda_u(P_{uc} - C) + \lambda_t(P_{tc} - P_{EW} - \frac{1}{3}P_{EW}^C) \ , \tag{2.94}$$

$$A\left(K^0\pi^+\right) = \lambda_u(P_{uc} + A) + \lambda_t(P_{tc} - \frac{1}{3}P_{EW}^C) \ , \tag{2.95}$$

where we have defined $\lambda_q \equiv V_{qb}^* V_{qs}$, as well as $P_{tc} \equiv P_t - P_c$ and $P_{uc} \equiv P_u - P_c$.[6] Illustrations of the above Feynman diagrams are given in Appendix A.1.

Alternatively, these four amplitudes can be decomposed into isospin amplitudes, according to the respective relative change ΔI of the transition. Analogous to the derivation of the isospin relation in the previous section, for $B \to K\pi$ decays there is a single $\Delta I = 1/2$ contribution from the iso-singlet part of the Hamiltonian, $B_{1/2}$, and two allowed amplitudes, $A_{1/2}$ and $A_{3/2}$ from the iso-triplet part [50]. The physical amplitudes are then given by

$$-A\left(K^+\pi^-\right) = B_{1/2} - A_{1/2} - A_{3/2} \ , \tag{2.96}$$

$$-\sqrt{2}A\left(K^+\pi^0\right) = B_{1/2} + A_{1/2} - 2A_{3/2} \ , \tag{2.97}$$

$$\sqrt{2}A\left(K^0\pi^0\right) = B_{1/2} - A_{1/2} + 2A_{3/2} \ , \tag{2.98}$$

$$A\left(K^0\pi^+\right) = B_{1/2} + A_{1/2} + A_{3/2} \ . \tag{2.99}$$

[6] Here, the index of P_q indicates the flavor of the virtual quark in the loop of the penguin diagram.

Next, we define the rate difference between CP-conjugates transitions as,

$$\Delta(B \to f) \equiv \Gamma(\bar{B} \to \bar{f}) - \Gamma(B \to f) \ , \tag{2.100}$$

which, using Eqs. 2.92-2.95, yields

$$\Delta\left(K^+\pi^-\right) = \mathrm{Im}\left[(P_{tc} + \frac{2}{3}P_{EW}^C)(T + P_{uc})^*\right] 4\mathrm{Im}(\lambda_t\lambda_u^*) \ , \tag{2.101}$$

$$2\Delta\left(K^+\pi^0\right) = \mathrm{Im}\left[(P_{tc} + P_{EW} + \frac{2}{3}P_{EW}^C)(T + C + A + P_{uc})^*\right] 4\mathrm{Im}(\lambda_t\lambda_u^*) \ , \tag{2.102}$$

$$2\Delta\left(K^0\pi^0\right) = \mathrm{Im}\left[(P_{tc} - P_{EW} - \frac{1}{3}P_{EW}^C)(-C + P_{uc})^*\right] 4\mathrm{Im}(\lambda_t\lambda_u^*) \ , \tag{2.103}$$

$$\Delta\left(K^0\pi^+\right) = \mathrm{Im}\left[(P_{tc} - \frac{1}{3}P_{EW}^C)(A + P_{uc})^*\right] 4\mathrm{Im}(\lambda_t\lambda_u^*) \ . \tag{2.104}$$

The rate differences can be combined by defining the difference $\delta_{K\pi}$,

$$\delta_{K\pi} \equiv \Delta\left(K^+\pi^-\right) + \Delta\left(K^0\pi^+\right) - 2\Delta\left(K^+\pi^0\right) - 2\Delta\left(K^0\pi^0\right) . \tag{2.105}$$

Following isospin symmetry, the dominant terms P_{tc} vanish in a combination, that can be obtained from Eqs. 2.96-2.99,

$$-A\left(K^+\pi^-\right) + A\left(K^0\pi^+\right) + \sqrt{2}A\left(K^+\pi^0\right) - \sqrt{2}\left(K^0\pi^0\right) = 0 \ . \tag{2.106}$$

The difference then contains only the following subleading terms,

$$\begin{aligned}\delta_{K\pi} = -\mathrm{Im}\Big[&(P_{EW} + P_{EW}^C)(T + C)^* + (P_{EW}C^* - P_{EW}^C T^*) \\ &+ (P_{EW} + P_{EW}^C)A^*\Big] 4\mathrm{Im}(\lambda_t\lambda_u^*) \ .\end{aligned} \tag{2.107}$$

In the $SU(3)$ and heavy quark limits, all terms vanish or are doubly suppressed relative to $\Delta\left(K^+\pi^-\right)$, and therefore $\delta_{K\pi} \approx 0$ is expected to hold within several percent. Finally, we can rewrite the sum rule (Eq. 2.105) in terms of direct CP-violation parameters $\mathcal{A}_{CP}$ and CP-averaged branching ratios $\mathcal{B}$:

$$I_{K\pi} = \mathcal{A}_{CP}^{K^+\pi^-} + \mathcal{A}_{CP}^{K^0\pi^+}\frac{\mathcal{B}_{K^0\pi^+}}{\mathcal{B}_{K^+\pi^-}}\frac{\tau_{B^0}}{\tau_{B^+}} - 2\mathcal{A}_{CP}^{K^+\pi^0}\frac{\mathcal{B}_{K^+\pi^0}}{\mathcal{B}_{K^+\pi^-}}\frac{\tau_{B^0}}{\tau_{B^+}} - 2\mathcal{A}_{CP}^{K^0\pi^0}\frac{\mathcal{B}_{K^0\pi^0}}{\mathcal{B}_{K^+\pi^-}} \ , \tag{2.108}$$

where τ_0 and τ_+ represent the lifetimes of neutral and charged B mesons.

At Belle II, all parameters can be accessed experimentally within a consistent framework. The measurement of $B^0 \to K^0\pi^0$ decays is particularly challenging for

several reasons. For one, this mode is a CP eigenstate and, as such, requires the determination of the B-meson flavor using a flavor-tagging algorithm. The efficiency of the Belle II flavor tagger is in the order of 30%, significantly limiting the statistical precision of such measurements [51]. Additionally, both daughter particles further decay within the detector volume and need to be reconstructed from long-lived final-state particles, reducing both efficiency and vertexing accuracy. To address these challenges, the Belle II experiment was designed to record approximately 50 times the amount of B-meson decays compared to its predecessor at unparalleled precision. The following chapter will present some of the most crucial aspects of the experiment.

3. The Belle II Experiment

The Belle II experiment is a next-generation flavor factory experiment located at the SuperKEKB collider in Tsukuba, Japan. Building upon the successes of its predecessor, the Belle experiment, Belle II has been designed to explore the phenomenon of CP violation in B-meson decays with unparalleled precision and search for physics beyond the Standard Model at the intensity-frontier. The collaboration leading this effort consists of an international group of approximately 1000 physicists and engineers from universities and institutes in currently 26 countries around the globe.

The experiment is centered around the Belle II detector, a sophisticated apparatus consisting of various subdetectors, developed to efficiently and accurately reconstruct the decay products of B-meson decays. As the successor of the Belle experiment, which made crucial contributions towards establishing the CKM mixing mechanism in the SM, the primary focus of Belle II has shifted towards the exploration of new physics scenarios. Next to overconstraining the CKM triangle, precision measurements of rare processes with branching ratios of $\mathcal{O}(10^{-6})$ or less play a pivotal role in this pursuit. With the upgraded detector, the clean environment of e^+e^- collisions can also be leveraged to perform missing-mass analyses to look for low multiplicity and dark sector signatures. The integrated luminosity, targeted to exceed that of the original Belle experiment by approximately 50-fold, is of vital importance for the success of this rich physics program.

To produce the large number of B mesons, the SuperKEKB accelerator provides high-intensity electron and positron beams at center-of-mass energies in the region of the $\Upsilon(4S)$ resonance, colliding them precisely at the center of the Belle II detector. The accelerator has undergone significant upgrades to achieve higher collision rates and beam intensities, currently holding the world record of $4.71 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in peak luminosity [52].

In the following sections, we give an outline of the SuperKEKB facility and the Belle II detector, highlighting the key technologies of the individual components. More in-depths descriptions can be found in [53, 54, 55, 56].

An equally crucial driver of the experiment is the dedicated Belle II Analysis Software Framework, or short basf2. It encompasses all key elements of the data-processing chain, including the generation and simulation of Monte-Carlo datasets, the processings of raw simulated and real detector signals, as well as the reconstruction of high-level analysis objects, needed to extract the desired physics results. The

description of all these software aspects is beyond the scope of this thesis and can be found in [57, 55].

3.1. SuperKEKB Accelerator

The SuperKEKB accelerator is an asymmetric e^+e^- collider built as an upgrade to its predecessor, the KEKB B-factory, which was in operation between 1998 until 2010 [58]. The facility comprises a linear accelerator and two storage rings, of about 3 km circumference each. After their production, bunches of positrons and electrons are accelerated to their final energies, before being injected into their respective storage rings. The high-energy ring (HER) containing electrons and the low-energy ring (LER) containing positrons intersect in the center of the Belle II detector, where the bunches are focused to collide for the efficient production of B mesons. A schematic drawing of the facility is shown in Figure 3.1.

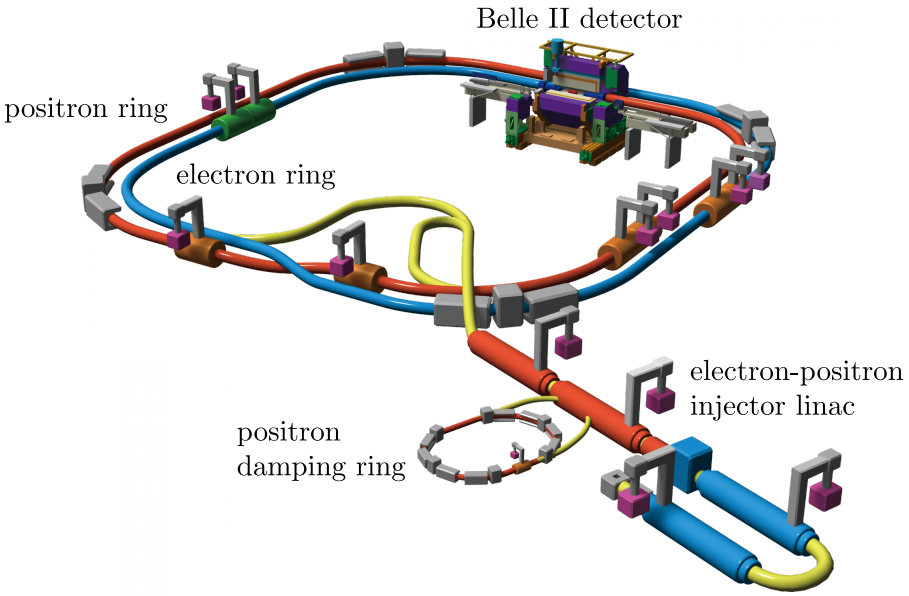


Figure 3.1.: Schematic drawing of the SuperKEKB facility, consisting of the linear accelerator and the two storage rings, along with the Belle II detector [56].

A key parameter in collider experiments is the instantaneous luminosity, $\mathcal{L}$, which, together with the cross-section σ for a given process, determines its event rate,

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \mathcal{L} \cdot \sigma \propto \frac{I}{\beta_y^*} \,. \tag{3.1}$$

The term on the right-hand-side expresses the dependence on the beam current, I, and the vertical beta function at the interaction point (IP), β_y^*, which characterizes the focusing strength of the particle beams in the vertical plane. To achieve the production rates needed for the success of the Belle II experiment, SuperKEKB aims to increase the luminosity by a factor of 40 compared to its predecessor [56].

An increase by a factor of two compared to KEKB is obtained by a doubling of the effective beam currents. Further, the novel *nano beam scheme*, proposed by P. Raimondi, is designed to significantly reduce the vertical beta function by another factor of 20 by minimizing the overlap of the bunches at the IP [54]. This is achieved by focusing the vertical component of both beams to widths of only around $\sigma_y^* \approx 50\,\text{nm}$ in the overlap region. Additionally, opposed to a more traditional head-on collision design, in the nano beam scheme, a large half-crossing angle of $\theta_y = 41.5\,\text{mrad}$ is chosen, yielding a significant reduction in the size of the interaction region along the z-direction (see Figure 3.2).

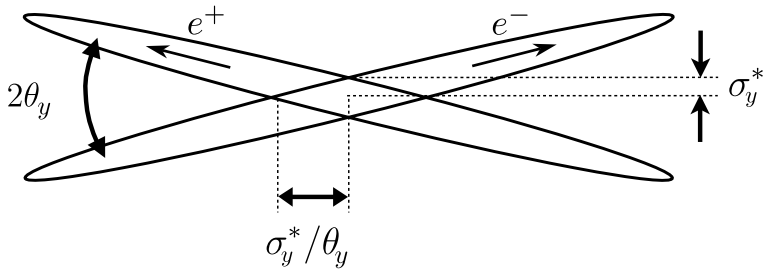


Figure 3.2.: Illustration of the beam-beam crossing in the nano beam scheme, used at SuperKEKB to minimize the effective overlap of bunches at the interaction point [59].

The accelerator is capable of delivering collisions in the center-of-mass energy range from slightly below the $\Upsilon(1S)$ resonance at 9.46 GeV to just above the $\Upsilon(6S)$ resonance at 11.24 GeV. In order to maximize the production cross-section for $B\bar{B}$ pairs, the main running mode of SuperKEKB lies on the peak of the $\Upsilon(4S)$ resonance at 10.58 GeV, which decays almost exclusively to a pair of B mesons.

For time-dependent decay rate measurements, the center-of-mass system is boosted along the beam direction (cf. Section 2.3.2). With beam energies of 7 GeV for electrons and 4 GeV for positrons, the boost is reduced by a factor of approximately 2/3 compared to KEKB, leading to an average distance between the decay vertices of two B mesons of around 130 µm. Although it comes with some advantages like a reduced Touschek scattering within the positron bunches and an increased solid angle acceptance for missing-energy analyses, this choice demands an improved vertexing resolution. The newly developed pixel vertex detector at Belle II effectively

addresses this challenge.

3.2. Belle II Detector

The Belle II detector is a magnetic spectrometer with large solid angle coverage, built around the interaction point. It is symmetric around the beam axis, but has an asymmetric acceptance to account for the boost of the $\Upsilon(4S)$, covering polar angles between 17° and 150°. The detector design is centered around the goal of efficiently reconstructing and identifying final state particles of B-meson decays with high accuracy. At Belle II, the following final-state particles are considered:

- charged particles: $e^{\pm}, \mu^{\pm}, \pi^{\pm}, K^{\pm}, p^{+}, d^{+}$
- neutral particles: γ, K_L^0 .

In order to precisely reconstruct the decay properties of the decaying B mesons, a number of subdetectors is employed, each specifically developed to deliver complementary information. From this, important parameters like charge, energy or momentum of the decay products can be reconstructed and used in further steps of the analysis.

Figure 3.3 shows a schematic drawing of the cross-section through the Belle II detector. Though the overall architecture closely resembles that of the Belle detector, with the significantly higher luminosity of SuperKEKB, increased event and background rates required substantial upgrades to the detector components, including the trigger and data acquisition system. In the following sections, a brief overview of each subsystem will be given.

3.2.1. Vertex detector (VXD)

The innermost detector component is the new Vertex Detector (VXD), which consists of the silicon Pixel Detector (PXD) and the Silicon Vertex Detector (SVD). The main purpose of the VXD is to enable the precise vertex reconstruction of short-lived decay products for time-dependant measurements, which require accurate vertexing resolutions in the order of $\mathcal{O}(10\mu m)$.

PXD

Due to the increased beam current and a reduced distance to the interaction point, at Belle II there is a substantial increase of background radiation at the first layer of the vertex detector. To reduce the occupancy to a manageable level, the first two layers, at radii of only 14 mm and 22 mm, consist of pixel sensors. They comprise a

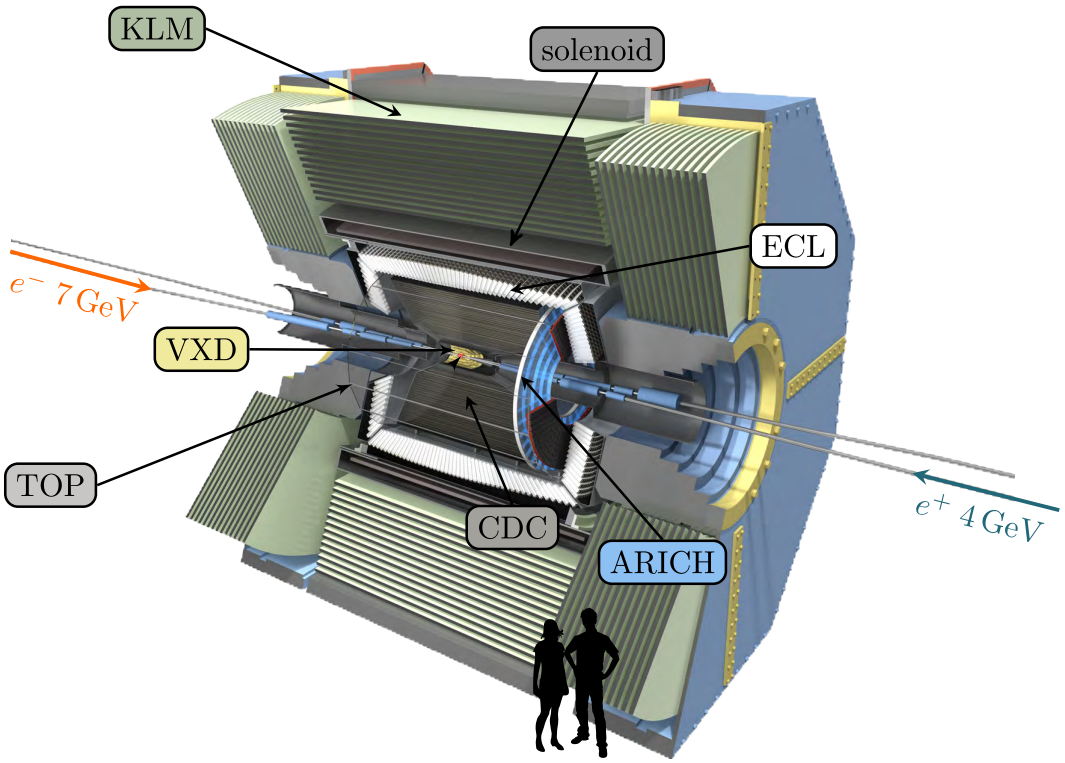


Figure 3.3.: Cross-section through the Belle II detector, composed of a number of subdetectors for the efficient reconstruction of B-meson decays. The outline of two humans is shown for scale.

total of 20 ladders, cylindrically arranged around the beam pipe and make up the PXD, depicted in Figure 3.4.

The sensors are based on the Depleted p-channel Field Effect Transistor (DEPFET) technology [62], which allows for an internal, primary amplification of charges created by traversing ionizing particles. The sensors are made from single silicon wavers without the need for additional structural support or active cooling in the active area. In order to minimize multiple scattering for an improved tracking of low transverse momentum particles, the monolithic structures are thinned down to a thickness of 75 µm in the sensitive area, corresponding to only 0.2% of a radiation length. With varying pixel sizes from $(50\times55)\,\mu m^2$ in the center of the inner layer to $(50\times85)\,\mu m^2$ in the outer layer, the PXD yields an impact parameter resolution of less than 15 µm, enabling precise tracking of charged particles and vertex reconstruction [61]. Additionally, the sensors exhibit low electronic noise levels, ensuring a high signal-to-noise ratio and the efficient detection of low-energy particles. The DEPFET technology further demonstrates fast readout capabilities to handle the high event rates in the harsh collision environment at Belle II. The precise timing information,

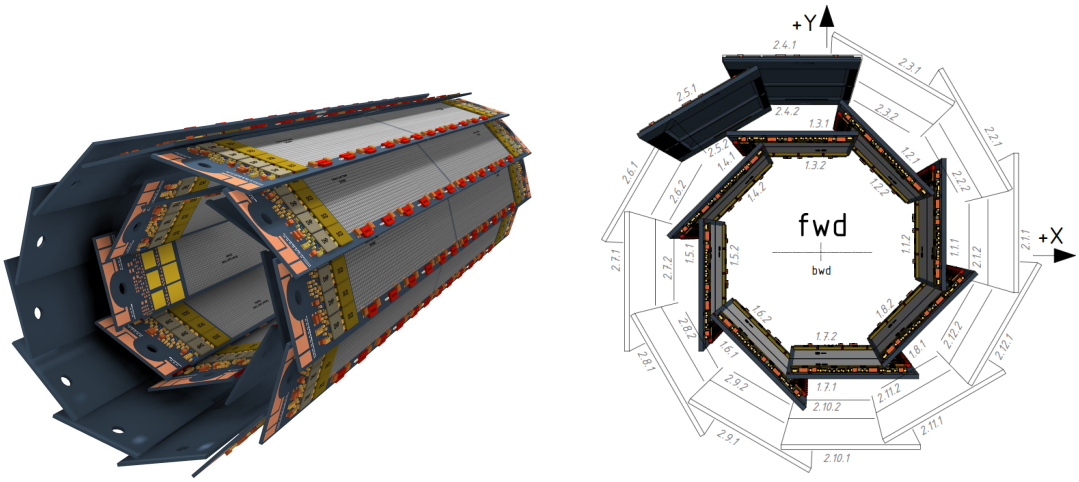


Figure 3.4.: Illustration of the silicon pixel detector (PXD), cylindrically arranged around the beam pipe. In the nominal design, the PXD consists of 8 inner and 12 outer ladders (left) [60]. Due to manufacturing problems, only 2 ladders had been implemented in the outer layer for the data taking period studied in this work (right) [61]. During the 2023 shutdown, the PXD was completed.

enabled by the PXD, is used to aid background rejection [53].

SVD

The VXD is completed by four layers of double-sided silicon strip sensors at radii of 39 mm, 80 mm, 104 mm, and 135 mm, concentrically arranged around the beam pipe. To accommodate the large solid-angle coverage and the boost of the center-of-mass system, the forward sensors in the outer three layers are mounted at angles towards the IP, improving the spatial resolution (see Figure 3.5).

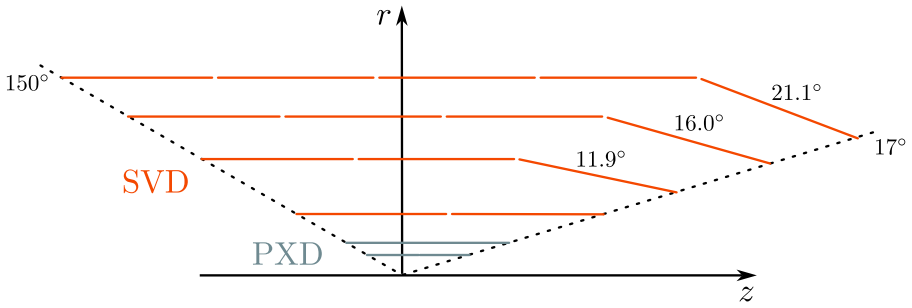


Figure 3.5.: Alignment of the Belle II vertex detector composed of two pixel sensor layers (PXD), and four strip sensor layers (SVD).

The sensor thickness varies between 320 µm for the slanted sensors, and 300 µm for all others, which corresponds to 0.7% of a radiation length. The strips measuring the distance along the beam axis (z) have a pitch size of $160 - 240$ µm, while those measuring the position in the $r - \phi$ plane have a pitch size of $50 - 75$ µm, depending on the layer.

With the small overall material budget, the VXD allows for a reconstruction of low transverse momentum tracks down to few tens of MeV/c. Furthermore, the SVD provides data to extrapolate tracks reconstructed in the outer detectors to hits in the PXD with high efficiency, aiding an efficient background rejection, required at high luminosities [63].

3.2.2. Central Drift Chamber (CDC)

The Central Drift Chamber (CDC) serves as the main tracking detector at Belle II. In combination with the 1.5 T magnetic field (see Section 3.2.5), it allows to determine the momentum of charged particles from their track curvature. It is also used for particle identification (PID) and background rejection at the trigger level.

The CDC consists of a gas-filled volume of cylindrical shape with an inner radius of 16 cm and an outer radius of 113 cm, interlaced with a large number of sense wires to measure ionization charges, caused by traversing ionizing particles. To account for the boost of the $\Upsilon(4S)$, it is asymmetric along the beam axis. The detector is composed of 56 layers of sense wires, further arranged into nine so-called super-layers. These super-layers alternate in their orientation between axial (A) orientation, which coincides with that of the magnetic field, and stereo orientation. The latter corresponds to slightly skewed wires at small positive (U) or negative (V) angles relative to the beam axis. In combination, this arrangement allows for a full three-dimensional track reconstruction. A schematic of the layer configuration is shown in Figure 3.6.

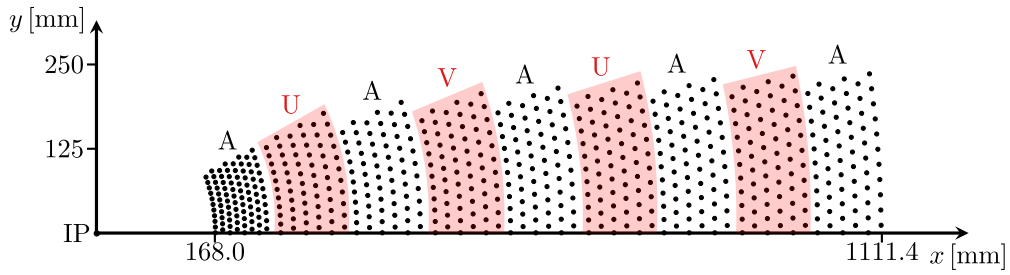


Figure 3.6.: Arrangement of the super-layers of the Belle II CDC, alternating between axial (A) and stereo (U, V) orientations to allow for a three-dimensional track reconstruction [40].

The wires are arranged in cells, each composed of a single sense wire surrounded by eight field wires, that generate strong electric fields. Traversing charged particles cause ionization charges in the C_2H_6-He gas mixture, which subsequently drift towards the surrounding sense wires. The resulting succession of time-resolved hits can be fitted to yield a precise reconstruction of the particle trajectory.

Next to timing and spatial information, the sense wires also provide signal amplitudes proportional to the collected charge, which can be used to measure the energy loss per unit length dE/dx. As the latter varies for different particle species, this information can be employed in the particle identification, which is especially important for low-momentum particles that do not reach the outer PID detectors. Figure 3.7 shows the energy deposition per unit length for different charged particles as a function of their momentum.

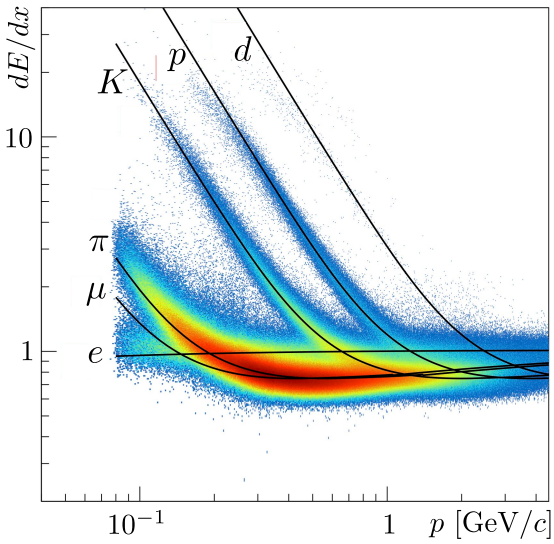


Figure 3.7.: Energy deposition per unit length, dE/dx, in the CDC for different charged particles as function of their momentum.

Lastly, the CDC is used as part of the trigger system at Belle II to suppress backgrounds originating from outside of the interaction region. To this end, a novel z-vertex trigger, based on a neural network approach, is employed to reconstruct the position along the beam axis using hit information of the CDC. This estimation is determined on dedicated Field Programmable Gate Arrays (FPGAs), without explicit track reconstruction and within the latency of the first level trigger [64].

3.2.3. Particle Identification with ARICH and TOP

Belle II employs two dedicated subdetectors for particle identification, the Time Of Propagation (TOP) detector and the Aerogel Ring-Imaging Cherenkov (ARICH)

detector. They both exploit the Cherenkov radiation emitted by charged particles passing through a dielectric medium at speeds greater than the speed of light in the medium,

$$v > c_n = \frac{c}{n} \ , \tag{3.2}$$

with n as the refractive index, and c as the speed of light in vacuum. The radiation is emitted as a light cone, where the emission angle, θ_C, depends on the particle's velocity as [65]

$$\cos\theta_C = \frac{c}{nv} = \frac{1}{n\beta} \ . \tag{3.3}$$

In combination with the momentum measurement provided by the CDC, the mass of the particle can be determined.

TOP

The TOP detector is arranged in a barrel-like geometry around the interaction point and uses quartz bars as Cherenkov radiators. Each of the 16 detector modules is composed of a 45 cm wide, 2 cm thick and 260 cm long quartz bar, with two rows of 16 multi-anode photon detectors mounted on one end and a focusing mirror on the other. The detectors are specially designed micro-channel-plate photo-multiplier tubes (MCP-PMT), offering a single-photon time resolution of $\mathcal{O}(50\,\text{ps})$. To achieve this precision, custom waveform-sampling readout electronics were developed [53].

After creation, some Cherenkov photons undergo total internal reflection until captured by fast multi-anode photon detectors. For a given momentum, the velocity-dependant Cherenkov angle leads to different path lengths for photons created by different particles (see Figure 3.8). The Cherenkov emission angle is determined by measuring the propagation time of individual photons from their emission point to the sensor plane. As there remain unavoidable ambiguities, PID information is obtained in terms of likelihoods by analyzing the timing distribution of photons in each of the 512 detector channels and comparing it with the expected Probability Density Functions (PDFs) for six different particle hypotheses (e, μ, π, K, p, d) [66].

ARICH

The Aerogel Ring Imaging Cherenkov (ARICH) counter provides PID information in the forward end-cap region of the Belle II detector. By detecting Cherenkov cones and measuring the corresponding Cherenkov angle, its main purpose is to distinguish between kaons and pions in the kinematic region between $0.4 - 4.0\,\text{GeV}/c$.

The detector is shaped like a rectangular torus with an inner radius of 420 mm and an outer radius of 1145 mm. It is composed of 124 pairs of wedge-shaped silica aerogel tiles, serving as radiator material, arranged in four concentric rings.

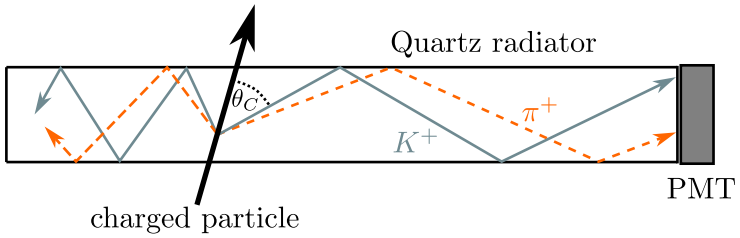


Figure 3.8.: Total internal reflection of Cherenkov photons inside a TOP quartz bar. The effective path length depends on the direction of the traversing particle and its velocity, which determines the Cherenkov angle θ_C.

Each of these pairs consists of two radiator layers with different refractive indices, $n_1 = 1.045$ and $n_1 = 1.055$, to increase the photon yield without degradation of the angular resolution [67]. For the detection, a total of 420 single-photon-sensitive hybrid avalanche photon detectors (HAPDs) are arranged in seven concentric rings. An illustration of the detection mechanism is shown in Figure 3.9.

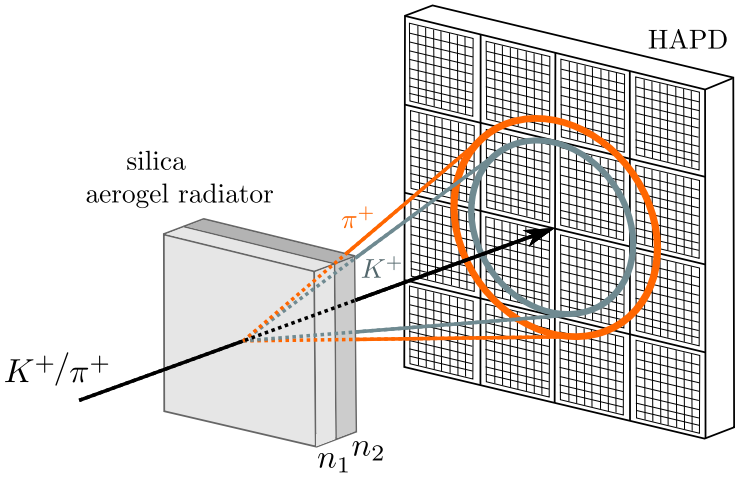


Figure 3.9.: Imaging of Cherenkov cones with the ARICH subdetector using silica aerogel radiator material and Hybrid Avalanche Photo Detectors (HAPD).

3.2.4. Electromagnetic Calorimeter (ECL)

The Electromagnetic Calorimeter (ECL) is composed of a highly segmented array of 8736 thallium-doped caesium iodide CsI(Tl) scintillator crystals. Its main purpose is the detection of photons and high-energy neutral pions in the energy range from few MeV to few GeV. It is further used to discriminate between electrons and hadrons.

Electrons and photons traversing the crystals deposit their energies in electromagnetic showers through the processes of bremsstrahlung or pair production. The resulting scintillation light is proportional to the deposited energy and is measured by photo-sensitive detectors. The E/p ratio, derived from the energy measurement of an ECL cluster relative to the momentum of its associated particle, serves as a valuable input for PID.

The ECL is made up of a 3 m long barrel section with an inner radius of 1.25 m, as well as two endcaps, covering polar angles between 12.4° and 155.1° (Figure 3.10). The CsI(Tl) crystals are oriented towards the interaction point and have typical dimensions of $6 \times 6\,\mathrm{cm}^2$ in cross-section and 30 cm in length, which corresponds to 16.2 radiation lengths. The scintillation is read out by two $10 \times 20\,\mathrm{mm}^2$ silicon PIN photodiodes attached to the rear surface of the crystals. To mitigate considerable pile-up noise due to the relatively long decay time of scintillations in the crystals, at Belle II new wave-form-sampling readout electronics are used.

3.2.5. K_L^0 and Muon Detector (KLM)

The ECL is encased by a superconducting solenoid that generates a uniform magnetic field of 1.5 T along the beam direction and serves to bend trajectories of charged particles via the exerted Lorentz force. The magnetic flux is returned by a return yoke, which consists of several iron plates that guide the magnetic flux to contain a homogeneous field inside the detector volume.

The iron plates further serve as additional absorber material for muons and long-lived neutral kaons, providing additional 3.9 interaction lengths. Whereas muons interact electromagnetically with the material and can be linked to their generated tracks, neutral kaons shower hadronically in the dense material, resulting in distinct cluster shapes. To measure their energy deposition, the iron plates are interleaved with active detector elements. In the forward and backward endcaps, as well as the first two innermost layers in the barrel region, the sensors consist of scintillator strips with wavelength-shifting fibers attached to silicon photo multipliers (SiPM) for light read-out. The remaining 13 outer barrel layers are equipped with Resistive Plate Counters (RPC). This multi-layer structure makes up the K_L^0 and Muon (KLM) detector and is the outermost subdetector at Belle II.

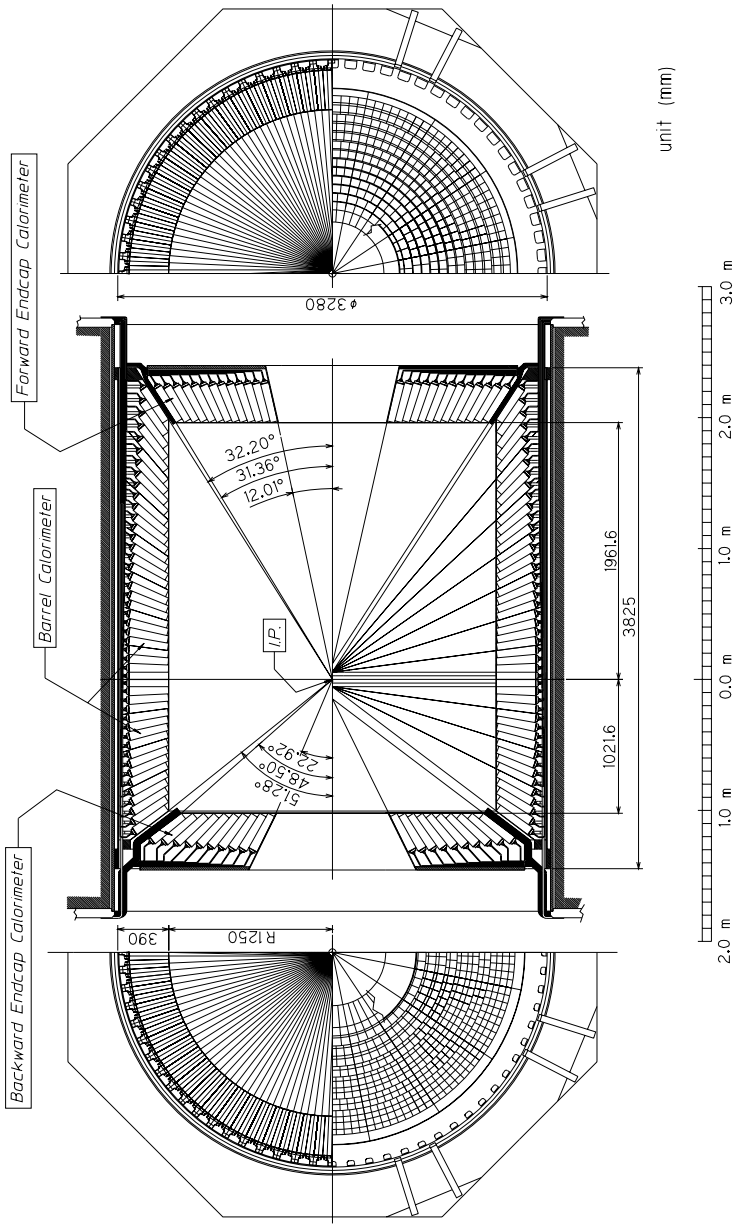


Figure 3.10.: Geometry of the Electromagnetic Calorimeter (ECL) [68].

4. Analysis of $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ Decays

4.1. Overview and Strategy

The goal of this study is to determine the branching ratios of $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ decays as well as the direct CP-violation parameter in $B^0 \to K^+\pi^-$ decays. The samples used for this analysis comprise the full dataset obtained with the Belle II detector between the start of the experiment in 2019 and 2022, recorded at a center-of-mass collision energy just above the production threshold of the $\Upsilon(4S)$ resonance at SuperKEKB. We dub this the *LS1* dataset, referring to the first long shutdown of the experiment, which marks the end of this data-taking period.

The reconstruction of candidate B mesons starts with the reconstruction of long-lived charged final state particles from tracks in the CDC. Using the respective mass-hypothesis, we fill lists of pions and kaons for each event and combine their four-vectors to create B-meson candidates. In this exclusive[1] decay analysis, it is sufficient to only reconstruct the signal-side of the decay tree, though some information of the tag-side is used to suppress background events.

At this stage of reconstruction, the majority of the data sample is composed of background events. In order to obtain a sample with sufficiently high signal-to-background ratio for a precise determination of the number of signal events, we impose quantitative requirements (*cuts*) on variables with high discrimination power between signal and background events.

One of the main challenges is posed by the fact that a large proportion of the e^+e^- collisions produce so-called continuum events, where pairs of lighter quarks $q\bar{q} = u\bar{u}, d\bar{d}, c\bar{c}, s\bar{s}$ are created. These events make up around 3/4 of the the total production, even at resonance. Consequently, such continuum events make up a dominant fraction of the total sample and present the major background contribution in many B-decay analyses. Especially in analyses of hadronic charmless decays – that is decays without any c quarks in the final state – the small decay rates of order

[1] Here, *exclusive* refers to the reconstruction of specific decay modes. In contrast, in *inclusive* analyses, a wide range of transitions is considered, generally requiring a reconstruction of the entire decay tree.

10^{-5} or lower amplify this circumstance. To reduce the amount to a manageable degree, we employ machine learning techniques trained on simulated Monte-Carlo events to better distinguish between B-meson decays and continuum events. After a loose cut, we further use the resulting classifier variable in the determination of signal yields.

After the final event selection, we identify three different components in our sample: correctly reconstructed signal events, misidentified events, in which one of the light hadrons was reconstructed with an incorrect mass-hypothesis, and continuum background events. In order to determine their respective yields, we fit the distributions of two variables in data using a maximum-likelihood fit. We simultaneously fit the energy-difference between the reconstructed energy and half of the center-of-mass energy, $\Delta E = E - \sqrt{s}/2$, and a transformed version of the continuum suppression classifier, C'. These variables provide good discrimination between all three components and are approximately uncorrelated.

The following sections detail all steps of the analysis, starting with the reconstruction of B-meson candidates from final-state particles and the baseline selection requirements. This is followed by a detailed explanation of the multivariate technique employed to efficiently suppress continuum background events. The remainder of this chapter describes the fitting methods used to extract the physics parameters.

Lastly, it is important to mention that all analysis procedures are first applied to simulated Monte Carlo (MC) experiments. This serves several purposes. For one, the extra information available in MC allows to unambiguously identify the entire decay chain for each reconstructed event. This is essential in finding variables, that can be employed to separate signal from background events and to determine the final reconstruction efficiencies. Secondly, we use simulated events to empirically extract the probability density functions (PDFs), describing the distribution of fit variables — a process often called model building. Lastly, we validate our entire analysis first on MC and perform additional checks on a control channel before applying it to data. This methodology is referred to as blind analysis and has become a standard practice in the high-energy physics community. The rationale is to develop the analysis without any biases that could, even unintentionally, arise if the outcome were known. Therefore, to avoid any tuning that could bias the results, the full analysis has to undergo a number of validation steps in MC as part of an internal review before the approval is given to unblind the full dataset. After the unblinding, the results are final and no changes can be made to any step of the analysis.

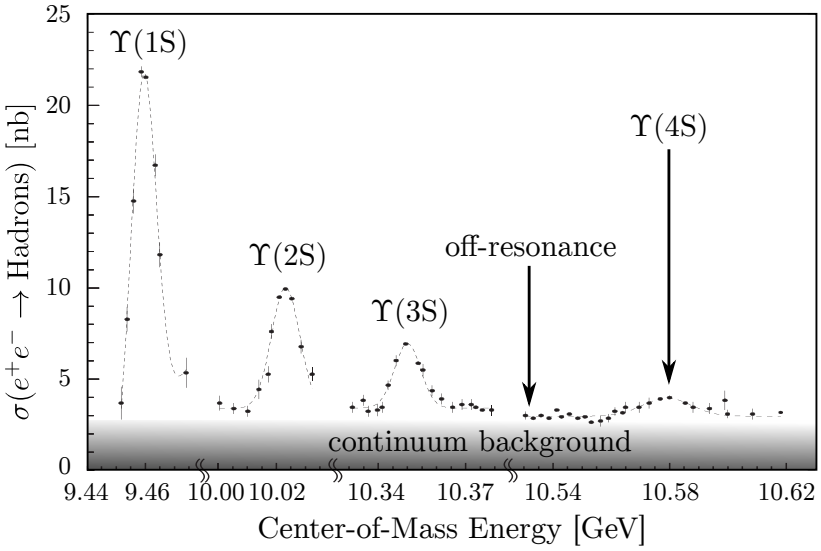


Figure 4.1.: Production cross-section as function of the center-of-mass energy for several Υ resonances. Adapted from [69].

4.2. Data Processing

4.2.1. Data Samples

For the measurements of the branching ratios and CP-violation parameter, we use the full LS1 data, collected with the Belle II detector between 2019 and 2022. The sample contains $(387 \pm 6) \times 10^6$ $B\bar{B}$ pairs and corresponds to an integrated luminosity of 362 ± 2 fb^{-1} recorded at the $\Upsilon(4S)$ resonance [70]. We employ a machine learning algorithm, trained on large samples of simulated events, to suppress the dominant background contribution from continuum events. To verify the agreement between data and simulation in the distributions of the training variables, we additionally reconstruct data containing only continuum background. This data is obtained by lowering the collision energy slightly below the production threshold of the $\Upsilon(4S)$ resonance. Whereas the production remains approximately unchanged for continuum, the energy is not sufficient to create pairs of B mesons. The off-resonance data sample corresponds to a total integrated luminosity of 42.3 ± 0.3 fb^{-1}.

Samples of simulated MC events are produced, modelling the physics processes resulting from e^+e^- collisions and their subsequent interactions with the detector. We employ a series of software packages to produce the simulated data: KKMC to generate continuum background [71], PYTHIA8 to simulate hadronization [72], EVTGEN to simulate particle decays [73], PHOTOS to simulate final state radiation [74], and GEANT4 to model interactions with the detector material [75].

The generic simulation sample consists of centrally produced, run-independent MC samples and include $B^0\bar{B}^0$, B^+B^-, $u\bar{u}$, $d\bar{d}$, $c\bar{c}$, $s\bar{s}$, $\tau^+\tau^-$, and simulated beam backgrounds in adequate proportions. The sample size corresponds to an integrated luminosity of $1\,\text{ab}^{-1}$. In addition, 2×10^6 events are generated for each signal channel, $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$. For these, one side of the generated $B\bar{B}$ events is forced to decay according to the desired decay mode, while leaving the B-meson decay on the tag-side unconstrained.

4.2.2. Reconstruction and Baseline Selection

Starting with recorded or simulated raw data, the signal provided by the different detector components is processed and further combined to a format more suitable for statistical analysis. The underlying principle is to reconstruct the particles responsible for recorded detector signals and successively combine them to reconstruct the original four-vectors of initial particles produced in the collision.

One crucial part of this process is the track reconstruction. Here, information provided by the tracking detectors are fed into tracking algorithms. In the first step, called track finding, these algorithms perform pattern recognition to identify

collections of hits stemming from individual trajectories of charged particles, which are referred to as tracks. Figure 4.2 illustrates this process of track finding in a slice of CDC hits for a single simulated event. Patterns of hits originating from the same particle are shown in colors, with a large fraction of remaining CDC hits rejected as background from either detector noise or beam-induced backgrounds shown in gray. In the subsequent step of track fitting, the best estimate for kinematics describing the particle trajectories, corresponding to the recorded hit pattern, is determined. This way, for each track, the particle's position and momentum near the interaction region can be precisely computed.

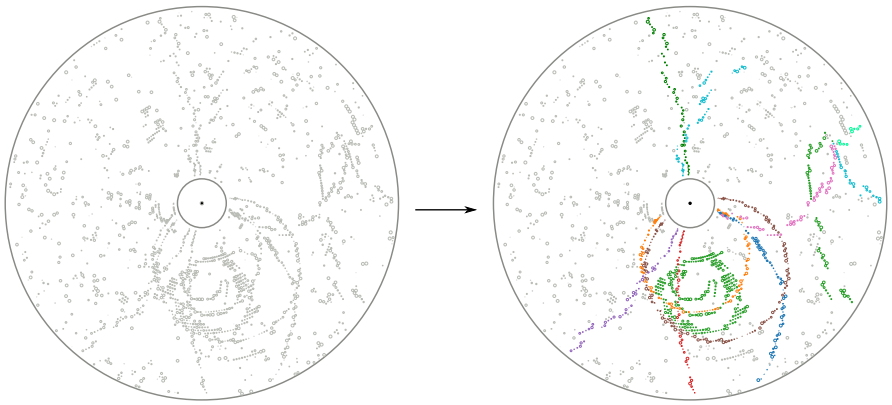


Figure 4.2.: Cross-section of the CDC in the x-y-plane for a single simulated event. Gray dots indicate registered hits and originate either from charged particles or detector noise. In the track finding process, patterns of connected hits form tracks, as indicated by the coloring on the right. All remaining hits are rejected as background.

The second integral part of the reconstruction of decay products is the clustering. Subdetectors such as the ECL register hits in multiple neighboring cells, which can be caused by the same traversing particle. In order to group them together, clustering algorithms are used, which allow to calculate properties like cluster energy, shape or center position. As the characteristics of the cluster shapes depend on the particle type causing it, we can make use of this information to help identify the particle.

The identification of the particle type causing a detector signal is in most cases ambiguous. However, the different subdetectors each provide valuable information, that are used in combination to determine the likelihood for each type of observable particle to have caused the recorded track and/or cluster. For example, the total energy loss over the track length in the CDC is compared to the expected values for different particle types. Likewise, the ARICH is checked for a corresponding detection of Cherenkov radiation, which can be connected to the same trajectory

and carries complementary information. In this way, the combined likelihood of all subdetectors is computed for a given track or cluster and can be used to select the desired particle type.

In this study, we analyze decays to two charged hadrons in the final state. To reconstruct signal B-meson candidates, we fill lists of pion and kaon candidates, using their respective mass-hypothesis and applying loose baseline selection criteria. For charged pions and kaons, we require the particle's polar angle to be within the range of the CDC, that is $17° < \theta < 150°$, with at least 20 registered CDC hits. To reduce the amount of beam backgrounds, we impose loose criteria on the distance to the interaction region. We require the distance of the point-of-closest-approach (POCA) to the interaction point in the $r - \phi$ plane (dr) to be less than 0.5 cm and less than 2 cm in the direction of the beam (dz). We then combine their four-vectors in kinematic fits consistent with the topologies of the desired decay.

One variable providing good discrimination power between B-meson decays and continuum events is the so-called beam-constrained mass, M_{bc}. In $B\bar{B}$ events, the two B mesons originate from a two-body decay of the $\Upsilon(4S)$, hence carrying equal momenta and energies in the center-of-mass frame. The resolution of the invariant B-meson mass is significantly improved by substituting the reconstructed energy with the precisely known beam energy: $M_{\mathrm{bc}} = \sqrt{s/4 - \mathbf{p}^2}$. For neutral B mesons, this variable peaks around the B-meson mass at 5.280 GeV/c^2. We impose a signal enriching window of $5.272 < M_{\mathrm{bc}} < 5.288$ GeV/c^2. The width of this window corresponds to three times the detector mass resolution as obtained by a fit to simulated signal events.

Further, we restrict the energy difference between the reconstructed energy of the B meson and half of the center-of-mass energy, $-0.1 < \Delta E < 0.2$ GeV. An asymmetric window is chosen to reduce B-decay backgrounds peaking at the low end of the spectrum while retaining all signal events. As this variable provides excellent distinction power between signal and background decays, we later fit the remaining distribution in ΔE.

Lastly, we impose a loose cut on an event shape variable, calculated using the so-called Fox-Wolfram moments. The latter are constructed to carry topological information such as momenta, angular and energy distributions of reconstructed jets and are expressed in terms of spherical harmonics [76]. We restrict R2, defined as the ratio of the second and the zeroth Fox-Wolfram moment, to be smaller than 0.9. This cut removes a significant amount of continuum events without impacting signal efficiencies, before a dedicated continuum suppression based on a multivariate technique is applied at a later stage.

We use vertexing and flavor-tagging information for the classification of continuum-background events. To reduce beam backgrounds, we apply loose selection requirements, similar to those on the signal side, on tracks and clusters used in the

tag-side reconstruction. The selection is summarized in Table 4.1. We subsequently reconstruct the vertex of the accompanying tag-side B mesons using a Kalman-based vertexing algorithm, implemented in the `RAVE` fit package [77]. We use a category-based algorithm to tag the flavor of the tag-side B mesons [78].

Table 4.1.: Selection requirements for tracks and clusters used in the tag-side reconstruction. Next to cuts on impact parameters, we impose loose requirements on the transverse momentum of tracks, as well as on the reconstructed cluster timing and energy, to reduce beam-background contributions.

tracks	**clusters**
$17° < \theta < 150°$	$17° < \theta < 150°$
$\mathrm{p_t} > 75\,\mathrm{MeV}/c$	$\lvert\mathrm{clusterTiming}\rvert < 200\,\mathrm{ns}$
$\mathrm{dr} < 10\,\mathrm{cm}$	$\mathrm{E} > 0.05\,\mathrm{GeV}$
$\mathrm{dz} < 20\,\mathrm{cm}$	

4.2.3. Continuum Suppression

The B-meson pairs are created in e^+e^- collisions produced by the SuperKEKB collider at a center-of-mass energy of 10.58 GeV. Though this energy corresponds to the peak of the $\Upsilon(4S)$, the production of the resonance makes up only around $1/4$ of the total hadronic production cross section, with the remaining $3/4$ producing $e^+e^- \to q\bar{q}$ (where $q = u, d, s, c$) events, referred to as continuum (cf. Figure 4.1). The latter presents the dominant source of combinatorial background where random combinations of final state particles mimic the kinematic signatures of the signal decay.

After the baseline selection is applied, the majority of candidates are misreconstructed continuum events. To effectively suppress this background, we construct a multivariate classifier using a Boosted Decision Tree (BDT), which exploits the differences in event topologies between B-meson decays and continuum events. The lower masses of light quark pairs lead to a jet-like event shape in the hadronization process, compared to a more spherical topology in the case of $B\bar{B}$ events, which carry very little momentum in the center-of-mass frame (Figure 4.3).

We compute an array of event shape variables quantifying this difference. One illustrative example is the so-called thrust, which gauges the coherence of the set of

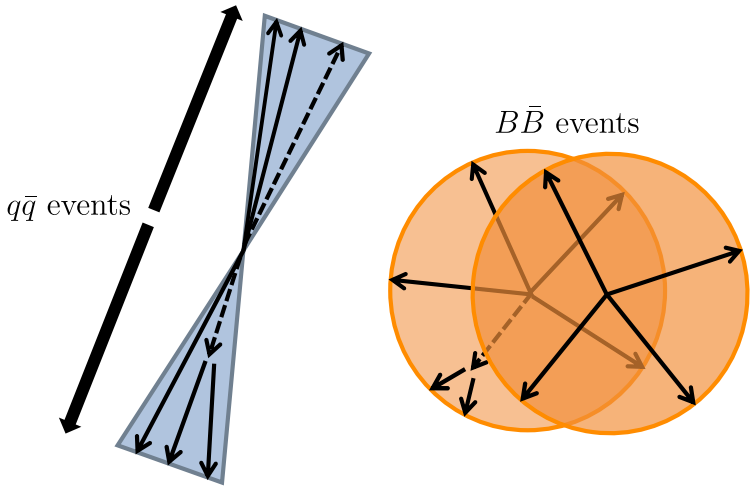


Figure 4.3.: Illustration of differences in the event topologies between continuum events, exhibiting jet-like event shapes, and $B\bar{B}$ events, carrying little momentum in the center-of-mass frame, leading to approximately spherical topologies.

detected particles created in the hadronization process of quarks. It is defined as:

$$T = \max_{|\mathbf{n}|=1} \frac{\sum_i |\mathbf{p}_i \cdot \mathbf{n}|}{\sum_i |\mathbf{p}_i|} , \tag{4.1}$$

where $\mathbf{p}_i$ is the momentum of particle i in the , $\mathbf{n}_i$ is a unit vector that maximizes the thrust, and we sum over all particles i. In the case of $B\bar{B}$ events, both B mesons are produced almost at rest in the $\Upsilon(4S)$ rest frame, leading to an isotropic distribution of their decay products. For continuum events events on the other hand, where all trajectories are approximately aligned with the thrust axis defined by $\mathbf{n}$, the scalar product evaluates to values close to 1.

We use the stochastic gradient-boosted decision algorithm `FBDT` [79], implemented natively within basf2, to combine a number of variables, known to provide statistical discrimination between signal and continuum background events. Starting from a large set of observables, we successively reduce the set of training variables, keeping only those with an importance for the distinction (or *feature importance*) greater than 0.1%, as evaluated by `FBDT`. We choose a final subset of 35 variables, whose absolute correlations with ΔE are below 5% in each component. This allows to factorize the two-dimensional PDFs and minimizes sculpting effects when cutting on the classifier later on. Figure 4.4 lists all input variables used for the training in order of their feature importance.

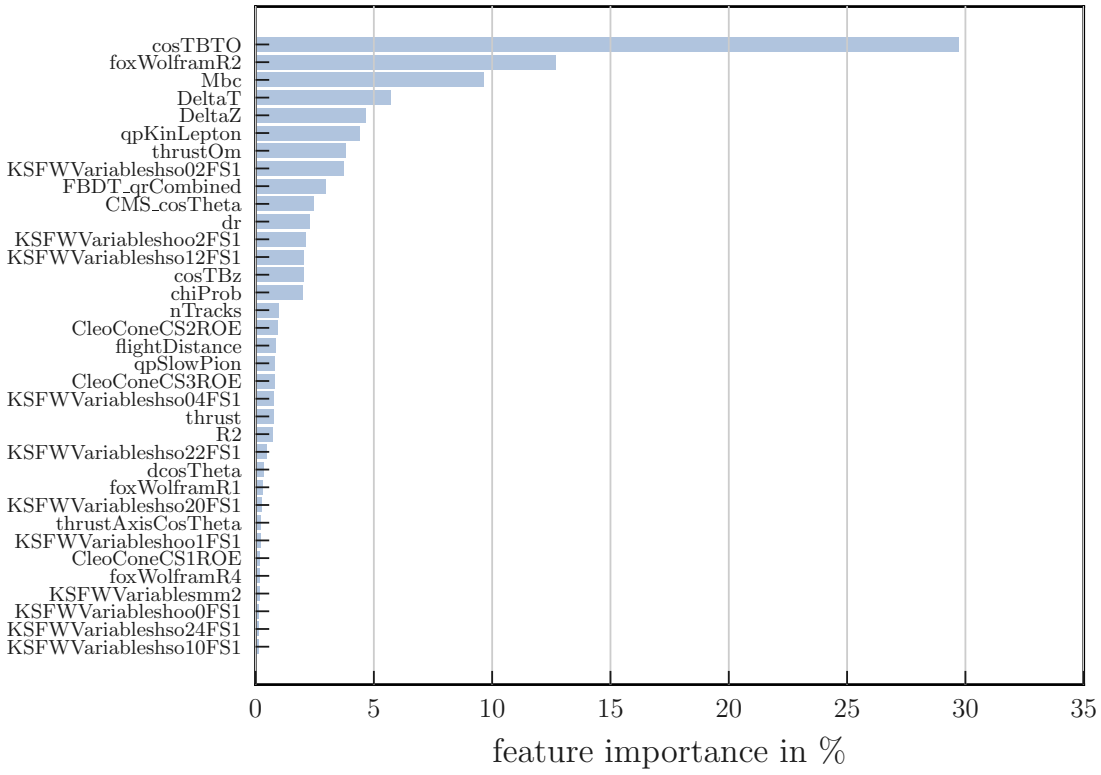


Figure 4.4.: List of variables used in the BDT training for the continuum suppression in order of their feature importance.

The following provides short descriptions for the training variables. More detailed definitions can be found in [55].

- `cosTBTO` – cosine of angle between thrust axis of signal B meson and thrust axis of tag-side B meson. The isotropic distribution of decay particles in $B\bar{B}$ events yields an approximately uniform distribution in the range $[0, 1]$. In continuum events, the momenta of decay particles are highly collimated, resulting in a distribution strongly peaked around values near 1.

- `foxWolframR1`, `foxWolframR2`, `foxWolframR4` – ratio of n-th to 0-th order Fox-Wolfram moments. The latter quantify event shapes by expanding the geometric distribution of particle momenta in terms of Legendre polynomials [76].

- `Mbc` – beam-constrained mass. B-meson decays peak at the nominal B-meson mass, while continuum events follow the shape of an ARGUS distribution.

- `DeltaT` – decay time difference between signal and tag-side B-meson decay. The finite lifetime of B-mesons results in a wider distribution for $B\bar{B}$ events,

compared to lighter quark pairs in continuum events, which hadronize immediately.

- `DeltaZ` – spatial difference between signal and tag-side B-meson decay vertices in beam direction
- `qpKinLepton` – charge of tag-side track with highest probability of being a primary lepton multiplied by respective probability
- `thrustOm` – magnitude of thrust of tag-side B meson
- `KSFWVariables` – modified Kakuno-Super-Fox-Wolfram moments, designed to differentiate between various particle classes in computation
- `FBDT_qrCombined` – tagged flavor, q, multiplied by respective dilution factor r, provided by flavor-tagging algorithm. In $B\bar{B}$ events, q typically corresponds to the opposite flavor of the reconstructed signal B meson, while it is random for continuum events.
- `CMS_cosTheta` – cosine of polar angle of signal B meson in center-of-mass frame. To conserve the total angular momentum of the spin-1 $\Upsilon(4S)$, polarized parallel to the beam axis, the spinless B mesons are created in a p-wave configuration. As a result, $B\bar{B}$ events exhibit a $1 - \cos^2\theta$ distribution, whereas continuum events follow a uniform distribution.
- `dr` – transverse distance of signal B meson with respect to interaction point
- `cosTBz` – cosine of angle between thrust axis of signal B meson and beam axis
- `chiProb` – vertex fit p-value of signal B-meson decay
- `nTracks` – total number of tracks in event
- CLEO cones – event shape variables designed to quantify flow of momentum within concentric volumes around thrust axis in intervals of 10°, using only tag-side particles [80]
- `flightDistance` – signal B-meson flight distance with respect to interaction point
- `qpSlowPion` – charge of tag-side track with highest probability of being a slow pion from decay of a primary D^* multiplied by its respective probability
- `thrust` – magnitude of thrust of event

- `dcosTheta` – cosine of polar angle of B-meson vertex with respect to interaction point
- `thrustAxisCosTheta` – cosine of polar angle of thrust axis
- `KSFWVariables(mm2)` – missing mass squared

In the context of supervised learning, the training of a BDT requires labeled data with prior knowledge of the true output value for each input. In our case, we face a binary classification problem, categorizing events as either signal or continuum background. For an effective continuum suppression, an abundant dataset containing both signal and background classes, each labeled with their true output values is needed.

With no way of accessing a sufficiently large sample of pure signal events in data, this limitation is overcome by relying on simulated Monte Carlo events. We train the BDT on a set of 5×10^5 correctly reconstructed signal events and 5×10^5 continuum events ($u\bar{u}$, $d\bar{d}$, $c\bar{c}$, $s\bar{s}$ and $\tau^+\tau^-$ in the expected proportions) surviving the baseline selection. The signal sample is composed of equal parts of $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ events. Subsequently, the total sample is divided into subsamples of 80% for training and 20% for testing.

Starting with the default hyperparameters of `FBDT`, we iteratively optimize the model architecture. A common challenge in machine learning is overtraining or overfitting, which occurs when the model is overly complex and closely fits the training data, losing its ability to generalize to new, unseen data. Therefore, determining the optimal model architecture involves finding the right balance between the complexity required for good discrimination and generalizability to new data. The final BDT architecture comprises 300 trees arranged in 5 layers.

Following training, the BDT algorithm is applied to data, generating a single output value, denoted C, for each event, which ranges continuously from 0 to 1. A value of 0 corresponds to a definite classification of continuum, while a value of 1 signifies a signal event. To check for overtraining, we assess the performance of the trained BDT model on both the training and test datasets by comparing the output distributions of the signal and background events (Figure 4.5). We observe excellent agreement.

A standard performance metric is provided by the ROC (receiver operating characteristic) curve, a graphical representation that illustrates the performance of a classification model across all possible classification thresholds. The ROC curve plots the true positive rate against the false positive rate. The area under the curve (AUC) is used as a measure of the classifier's discrimination power. Figure 4.6 shows the ROC curve for our BDT evaluated on the test sample, along with that of an alternative BDT, where we excluded M_{bc} from the training variables. Both

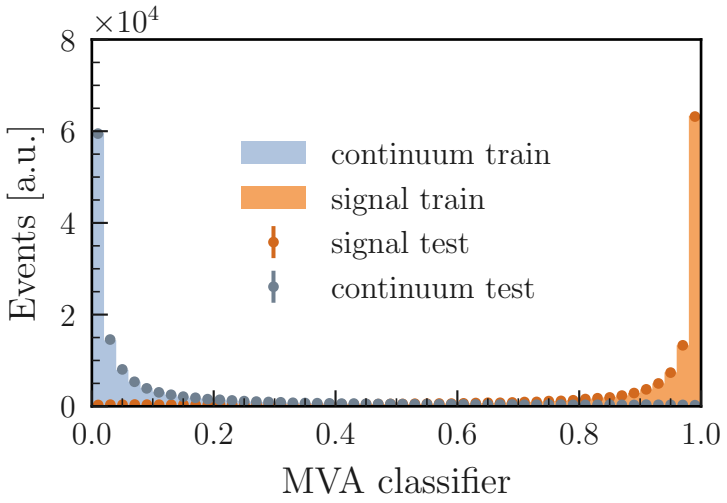


Figure 4.5.: Overtraining test. Comparing the BDT output in the training and test samples, we observe good agreement in both signal and continuum event distributions. Continuum events peak near 0, signal events near 1, illustrating the excellent separation power of the classifier.

classifiers demonstrate AUCs close to the maximum value of 1, illustrating the excellent performance of the classifiers.

After training, the BDT is applied to real data to classify the signal and background events. However, in order to ensure the reliability of the classification, it is essential to ensure that the simulation reproduces data sufficiently well. To this end, we compare the distribution of all training variables in the simulated data with those in the real data. Figs. 4.7–4.11 show the normalized distributions of all BDT input variables in real and simulated off-resonance data. To illustrate the distinction power of each variable, the distribution of simulated signal events is superimposed.

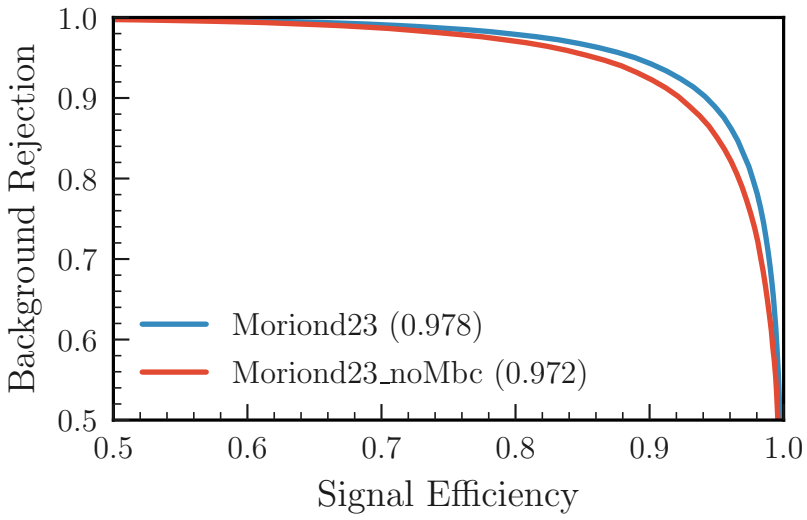


Figure 4.6.: Receiver Operating Characteristic (ROC) curve of the continuum-suppression classifier used for selecting signal-event candidates, evaluated on the test sample. The two ROC curves refer to different sets of training variables: `Moriond23` includes all training variables, `Moriond23_noMbc` excludes the beam-constrained mass M_{bc}. The former is used for the work presented here, motivated by the slight enhancement in the Area Under the Curve (AUC), shown in parentheses.

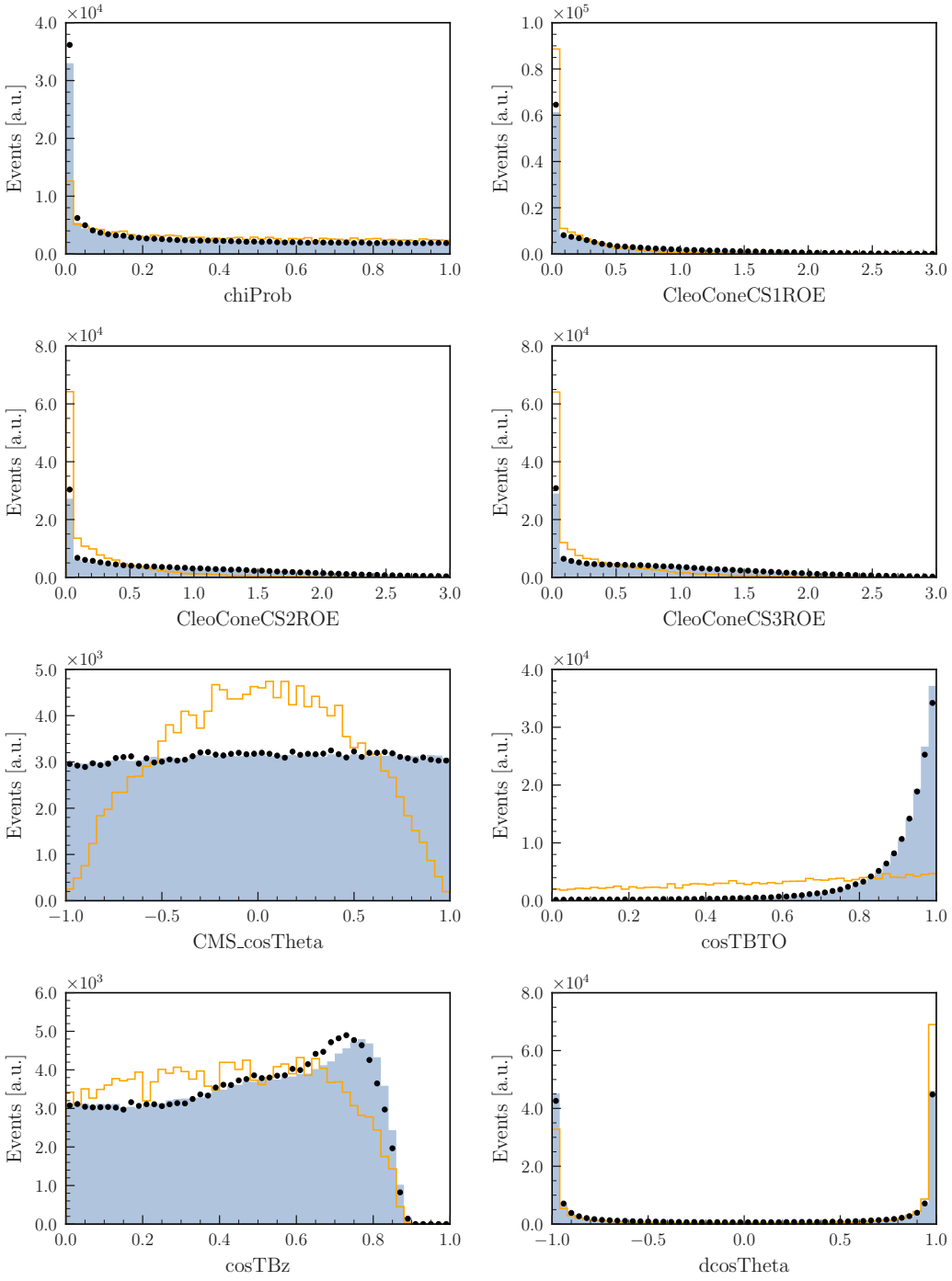


Figure 4.7.: Normalized distributions of training variables used in the continuum suppression classifier in real (black) and simulated (light blue) off-resonance data. To illustrate the distinction power, the distribution of simulated signal events is superimposed (orange).

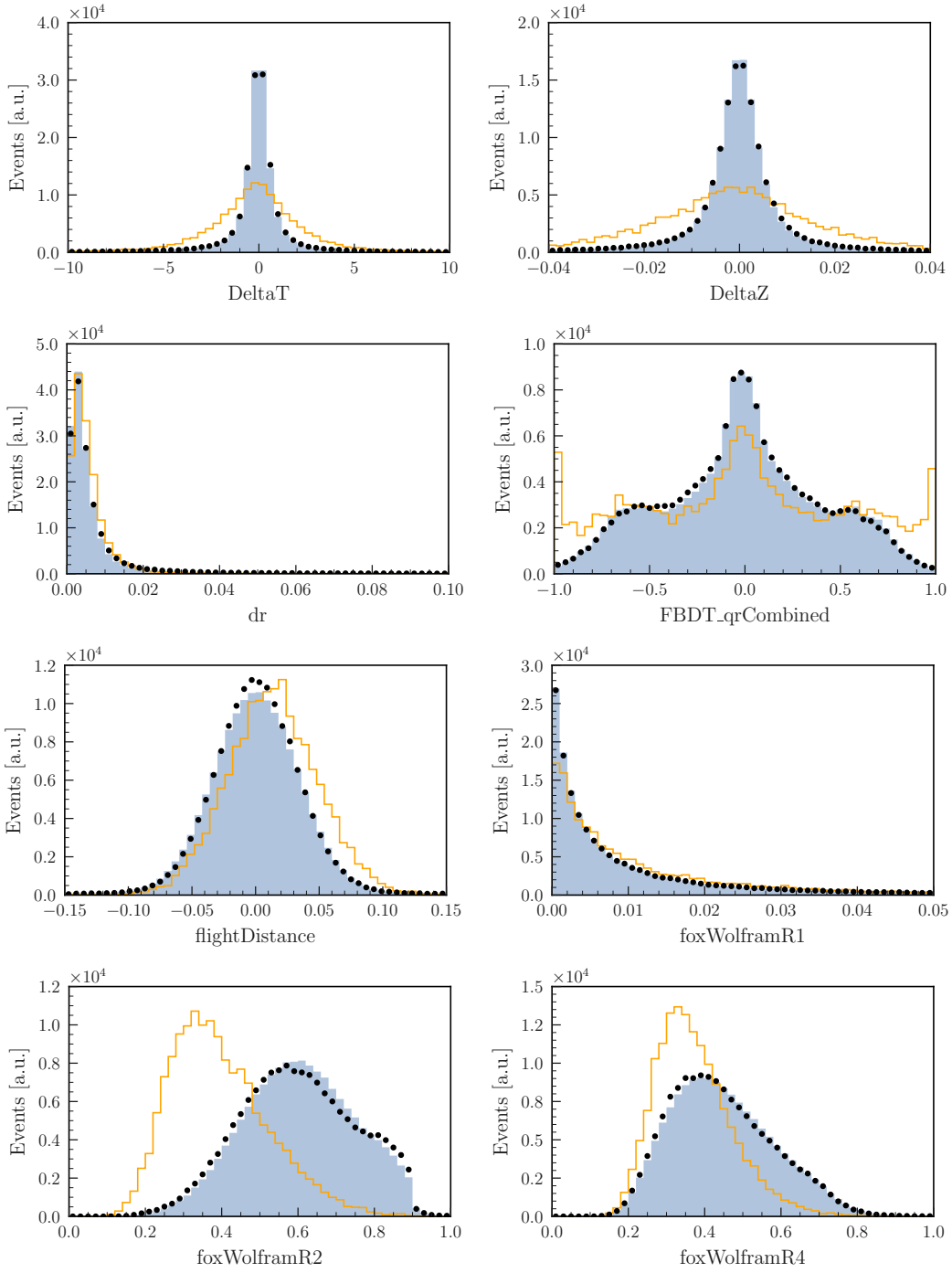


Figure 4.8.: Normalized distributions of training variables used in the continuum suppression classifier in real (black) and simulated (light blue) off-resonance data. To illustrate the distinction power, the distribution of simulated signal events is superimposed (orange).

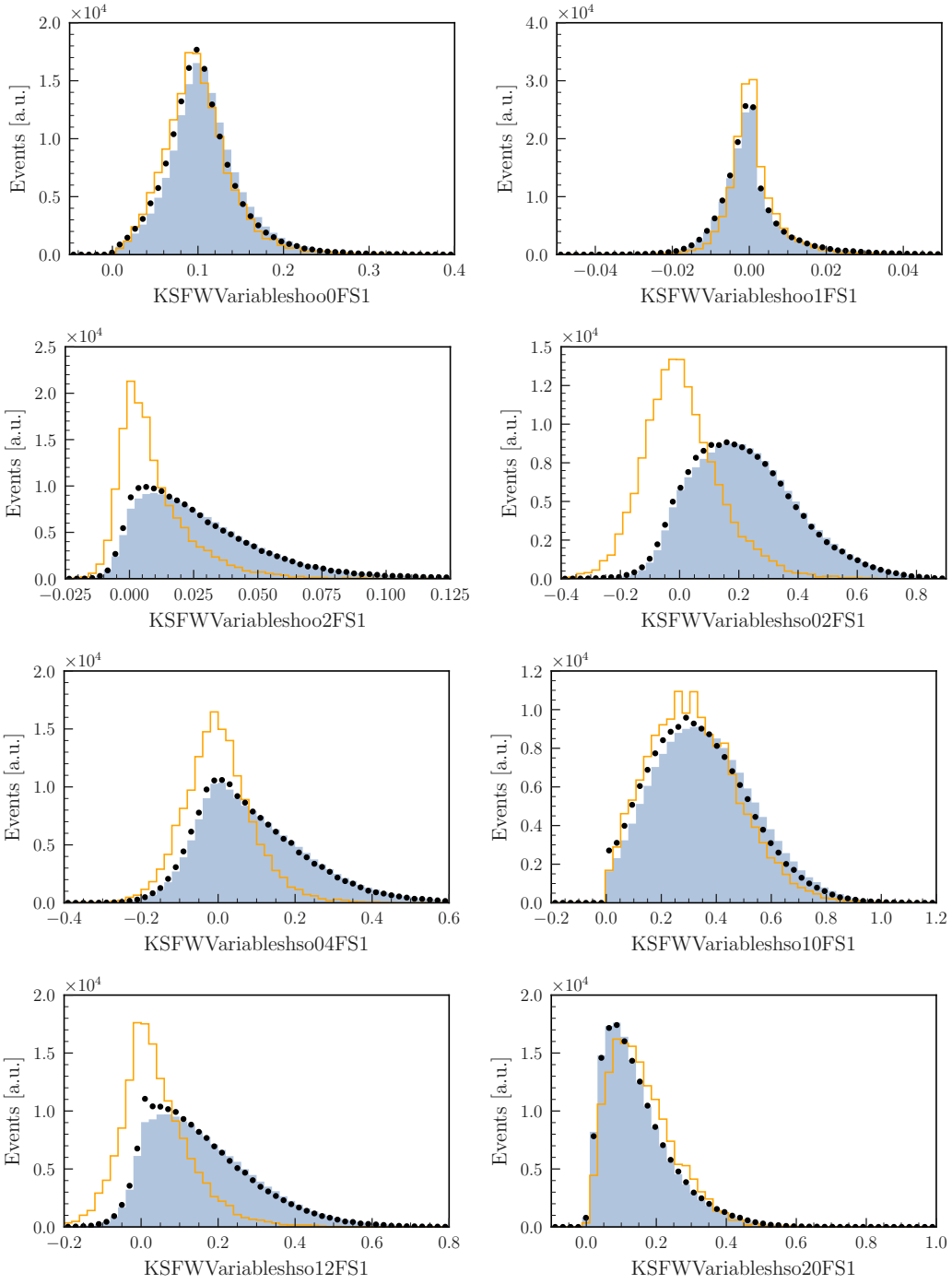


Figure 4.9.: Normalized distributions of training variables used in the continuum suppression classifier in real (black) and simulated (light blue) off-resonance data. To illustrate the distinction power, the distribution of simulated signal events is superimposed (orange).

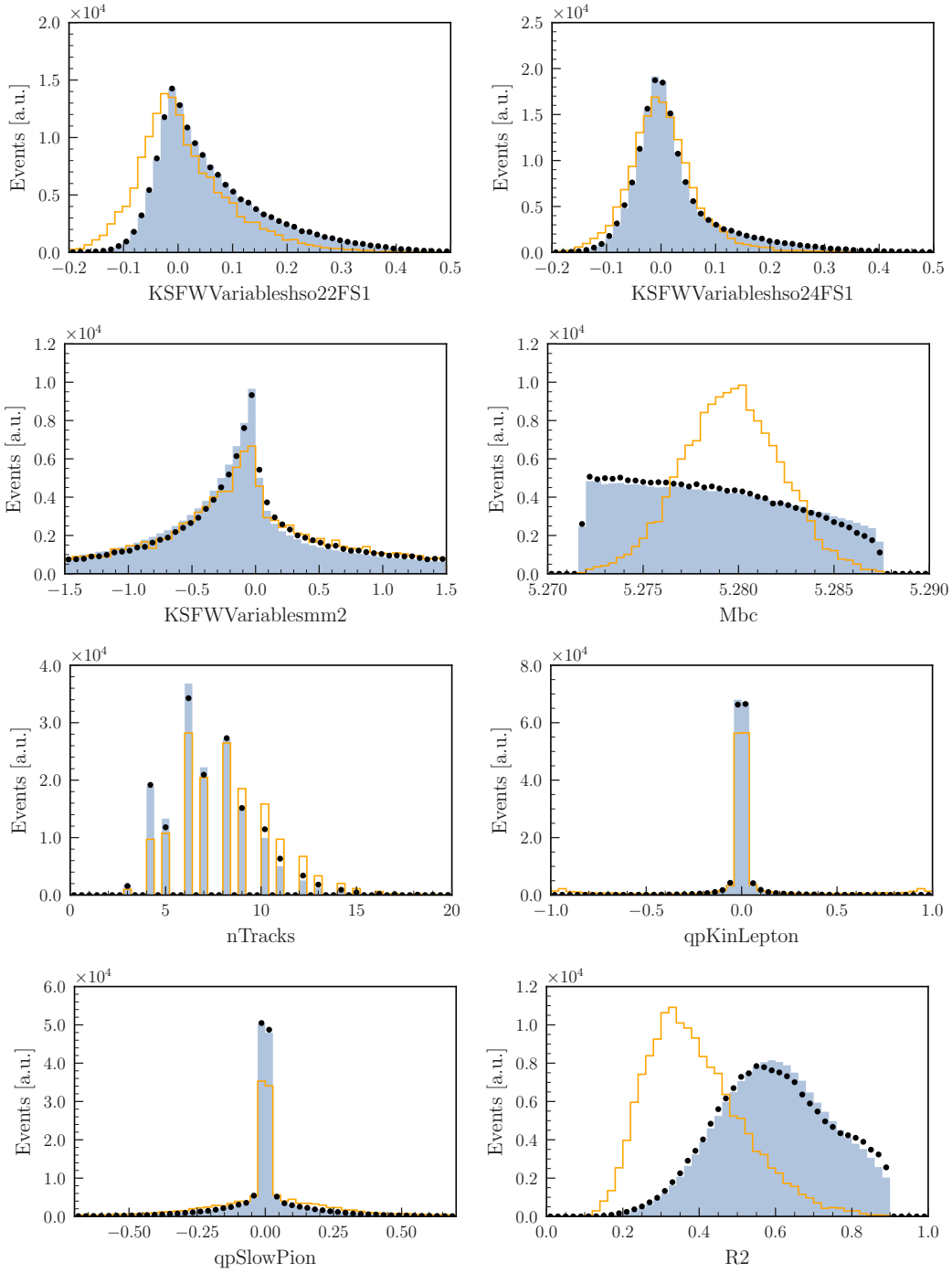


Figure 4.10.: Normalized distributions of training variables used in the continuum suppression classifier in real (black) and simulated (blue) off-resonance data. To illustrate the distinction power, the distribution of simulated signal events is superimposed (orange).

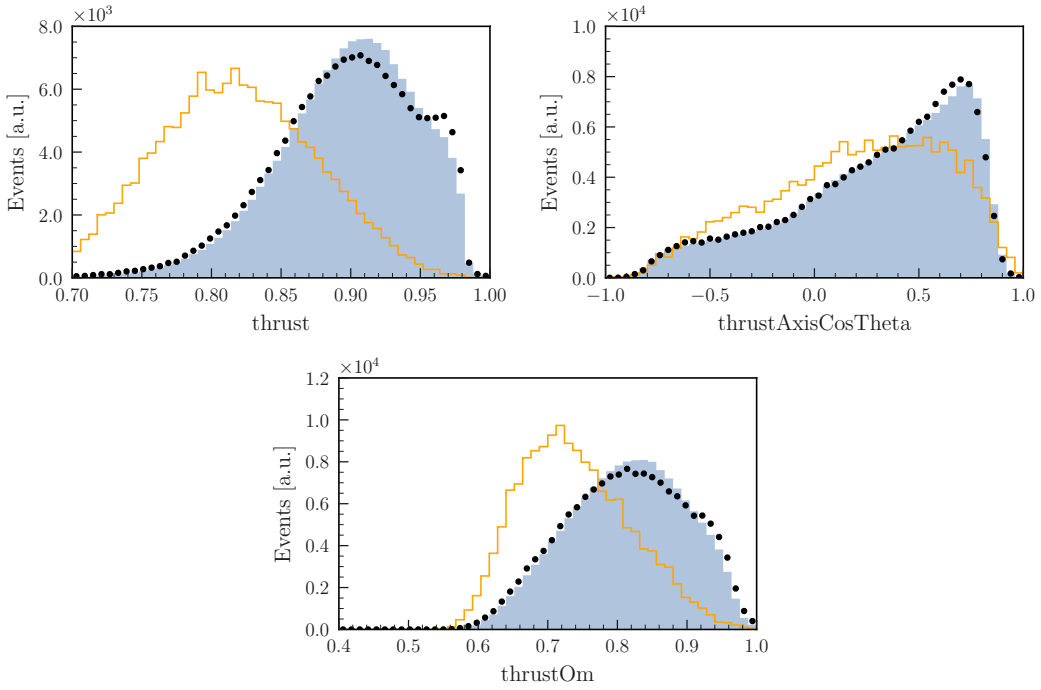


Figure 4.11.: Normalized distributions of training variables used in the continuum suppression classifier in real (black) and simulated (light blue) off-resonance data. To illustrate the distinction power, the distribution of simulated signal events is superimposed (orange).

Finally, we perform a cut on the continuum suppression classifier to reduce background from continuum events to a manageable level. As we include the classifier variable in the fit, assuming a perfect modeling, more restrictive cuts on this variable would reduce the statistical power of the data, leading to a decreased precision. This trend can be seen in Figure 4.12, which shows the mean uncertainties of the physics parameters for varying cuts, obtained with the nominal fitting algorithm on large sets of pseudo experiments. For this, we first create samples with different CS selection cuts, from which we empirically determine the fit shapes of all components. Sampling from these fit shapes, we generate distributions for each component in their expected proportion, scaled to an integrated luminosity of the LS1 dataset and fit the samples using our nominal fit model. We repeat this process 300 times for each CS cut and evaluate the mean uncertainties on the physics parameters.

In this study, we generate and fit using the same PDF models, which is in general an oversimplification. In order to reduce our accuracy requirements on the background modeling to a reasonable level, we chose to impose loose cuts on the BDT output, which remove 90% of continuum events as determined in MC: $C > 0.41$ for $B^0 \to K^+\pi^-$ and $C > 0.34$ for $B^0 \to \pi^+\pi^-$. For both modes, this selection retains more than 95% of signal events.

We improve the fitting of the classifier distribution by applying the so-called μ-transformation (or probability integral transformation) [81]. Here, we use the cumulative distribution function $F_c(C)$ of correctly reconstructed signal events to transform the classifier variable C to $C' = F_c(C)$. This allows for a simple description of the PDFs using analytical functions, where the signal component is per-definition uniformly distributed on the segment [0, 1] and continuum events follow an exponential function.

For the extraction of signal yields, we fit simultaneously in ΔE and the transformed classifier C'. We find the Pearson-correlation coefficients to be small in all components, supporting a factorization of the two-dimensional PDFs. Figure 4.13 shows the 2D-histograms in (C', ΔE), demonstrating negligible correlations in all components.

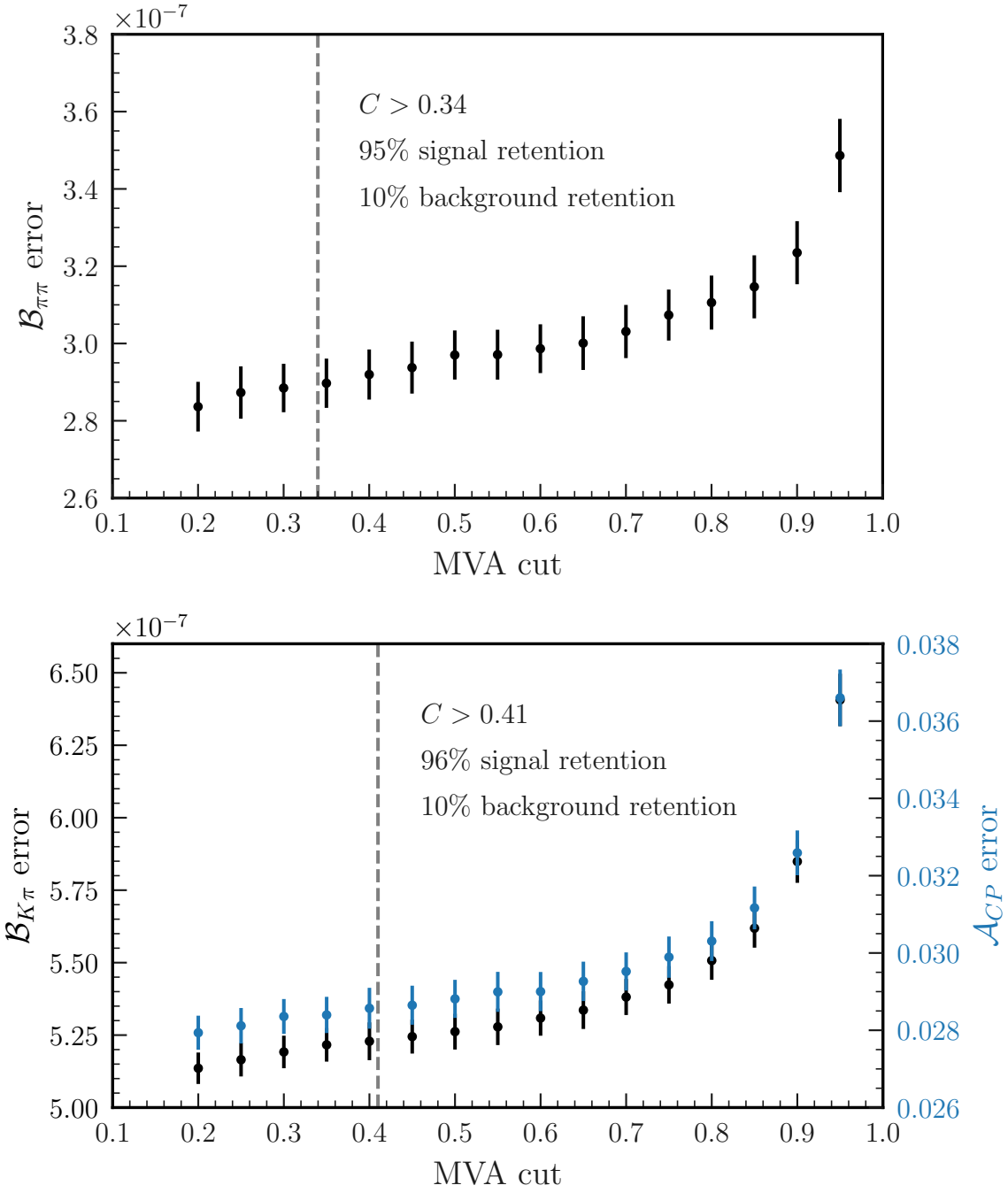


Figure 4.12.: Mean uncertainties of $\mathcal{B}_{\pi\pi}$ (top), and $\mathcal{B}_{K\pi}$ and $\mathcal{A}_{CP,K\pi}$ (bottom) for varying cuts on the continuum suppression classifier variable, evaluated on large sets of pseudo experiments. The dashed line represents the cut value chosen to reject 90% of continuum events.

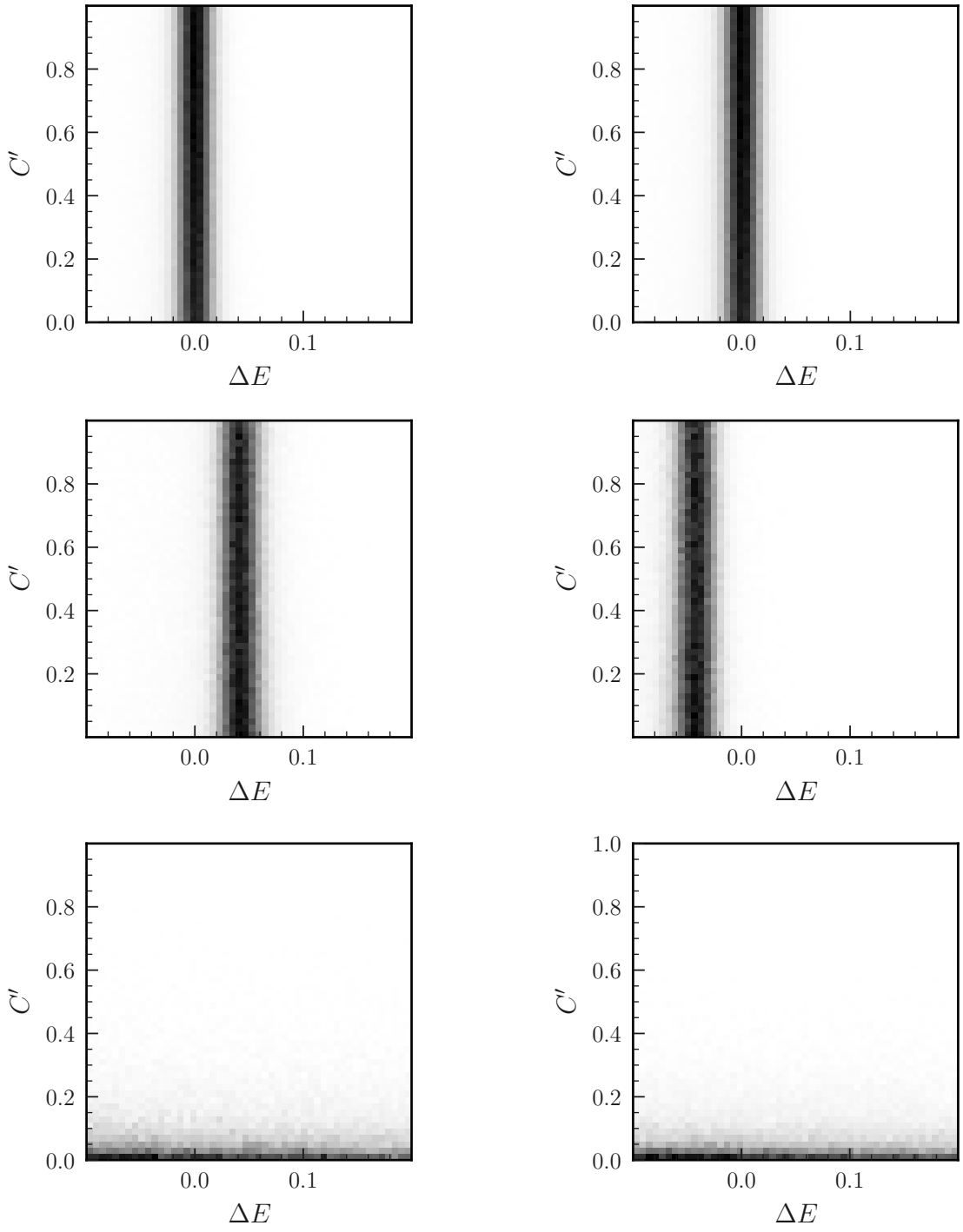


Figure 4.13.: 2D-histograms in (C', ΔE) for signal (top), cross-feed (center) and combinatorial background (bottom) events reconstructed as $B^0 \to K^+\pi^-$ (left) and $B^0 \to \pi^+\pi^-$ (right).

4.2.4. PID Selection

We use particle identification information (PID) to assign each reconstructed B-meson candidate to one of the two possible final-state samples, $B^0 \to K^+\pi^-$ or $B^0 \to \pi^+\pi^-$. To this end, we evaluate the binary PID of reconstructed kaon and pion candidates, defined as:

$$\mathcal{L}_{K/\pi} \equiv \frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi} \, .$$

For a given track, it compares specifically the two likelihoods of being produced by either a kaon or a pion. By definition, assuming an ideal construction, kaons will be assigned $\mathcal{L}_{K/\pi} > 0.5$ and analogously pions accumulate in the lower region $\mathcal{L}_{K/\pi} < 0.5$. In some cases this assignment fails, leading to non-negligible fake rates of misidentified candidates. In this analysis, we simultaneously fit reconstructed $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$ events, which exhibit cross-feed (CF) due to named misidentification. The cut on the binary likelihood should therefore be chosen in such a way, that fake rates are minimized to allow for clean signal peaks, while at the same time keeping the selection efficiencies as high as possible in both samples.

Since we also use the cross-feed components in the determination of the branching ratios, it is not trivial to determine the ideal PID cut. We therefore check the impact of different choices on the measurement of our physics parameters by means of pseudo experiments. For this, we first create multiple samples, differing in PID selection, from which we determine the fit shapes for all components. Sampling from these fit shapes, we then generate distributions for each component in their expected proportion, scaled to an integrated luminosity of the LS1 dataset and fit the samples using our nominal fit model. We repeat this process 300 times for each PID selection and compare the mean uncertainties on the extracted physics parameters.

Figure 4.14 shows the uncertainty distributions obtained with sets of three different PID selection cuts used to classify kaons, namely $\mathcal{L}_{K/\pi} > 0.4$, $\mathcal{L}_{K/\pi} > 0.5$ and $\mathcal{L}_{K/\pi} > 0.6$. Tracks failing these cuts, are assigned as pions. The difference in mean fit uncertainty of $\mathcal{B}(B^0 \to K^+\pi^-)$ between $\mathcal{L}_{K/\pi} > 0.4$ and $\mathcal{L}_{K/\pi} > 0.6$ measures 0.05% relative to the generated input value. Compared to the total uncertainty of 1.8%, this difference is deemed negligible. For $\mathcal{B}(B^0 \to \pi^+\pi^-)$ and $\mathcal{A}_{CP}(B^0 \to K^+\pi^-)$, we observe no significant difference. We therefore adhere to the originally chosen requirement of $\mathcal{L}_{K/\pi} > 0.5$ to select charged kaons. This requirement correctly identifies 90% of pions and 84% of kaons.

We perform a correction of possible mismodeling of the PID selection efficiencies and misidentification probabilities in the simulation using the Systematic Corrections Framework [82]. This procedure is described in detail in Section 5.1.

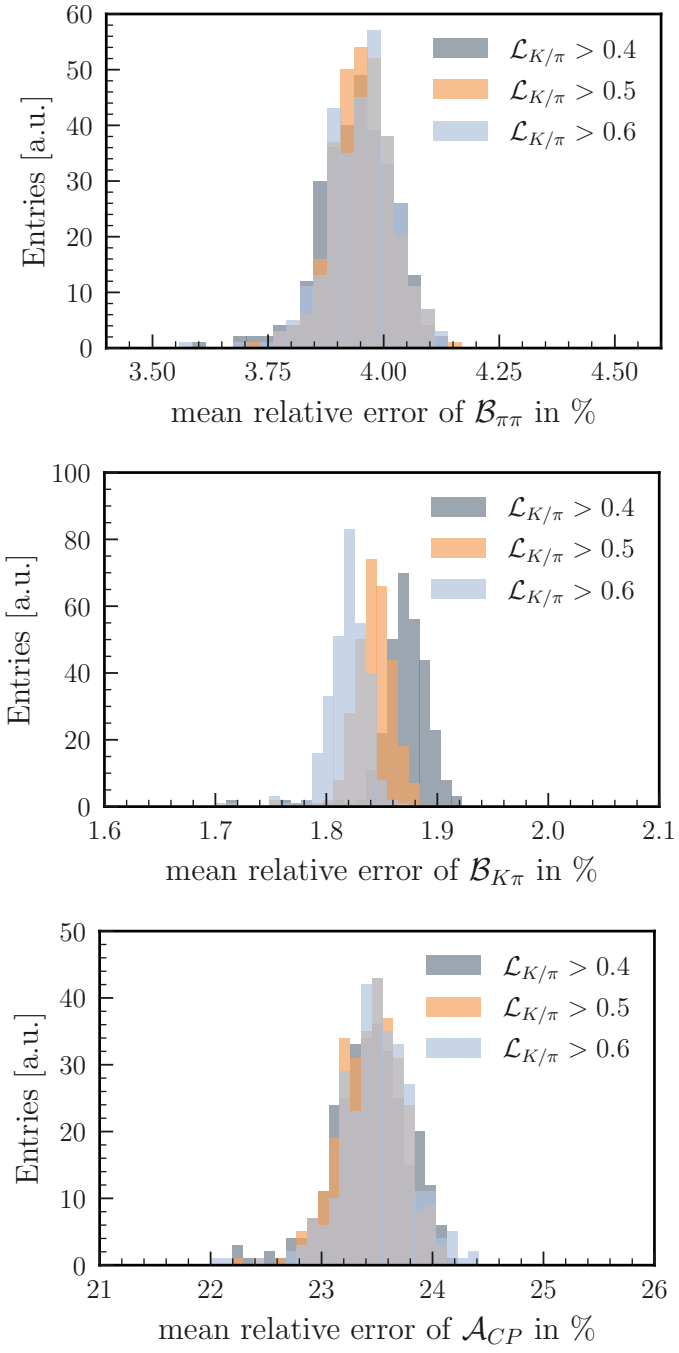


Figure 4.14.: Distributions of relative fit uncertainties for three different PID selection requirements to select kaons and pions for $\mathcal{B}(B^0 \to \pi^+\pi^-)$ (top), $\mathcal{B}(B^0 \to K^+\pi^-)$ (center), and $\mathcal{A}_{CP}(B^0 \to K^+\pi^-)$ (bottom), obtained using fits to pseudo experiments.

4.2.5. Final Selection and Composition

After all selection requirements, summarized in Table 4.2, the average number of signal candidates per event ranges between 1.000 and 1.001. We do not apply a best candidate selection and keep all candidates.

We use MC truth information to identify correctly reconstructed signal events and calculate charge-dependent selection efficiencies. The efficiency and self-cross-feed (SCF) fraction are defined as:

$$\epsilon \equiv \frac{S}{G} \quad \text{and} \quad \text{SCF} \equiv \frac{B}{S+B} \ ,$$

where S is the number of correctly reconstructed signal events, G is the total number of generated signal events and B is the number of misreconstructed events in the signal MC sample. The final charge-averaged signal efficiencies for correctly reconstructed events are 52.4% for $B^0 \to K^+\pi^-$ decays and 56.4% for $B^0 \to \pi^+\pi^-$ decays. Cross-feed events, in which one of the final-state hadrons is misreconstructed with an incorrect mass-hypothesis are selected with efficiencies of 9.7% for $B^0 \to K^+\pi^-$ decays and 13.1% for $B^0 \to \pi^+\pi^-$ decays. We include the non-negligible self-cross-feed of 1.9% in $B^0 \to K^+\pi^-$ in the signal component, as these events share a similar distribution with correctly reconstructed events in both fit variables. In the case of $B^0 \to \pi^+\pi^-$, we find a negligible SCF fraction of $< 0.1\%$.

Table 4.2.: Summary of selection requirements for reconstructed tracks and B-meson candidates.

tracks	$\boldsymbol{B^0}$
$17° < \theta < 150°$	$5.272 < M_{\text{bc}} < 5.288\,\text{GeV}/c^2$
nCDCHits > 20	$-0.1 < \Delta E < 0.2\,\text{GeV}$
dr $< 0.5\,\text{cm}$	CS classifier > 0.41 for $K\pi$
dz $< 2\,\text{cm}$	CS classifier > 0.34 for $\pi\pi$
$\mathcal{L}_{K/\pi} > 0.5$ for kaons	R2 < 0.9
$\mathcal{L}_{K/\pi} < 0.5$ for pions	

Figures 4.15 and 4.16 show the contributions of different components making up the total samples in generic, simulated data. We identify three components: signal events, cross-feed events due to misidentified tracks, and combinatorial background events, dominated by continuum. In samples scaled to the size of the LS1 dataset, containing around 16000 candidates, background events from B-meson decays amount to only 16 and 5 events for the $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$ samples,

respectively. Therefore, we include these in the combinatorial background instead of assigning dedicated models.

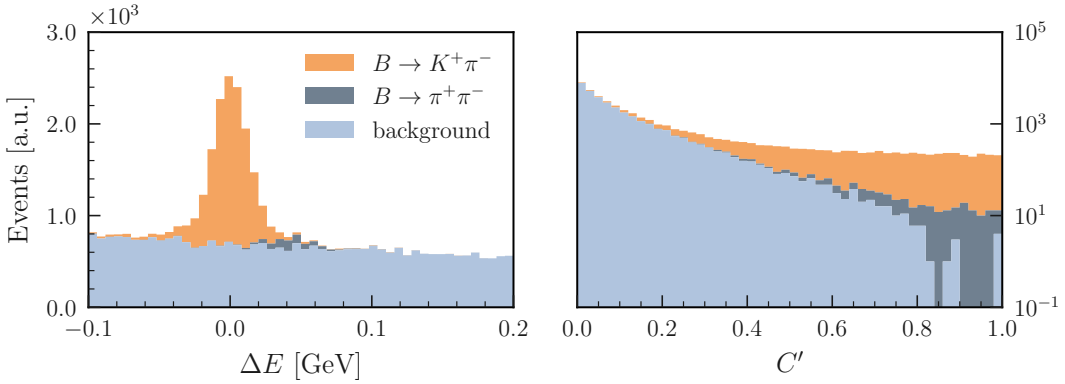


Figure 4.15.: Fit variable distributions in ΔE (left) and C' (right) of $B^0 \to K^+\pi^-$ candidates reconstructed in 1 ab^{-1} of generic MC. The sample contains contributions from $B^0 \to K^+\pi^-$ decays, misidentified $B^0 \to \pi^+\pi^-$ events and combinatorial background events.

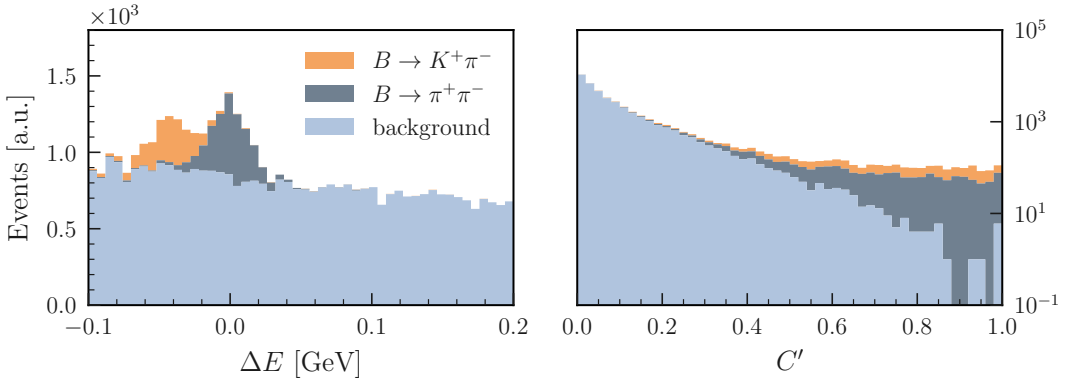


Figure 4.16.: Fit variable distributions in ΔE (left) and C' (right) of $B^0 \to \pi^+\pi^-$ candidates reconstructed in 1 ab^{-1} of generic MC. The sample contains contributions from $B^0 \to \pi^+\pi^-$ decays, misidentified $B^0 \to K^+\pi^-$ events and combinatorial background events.

4.3. Extraction of Physics Parameters

We determine the branching ratios and the CP-violation parameter using an extended two-dimensional maximum-likelihood fit of the unbinned energy-difference (ΔE) and transformed CS classifier (C') distributions.

To perform a maximum-likelihood fit, one first needs to define a model that describes the expected shape of the relevant signal and background distributions. This model is typically a probability density function (PDF) that depends on one or more parameters. The likelihood is given by the value of that probability density function evaluated at the measured values of the observables. The aim of the fit is to find the set of parameter values that maximizes the likelihood function, i.e., the set of parameter values that best describes the data.

For a set of uncorrelated variables, the likelihood function for two-dimensional probability distribution function $f(x, y; \boldsymbol{\theta})$ can be factorized into a product of one-dimensional PDFs $f_1(x; \boldsymbol{\theta})$ and $f_2(y; \boldsymbol{\theta})$:

$$\mathcal{L}(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_{i=1}^{N} f(x_i, y_i|\boldsymbol{\theta}) = \prod_{i=1}^{N} f_1(x_i|\boldsymbol{\theta}) \cdot f_2(y_i|\boldsymbol{\theta}) \ . \tag{4.2}$$

Here, $\boldsymbol{x}$ is the set of observed data points, $\boldsymbol{\theta}$ represents the parameters of the PDF that are being fitted, N is the total number of data points, and x_i and y_i are the values of the two variables for the i-th data point. In our case, f_1 and f_2 will be given by the PDFs of our fit variables, ΔE and C', which will be derived in detail for all components in the following section.

In the context of this study, the detection of events can be modeled as a Poisson process, where the number of events detected in a given interval of time or phase space follows a Poisson distribution, assuming that the events occur independently of each other with a constant average rate. The likelihood for a dataset consisting of N events with expected number of events λ is given by the Poisson probability mass function:

$$P(N|\lambda) = \frac{\lambda^N e^{-\lambda}}{N!} \ . \tag{4.3}$$

We include it in the likelihood function as an *extension* term:

$$\mathcal{L}(\lambda, \boldsymbol{\theta}|\boldsymbol{x}) = P(N|\lambda) \prod_{i=1}^{N} f_1(x_i|\boldsymbol{\theta}) \cdot f_2(y_i|\boldsymbol{\theta}) \ . \tag{4.4}$$

The likelihood function is then maximized with respect to the parameters of the signal and background distributions $\boldsymbol{\theta}$, as well as the expected number of signal events λ, in order to determine the best-fit values and uncertainties of the model

parameters:

$$\hat{\lambda}, \hat{\boldsymbol{\theta}} = \text{argmax}_{\lambda,\boldsymbol{\theta}} \mathcal{L}(\lambda, \boldsymbol{\theta}|\boldsymbol{x}) \ . \tag{4.5}$$

In practice, this optimization problem is typically solved using numerical minimization algorithms, which iteratively adjust the parameter values to find the maximum of the likelihood function. One popular algorithm for this purpose is the gradient descent method, which uses the gradient of the likelihood function with respect to the parameters to update the parameter values at each iteration, and is used in this study.

For the error estimation, the Hessian matrix is calculated to determine the corresponding covariance matrix. This involves computing the second derivatives of the likelihood function with respect to the fit parameters at the minimum, assuming that the likelihood is locally parabolic.

In this analysis, we simultaneously fit the charge-integrated $B^0 \to \pi^+\pi^-$ sample and the two charge-specific $B \to K^\pm\pi^\mp$ samples. As the $B \to K^\pm\pi^\mp$ decays are flavor-specific, the measurement of $\mathcal{A}_{CP}$ is practically obtained by assigning the candidates to samples according to the charge of the kaon and combining the measured signal yields in an appropriate manner. In each of the three samples, both the signal component and the peaking background of misidentified cross-feed events are used in the determination of the branching ratios. In total, we constrain the physics parameters by fitting six yields:

$$N^{\text{sig}}_{K^+\pi^-} = N_{B^0\bar{B}^0} \times \mathcal{B}(K^+\pi^-) \times \varepsilon_{K^+\pi^-} \times \varepsilon_K^{corr} \times \varepsilon_\pi^{corr} \times \frac{1}{2}(1 - \mathcal{A}_{CP}) \ , \tag{4.6}$$

$$N^{\text{CF},+}_{\pi^+\pi^-} = N_{B^0\bar{B}^0} \times \mathcal{B}(\pi^+\pi^-) \times \varepsilon^{\text{CF}}_{K^+\pi^-} \times f_\pi^{corr} \times \varepsilon_\pi^{corr} \times \frac{1}{2} \ , \tag{4.7}$$

$$N^{\text{sig}}_{K^-\pi^+} = N_{B^0\bar{B}^0} \times \mathcal{B}(K^+\pi^-) \times \varepsilon_{K^-\pi^+} \times \varepsilon_K^{corr} \times \varepsilon_\pi^{corr} \times \frac{1}{2}(1 + \mathcal{A}_{CP}) \ , \tag{4.8}$$

$$N^{\text{CF},-}_{\pi^+\pi^-} = N_{B^0\bar{B}^0} \times \mathcal{B}(\pi^+\pi^-) \times \varepsilon^{\text{CF}}_{K^-\pi^+} \times f_\pi^{corr} \times \varepsilon_\pi^{corr} \times \frac{1}{2} \ , \tag{4.9}$$

$$N^{\text{sig}}_{\pi^+\pi^-} = N_{B^0\bar{B}^0} \times \mathcal{B}(\pi^+\pi^-) \times \varepsilon_{\pi^+\pi^-} \times \varepsilon_\pi^{corr} \times \varepsilon_\pi^{corr} \ , \tag{4.10}$$

$$N^{\text{CF}}_{K\pi} = N_{B^0\bar{B}^0} \times \mathcal{B}(K^+\pi^-) \times \varepsilon^{\text{CF}}_{\pi^+\pi^-} \times f_{K^+}^{corr} \times \varepsilon_{\pi^-}^{corr} \ , \tag{4.11}$$

where N_i^j are the fitted signal yields, $N_{B^0\bar{B}^0}$ is the number of neutral B-meson events in the sample, $\mathcal{B}(i)$ are the respective branching fractions, ε_i and $\varepsilon_i^{\text{CF}}$ are the selection efficiencies for correctly and misreconstructed events, and ε_k^{corr} and f_k^{corr} are PID efficiency and fake rate correction factors. Note, that the latter are

determined for each yield separately, as the data-MC corrections are dependant on the momentum and angular distribution of the respective charged final-state particles, which can differ significantly among these components. The extraction and application of these corrections is detailed in Section 5.1.

4.3.1. Modeling

The PDFs of the signal and background components are modeled using analytical functions that are chosen to accurately represent the expected shape of the distributions. We empirically obtain the PDF shape parameters by fitting to large MC samples. To improve statistical power, pure signal Monte Carlo samples are generated and reconstructed for the signal and cross-feed components. We find the shapes of the signal and cross-feed components in $B \to K^\pm\pi^\mp$ decays to be independent of the charge and use the same parameters for both charges.

Signal models for $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$

ΔE

We use a combination of a Gaussian function and a Crystal Ball shape function (CB) to model the ΔE distributions of the signal components in both $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$. The Gaussian function is defined as:

$$f_G(\Delta E; \mu_G, \sigma_G) \equiv N \cdot \frac{1}{\sqrt{2\pi}\sigma_G} \cdot \exp\left(-\frac{(\Delta E - \mu_G)^2}{2\sigma_G^2}\right) , \tag{4.12}$$

where N is a normalization constant, and σ_G and μ_G are the mean and width of the distribution.

The Crystal Ball function combines a Gaussian function for the core of the distribution with a power-law function to model a one-sided tail [83]. The latter is needed to account for final-state radiation of soft photons, which leads to a higher population towards negative values in ΔE. The Crystal ball shape function is defined as:

$$f_{CB}(\Delta E; \alpha, n, \mu_{CB}, \sigma_{CB}) \equiv N \cdot \begin{cases} \exp\left(-\frac{(\Delta E-\mu_{CB})^2}{2\sigma_{CB}^2}\right), & \text{for } \frac{\Delta E-\mu_{CB}}{\sigma_{CB}} > -\alpha \\ A \cdot \left(B - \frac{\Delta E-\mu_{CB}}{\sigma_{CB}}\right)^{-n}, & \text{for } \frac{\Delta E-\mu_{CB}}{\sigma_{CB}} \leq -\alpha \end{cases} \tag{4.13}$$

with

$$A = \left(\frac{n}{|\alpha|}\right)^n \mathrm{e}^{-|\alpha|^2/2} , \quad B = \frac{n}{|\alpha|} - |\alpha| , \tag{4.14}$$

where N is a normalization constant, μ_{CB} and σ_{CB} are the mean and width of the distribution, α determines the transition point between the Gaussian and the tail, and n is a parameter controlling the power-law behavior of the tail.

The two functions are combined with a weight parameter $w \in [0, 1]$, that is also determined in the fit to simulated events:

$$\begin{aligned} f_{\text{sig}}(\Delta E) \equiv w \cdot f_{CB}(\Delta E; \alpha, n, \mu_{CB}, \sigma_{CB}) \\ + (1 - w) \cdot f_G(\Delta E; \mu_G, \sigma_G) \ . \end{aligned} \tag{4.15}$$

We determine the best-fit parameters by means of maximum-likelihood fits to samples of correctly reconstructed signal events, surviving all selection criteria. We do this for each signal mode separately and fix all shape parameters.

C′

By construction, the μ-transformation of the CS classifier variable results in a perfectly flat distribution in the case of correctly reconstructed signal events, i.e.,

$$f_{\text{sig}}(C') \equiv 1 = \text{const.} \tag{4.16}$$

Thus, no parameters are extracted from the fit to MC events.

Figure 4.17 shows the fit variable distributions for correctly reconstructed signal events alongside the corresponding fitted PDFs for both modes. We combine the two one-dimensional PDFs according to Eq. 4.4 to obtain the two-dimensional PDFs for the signal components. For $B^0 \to K^+\pi^-$, we include the small fraction of 1.9% of self-cross-feed events, formed by misreconstructed signal decays, in the signal component, as their distributions resemble those of correctly reconstructed signal events in both fit variables.

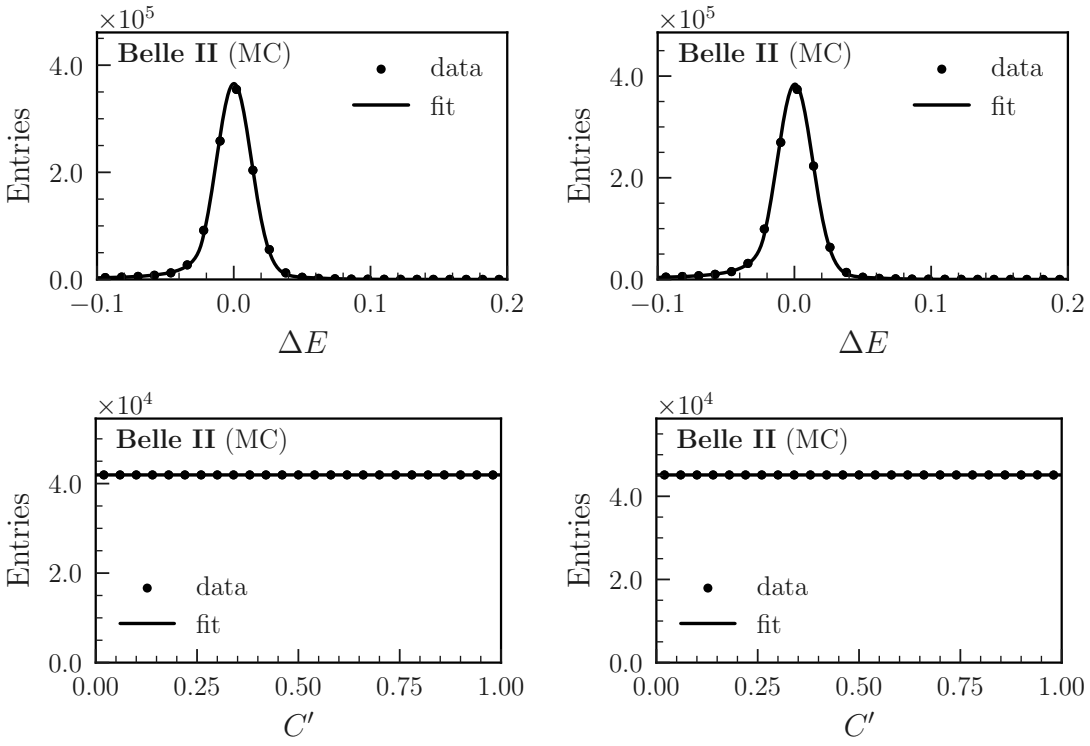


Figure 4.17.: Signal event distributions reconstructed from abundant signal MC samples, alongside fitted PDFs for $B^0 \to K^+\pi^-$ (left) and $B^0 \to \pi^+\pi^-$ (right).

Cross-feed models for $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$

ΔE

We use a combination of two Gaussian functions to model the ΔE distributions of the cross-feed components in both $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$. The two functions are combined with a weight parameter w, that is determined in the fit to simulated events:

$$\begin{aligned} f_{\mathrm{CF}}(\Delta E) \equiv\ & w \cdot f_{G,1}(\Delta E; \mu_{G,1}, \sigma_{G,1}) \\ & + (1-w) \cdot f_{G,2}(\Delta E; \mu_{G,2}, \sigma_{G,2})\ . \end{aligned} \tag{4.17}$$

C′

We use a second-order polynomial to model the C' distributions of the cross-feed components in both $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$:

$$f_{\mathrm{CF}}(C') \equiv N \cdot (1 + a \cdot C' + b \cdot C'^2). \tag{4.18}$$

We determine the best-fit parameters by means of maximum-likelihood fits to samples of misidentified signal events, surviving all selection criteria. We do this for each signal mode separately and fix all shape parameters. Figure 4.18 shows the fit variable distributions for misidentified signal events alongside the corresponding fitted PDFs for both modes. We combine the two one-dimensional PDFs according to Eq. 4.4 to obtain the two-dimensional PDFs for the cross-feed components.

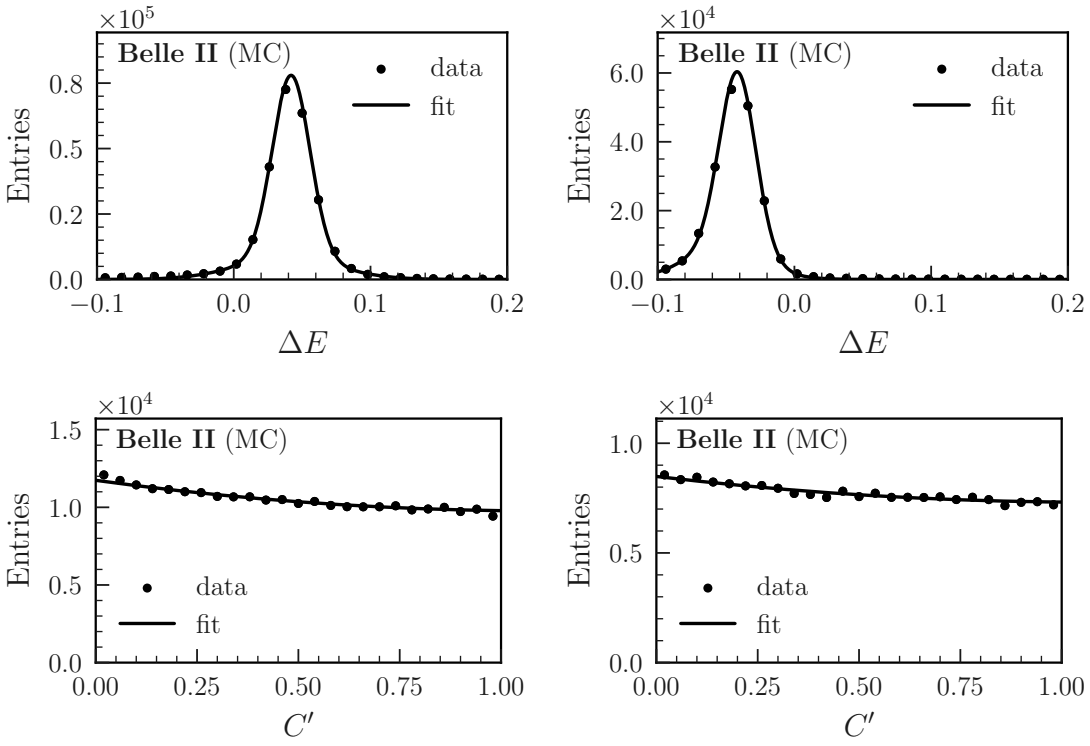


Figure 4.18.: Misidentified signal event distributions reconstructed from abundant signal MC samples, alongside fitted PDFs for $B^0 \to K^+\pi^-$ (left) and $B^0 \to \pi^+\pi^-$ (right).

Continuum models for $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$

ΔE

We use an exponential function to model the ΔE distributions of combinatorial background, dominated by continuum events, in both $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$:

$$f_{\text{bkg}}(\Delta E) \equiv N \cdot \exp\left(c \cdot \Delta E\right). \tag{4.19}$$

C'

We use a combination of two exponential functions to model the C' distributions of the combinatorial background components in both $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$:

$$\begin{aligned} f_{\text{bkg}}(C') \equiv\ & w \cdot \exp\left(a \cdot \Delta E\right) \\ & + (1-w) \cdot \exp\left(b \cdot \Delta E\right)) \ . \end{aligned} \tag{4.20}$$

Due to the abundance of continuum events, we leave all background shape parameters floating in the nominal fit. For completeness, we show the result of a fit to a simulated continuum background sample in Figure 4.19. We combine the two one-dimensional PDFs according to Eq. 4.4 to obtain the two-dimensional PDFs for the combinatorial background components.

Table 4.3 summarizes the fit models for both signal modes. In order to account for potential data-MC mismodeling, we obtain correction factors using the control channel $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$. From fits to data and MC samples, we extract parameters quantifying any mismodeling of the widths and means in the ΔE distributions, σ_{corr} and μ_{corr}. We apply these corrections to the signal and cross-feed models in the fit to data. Section 5.3 details the extraction of the correction factors.

$$\begin{aligned} \mu_G \to \mu_G + \mu_{corr} \quad , \quad \mu_{CB} \to \mu_{CB} + \mu_{corr} \\ \sigma_G \to \sigma_G + \sigma_{corr} \quad , \quad \sigma_{CB} \to \sigma_{CB} + \sigma_{corr} \end{aligned}$$

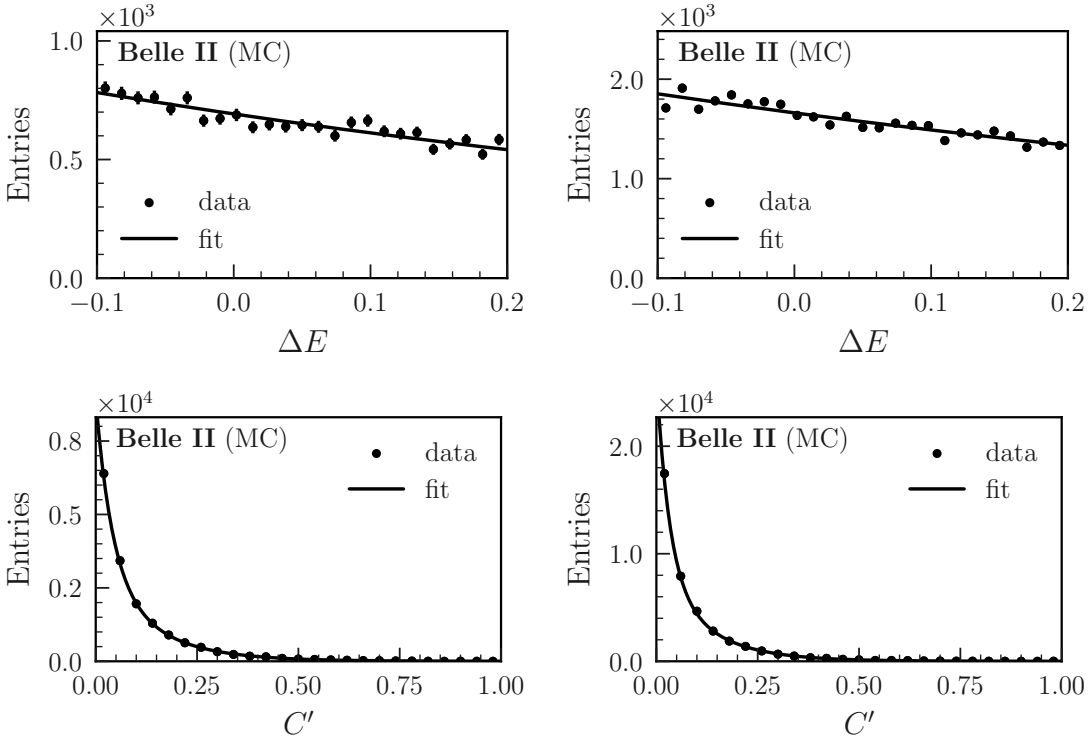


Figure 4.19.: Continuum background event distributions reconstructed from generic MC samples, alongside fitted PDFs for $B^0 \to K^+\pi^-$ (left) and $B^0 \to \pi^+\pi^-$ (right).

Table 4.3.: Summary of shape functions used to model the PDFs

Decay	Component	ΔE model	C' model
$B^0 \to K^+\pi^-$	signal	Gaussian + CB	flat
	cross-feed	2 Gaussians	2nd order poly.
	comb. background	exp.	2 exp.
$B^0 \to \pi^+\pi^-$	signal	Gaussian + CB	flat
	cross-feed	2 Gaussians	2nd order poly.
	comb. background	exp.	2 exp.

4.3.2. Fitter Validation

In this section, we present the results of several fitter tests performed on simulated data to assess the reliability and accuracy of our fit procedure. The fitter validation tests are designed to identify any biases or systematic errors that may arise during the fitting process, as well as to estimate the statistical uncertainties of the fitted parameters. In addition, ensemble tests are employed to test the robustness of the fitter in different regions of the parameter space and to quantify the effect of statistical fluctuations on the results.

Fit to simulated data

We apply the fitter on a simulated sample corresponding to an integrated luminosity of 1 ab^{-1} and compare the fit results with MC truth information. Figures 4.20 and 4.21 show the fit variable distributions for reconstructed $B^0 \to \pi^+\pi^-$, $B^0 \to K^+\pi^-$ and $\bar{B}^0 \to K^-\pi^+$ candidates alongside projections of an extended unbinned maximum-likelihood fit. We observe good agreement between true and fitted values in all components within their respective uncertainties (see Table 4.4).

Table 4.4.: Comparison of fit results with MC truth information. The values are obtained by a simultaneous fit to $B^0 \to \pi^+\pi^-$, $B^0 \to K^+\pi^-$ and $\bar{B}^0 \to K^-\pi^+$ candidates, reconstructed in simulated data corresponding to an integrated luminosity of 1 ab^{-1}. We also list the normalized residuals (NR) of the individual parameters.

Parameter	Fit Result	Simulation Truth	NR
$\mathcal{B}(B^0 \to \pi^+\pi^-)$	$(5.15 \pm 0.12) \times 10^{-6}$	5.12×10^{-6}	0.25
$\mathcal{B}(B^0 \to K^+\pi^-)$	$(19.39 \pm 0.21) \times 10^{-6}$	19.58×10^{-6}	-0.90
$\mathcal{A}_{CP}(B^0 \to K^+\pi^-)$	0.010 ± 0.012	0.006	0.33
$K^+\pi^-$ background	16462 ± 136	16365	0.71
$K^-\pi^+$ background	16488 ± 136	16495	-0.05
$\pi^+\pi^-$ background	39497 ± 206	39489	0.04

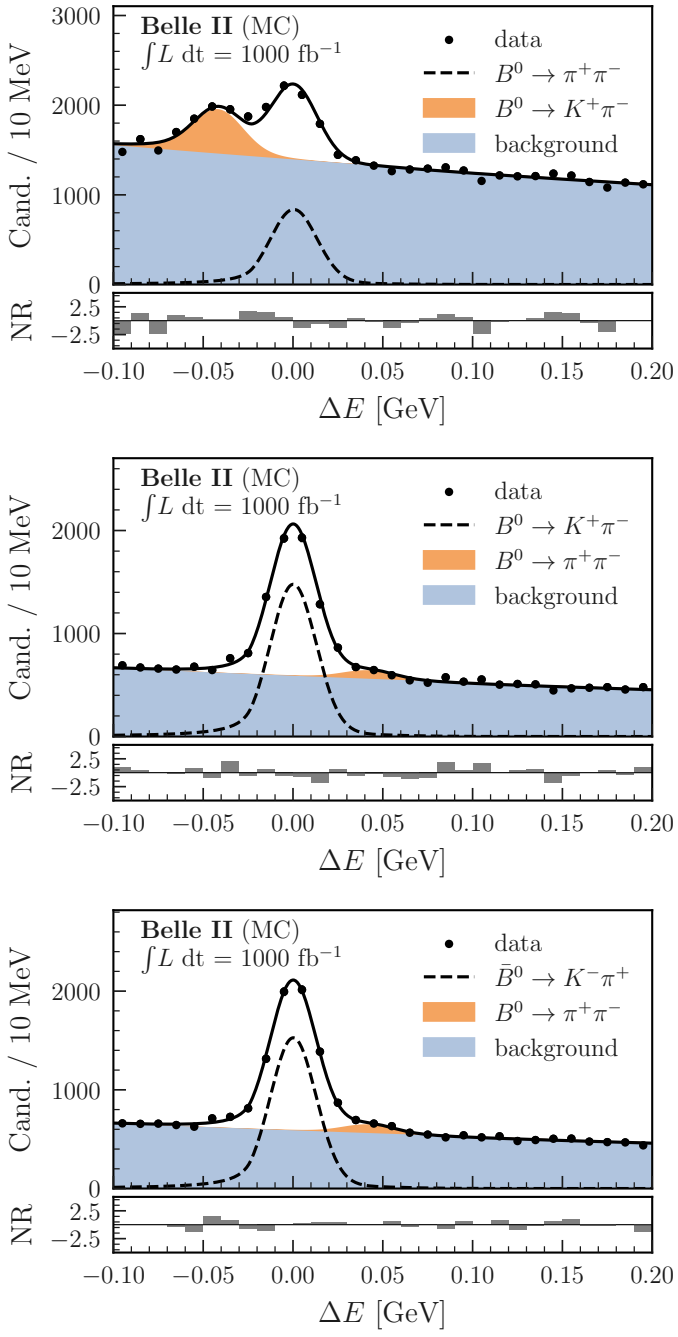


Figure 4.20.: Distributions in ΔE for $B^0 \to \pi^+\pi^-$ (top), $B^0 \to K^+\pi^-$ (center), and $\bar{B}^0 \to K^-\pi^+$ (bottom) candidates reconstructed in an MC sample corresponding to an integrated luminosity of $1\,\mathrm{ab}^{-1}$. The projections of an extended unbinned maximum-likelihood fit are overlaid.

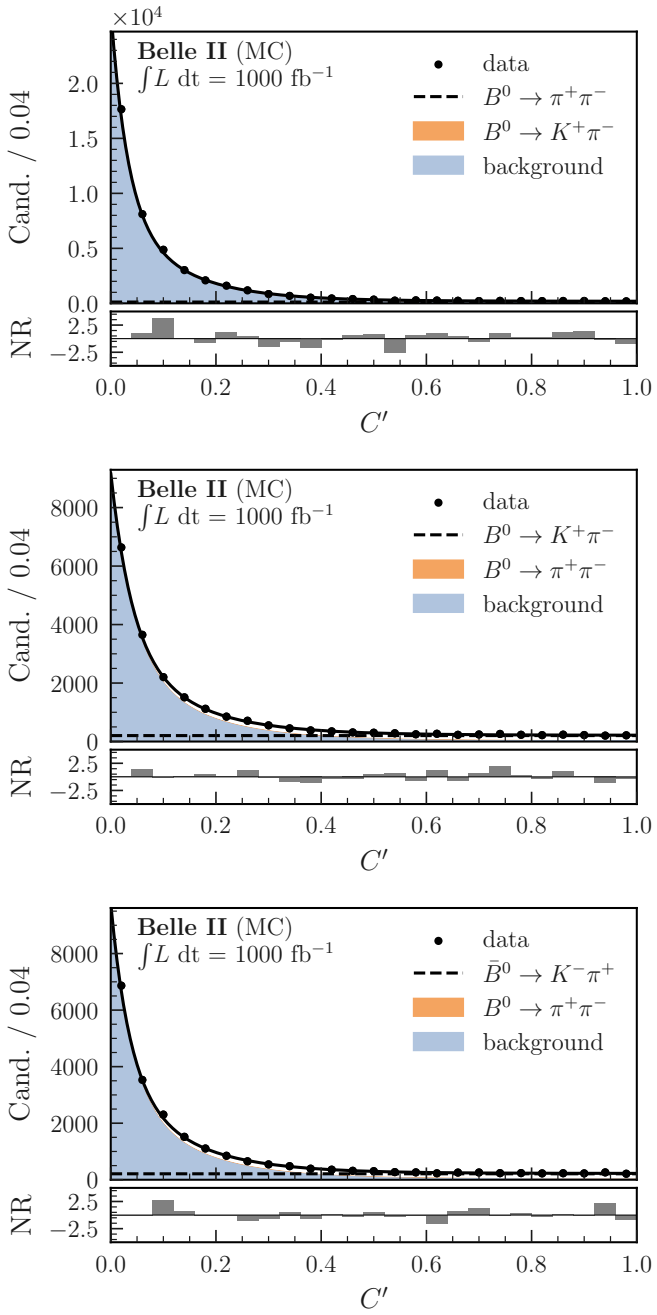


Figure 4.21.: Distributions in C' for $B^0 \to \pi^+\pi^-$ (top), $B^0 \to K^+\pi^-$ (center), and $\bar{B}^0 \to K^-\pi^+$ (bottom) candidates reconstructed in an MC sample corresponding to an integrated luminosity of 1 ab^{-1}. The projections of an extended unbinned maximum-likelihood fit are overlaid. In the background-dominated samples, the cross-feed component (orange) is only faintly discernible.

Ensemble tests

We validate our fit procedure using a large ensemble of 1000 artificially generated datasets. Sampling from the nominal fit shapes, we generate distributions for each fit variable and component in their expected proportion. For the physics parameters of interest, we use the current average values reported by the Particle Data Group (PDG) [20]:

$$\begin{aligned} \mathcal{B}(B^0 \to \pi^+\pi^-) &= 5.12 \times 10^{-6} \ , \\ \mathcal{B}(B^0 \to K^+\pi^-) &= 19.58 \times 10^{-6} \ , \\ \mathcal{A}_{CP}(B^0 \to K^+\pi^-) &= -0.0834 \ . \end{aligned}$$

We calculate the expected event yields assuming the integrated luminosity of the LS1 dataset and the reconstruction efficiencies described in Section 4.2.2. We randomly fluctuate the yields sampling from Poisson distributions around the expected values. We then fit these artificial datasets and calculate the normalized residual (NR) for each fit parameter. The latter is defined as:

$$\mathrm{NR} \equiv \frac{x_{fit} - x_{true}}{\sigma_x} \ , \tag{4.21}$$

where x_{fit} is the measured fit parameter, x_{true} is the true input value of this parameter and σ_x is the uncertainty on x_{fit}, provided by the minimizer. Figure 4.22 shows the NR distributions for the three physics parameters, $\mathcal{B}(B^0 \to \pi^+\pi^-)$, $\mathcal{B}(B^0 \to K^+\pi^-)$ and $\mathcal{A}_{CP}(B^0 \to K^+\pi^-)$. We fit the distributions with a Gaussian function. The fits yield unit widths and means consistent with zero, indicating an unbiased fitter with accurate uncertainty estimation.

The expected statistical uncertainties on the physics parameters can be estimated by computing the means of fit uncertainties across the ensemble of datasets. We obtain the following values:

$$\begin{aligned} \langle \sigma_{\mathcal{B}(\pi\pi)} \rangle &= 0.20 \times 10^{-6} \ , \\ \langle \sigma_{\mathcal{B}(K\pi)} \rangle &= 0.35 \times 10^{-6} \ , \\ \langle \sigma_{\mathcal{A}_{CP}(K\pi)} \rangle &= 0.019 \ . \end{aligned}$$

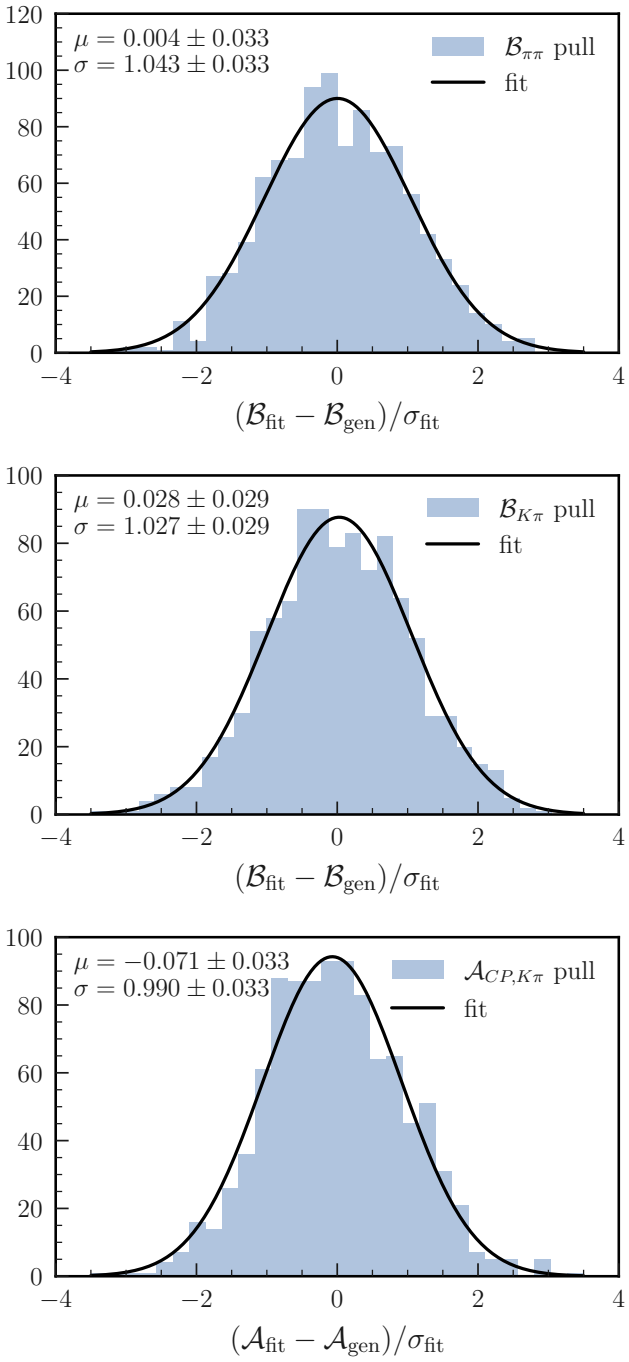


Figure 4.22.: Distributions of normalized residuals for the physics parameters $\mathcal{B}(B^0 \to \pi^+\pi^-)$, $\mathcal{B}(B^0 \to K^+\pi^-)$ and $\mathcal{A}_{CP}(B^0 \to K^+\pi^-)$ obtained by fits to 1000 artificially generated datasets, sampled from PDF shapes. Projections of fits with a Gaussian function are shown alongside the numeric fit results.

Linearity Tests

We perform linearity tests to check for any biases in the fit procedure in case of physics parameters deviating significantly from the reported average values. Fitting large ensembles of artificial datasets, generated with varying physics parameters, we test the accuracy of the fitter in different regions of the parameter space. We follow two different approaches. In the first, we generate distributions from the PDFs of the fit variables in each component. In the second approach, we sample from large MC datasets with replacement. For both cases, we calculate the expected event yields of each component assuming the integrated luminosity of the LS1 dataset and the reconstruction efficiencies described in Section 4.2.2. We randomly fluctuate the yields for each component sampling from Poisson distributions around the expected values, estimated on MC.

(i) Sampling from PDFs

Sampling from the nominal fit shapes, we generate distributions for each fit variable and component in their expected proportion. We vary the physics parameters from 70% to 130% of their nominal values in increments of 10%. We subsequently fit the combined distributions, containing signal, cross-feed and combinatorial background components. For each fraction, we perform 300 such pseudo experiments, varying only one physics parameter at a time, and compare the average of the fit result with the true input value. The results of the linearity tests are shown in Figure 4.23. We fit these distributions with a linear function. We observe unit slopes and intercepts consistent with zero, indicating an unbiased linear relationship across the tested parameter space for all three physics parameters.

(ii) Sampling from MC datasets

In addition to drawing the samples from the modeling shapes, we perform linearity tests following the principles of a method referred to as bootstrapping [84]. Bootstrapping is a statistical technique used to create large numbers of artificial datasets by randomly sampling from the original data with replacement. Assuming the original dataset represents a random sample of the underlying population and individual observations are uncorrelated, the statistics of the generated datasets closely approximate those of the original dataset. One advantage of this method is that it preserves correlations that might not have been accounted for in the modeling.

To create ensembles of bootstrap samples, we randomly select entries from large MC datasets. For the signal and cross-feed components, we use abundant signal MC samples, avoiding any resampling of same events. Regarding combinatorial background, we use a simulation sample roughly three times the size of the generated bootstrap samples. As a result, there is substantial overlap among the sets in our

ensembles, with a significant fraction of them containing the same background events. However, considering the quantity of such events numbering around 30 000 in the original datasets, we can safely assume them to provide a proper approximation of the underlying distribution.

Keeping the number of combinatorial background events constant across samples, we vary the physics parameters from 80% to 120% of their nominal values in increments of 10%. For each fraction, we generate an ensemble of 300 bootstrap samples, varying only one physics parameter at a time. We extract the physics parameters using our nominal fitter and compare the average over all fit results across each ensemble with the true input value. The results of the linearity tests are shown in Figure 4.24. We fit the resulting distributions with a linear function. We observe unit slopes and intercepts consistent with zero, confirming an unbiased linear relationship across the tested parameter space for all three physics parameters.yout

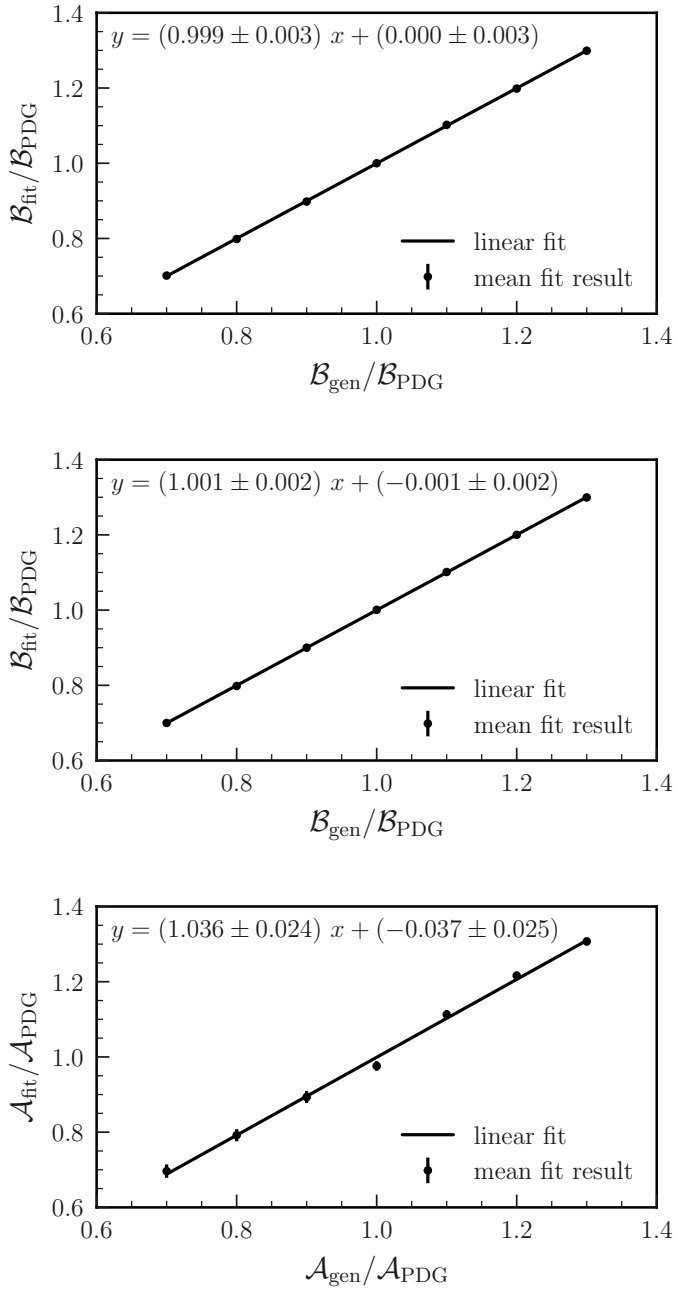


Figure 4.23.: Linearity test performed on large ensembles of artificial datasets, sampled from PDF shapes, for varying generated values of $\mathcal{B}(B^0 \to \pi^+\pi^-)$ (top), $\mathcal{B}(B^0 \to K^+\pi^-)$ (center) and $\mathcal{A}_{CP}(B^0 \to K^+\pi^-)$ (bottom). The mean of the fitted parameter plotted against the generated value follows a linear relationship with unit slope and an intercept consistent with zero.

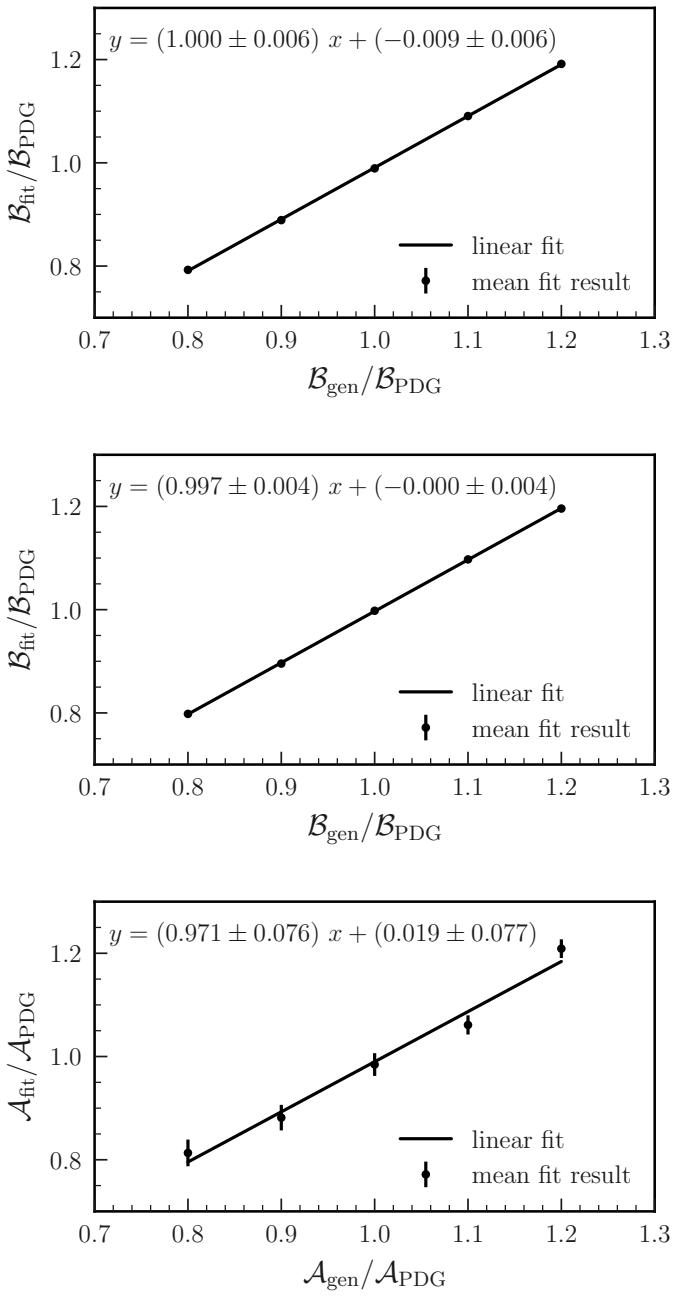


Figure 4.24.: Linearity test performed on large ensembles of bootstrap datasets, resampled from large MC datasets with replacement, for varying generated values of $\mathcal{B}(B^0 \to \pi^+\pi^-)$ (top), $\mathcal{B}(B^0 \to K^+\pi^-)$ (center) and $\mathcal{A}_{CP}(B^0 \to K^+\pi^-)$ (bottom). The mean of the fitted parameter plotted against the generated value follows a linear relationship with unit slope and an intercept consistent with zero.

5. Differences between MC and Data

Monte Carlo simulations can suffer from imperfections in the modeling of various physical processes, like the generation of particle decays or the interactions with detector material, to name only a few. To mitigate any resulting discrepancies between MC samples and real data, we perform correction procedures, which are instrumental in aligning the simulations more closely with the actual observations.

In this process, we turn to a number of control channel decays. Ideally, these transitions share similar kinematic and topological properties with the targeted $B \to hh'$ ($h = \pi^{\pm}, K^{\pm}$) decays. By comparing fits to simulated and real data, we extract correction factors to address mismodelings of the probability density functions, used in the targeted signal analysis. Additionally, control modes allow us to evaluate and correct the efficiency of continuum suppression requirements. We further extract weights to rectify mismodelings of particle identification variables. Lastly, for the measurement of the CP-violation parameter $\mathcal{A}_{CP}$, we obtain a correction factor to account for any discrepancies in the instrumental charge asymmetry.

Beyond calibration, control modes play a crucial role in validating our analysis methodologies. With significantly larger branching ratios compared to our targeted signal modes, they allow for the reconstruction of abundant, signal-dominated samples. These are utilized to verify the accurate extraction of efficiencies and ensure the correct consideration of external input parameters, such as the number of $B\bar{B}$ events in the sample. We develop a full analysis procedure akin to that used for our signal modes for the decay mode $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$. We measure the branching ratio in the LS1 dataset using a maximum-likelihood fit, and compare our result with the world-average.

5.1. Particle Identification (PID)

An essential component in the reconstruction of B-meson candidates is the efficient selection of daughter particles through cuts on their particle identification distributions. As eluded to in Chapter 4, we assign charged tracks with binary PID likelihood $\mathcal{L}_{K/\pi} < 0.5$ as pions. Accordingly, tracks with $\mathcal{L}_{K/\pi} > 0.5$ are assigned as kaons. The above likelihood, computed from an array of subdetector information, largely depends on the momentum and the polar angle of a track, which are prone to mismodeling in the simulation. Thus, a selection cut on the PID may induce a large systematic error in the estimation of the actual selection efficiency.

This is mitigated by using a control sample, where tracks can be unambiguously selected without requirements on the PID variable. Here, we use the decay $D^{*+} \to D^0[\to K^-\pi^+]\pi^+$, where the charge of the prompt pion unambiguously tags both decay products of the D meson. By fitting the invariant mass distribution of signal candidates, we compute so-called sWeights, which in turn are used to derive distributions of PID, track momentum and polar angle for tracks in signal events. This is achieved by effectively subtracting the contribution of background events, under the assumption of negligible correlations between the discriminant variable, i.e., the invariant mass, and the variables of interest [85]. We use the Belle II Systematic Corrections Framework developed for this purpose [82].

We obtain those distributions for a simulated sample corresponding to an integrated luminosity of $1\,\mathrm{ab}^{-1}$ as well as for the LS1 dataset. Comparing the efficiencies of our PID requirements in bins of momentum and the cosine of the polar angle, we calculate correction weights, corresponding to the efficiency ratio between data and MC. As we also fit the misidentified cross-feed events in our analysis, we additionally extract such correction weights for the fake rates. Figures 5.1 and 5.2 show the correction tables for charge-integrated kaon and pion efficiencies and fake rates in %. Uncertainties are computed by taking the relative difference between the nominal sWeight distribution and a second distribution obtained using an alternative PDF for the invariant mass fit.

Using these lookup tables, we subsequently retrieve correction weights for each charged track in reconstructed signal candidates and calculate the mean correction over a large set. For this, we use large samples of simulated signal events and do so for correctly reconstructed as well as for misidentified events, to account for the cross-feed contribution. We obtain the correction factors independently for the different components, as the momentum and angular distribution can differ substantially between them. Figure 5.3 shows the normalized momentum and $\cos\theta$ distributions of pions in the different samples in MC. While pions in the correctly reconstructed $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$ events follow almost identical distributions, pions in the cross-feed components significantly deviate from this.

Table 5.1 lists the final PID correction weights computed for all categories of charged particles used in the calculation of the branching ratios (cf. Equations 4.6–4.11). In the fit to data, we fix the correction factors and assign a systematic uncertainty taking into account their uncertainties.

Table 5.1.: PID correction factors to account for mismodeling of selection efficiencies (ε_i) and fake-rates (f_i), obtained for different categories i of charged pions and kaons present in $B \to hh'$ samples. The suffix CF refers to the cross-feed components containing misidentified particles.

Parameter	Value
ε_K	0.9704 ± 0.0013
ε_π	0.9811 ± 0.0011
$f_{\pi,\pi\pi-\mathrm{CF}}$	1.157 ± 0.010
$\varepsilon_{\pi,\pi\pi-\mathrm{CF}}$	0.9851 ± 0.0012
$f_{K,K\pi-\mathrm{CF}}$	1.192 ± 0.008
$\varepsilon_{\pi,K\pi-\mathrm{CF}}$	0.9861 ± 0.0011

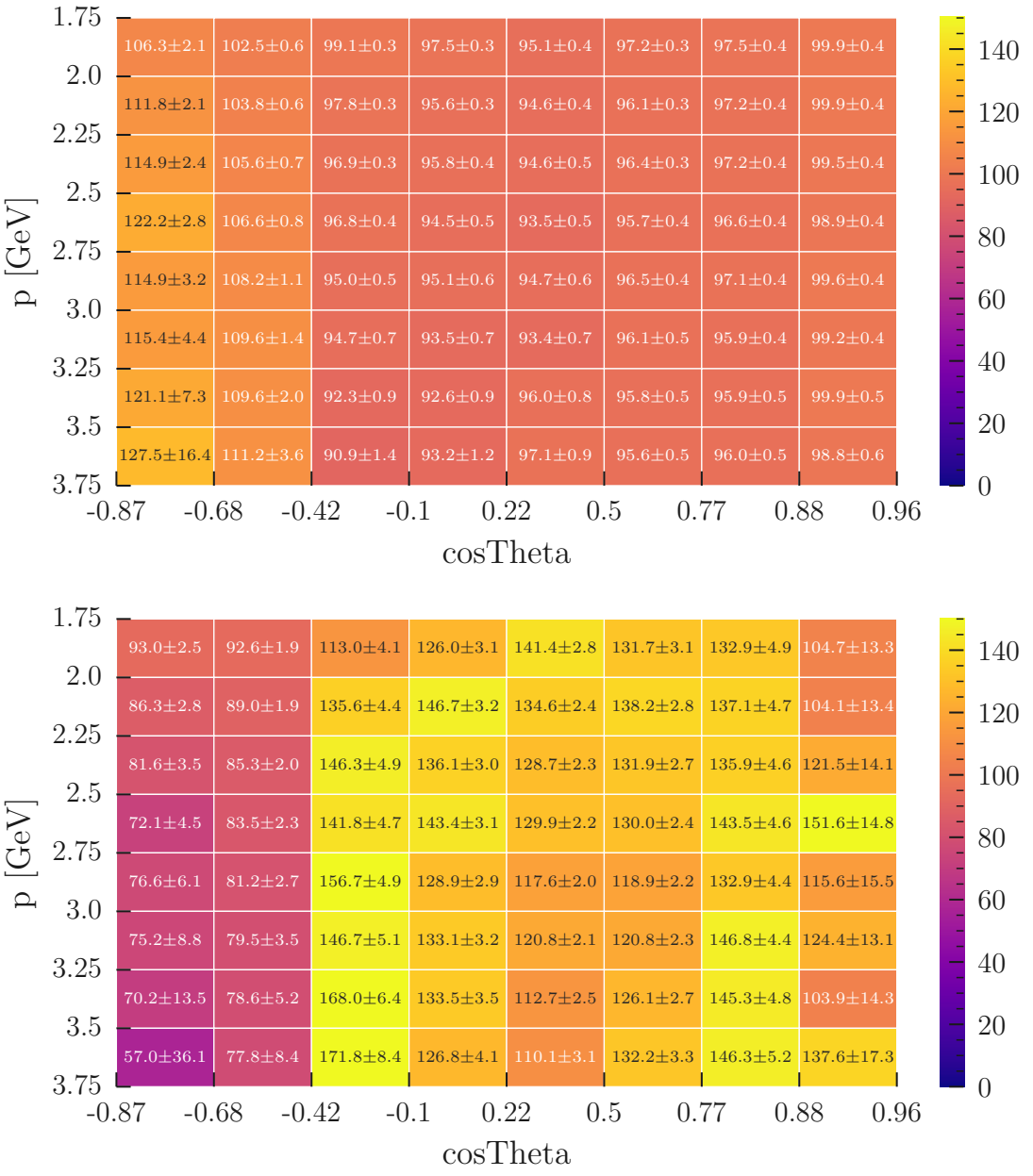


Figure 5.1.: PID correction table for charge-integrated kaon efficiencies (top) and fake rates (bottom) in % as function of momentum and polar angle.

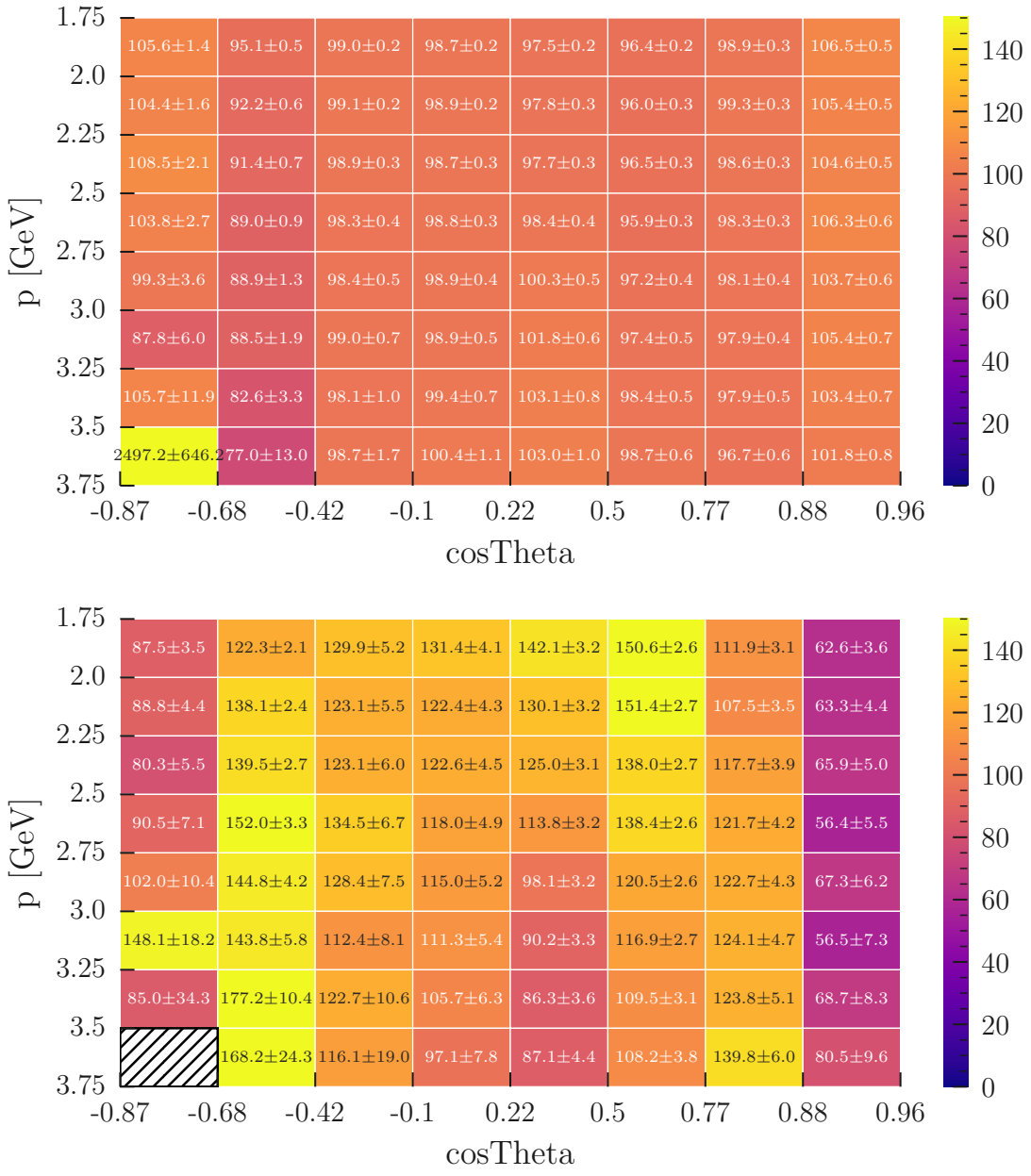


Figure 5.2.: PID correction table for charge-integrated pion efficiencies (top) and fake rates (bottom) in % as function of momentum and polar angle.

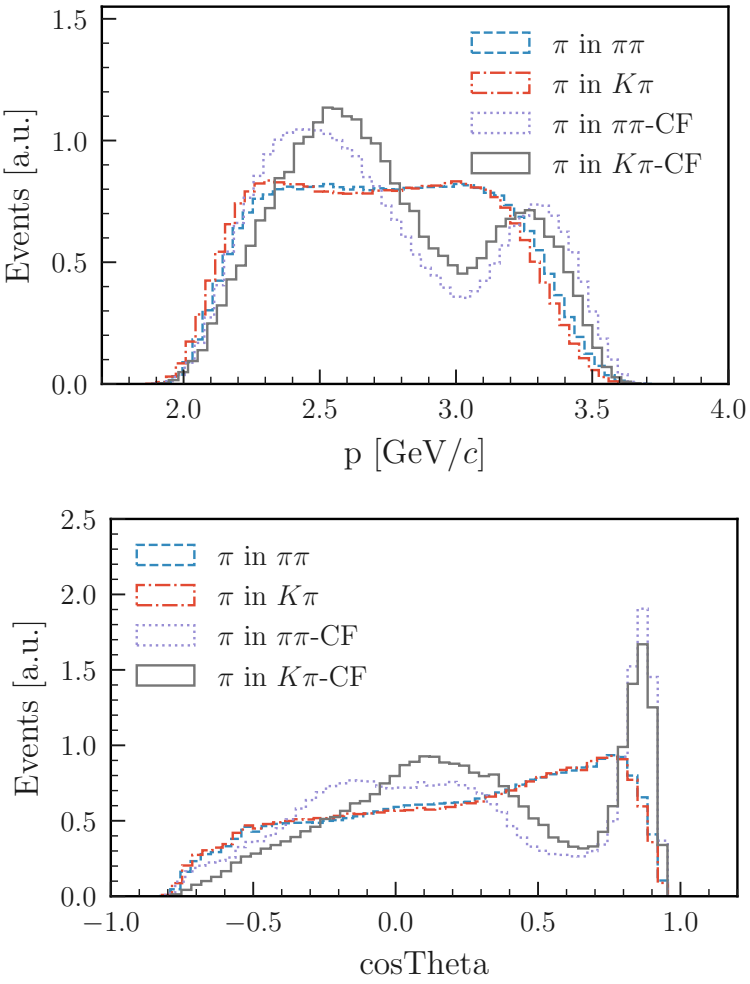


Figure 5.3.: Normalized momentum (top) and polar angle (bottom) distributions of pions in samples of correctly reconstructed and misidentified signal events, obtained in simulated data. To address the observed differences, we compute separate particle identification corrections for each component.

5.2. Correction for Instrumental Charge-Asymmetry

To measure the direct CP-violation parameter, $\mathcal{A}_{CP}$, we determine separate signal yields for $B^0 \to K^+\pi^-$ and $\bar{B}^0 \to K^-\pi^+$ transitions. Yield asymmetries can be biased by charge-dependent reconstruction efficiencies of final state particles. A known example is the detection asymmetry due to differing interaction probabilities of charged kaons with the detector material. Simulation might not perfectly replicate such effects. To mitigate potential discrepancies, supplementary measurements on abundant control mode samples are employed.

The measured charge-dependent yield asymmetry can be written as the sum of the physical CP asymmetry of the decay, $\mathcal{A}_{CP}$, and the instrumental asymmetry due to differences in interaction and reconstruction probabilities between particles and antiparticles, $\mathcal{A}_{\text{det}}$:

$$\mathcal{A}_{\text{raw}} = \mathcal{A}_{CP} + \mathcal{A}_{\text{det}} \ .$$

With a branching ratio of $3.947 \pm 0.030\%$, $D^0 \to K^-\pi^+$ decays occur sufficiently frequently to allow for a precise measurement of the yield asymmetry between the two charge cases. While sharing the same particles in the final state as our signal decay, the transitions are flavor-specific, which allows for an unambiguous charge assignment. In $D^0 \to K^-\pi^+$ transitions, direct CP violation is expected to be smaller than 0.1% [20]. We can therefore attribute any excess asymmetry to instrumental charge asymmetries.

In a separate Belle II study, $D^0 \to K^-\pi^+$ decays are reconstructed in MC and LS1 data to determine $\mathcal{A}_{\text{det}}$. The selection requirements are kept similar to those employed in the $B \to hh'$ analysis. The resulting ΔE distributions of candidates, dominated by signal events, are subsequently used to extract the charge-specific signal yields in maximum-likelihood fits. We obtain a correction of $\mathcal{A}_{\text{det}} = (0.62 \pm 0.67)\%$, which is subtracted from the raw rate asymmetry measured between $B^0 \to K^+\pi^-$ and $\bar{B}^0 \to K^-\pi^+$ decays.

5.3. Control Channel Analysis

We reconstruct the abundant decay mode $B^+ \to \bar{D}^0[K^+\pi^-]\pi^+$ and develop a fit procedure to measure the branching ratio in the LS1 data. A comparison with the precisely determined value reported by the PDG allows for a rigorous validation of our analysis methodologies. Further, we use this control channel to extract calibration parameters for signal mode PDF shapes employed in the target analysis, and to test the MC-data agreement in the efficiencies of continuum suppression requirements.

5.3.1. Reconstruction and Selection

Keeping the requirements in the control channel analysis similar to those used for $B \to hh'$ decays reduces the risk of introducing selection-related biases and ensures consistency in the estimation of systematic uncertainties. For final-state particles, we start with the same selection requirements as used in the target analysis. Subsequently, pairs of kaon and pion candidates are combined in kinematic fits to reconstruct D-meson candidates. We constrain the D^0 mass to be centered around its nominal mass $1.853 < m(K^-\pi^+) < 1.876\,\mathrm{GeV}/c^2$, which significantly reduces combinatorial background. The width of this window corresponds to three times the detector mass resolution as measured in a fit to simulated signal events. We further combine the D-meson candidates with reconstructed pions using vertex fits. The resulting B-meson candidates are required be fulfill $-0.1 < \Delta E < 0.1$ GeV. The final selection step is a loose cut on the continuum suppression classifier, for which we use the BDT trained on $B \to hh'$ decays. We use the same requirement as for the $B^0 \to K^+\pi^-$ analysis, $C > 0.41$, which reduces continuum significantly while retaining around 95% of signal events. The selection requirements are summarized in Table 5.2.

Figure 5.4 shows the ΔE distribution of $B^+ \to \bar{D}^0[K^+\pi^-]\pi^+$ candidates reconstructed in 1 ab^{-1} of generic simulation after all selection steps. We identify three components, namely signal, B-decay background and continuum background events.

Table 5.2.: Selection requirements for tracks, D- and B-meson candidates in the reconstruction of $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ events.

tracks	$\boldsymbol{D^0}$ / $\boldsymbol{B^+}$
$17° < \Theta < 150°$	$1.853 < m(K^-\pi^+) < 1.876\,\mathrm{GeV}/c^2$
nCDCHits > 20	$5.272 < M_{\mathrm{bc}} < 5.288\ \mathrm{GeV}/c^2$
dr < 0.5 cm	$-0.1 < \Delta E < 0.1\ \mathrm{GeV}$
$\lvert\mathrm{dz}\rvert < 2$ cm	$C > 0.41$
$\mathcal{L}_{K/\pi} > 0.5$ for kaons	
$\mathcal{L}_{K/\pi} < 0.5$ for pions	

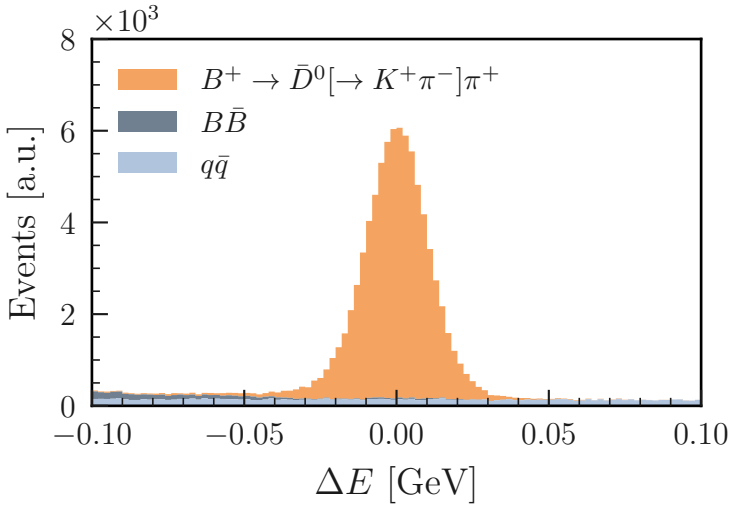


Figure 5.4.: ΔE distribution of $B^+ \to \bar{D}^0[K^+\pi^-]\pi^+$ candidates reconstructed in 1 ab^{-1} of generic simulation after all selection requirements.

5.3.2. Modeling

To extract the yields in each of the three contributions, we fit the distribution of the energy-difference (ΔE) of reconstructed candidates using an extended unbinned maximum-likelihood fit. Analogous to the target analysis, we first extract shapes for each component by fitting to abundant, truth-matched simulation samples. For the signal shape, we generate two million signal events. For the B-decay background and the continuum background, we make use of a generic MC sample corresponding to an integrated luminosity of 1 ab^{-1}.

Signal model

We use a combination of a Gaussian function and a Crystal Ball shape function to model the ΔE distribution of the signal component. The two functions are combined with a weight parameter $w \in [0, 1]$, that is determined in the fit to simulated events:

$$\begin{aligned} f_{sig}(\Delta E) \coloneqq w \cdot f_{CB}(\Delta E; \alpha, n, \mu_{CB}, \sigma_{CB}) \\ + (1 - w) \cdot f_G(\Delta E; \mu_G, \sigma_G) \ . \end{aligned} \tag{5.1}$$

We determine the best-fit parameters by means of a maximum-likelihood fit to a sample composed of correctly reconstructed signal events, surviving all selection criteria. Figure 5.5 shows the ΔE distribution for correctly reconstructed signal events alongside the fitted PDF.

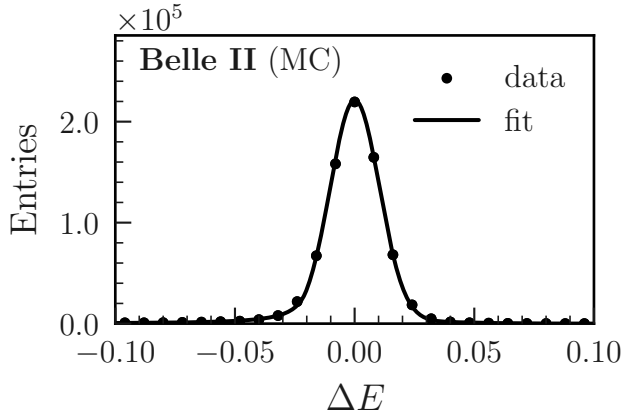


Figure 5.5.: ΔE distribution of correctly reconstructed $B^+ \to \bar{D}^0[K^+\pi^-]\pi^+$ signal events alongside the fitted signal PDF, composed of a Crystal Ball shape and a Gaussian function.

B-decay background model

The B-decay background is composed of a multitude of B-meson decays, where one or more particles were misreconstructed, leading to a number of pronounced features in the ΔE distribution. Since the relative contribution of this component is small ($\mathcal{O}(10^3)$) compared to the amount of signal decays ($\mathcal{O}(10^5)$), we chose to model this component using a single PDF. We opt for a kernel density estimation (KDE) PDF, which models the distribution as a superposition of multiple Gaussian functions. In this case, the widths of these Gaussian functions are adaptively calculated, taking the population density of the respective region into account.

We determine the shape of the PDF on a sample composed of simulated B-decay background events surviving the selection requirements. Figure 5.6 shows the ΔE distribution of B-decay background events along with the obtained PDF. A pronounced peaking structure around $\Delta E = 0$ is identified with misreconstructed $B^- \to J/\psi[\to l^+l^-]K^-$ events. Given that this contribution is estimated to yield only around 90 events under a signal peak of ∼27 000 events in the LS1 data, we choose to include it in the KDE PDF, instead of assigning a dedicated model.

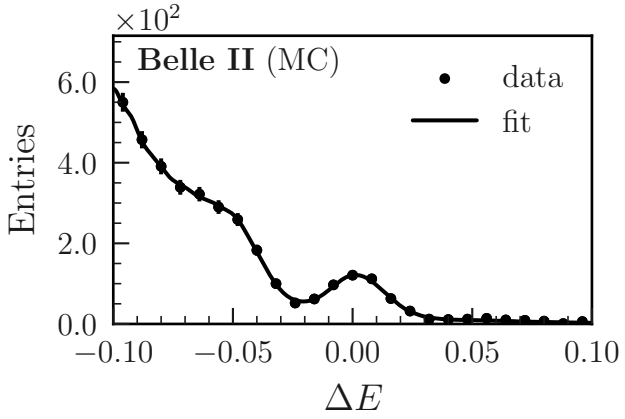


Figure 5.6.: ΔE distribution of simulated B-decay background events alongside the resulting KDE PDF. A pronounced peaking structure around $\Delta E = 0$ is identified with misreconstructed $B^- \to J/\psi[\to l^+l^-]K^-$ events.

Continuum background model

We use an exponential function to model the ΔE distribution of continuum background events. We leave the shape parameter floating. For completeness, we show the result of a fit to a simulated continuum background sample in Figure 5.7.

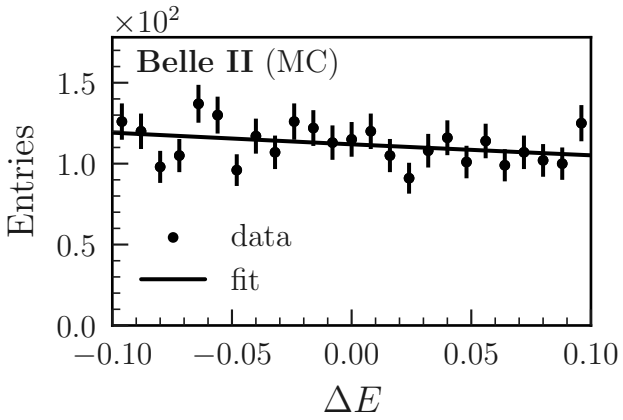


Figure 5.7.: ΔE distribution of simulated continuum background events alongside the fitted exponential function PDF.

5.3.3. Fits to MC and Data

We combine the PDFs describing the shapes of signal, B-decay background and continuum background events to build a single likelihood function by multiplying them with their respective yields. To account for the Poissonian nature of the measurement, an extension term is included. We leave the yields floating in the fit to data. To ensure a sufficient fit stability, we use a Gaussian constraint on the B-decay background yield, centered around the expected value estimated on MC. The resulting signal yield is used to measure the branching ratio of the control mode. For this, we extract the selection efficiency from a large simulation sample, consisting of two million signal events.

Fit to MC

Before fitting the LS1 data, we validate the fitting procedure on a large, generic MC sample corresponding to an integrated luminosity of 1 ab^{-1} and compare the fit results with MC truth information. Figure 5.8 shows the distribution in ΔE along with the fit projection. Table 5.3 lists the best-fit parameters along with their true values. We observe excellent agreement between fit result and true values in all three contributions.

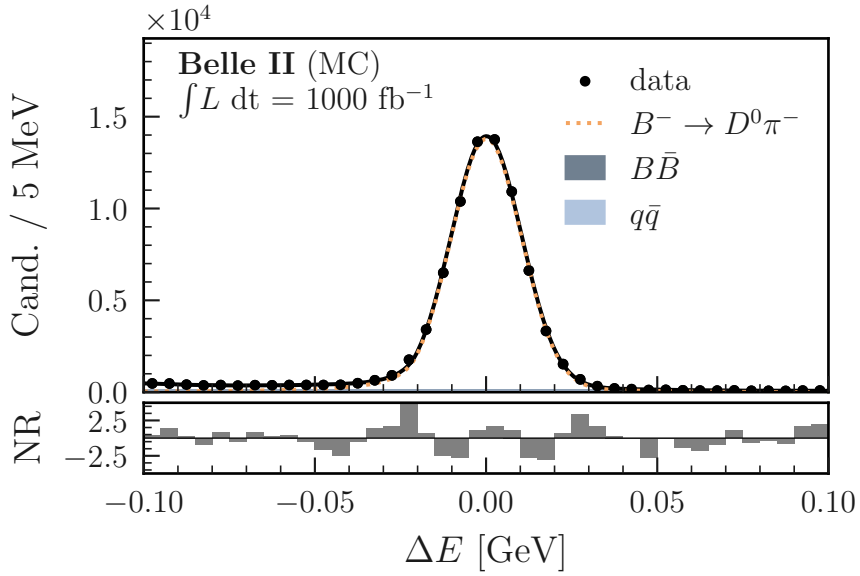


Figure 5.8.: Distribution of ΔE for $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ candidates reconstructed in 1 ab^{-1} of generic simulation. The projection of an extended unbinned maximum-likelihood fit is overlaid. In the signal-dominated sample, the background components are only faintly discernible.

Table 5.3.: Best-fit parameters, obtained in a fit to ΔE of $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ candidates reconstructed in simulated data corresponding to an integrated luminosity of 1 ab^{-1}, along with their true values and the corresponding normalized residuals.

Parameter	Fit Result	Simulation Truth	NR
signal yield	74653 ± 287	74545	0.38
$B\bar{B}$ background	3492 ± 59	3501	0.15
$q\bar{q}$ background	2865 ± 143	2797	0.48

Fit to data

For the fit to data, the selection efficiency needs to be corrected to account for mismodeling of the PID variables, as laid out in detail in the previous section. We follow the same procedure as for the $B \to hh'$ analysis and compute correction parameters for all three final state particles of the decay, taking their momentum and polar angle distributions into account. We obtain a final selection efficiency of $\varepsilon = 37.4\% \times 94.4\% = 35.3\%$, where the second factor refers to the combined PID correction.

In order to account for possible mismodeling of the signal PDF shape, we allow the width and mean of the signal peak to vary from the nominal values determined on the simulation sample. We assume that the measurement effects leading to any deviations are also applicable to the signal mode analysis. The deviations labeled as σ_{corr} and μ_{corr} correspond to the correction factors, which we employ in the $B \to hh'$ analysis to adjust the widths and means of all peaking components (cf. Section 4.3.1).

$$\mu_G \to \mu_G + \mu_{corr} \quad , \quad \sigma_G \to \sigma_G + \sigma_{corr}$$

Figure 5.9 shows the ΔE distribution in the full LS1 dataset along with the projection of an extended unbinned maximum-likelihood fit. The data sample of reconstructed $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ candidates is largely dominated by signal events with only minuscule contributions from B-decay and continuum background events. After the PID correction, the branching ratio evaluates to $(1.829 \pm 0.012) \times 10^{-4}$, which agrees with the current PDG average of $(1.847 \pm 0.053) \times 10^{-4}$. Table 5.4 lists all best-fit parameters, including the shape correction factors.

Table 5.4.: Best-fit parameters, obtained by a fit to the ΔE distribution of $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ candidates reconstructed in the LS1 data.

Parameter	Fit result
$\mathcal{B}$	$(1.829 \pm 0.012) \times 10^{-4}$
ΔE mean shift	(-0.29 ± 0.08) MeV
ΔE width scaling	1.115 ± 0.007
signal	$25\,754 \pm 170$
$q\bar{q}$ background	$1\,042 \pm 89$
$B\bar{B}$ background	$1\,246 \pm 35$

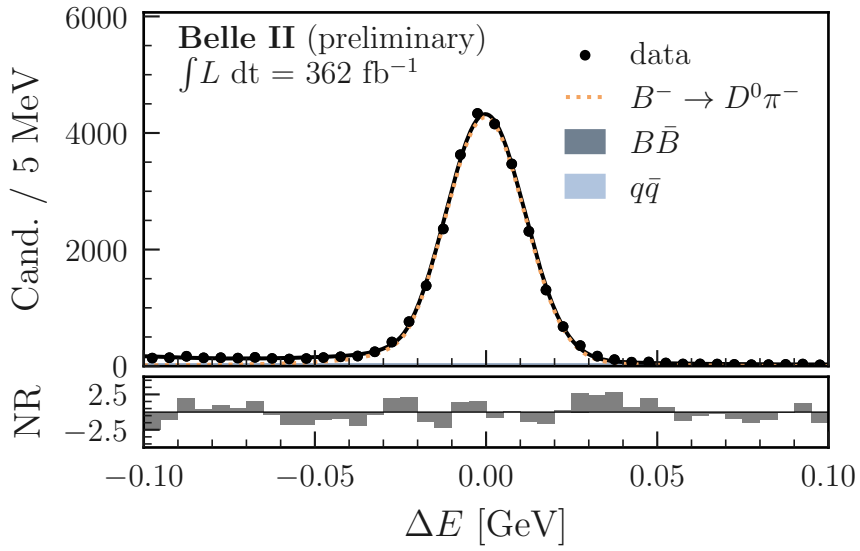


Figure 5.9.: Distribution of ΔE for $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ candidates reconstructed in the LS1 dataset corresponding to an integrated luminosity of 362 fb^{-1}. The projection of an extended, unbinned maximum-likelihood fit is overlaid. In the signal-dominated sample, the background components are only faintly discernible.

5.3.4. Continuum Suppression Classifier

We compare the distributions of the CS classifier after applying the boosted decision tree of the $B \to hh'$ analysis to the control mode samples in MC and data. We divide the samples into signal-enhanced and background-enhanced subsets to assess its applicability in both components. Within a $|\Delta E| < 50\,\text{MeV}$ window, signal events dominate, while candidates outside this window primarily consist of background events. Figure 5.10 illustrates the distributions in data and MC for both scenarios. To facilitate the comparison, we normalize the MC distribution to match the number of data points. We observe excellent agreement between the signal-enhanced samples. In the background-enhanced samples a slight discrepancy is noticeable. This discrepancy can be attributed to the wider peak width observed in data compared to MC ($\sigma_{corr} = 1.115 \pm 0.007$), resulting in a greater number of signal events contributing to this sample at higher values.

Additionally, we utilize the control mode to check for discrepancies between data and simulation concerning the efficiencies of the cuts imposed on the continuum suppression classifier. To obtain the selection efficiency of this requirement, we perform a simultaneous fit to two disjoint samples: the first comprises candidates that satisfy the CS selection cut used in the signal mode analysis, while the sec-

ond comprises candidates that fail this selection. We employ the identical fitting procedure as described in the previous section, allowing for varying shapes of the signal peaks between the two samples. Figure 5.11 presents the ΔE distributions of $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ candidates that satisfy and fail the CS requirement used to select $B^0 \to K^+\pi^-$ candidates, reconstructed in MC. The figure also displays the projections resulting from a simultaneous fit to these two samples. The selection efficiency is determined by the ratio of their respective signal yields. Figure 5.12 shows the distribution and fit projections for data. In Figures 5.13 and 5.14, we present analogous studies for the CS requirement used to select $B^0 \to \pi^+\pi^-$ candidates. We find compatible efficiencies between data and MC for both requirements (see Table 5.5). We include a systematic uncertainty related to the CS selection quoting the statistical uncertainty on the efficiency ratio of data and MC.

Table 5.5.: Selection efficiencies of continuum suppression requirements used in the $B \to hh'$ analysis, extracted from simultaneous fits to $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ candidates passing and failing the cuts. The values obtained on data and MC show good agreement.

CS cut	ε(data)	ε(MC)	ε(data)/ε(MC)
>0.34	95.56 ± 0.18 %	95.87 ± 0.10 %	99.68 ± 0.21 %
>0.41	94.64 ± 0.19 %	94.78 ± 0.11 %	99.85 ± 0.23 %

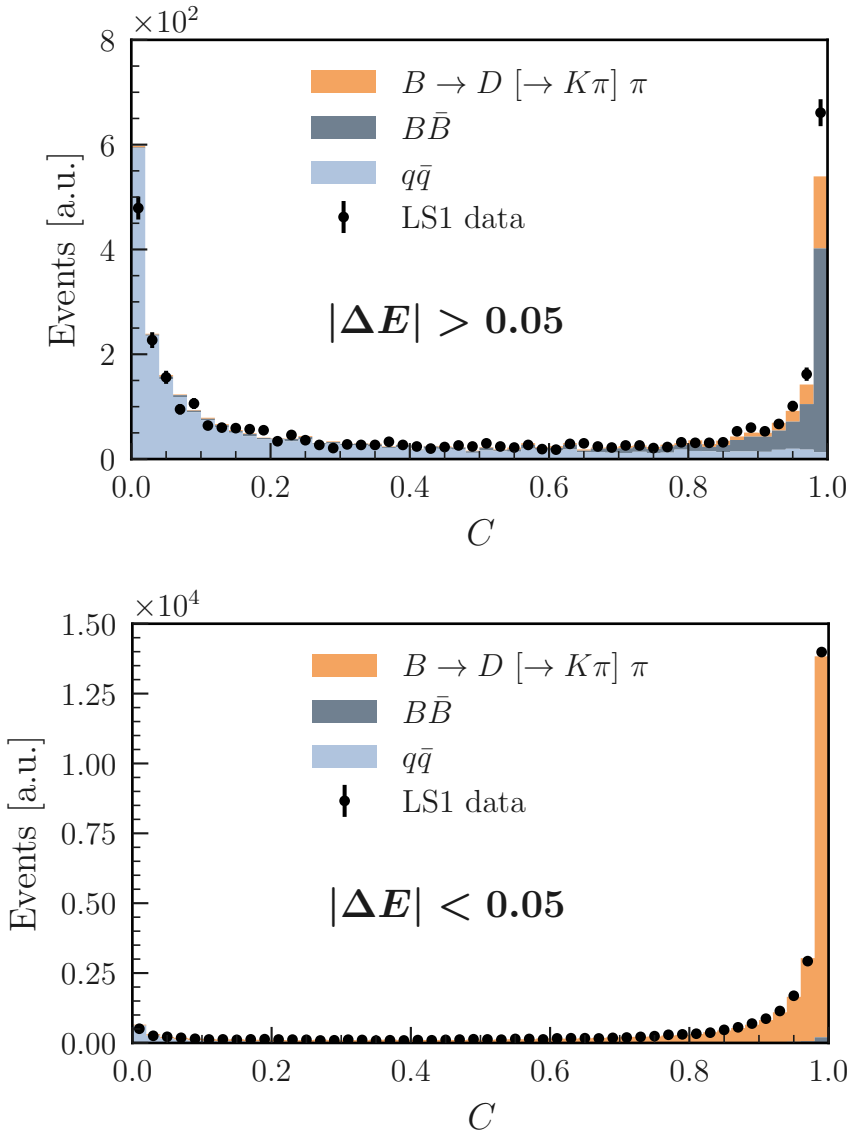


Figure 5.10.: Data-MC comparison of the CS classifier for background-enhanced (top) and signal-enhanced (bottom) $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ candidates. MC is normalized to the total number of data events.

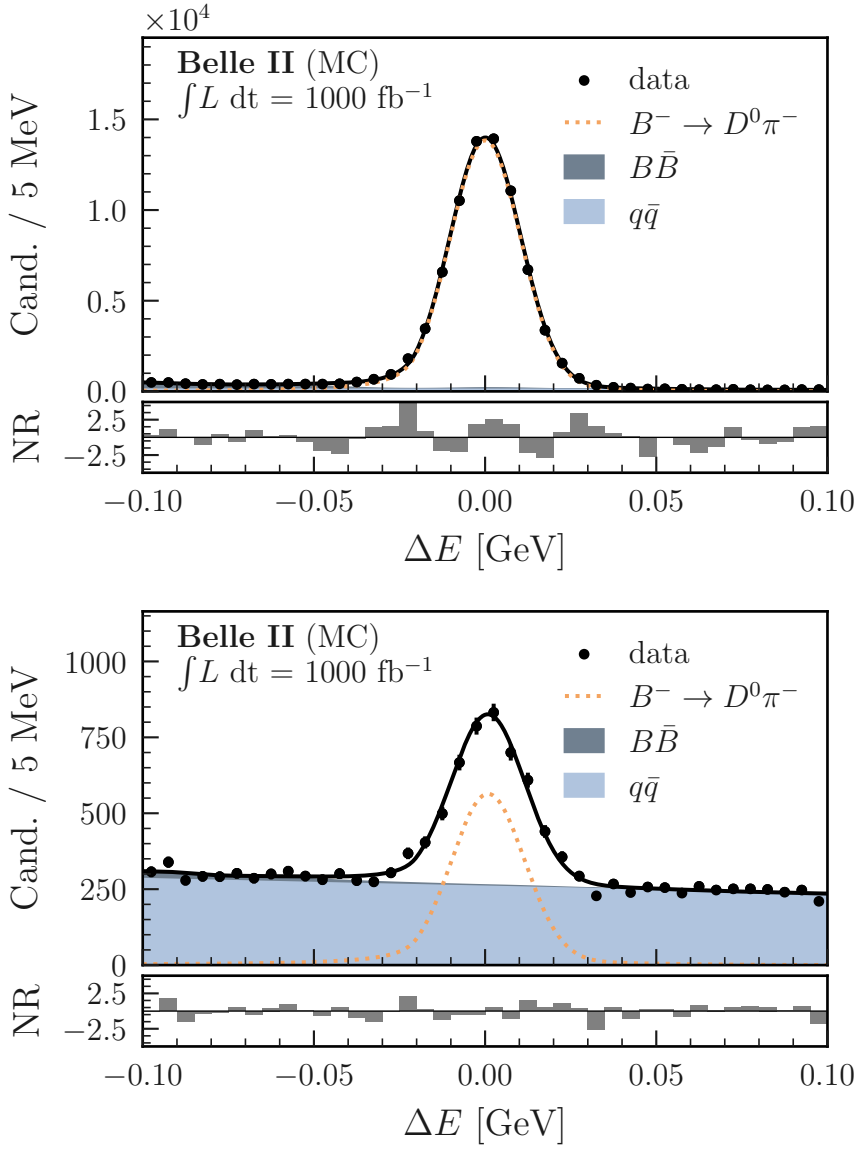


Figure 5.11.: ΔE distributions of $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ candidates that satisfy (top) and fail (bottom) the CS requirement, $C > 0.41$, used to select $B^0 \to K^+\pi^-$ candidates, reconstructed in a simulated sample corresponding to an integrated luminosity of 1 ab^{-1}.

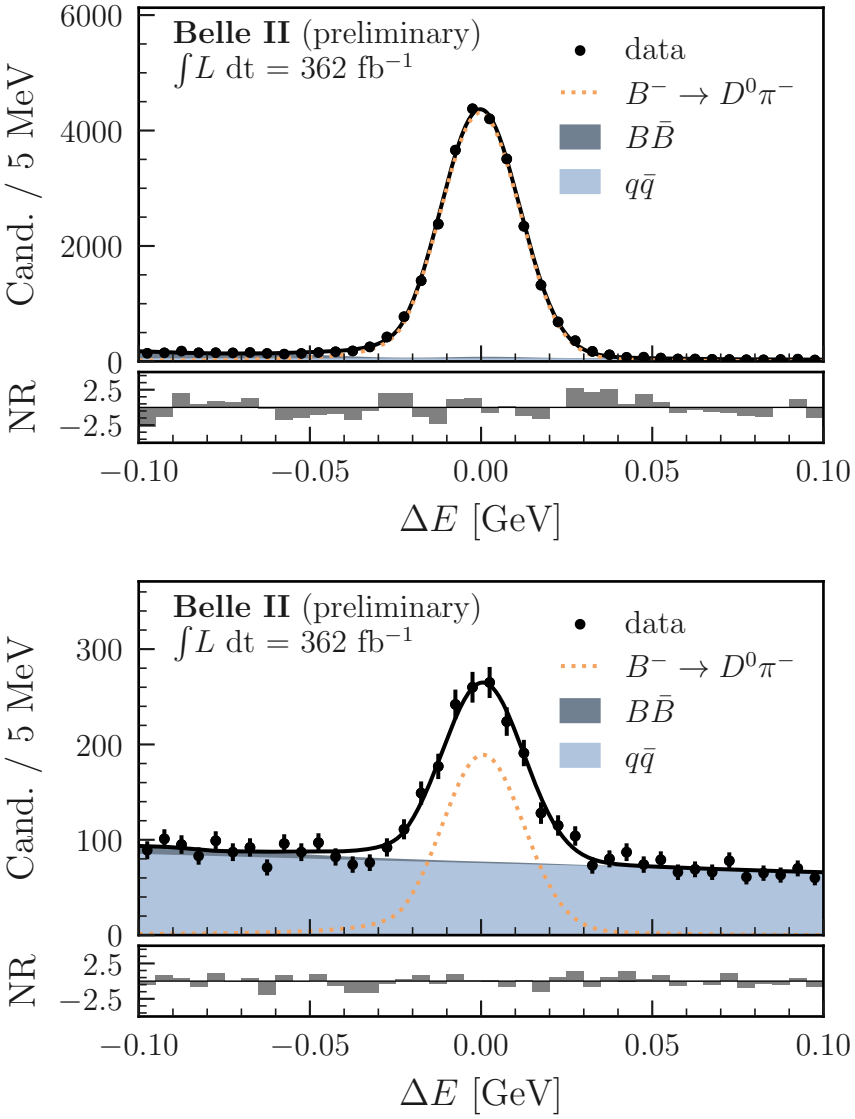


Figure 5.12.: ΔE distributions of $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ candidates that satisfy (top) and fail (bottom) the CS requirement, $C > 0.41$, used to select $B^0 \to K^+\pi^-$ candidates, reconstructed in the LS1 dataset.

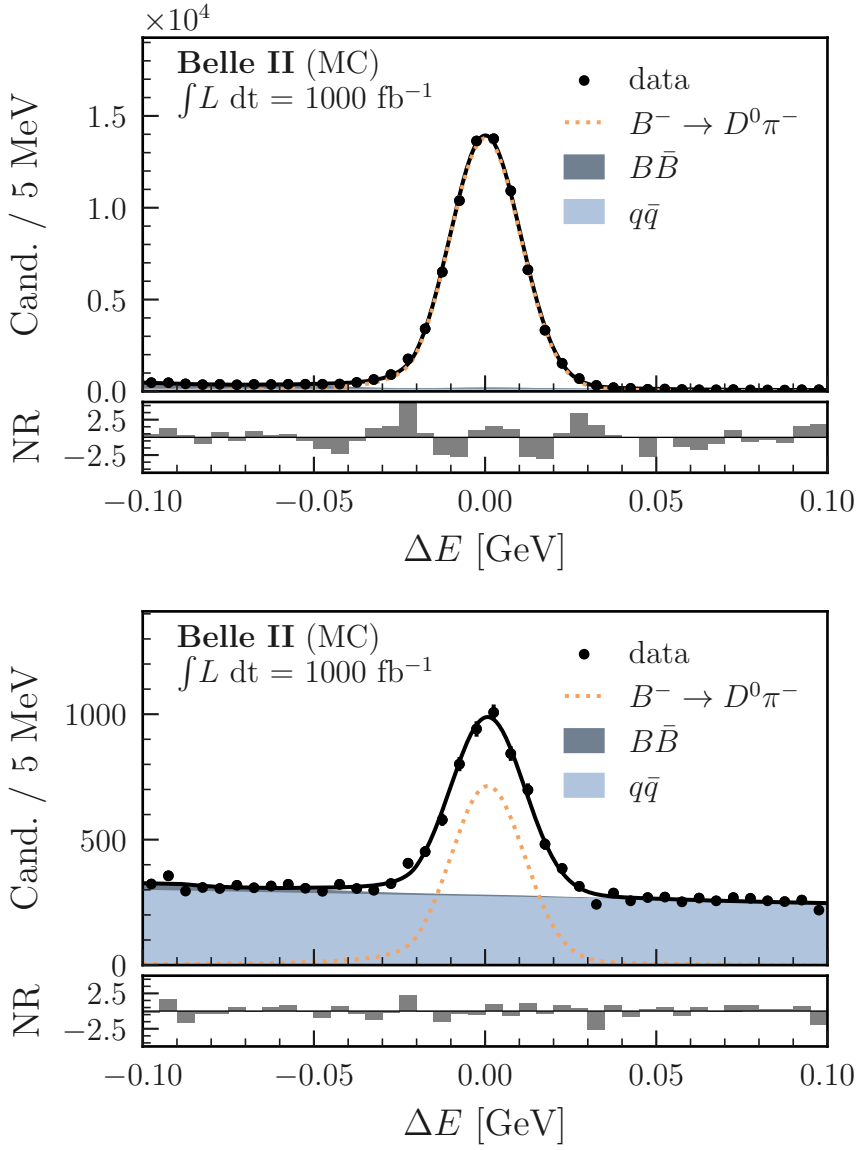


Figure 5.13.: ΔE distributions of $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ candidates that satisfy (top) and fail (bottom) the CS requirement, $C > 0.34$, used to select $B^0 \to \pi^+\pi^-$ candidates, reconstructed in a simulated sample corresponding to an integrated luminosity of 1 ab^{-1}.

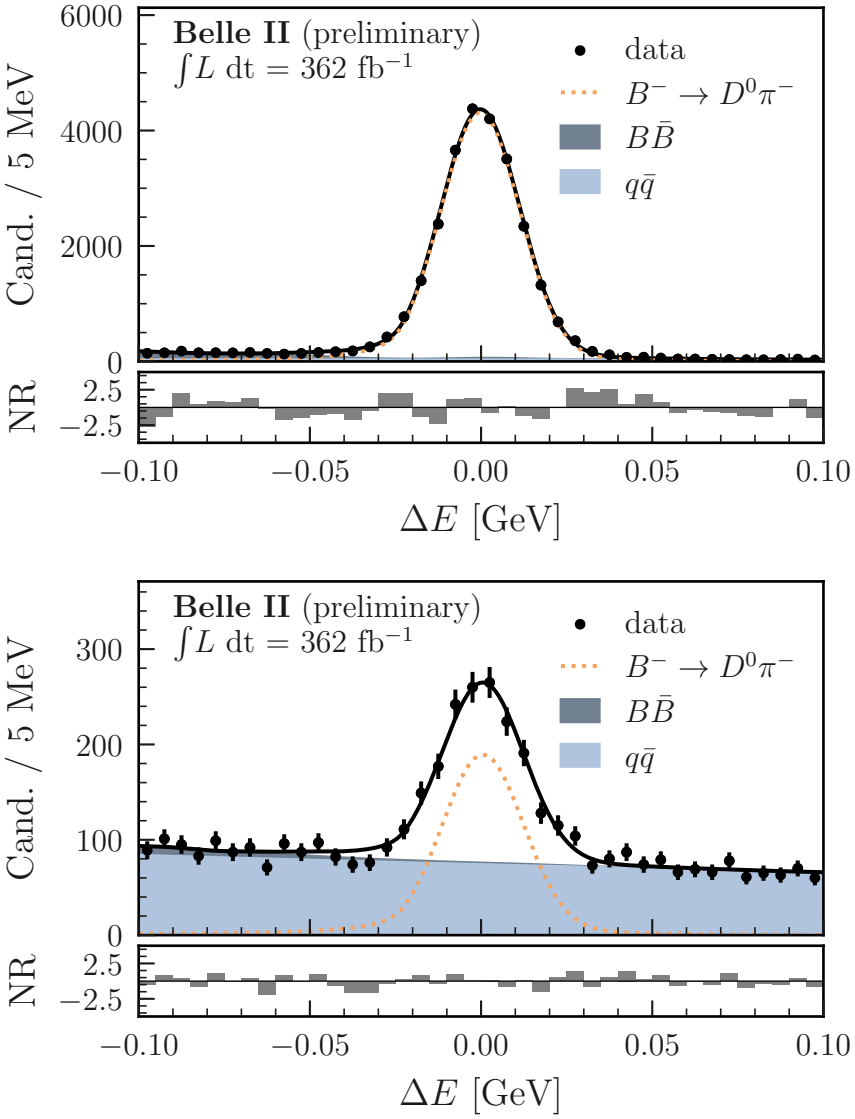


Figure 5.14.: ΔE distributions of $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ candidates that satisfy (top) and fail (bottom) the CS requirement, $C > 0.34$, used to select $B^0 \to \pi^+\pi^-$ candidates, reconstructed in the LS1 dataset.

6. Results

In the following, we present the results of a simultaneous fit to the LS1 data to measure the branching ratios $\mathcal{B}(B^0 \to K^+\pi^-)$ and $\mathcal{B}(B^0 \to \pi^+\pi^-)$, as well as the direct CP-violation parameter, $\mathcal{A}_{CP}$, in $B^0 \to K^+\pi^-$ decays. We estimate the systematic uncertainties by investigating various sources and compare the final results with previous studies performed at Belle and Belle II. Lastly, we incorporate the new measurement into a model-independent isospin sum rule along with other updated or new measurements of related $B \to K\pi$ decay modes and compare the result with SM expectation.

6.1. Fit to LS1 Data

We simultaneously fit the distributions in the energy-difference, ΔE, and the transformed continuum suppression classifier, C', in three disjoint samples, composed of $B^0 \to \pi^+\pi^-$, $B^0 \to K^+\pi^-$ and $\bar{B}^0 \to K^-\pi^+$ candidates. After all selection requirements, the total number of candidates in these samples are 10124, 6511 and 6240, respectively. There are a total of 18 free parameters in the fit: two branching ratios, the charge asymmetry, three background yields and 12 background shape parameters.

Figures 6.1 and 6.2 display the projections onto the two fit dimensions for all three samples, along with the corresponding normalized residuals. The fit result exhibits excellent agreement with data. The observed discrepancies between the fitted model and the data points are well within the expected statistical fluctuations across the parameter space, following approximately a standard normal distribution. After applying all corrections, we obtain the following branching ratios and CP asymmetry:

$$\begin{aligned} \mathcal{B}(B^0 \to \pi^+\pi^-) &= (5.83 \pm 0.22) \times 10^{-6} \ , \\ \mathcal{B}(B^0 \to K^+\pi^-) &= (20.67 \pm 0.37) \times 10^{-6} \ , \\ \mathcal{A}_{CP}(B^0 \to K^+\pi^-) &= -0.072 \pm 0.019 \ . \end{aligned}$$

The listed uncertainties are statistical and correspond to the symmetric errors obtained by the covariance matrix. The systematic uncertainties are discussed in

the following section. Table 6.1 lists the best-fit values for the full set of floating parameters. The statistical correlation coefficients between all fit parameters are listed in Appendix A.3.

Table 6.1.: Best-fit values of floating parameters obtained in an extended unbinned maximum-likelihood fit to the LS1 data. The listed CP-violation parameter already includes the correction for instrumental charge asymmetries. a, b, c and w represent PDF shape parameters for the combinatorial background components (see Section 4.3.1 for details).

Parameter	Fit Result
$\mathcal{B}(B^0 \to \pi^+\pi^-)$	$(5.83 \pm 0.22) \times 10^{-6}$
$\mathcal{B}(B^0 \to K^+\pi^-)$	$(20.67 \pm 0.37) \times 10^{-6}$
$\mathcal{A}_{CP}(B^0 \to K^+\pi^-)$	-0.072 ± 0.019
background yield $(\pi^+\pi^-)$	8093 ± 97
background yield $(K^+\pi^-)$	4286 ± 72
background yield $(K^-\pi^+)$	4231 ± 71
a $(C'; \pi^+\pi^-)$	-6.68 ± 0.29
b $(C'; \pi^+\pi^-)$	-24.6 ± 1.9
w $(C'; \pi^+\pi^-)$	0.566 ± 0.038
c $(\Delta E; \pi^+\pi^-)$	-0.951 ± 0.134
a $(C'; K^+\pi^-)$	-6.80 ± 0.42
b $(C'; K^+\pi^-)$	-24.9 ± 3.6
w $(C'; K^+\pi^-)$	0.662 ± 0.061
c $(\Delta E; K^+\pi^-)$	-1.15 ± 0.183
a $(C'; K^-\pi^+)$	-6.94 ± 0.46
b $(C'; K^-\pi^+)$	-22.6 ± 3.8
w $(C'; K^-\pi^+)$	0.669 ± 0.074
c $(\Delta E; K^-\pi^+)$	-0.979 ± 0.184

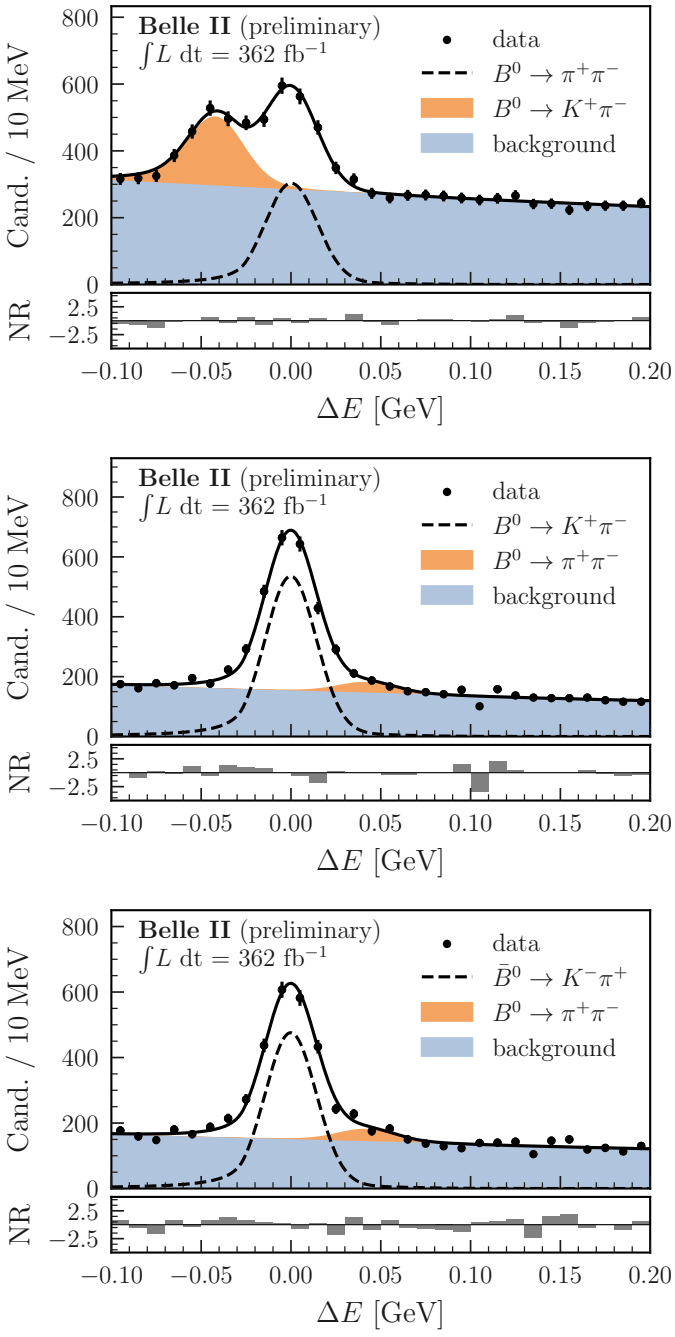


Figure 6.1.: Distributions in ΔE for $B^0 \to \pi^+\pi^-$ (top), $B^0 \to K^+\pi^-$ (center), and $\bar{B}^0 \to K^-\pi^+$ (bottom) candidates reconstructed in the LS1 dataset. The projections of an extended unbinned maximum-likelihood fit are overlaid. Normalized residuals are shown below.

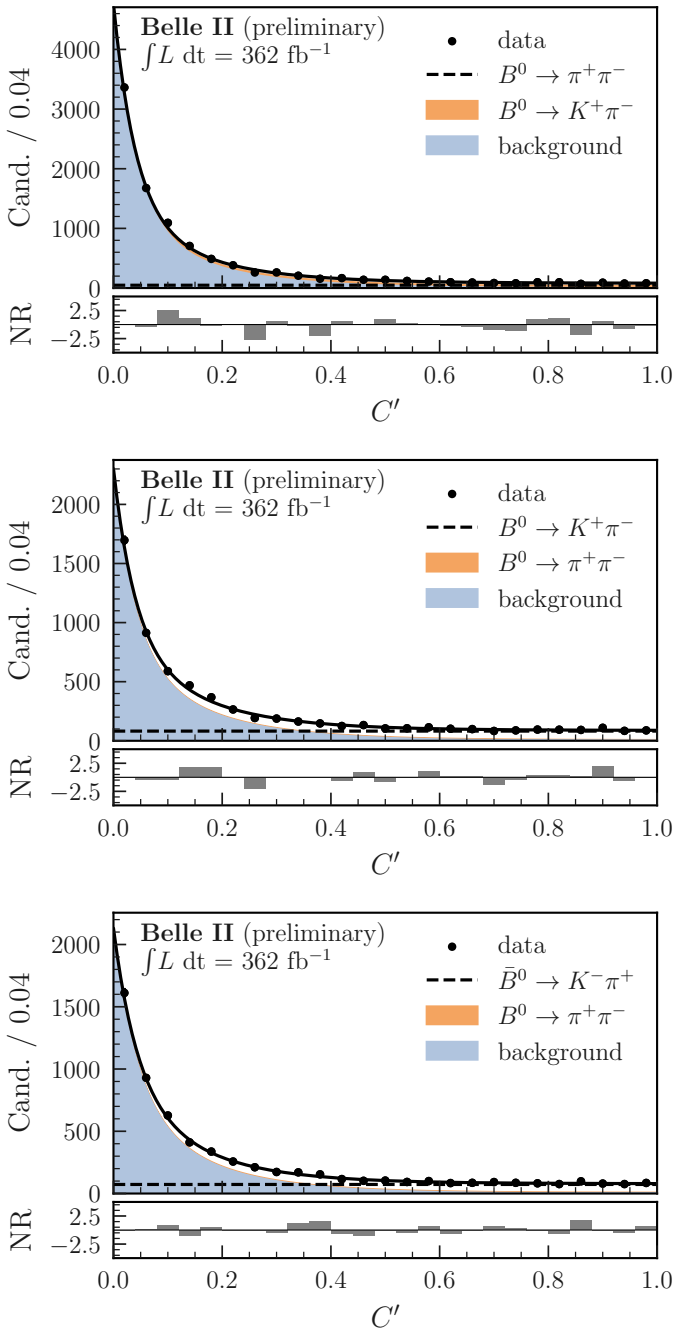


Figure 6.2.: Distributions in C' for $B^0 \to \pi^+\pi^-$ (top), $B^0 \to K^+\pi^-$ (center), and $\bar{B}^0 \to K^-\pi^+$ (bottom) candidates reconstructed in the LS1 dataset. The projections of an extended unbinned maximum-likelihood fit are overlaid. Normalized residuals are shown below.

6.2. Systematic Uncertainties

The proper estimation of systematic uncertainties plays a crucial role in quantifying potential biases and ensuring the reliability of a measurement. Combined with statistical uncertainties, provided by the minimizer, systematic uncertainties are an integral factor in determining the significance of the parameter estimation as a whole. In the following, we discuss the various sources of systematic uncertainties considered relevant to the level of statistical precision attained in our measurement. Table 6.2 lists the individual contributions along with the statistical uncertainties for measurements of branching ratios and the CP-violation parameter. The total systematic uncertainties are obtained by summing the individual contributions in quadrature.

Table 6.2.: Summary of estimated systematic uncertainties for a number of relevant sources. The uncertainties are given in percent for the branching ratios, and as absolute values for the CP-violation parameter.

Source	$\mathcal{B}(B^0 \to K^+\pi^-)$ [%]	$\mathcal{B}(B^0 \to \pi^+\pi^-)$ [%]	$\mathcal{A}_{CP}(B^0 \to K^+\pi^-)$
Tracking efficiency	0.5	0.5	–
$B\bar{B}$ pair counting	1.5	1.5	–
f^{00}	2.5	2.5	–
CS efficiency	0.2	0.2	–
PID correction	0.1	0.1	<0.001
ΔE shift and scale	0.1	0.2	<0.001
$K\pi$ signal model	0.1	0.2	<0.001
$\pi\pi$ signal model	<0.1	0.1	<0.001
$K\pi$ cross-feed model	<0.1	0.1	<0.001
$\pi\pi$ cross-feed model	0.1	0.2	<0.001
Instrumental asymmetry	–	–	0.007
Multiple candidates	<0.1	<0.1	<0.001
Systematic uncertainty	3.0	3.0	0.007
Statistical uncertainty	1.8	3.8	0.019

6.2.1. Tracking Efficiency

In a dedicated study on $e^+e^- \to \tau^+\tau^-$ events, the tracking efficiencies are determined in data and MC using a tag-and-probe method. To this end, events are reconstructed, where one τ decays leptonically, $\tau \to \ell^\pm \nu_\ell \bar{\nu}_\tau$ with $\ell = e, \mu$, and the other one

hadronically, $\tau \to 3\pi^{\pm}\nu_{\tau} + n\pi^0$. The measurement of two charged pions and the lepton allows to infer the additional pion from charge conservation. The reconstruction efficiency of this probe track is used as an approximation of the tracking efficiency. The discrepancy between data and MC is then used to estimate the corresponding uncertainty. We assign a systematic uncertainty on the branching ratios of 0.24% per track in the final state. The study is explained in detail in [86].

6.2.2. $B\bar{B}$ Pair Counting

The number of $B\bar{B}$ pairs contained in the LS1 data is determined in a dedicated study by the Belle II performance group. Comparing the number of recorded events between an off-resonance data sample recorded at a centre-of-mass energy of 10.519 GeV and that of an on-resonance sample, the total number of $B\bar{B}$ pairs is estimated.

To account for the different sample sizes, and energy-dependent selection efficiencies and cross-sections, the following equation is used:

$$N_{B\bar{B}} = \frac{N_{\text{on}} - R \cdot k \cdot N_{\text{off}}}{\epsilon_{B\bar{B}}} \,, \tag{6.1}$$

where N_{on} is the number of selected hadronic events in on-resonance data, $R{\cdot}k{\cdot}N_{\text{off}}$ is the estimated number of non-$B\bar{B}$ events in on-resonance data extrapolated from the off-resonance sample, and $\epsilon_{B\bar{B}}$ is the selection efficiency for $B\bar{B}$ events. $R = \mathcal{L}_{\text{on}}/\mathcal{L}_{\text{off}}$ is the luminosity ratio between the on-resonance and off-resonance data samples, and k takes into account the variation in non-$B\bar{B}$ efficiencies and cross sections due to differences in center-of-mass energy. We assign a systematic uncertainty of 1.5% to account for the uncertainty on the number of $B\bar{B}$ pairs, where the dominant contribution stems from systematic uncertainties in the luminosity measurements. The study is explained in detail in [87].

6.2.3. f^{00}

The production ratio of charged and neutral B mesons in decays of the $\Upsilon(4S)$ resonance is given by:

$$f^{+-}/f^{00} = \frac{\Gamma[\Upsilon(4S) \to B^+B^-]}{\Gamma[\Upsilon(4S) \to B^0\bar{B}^0]} \,. \tag{6.2}$$

To compute the number of neutral B mesons in our sample, we multiply the total number of B mesons with f^{00}. The most precise measurement of this parameter to date is provided by the measurement of $B^0 \to J/\psi(\to l^+l^-)K^0$ and

$B^+ \to J/\psi(\to l^+l^-)K^+$ decays using the full Belle dataset [88]. Assuming isospin invariance in the $B \to J/\psi(\to l^+l^-)K$ system and accounting for reconstruction efficiencies as well as lifetimes of the B mesons, the production ratio f^{+-}/f^{00} is determined. Taking the combined uncertainties on this measurement, we compute the uncertainty on f^{00} and assign it as a systematic uncertainty in our branching ratio measurement.

6.2.4. CS Selection

We suppress continuum background by imposing loose selection requirements on the outputs of a dedicated boosted decision tree, which is trained and tested on simulation samples. Possible differences between data and simulation in the distributions of training variables can lead to non-negligible differences in the resulting classifier distributions. We therefore investigate the data-MC agreement in an abundant control sample of $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ decays. To determine the efficiencies of the selection requirements on the CS classifier variable used in the $B \to hh'$ analysis, we measure the ratio of signal yields passing and failing these cuts in a simultaneous fit to both samples (cf. Section 5.3.4). We obtain the efficiencies for both, simulation and data samples. We find compatible efficiencies between the two and assign the statistical uncertainties on the efficiency ratios as systematic uncertainties.

6.2.5. PID Correction

The Systematic Corrections Framework provides uncertainties on the extracted PID correction tables, which include systematic as well as statistical uncertainties [82]. In the nominal correction, we use the central value for a given bin in momentum and polar angle to obtain a correction weight for the PID selection efficiency for each track on the signal side. Averaging over a large set of simulated events, we obtain efficiency corrections for each class of particles, i.e., correctly reconstructed and misidentified charged pions and kaons (cf. Section 5.1).

To estimate the systematic uncertainty due to uncertainties in the PID correction tables, we generate large sets of alternative corrections for each particle class. Varying the correction for each track within its uncertainty 300 times, and subsequently taking the average over each set, we compute an ensemble of 300 alternative corrections for each particle class. We use the standard deviation as a measure of the systematic uncertainty of the PID correction for particle class c,

$$\sigma_c = \mathrm{std}_k \left(\frac{1}{N} \sum_{i=1}^{N} \epsilon_{i,c,k}^{\mathrm{corr}} \right) \ ,$$

where std_k represents the standard deviation over $k \in \{1, 2, 3, \ldots, 300\}$, N is the number of events/tracks in the simulation sample, and $\epsilon^{\mathrm{corr}}_{i,c,k}$ is the PID efficiency correction for track i of particle class c in variation k. The resulting uncertainties are listed in Table 5.1.

We combine the uncertainties of two tracks according to the event type and compute the weighted average of signal and misidentified event yields to account for the fact that both yields are used in the determination of branching ratios.

6.2.6. Shape Correction

To estimate the systematic uncertainty associated to the ΔE mean shift and width scaling obtained in the control mode analysis, we generate 100 pseudo datasets. Sampling from the nominal fit shapes, we generate distributions in ΔE and C' for each component in their expected proportion. We randomly fluctuate the yields sampling from a Poisson distribution around the expected value, estimated on a large generic simulation sample, and appropriately scaled to match the integrated luminosity of the LS1 dataset.

We fit these pseudo datasets twice: once using our nominal fitter and once using alternative calibration parameters. The latter are obtained sampling from Gaussian distributions, centered around the nominal corrections with widths corresponding to their respective uncertainties. We calculate the relative differences in measured physics parameters for each pair of fit results across the 100 pseudo experiments. The resulting distribution is fitted with a Gaussian function and we quote the fitted width, σ, as systematic uncertainty. Figure 6.3 shows the distributions of the relative fit differences for the branching ratio measurements. We observe no significant difference in the measurement of $\mathcal{A}_{CP}$.

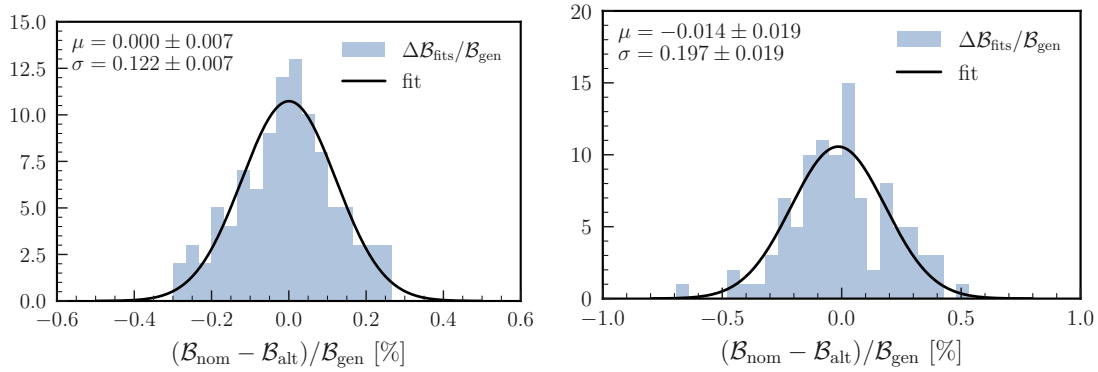


Figure 6.3.: Comparison of branching ratios for $\mathcal{B}(B^0 \to K^+\pi^-)$ (left) and $\mathcal{B}(B^0 \to \pi^+\pi^-)$ (right) obtained from pairs of fits across 100 pseudo experiments using both nominal and varied calibration parameters for the width and mean of peaking components. The distributions depict the relative differences between measured branching ratios. The width of a fitted Gaussian function, σ, provides a measure of the systematic uncertainty associated with the shape corrections.

6.2.7. Signal and Cross-Feed Modeling

The PDF shapes used for modeling the signal and cross-feed components are extracted from fits to large signal simulation samples. Since we are using analytical functions to describe these shapes, we use the uncertainties on the shape parameters, provided in the minimization step, to estimate systematic uncertainties related to modeling imperfections.

We evaluate the effect of the uncertainties on the shape parameters using ensembles of 500 pseudo experiments. Sampling from the nominal fit shapes, we generate distributions in ΔE and C' for each component in their expected proportion. We randomly fluctuate the yields sampling from a Poisson distribution around the expected value, estimated on a large generic simulation sample, and appropriately scaled to match the integrated luminosity of the LS1 dataset.

We subsequently fit these pseudo datasets twice: once using our nominal fitter and once using varied PDF shape parameters. The latter are obtained sampling from Gaussian distributions centered around the nominal shape parameters with widths corresponding to their respective uncertainties. We calculate the relative differences in measured physics parameters for each pair of fit results across the 500 pseudo experiments. The resulting distribution is fitted with a Gaussian function and we quote the width, σ, as systematic uncertainty related to the uncertainty on the PDF under consideration. We perform this study separately for each peaking component, that is, $B^0 \to K^+\pi^-$ signal, $B^0 \to \pi^+\pi^-$ signal, $B^0 \to K^+\pi^-$ cross-

feed, and $B^0 \to \pi^+\pi^-$ cross-feed. Figures 6.4–6.7 show the distributions of relative differences in branching ratios between nominal and alternative fit, for all four components. We observe no significant differences in the measurements of $\mathcal{A}_{CP}$.

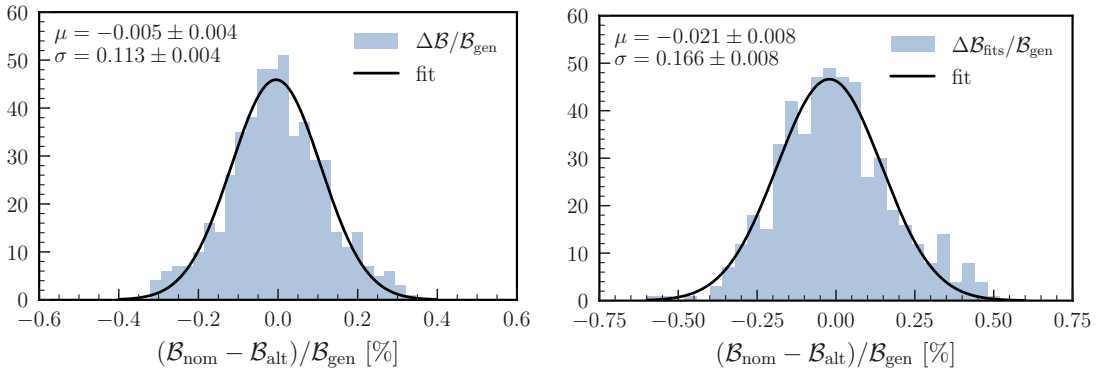


Figure 6.4.: Relative differences in $\mathcal{B}(B^0 \to K^+\pi^-)$ (left) and $\mathcal{B}(B^0 \to \pi^+\pi^-)$ (right) obtained from pairs of fits across 500 pseudo experiments using nominal and varied $B^0 \to K^+\pi^-$ signal PDF shapes. The width of a fitted Gaussian function, σ, provides a measure of the corresponding systematic.

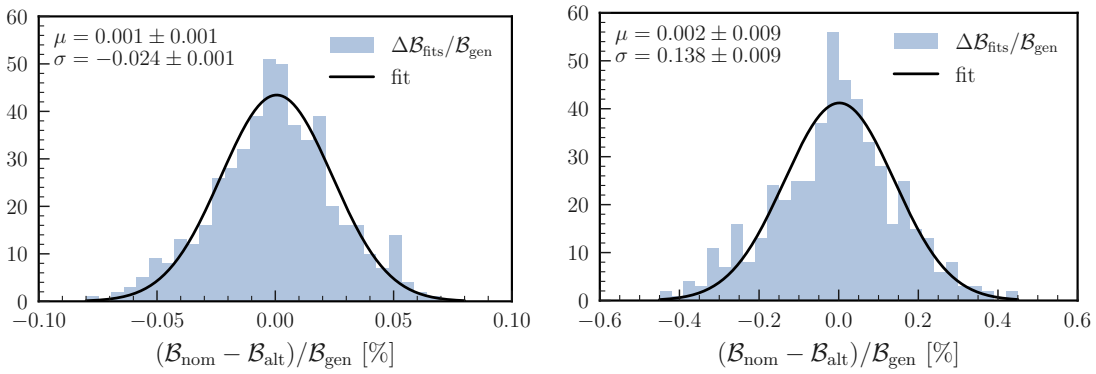


Figure 6.5.: Relative differences in $\mathcal{B}(B^0 \to K^+\pi^-)$ (left) and $\mathcal{B}(B^0 \to \pi^+\pi^-)$ (right) obtained from pairs of fits across 500 pseudo experiments using both nominal and varied $B^0 \to \pi^+\pi^-$ signal PDF shapes. The width of a fitted Gaussian function, σ, provides a measure of the corresponding systematic.

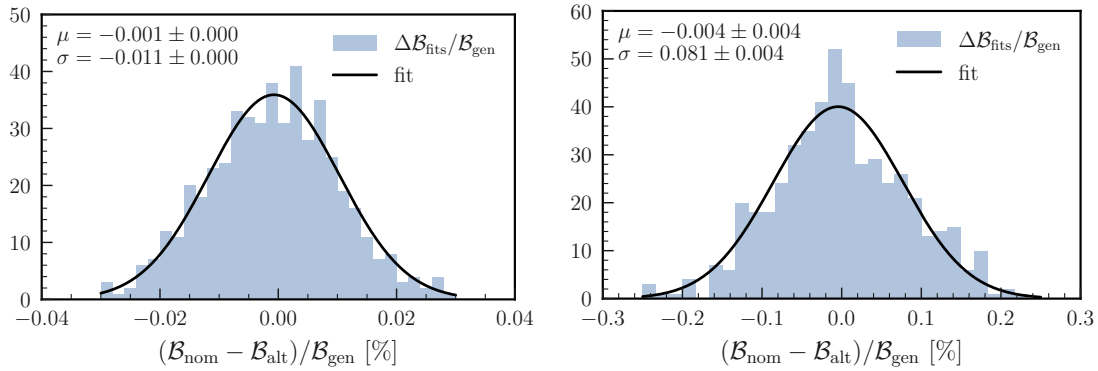


Figure 6.6.: Relative differences in $\mathcal{B}(B^0 \to K^+\pi^-)$ (left) and $\mathcal{B}(B^0 \to \pi^+\pi^-)$ (right) obtained from pairs of fits across 500 pseudo experiments using both nominal and varied $B^0 \to K^+\pi^-$ cross-feed PDF shapes. The width of a fitted Gaussian function, σ, provides a measure of the corresponding systematic.

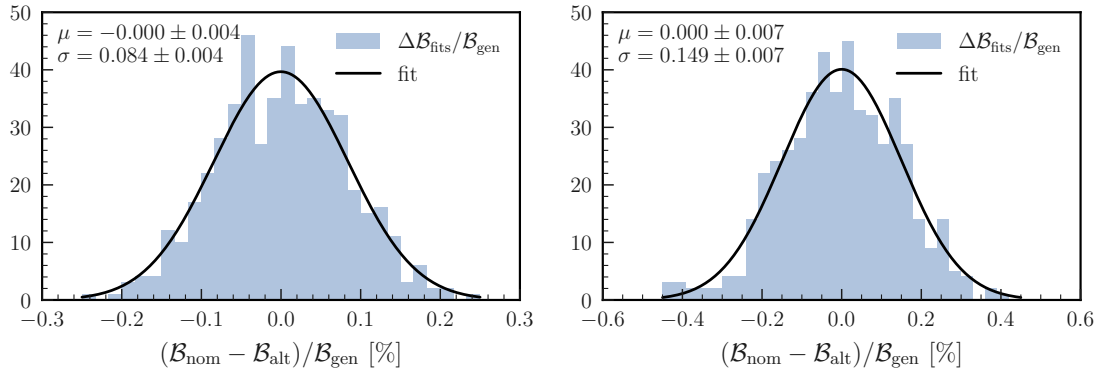


Figure 6.7.: Relative differences in $\mathcal{B}(B^0 \to K^+\pi^-)$ (left) and $\mathcal{B}(B^0 \to \pi^+\pi^-)$ (right) obtained from pairs of fits across 500 pseudo experiments using both nominal and varied $B^0 \to \pi^+\pi^-$ cross-feed PDF shapes. The width of a fitted Gaussian function, σ, provides a measure of the corresponding systematic.

6.2.8. Instrumental Asymmetry

To determine the instrumental asymmetry, $\mathcal{A}_{\text{det}}$, we measure the charge-asymmetry in $D^0 \to K^-\pi^+$ decays in the LS1 data. We expect the asymmetry to depend on the topology of the decay, which differs between the control mode and our targeted signal mode, $B^0 \to K^+\pi^-$. We assess a systematic uncertainty to account for any potential discrepancies in the measurement of $\mathcal{A}_{\text{det}}$ between these two decay modes. A dedicated study outlining the strategy in detail confirms that simulation

accurately models the dependencies of $\mathcal{A}_{\text{det}}$ [89]. We therefore measure the value of $\mathcal{A}_{\text{det}}$ in simulation for signal and control mode decays, and assign the difference, $\Delta\mathcal{A}_{\text{det}} = \mathcal{A}_{\text{det}}^{\text{sig,MC}} - \mathcal{A}_{\text{det}}^{\text{control,MC}}$, as systematic uncertainty.

6.2.9. Multiple Candidates

We perform no best candidate selection in events with more than one reconstructed candidates. To assess a systematic uncertainty associated with a possible data-simulation mismatch in candidate multiplicity, we repeat the fit to data after randomly selecting a single candidate in each event. We observe no differences in fit results for any physics parameter.

6.3. Comparison with Previous Results

We compare our findings with results of a prior Belle II analysis, which was conducted on a smaller subset of the LS1 data. Additionally, we draw comparisons with results from the final study of these decays at the Belle experiment, which utilized the full Belle dataset.

6.3.1. Previous Belle II Result

In the previous Belle II study, measurements of $\mathcal{B}(B^0 \to \pi^+\pi^-)$, $\mathcal{B}(B^0 \to K^+\pi^-)$ and $\mathcal{A}_{CP}(B^0 \to K^+\pi^-)$ were carried out on a dataset containing 68 million $B\bar{B}$ pairs [90]. Besides the substantial increase in integrated luminosity, our study introduces several improvements aimed at enhancing the overall precision.

Contrasting the previous approach, where all physics parameters were measured in separate fits, in this study, we simultaneously fit the $B^0 \to \pi^+\pi^-$ sample and both charge-separated $B \to K^\pm\pi^\mp$ samples. This enables the consideration of misidentified events, in which one of the final-state hadrons was reconstructed with an incorrect mass-hypothesis, in the computation of the branching ratios. In this manner, we attain a significant enhancement of the effective reconstruction efficiency and thereby the statistical precision of the measurement.

This improvement is augmented by our revised selection of fit variables. The previous analysis entailed fitting the distributions of the energy-difference, ΔE, and beam-constrained mass, M_{bc}, following an optimized cut on a continuum suppression classifier. To this end, the selection criteria for particle identification variables and the continuum suppression classifier were chosen to maximize a figure-of-merit in a signal-enriched region. In contrast, our current approach involves fitting the distribution of the continuum suppression classifier directly. This change enables us to significantly relax the selection requirements on the analogous classifier, which further enhances the efficiency for signal events.

We also refine the computation of the continuum suppression classifier itself. While the boosted decision tree to compute the classifier used in the previous analysis was trained on a sample composed of a wide variety of hadronic B-meson decays, we tailor our method to the task at hand. We revise the array of input variables and train the BDT specifically on simulated $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$ event samples. This modification further enhances the discriminative power of the classifier, allowing for improved signal efficiencies at given levels of background retention.

The enhanced efficiencies in the reconstruction and selection process become evident when comparing the yields per unit of integrated luminosity between the two studies. The normalized yields are tabulated in Table 6.3 for both Belle II

analyses as well as for the final Belle analysis. To account for the incorporation of cross-feed yields in the branching ratio determination, we have added the yields of correctly reconstructed and misidentified events in parentheses. This study surpasses all prior ones in terms of selection efficiencies, with an efficiency increase of 46% and 133% for $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$ events, respectively, compared to the preceding Belle II study.

Further, several improvements have been made to also reduce systematic uncertainties. In [91], the leading uncertainties were estimated to be due to imperfections in the modeling of peaking components at around 4% for both decay modes. Besides modifications to our fitting procedure, we adopt a novel approach for translating uncertainties on individual shape parameters into uncertainties associated with modeling inaccuracies. We estimate the relative systematic uncertainty of this effect to be as low as 0.1% for $\mathcal{B}(B^0 \to K^+\pi^-)$ and 0.3% for $\mathcal{B}(B^0 \to \pi^+\pi^-)$.

Similar improvements have been made for nearly all other sources, with the leading uncertainties in the branching ratio measurements now stemming from uncertainties in the determination of the number of neutral B mesons in the sample.

Table 6.3.: Comparison of signal yields per fb^{-1} of integrated luminosity with previous analyses. The values in parentheses are obtained when including the cross-feed yields in the calculation.

Decay mode	Belle II (2023)	Belle II (2021) [90]	Belle (2012) [91]
$B^0 \to K^+\pi^-$	10.7 (13.1)	9.0	10.6
$B^0 \to \pi^+\pi^-$	3.3 (4.2)	1.8	3.0

The measurements presented in this thesis share a significant fraction of the analyzed data with the previous Belle II study. As an additional validation measure, we apply our fit to data from the same runs, corresponding to an integrated luminosity of 62.8 fb^{-1}. Beyond statistical fluctuations, we anticipate minor differences due to different data processings with revised detector calibrations and updated reconstruction software. We apply the same corrections (PID, instrumental asymmetry, PDF shapes) as used in the nominal analysis, assuming negligible run-dependence of these parameters. Figures 6.8 and 6.9 depict the projections onto the two fit dimensions for all three samples, along with the corresponding normalized residuals. We list the fit results obtained on the partial dataset in Table 6.4, alongside the previous results reported in [91]. We observe good agreement between the measurements in all parameters within their uncertainties, with increased statistical precision in the updated analysis.

Table 6.4.: Comparison of fit results with previous measurements reported in [91], obtained on the partial dataset corresponding to an integrated luminosity of 62.8 fb^{-1}. We list only the statistical uncertainties.

Parameter	Belle II (2023)	Belle II (2021) [90]
$\mathcal{B}(B^0 \to K^+\pi^-)$	$(20.9 \pm 0.9) \times 10^{-6}$	$(19.9 \pm 1.0) \times 10^{-6}$
$\mathcal{B}(B^0 \to \pi^+\pi^-)$	$(6.1 \pm 0.5) \times 10^{-6}$	$(5.9 \pm 0.7) \times 10^{-6}$
$\mathcal{A}_{CP}(B^0 \to K^+\pi^-)$	-0.08 ± 0.04	-0.16 ± 0.05

6.3.2. Final Belle Result

We additionally compare our measurement with the final result reported by the Belle experiment [91]. Using a dataset containing 772 million $B\bar{B}$ events, the referenced study follows a similar analysis strategy as presented in this thesis, extracting the physics parameters from a simultaneous fit to the $B^0 \to \pi^+\pi^-$ and $B \to K^\pm\pi^\mp$ samples.

We list the results in Table 6.5 along with our measurements for better comparison. We observe good agreement for $\mathcal{B}(B^0 \to K^+\pi^-)$ and $\mathcal{A}_{CP}(B^0 \to K^+\pi^-)$. For $\mathcal{B}(B^0 \to \pi^+\pi^-)$, we compute a discrepancy of 2.0 standard deviations. However, this deviation is put into perspective when compared to an alternative result from Belle reported in [39], which measures a central value of $\mathcal{B}(B^0 \to \pi^+\pi^-) = 5.63 \times 10^{-6}$ on the same dataset. We also note the most recent result from the BaBar experiment, which reports 5.5×10^{-6} [92].

Next to the central values, it is noteworthy that our measurements exhibit comparable statistical significance to those of the Belle study, despite using a significantly smaller dataset ($\sim 50\%$). We can estimate the projected statistical uncertainties of the Belle analysis, scaled to an integrated luminosity of the LS1 data. The relative increase in precision measures 24% for $\mathcal{B}(B^0 \to \pi^+\pi^-)$, 23% for $\mathcal{B}(B^0 \to K^+\pi^-)$, and 5% for $\mathcal{A}_{CP}(B^0 \to K^+\pi^-)$.

Regarding the similar signal yields per unit of integrated luminosity, listed in Table 6.3), we attribute much of the improvement in precision to the advanced continuum suppression technique. The latter allows for a significant reduction of continuum background events at high signal efficiencies while still providing excellent discrimination power in the fit to the classifier variable distribution.

Table 6.5.: Comparison of results with previous measurements, reported by the Belle [91] and BaBar [92, 93] collaborations. Despite using a substantially smaller dataset, we achieve comparable statistical precision in the branching ratios measurements.

Parameter	Belle II (2023)	Belle (2012)	BaBar (2007, 2013)
$\mathcal{B}(\pi^+\pi^-)$ $[\times 10^{-6}]$	$(5.83 \pm 0.22 \pm 0.17)$	$(5.04 \pm 0.21 \pm 0.18)$	$(5.5 \pm 0.4 \pm 0.3)$
$\mathcal{B}(K^+\pi^-)$ $[\times 10^{-6}]$	$(20.67 \pm 0.37 \pm 0.62)$	$(20.00 \pm 0.34 \pm 0.60)$	$(19.1 \pm 0.6 \pm 0.6)$
$\mathcal{A}_{CP}(K^+\pi^-)$	$-0.072 \pm 0.019 \pm 0.007$	$-0.069 \pm 0.014 \pm 0.007$	$-0.107 \pm 0.016^{+0.006}_{-0.004}$

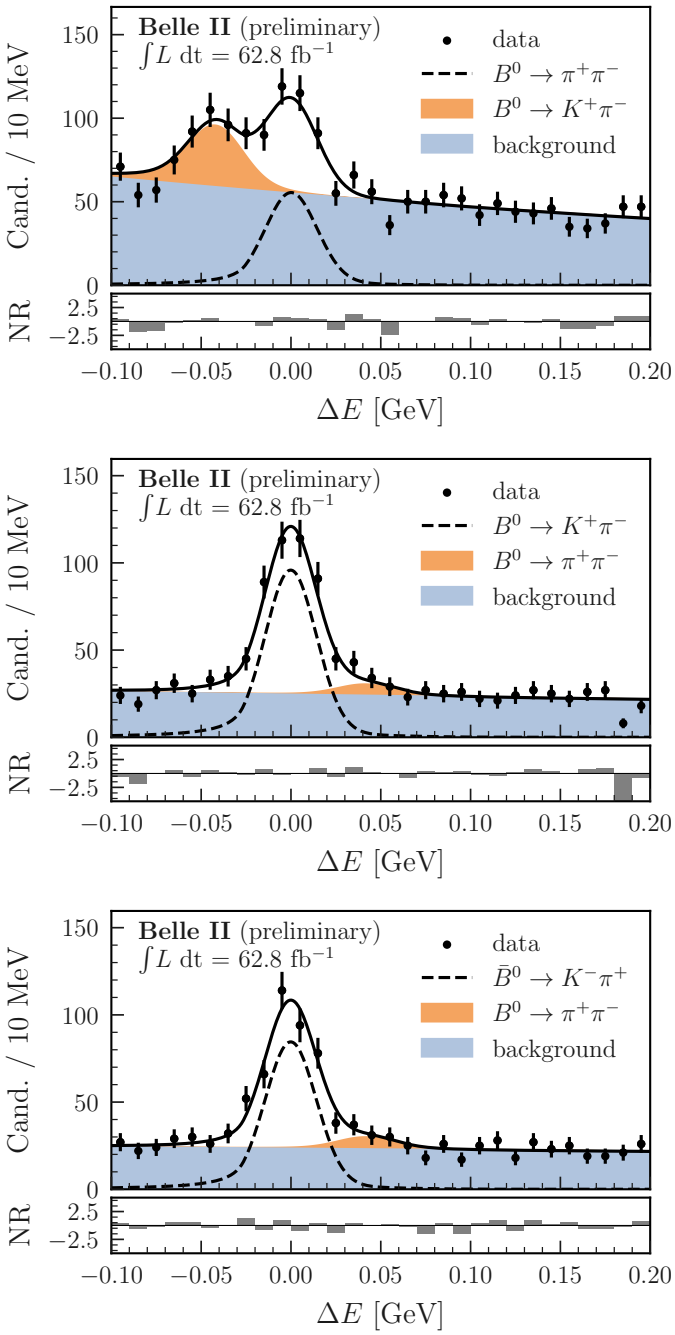


Figure 6.8.: Distributions in ΔE for $B^0 \to \pi^+\pi^-$ (top), $B^0 \to K^+\pi^-$ (center), and $\bar{B}^0 \to K^-\pi^+$ (bottom) candidates reconstructed in the partial dataset of 62.8 fb^{-1}, previously analyzed in [90]. The projections of an extended unbinned maximum-likelihood fit are overlaid. Normalized residuals are shown below.

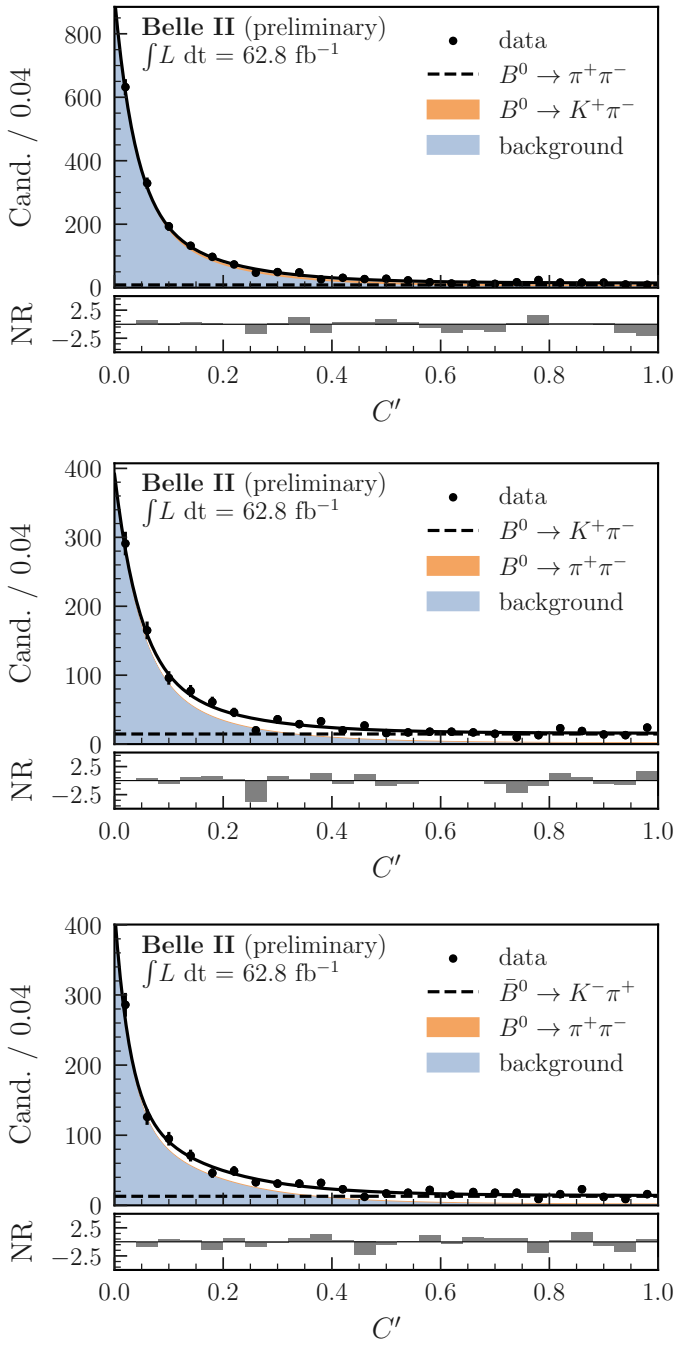


Figure 6.9.: Distributions in C' for $B^0 \to \pi^+\pi^-$ (top), $B^0 \to K^+\pi^-$ (center), and $\bar{B}^0 \to K^-\pi^+$ (bottom) candidates reconstructed in the partial dataset of 62.8 fb^{-1}, previously analyzed in [90]. The projections of an extended unbinned maximum-likelihood fit are overlaid. Normalized residuals are shown below.

6.4. Determination of $I_{K\pi}$

In the context of probing the Standard Model, charmless hadronic B-meson decays are particularly intriguing due to their substantial contributions from loop amplitudes. Consequently, the measurements of branching fractions and CP asymmetries in these decays offer valuable insight into potential non-Standard Model physics. However, interpreting these measurements is not straightforward, primarily due to considerable uncertainties inherent in the theoretical predictions. These uncertainties stem from non-factorizable hadronic amplitudes, which impede the application of perturbation theory.

The impact of these uncertainties can be mitigated by applying appropriate combinations of measurements from decays that are related by flavor symmetries. By exploiting isospin symmetry, sum rules can be constructed. These sum rules are linear combinations of branching fractions and CP asymmetries that help to reduce the influence of theoretical and experimental uncertainties.

In particular, for the set of $B \to K\pi$ decays, the following sum rule can be constructed[1]:

$$I_{K\pi} = \mathcal{A}_{CP}^{K^+\pi^-} + \mathcal{A}_{CP}^{K^0\pi^+} \frac{\mathcal{B}_{K^0\pi^+}}{\mathcal{B}_{K^+\pi^-}} \frac{\tau_0}{\tau_+} - 2\mathcal{A}_{CP}^{K^+\pi^0} \frac{\mathcal{B}_{K^+\pi^0}}{\mathcal{B}_{K^+\pi^-}} \frac{\tau_0}{\tau_+} - 2\mathcal{A}_{CP}^{K^0\pi^0} \frac{\mathcal{B}_{K^0\pi^0}}{\mathcal{B}_{K^+\pi^-}} \,. \tag{6.3}$$

Here, $\mathcal{A}_{CP}^{K\pi}$ and $\mathcal{B}_{K\pi}$ are the direct CP asymmetry and the CP-averaged branching fraction of the respective $B \to K\pi$ decay, and τ_0 and τ_+ are the lifetimes of the neutral and charged B mesons. The sum rule offers a robust test of the SM, which predicts $I_{K\pi} \approx 0$ with a precision of a few percent in the $SU(3)$ and heavy quark limits [47, 94, 95, 96].

To compute the value of the isospin sum rule we combine our measurements with those of the other decay modes determined in the LS1 data [97]. In order to allow for a consistent combination, the analyses were partly developed in collaboration, with similar selection requirements, fit strategies, and treatments of systematic errors. Due to this consistency, common systematic uncertainties cancel out in the ratios of branching fractions, increasing the precision of the evaluation. Specifically, these uncertainties relate to tracking efficiencies, the number of produced B mesons, and the production ratio for neutral and charged B mesons in $\Upsilon(4S)$ decays, $f^{+-/00}$. We also considered the anti-correlation of f^{00} and f^{+-} uncertainties for the ratio between B^+ and B^0 decays.

The resulting ratios of branching fractions and CP-violation parameters are listed in Table 6.6. We use the current averages reported by the PDG for the ratio

[1] For the derivation, see Section 2.3.3.

$\tau_{B^0}/\tau_{B^+} = 0.9273 \pm 0.0033$ [20]. We obtain a value of

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

in agreement with the SM expectation of zero. This result provides a useful constraint to new physics models with uncertainties comparable with those of the corresponding theoretical calculations. The precision is limited by sample size and is competitive with the determination using averages of measurements by the Belle, BaBar, and LHCb collaborations [20].

Table 6.6.: Ratios of branching fractions and CP-violation parameters, used as input for the determination of $I_{K\pi}$ at Belle II in the LS1 dataset [97].

Parameter	Value
$\mathcal{B}_{K^0\pi^+}/\mathcal{B}_{K^+\pi^-}$	$1.180 \pm 0.040 \pm 0.027$
$\mathcal{B}_{K^+\pi^0}/\mathcal{B}_{K^+\pi^-}$	$0.687 \pm 0.022 \pm 0.040$
$\mathcal{B}_{K^0\pi^0}/\mathcal{B}_{K^+\pi^-}$	$0.508 \pm 0.031 \pm 0.030$
$\mathcal{A}_{CP}^{K^+\pi^-}$	$-0.072 \pm 0.019 \pm 0.007$
$\mathcal{A}_{CP}^{K^+\pi^0}$	$0.013 \pm 0.027 \pm 0.005$
$\mathcal{A}_{CP}^{K^0\pi^+}$	$0.046 \pm 0.029 \pm 0.007$
$\mathcal{A}_{CP}^{K^0\pi^0}$	$-0.01 \pm 0.12 \pm 0.05$

7. Conclusion

The need for additional sources of CP violation, beyond the amount introduced by the CKM mixing matrix, is only one of many indications that call for physics beyond the Standard Model. The Belle II experiment is one of the leading experiments at the intensity frontier and searches for new physics contributions with precise measurements of quark-flavor dynamics in B-meson decays.

This thesis presents the measurements of branching ratios in $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ decays and the direct CP-violation parameter in $B^0 \to K^+\pi^-$ decays. In both decay modes, the tree diagram contribution to the transition amplitude is strongly CKM-suppressed, rendering the QCD penguin loop diagram to be of leading order. As a consequence, these rare decays are particularly sensitive to non-SM physics.

The analyzed data contains $(387 \pm 6) \times 10^6$ $B\bar{B}$-meson pairs, corresponding to an integrated luminosity of $(362 \pm 2)\,\mathrm{fb}^{-1}$ collected with the Belle II detector, using electron-positron collisions at the $\Upsilon(4S)$ resonance provided by the SuperKEKB collider. We obtain the following branching ratios and CP-violation parameter:

$$\begin{aligned}
\mathcal{B}(B \to \pi^+\pi^-) &= (5.83 \pm 0.22 \pm 0.17) \times 10^{-6}\ , \\
\mathcal{B}(B \to K^+\pi^-) &= (20.67 \pm 0.37 \pm 0.62) \times 10^{-6}\ , \\
\mathcal{A}_{CP}(B \to K^+\pi^-) &= -0.072 \pm 0.019 \pm 0.007\ .
\end{aligned}$$

The results agree with current world averages at a level of precision comparable to the current best results, despite using a substantially smaller dataset. The measurement of $\mathcal{B}(B \to \pi^+\pi^-)$ presents the most precise determination by a single experiment to date.

The measurements of $\mathcal{B}(B \to K^+\pi^-)$ and $\mathcal{A}_{CP}(B \to K^+\pi^-)$ are combined with those of related $B \to K\pi$ transitions, obtained on the same dataset, in an isospin-based sum rule that serves as a stringent test of the SM. We obtain a value of

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

in agreement with the Standard Model expectation of zero. The precision is competitive with that resulting from the average of measurements by the Belle, BaBar, and LHCb collaborations [20].

The small branching ratios pose major challenges in these measurements. We use a multivariate maximum-likelihood fit to extract signal yields from background-dominated samples. In this study, for the first time at Belle II, we simultaneously fit the $B^0 \to \pi^+\pi^-$ sample and both charge-separated $B \to K^\pm\pi^\mp$ samples. This enables the consideration of misidentified events, in which one of the final-state hadrons was reconstructed with an incorrect mass-hypothesis, in the computation of the branching ratios. In this manner, we attain a significant enhancement of the effective reconstruction efficiency compared to previous measurements.

A rather inclusive choice of selection requirements is chosen to retain a high efficiency for signal events, at the cost of large background yields. After the baseline selection, the majority of candidates are misreconstructed continuum events consisting of lighter quark pairs, $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$). To suppress this background, we develop a multivariate machine-learning algorithm which exploits the differences in event topologies between B-meson decays and continuum events. We train a stochastic gradient-boosted decision tree on the distributions of 35 discriminating variables in large samples of simulated signal and continuum events. The resulting classifier variable exhibits excellent distinguishing power and allows to remove 90% of background events at signal efficiencies of more than 95%. This improvement is augmented by including the continuum suppression classifier as additional fit dimension, which enables us to significantly relax the selection requirements, further enhancing the signal efficiency.

This study surpasses all prior ones in terms of selection efficiencies, with an increase of 46% and 133% for $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$ events, respectively, compared to the preceding Belle II study. We attribute much of the improvement to the efficient continuum background suppression. As machine learning techniques continue to evolve, we anticipate further enhancements in the near future. First studies exploring the potential of employing deep learning methods with low-level track variables for discriminating between signal and background events show promising results [98].

Even with the comparatively small data sample used in this study, the systematic uncertainty already limits the precision in the measurement of $\mathcal{B}(B^0 \to K^+\pi^-)$. As data collection is set to resume by the end of 2023, the other measurements will soon follow suit. We undertake an array of validation and alignment procedures to minimize analysis-specific systematic uncertainties.

Monte Carlo simulation samples are used in many steps of the analysis, such as continuum background suppression, model building or determination of selection efficiencies. Before unblinding data in the targeted signal mode analysis, large efforts are spent to mitigate the effects of potential data-MC discrepancies using abundant control mode decays. We reconstruct $B^+ \to \bar{D}^0[\to K^+\pi^-]\pi^+$ decays to validate our

reconstruction procedures and extract correction parameters to accurately align the MC simulation with recorded data. We further validate the parametric modeling of signal and background components on a number of MC ensemble tests and estimate the impact of potential imperfections on the physics observables.

The leading systematic uncertainty in the branching ratio measurements is estimated to be the uncertainty on the number of neutral B mesons in the sample. The latter is determined in external measurements of $N_{B\bar{B}}$ and f^{00} in dedicated control sample studies. With more data and refined analyses tools, we anticipate crucial improvements of these measurements in the future.

The accumulation of more data will also play a decisive role in further constraining the angle α/ϕ_2 of the CKM triangle using time-dependant measurements of mixing-induced CP violation in $B \to \pi\pi$ decays. With the excellent particle identification capabilities of Belle II, showcased in this study, and the state-of-the-art pixel vertex detector, enabling extraordinary vertexing resolution, Belle II will be able to measure all three $B \to \pi\pi$ transitions within a consistent framework at unprecedented precision. Though LHCb is expected to soon provide the leading measurements of CP asymmetries in fully-charged final-state transitions, due to the unmatched B-meson production yields offered by the LHC, the complex event topologies created in proton-proton collisions render the reconstruction of neutral pions challenging. Likewise, Belle II's measurements of the full set of $B \to K\pi$ transitions will be vital to achieve the precision needed for any conclusive statements about potential deviations from the SM expectation in the determination of $I_{K\pi}$. Currently, the uncertainties are strongly dominated by the measurement of $B^0 \to K^0\pi^0$ decays, to which Belle II has unique access.

With less than 1% of the projected integrated luminosity recorded so far, the search for new physics beyond the Standard Model at Belle II has only just begun. So far, the SM consistently holds up to all experimental tests, including the one presented in this study. It is left for future measurements to show whether new phenomena will finally emerge and lead the way to an improved understanding of particle physics.

A. Appendix

A.1. Feynman Diagrams

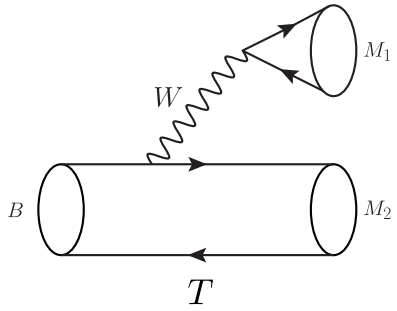
Figure A.1 shows the dominant Feynman diagrams relevant for hadronic decays of B mesons to pairs of light pseudoscalar mesons, M_1 and M_2.

A.2. Shape Parameters

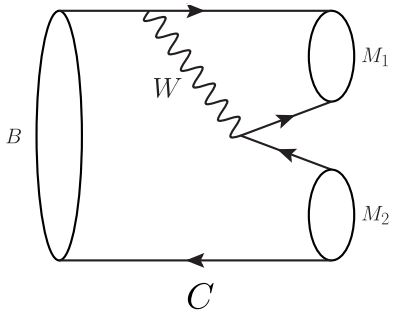
Table A.1 lists all shape parameters of the parametric models used to describe the different components in $B \to hh'$ samples. The definitions of the parameters can be found in Section 4.3.1. The values are obtained in fits to abundant simulation samples. For signal and cross-feed components, signal-only MC samples are used. The shape parameters for background components are determined in fits to generic MC samples, but left floating in the fit to data.

A.3. Correlations of Fit Parameters

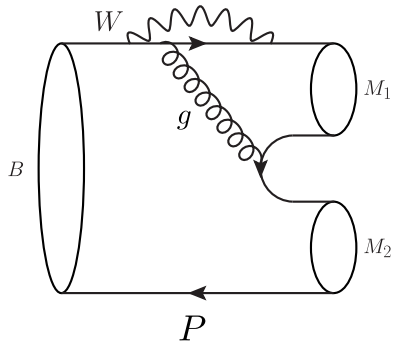
Table A.2 displays the correlation matrix for all floating parameters in the extended unbinned maximum-likelihood fit to $B \to hh'$ candidates, reconstructed in the LS1 dataset.



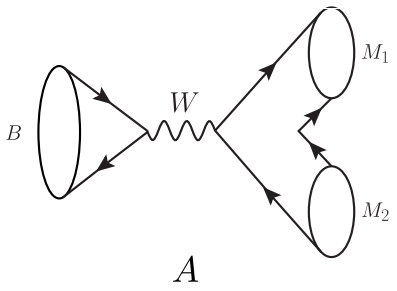
(a) Color-allowed tree diagram



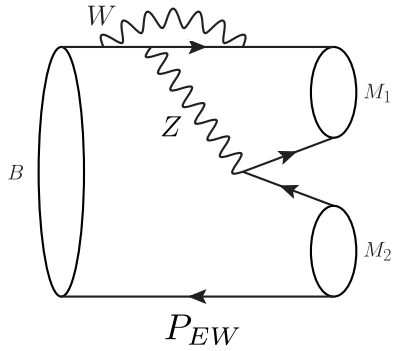
(b) Color-suppressed tree diagram



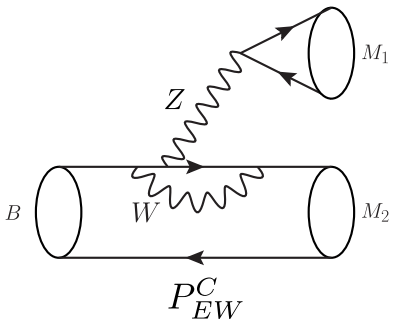
(c) QCD penguin diagram



(d) Annihilation diagram



(e) Color-allowed electroweak penguin diagram



(f) Color-suppressed electroweak penguin diagram

Figure A.1.: Feynman diagrams describing decays of B mesons to pairs of light pseudoscalar mesons.

Table A.1.: Shape parameters of parametric models obtained by fits to abundant simulated samples. The definitions of the parameters can be found in Section 4.3.1.

Component	Parameter	Value	Uncertainty
$B^0 \to \pi^+\pi^-$ signal	pipi_dE_alpha	1.7463	0.0109
	pipi_dE_frac	0.0435	0.0018
	pipi_dE_mean1	0.0159	0.0012
	pipi_dE_mean2	0.0003435	0.000016
	pipi_dE_n	1.4419	0.0238
	pipi_dE_sigma1	0.0322	0.0004
	pipi_dE_sigma2	0.01293	0.00002
$B^0 \to \pi^+\pi^-$ cross-feed	pipi_TCS_pb_a	-0.25595	0.0277
	pipi_TCS_pb_b	0.1184	0.0274
	pipi_dE_pb_frac	0.2492	0.0090
	pipi_dE_pb_mean1	-0.0530	0.0005
	pipi_dE_pb_mean2	-0.0416	0.0001
	pipi_dE_pb_sigma1	0.0282	0.0005
	pipi_dE_pb_sigma2	0.01326	0.00008
$B^0 \to \pi^+\pi^-$ background	pipi_TCS_qq_a	-8.2048	0.1064
	pipi_TCS_qq_b	-30.7866	0.9442
	pipi_TCS_qq_frac	0.6218	0.0132
	pipi_dE_qq_a	-1.0991	0.0583
$B^0 \to K^\pm\pi^\mp$ signal	kpi_dE_alpha	1.8306	0.0118
	kpi_dE_frac	0.0455	0.0017
	kpi_dE_mean1	0.0159	0.0011
	kpi_dE_mean2	0.000091	0.000016
	kpi_dE_n	1.4555	0.0259
	kpi_dE_sigma1	0.0320	0.0004
	kpi_dE_sigma2	0.0127	0.00002
$B^0 \to K^\pm\pi^\mp$ cross-feed	kpi_TCS_pb_a	-0.3040	0.0210
	kpi_TCS_pb_b	0.1376	0.0208
	kpi_dE_pb_frac	0.2328	0.0042
	kpi_dE_pb_mean1	0.0386	0.0002
	kpi_dE_pb_mean2	0.0420	0.00004
	kpi_dE_pb_sigma1	0.0369	0.0005
	kpi_dE_pb_sigma2	0.01375	0.00006
$B^0 \to K^+\pi^-$ background	kpi_TCS_qq_a_plus	-7.4649	0.1654
	kpi_TCS_qq_b_plus	-24.8956	1.3867
	kpi_TCS_qq_frac_p	0.6426	0.0254
	kpi_dE_qq_a_plus	-1.1940	0.0905
$\bar{B}^0 \to K^-\pi^+$ background	kpi_TCS_qq_a_minus	-7.5630	0.1636
	kpi_TCS_qq_b_minus	-26.4193	1.4656
	kpi_TCS_qq_frac_m	0.6410	0.0241
	kpi_dE_qq_a_minus	-1.1734	0.0902

Table A.2.: Correlations between fit parameters obtained in a fit to $B^0 \to h^+h^-$ candidates reconstructed in an MC sample corresponding to an integrated luminosity of $1\,\mathrm{ab}^{-1}$.

	$\mathcal{A}_{K\pi}$	$\mathcal{B}_{K\pi}$	C' $\mathrm{bkg}^-_{K\pi}$ a	C' $\mathrm{bkg}^+_{K\pi}$ a	C' $\mathrm{bkg}^-_{K\pi}$ b	C' $\mathrm{bkg}^+_{K\pi}$ b	C' $\mathrm{bkg}^-_{K\pi}w$	C' $\mathrm{bkg}^+_{K\pi}w$	ΔE $\mathrm{bkg}^-_{K\pi}$ c
$\mathcal{A}_{K\pi}$	1.00	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.14
$\mathcal{B}_{K\pi}$	-0.02	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14
C' $\mathrm{bkg}^-_{K\pi}$ a	0.00	0.00	1.00	0.00	-0.10	0.00	0.33	0.00	0.00
C' $\mathrm{bkg}^+_{K\pi}$ a	0.00	0.00	0.00	1.00	0.00	-0.11	0.00	0.31	0.00
C' $\mathrm{bkg}^-_{K\pi}$ b	0.00	0.00	-0.10	0.00	1.00	0.00	-0.30	0.00	0.00
C' $\mathrm{bkg}^+_{K\pi}$ b	0.00	0.00	0.00	-0.11	0.00	1.00	0.00	-0.34	0.00
C' $\mathrm{bkg}^-_{K\pi}w$	0.00	0.00	0.33	0.00	-0.30	0.00	1.00	0.00	0.00
C' $\mathrm{bkg}^+_{K\pi}w$	0.00	0.00	0.00	0.31	0.00	-0.34	0.00	1.00	0.00
ΔE $\mathrm{bkg}^-_{K\pi}$ c	0.14	0.14	0.00	0.00	0.00	0.00	0.00	0.00	1.00
ΔE $\mathrm{bkg}^+_{K\pi}$ c	-0.14	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00
yield $\mathrm{bkg}^-_{K\pi}$	-0.23	-0.24	0.00	0.00	0.00	0.00	0.00	0.00	-0.14
yield $\mathrm{bkg}^-_{K\pi}$	0.25	-0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\mathcal{B}_{\pi\pi}$	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.01
C' $\mathrm{bkg}_{\pi\pi}a$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C' $\mathrm{bkg}_{\pi\pi}b$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C' $\mathrm{bkg}_{\pi\pi}$ w	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ΔE $\mathrm{bkg}_{\pi\pi}$ c	0.00	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.02
yield $\mathrm{bkg}_{\pi\pi}$	0.00	-0.12	0.00	0.00	0.00	0.00	0.00	0.00	-0.02

	ΔE $\mathrm{bkg}^+_{K\pi}$ c	yield $\mathrm{bkg}^-_{K\pi}$	yield $\mathrm{bkg}^-_{K\pi}$	$\mathcal{B}_{\pi\pi}$	C' $\mathrm{bkg}_{\pi\pi}$ a	C' $\mathrm{bkg}_{\pi\pi}$ b	C' $\mathrm{bkg}_{\pi\pi}w$	ΔE $\mathrm{bkg}_{\pi\pi}$ c	yield $\mathrm{bkg}_{\pi\pi}$
$\mathcal{A}_{K\pi}$	-0.14	-0.23	0.25	0.00	0.00	0.00	0.00	0.00	0.00
$\mathcal{B}_{K\pi}$	0.14	-0.24	-0.24	0.03	0.00	0.00	0.00	0.12	-0.12
C' $\mathrm{bkg}^-_{K\pi}$ a	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C' $\mathrm{bkg}^+_{K\pi}$ a	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C' $\mathrm{bkg}^-_{K\pi}$ b	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C' $\mathrm{bkg}^+_{K\pi}$ b	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C' $\mathrm{bkg}^-_{K\pi}w$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C' $\mathrm{bkg}^+_{K\pi}w$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ΔE $\mathrm{bkg}^-_{K\pi}$ c	0.00	-0.14	0.00	0.01	0.00	0.00	0.00	0.02	-0.02
ΔE $\mathrm{bkg}^+_{K\pi}$ c	1.00	0.00	-0.13	0.01	0.00	0.00	0.00	0.02	-0.02
yield $\mathrm{bkg}^-_{K\pi}$	0.00	1.00	0.01	-0.09	0.00	0.00	0.00	-0.05	0.06
yield $\mathrm{bkg}^-_{K\pi}$	-0.13	0.01	1.00	-0.09	0.00	0.00	0.00	-0.05	0.06
$\mathcal{B}_{\pi\pi}$	0.01	-0.09	-0.09	1.00	0.00	0.00	0.00	0.24	-0.39
C' $\mathrm{bkg}_{\pi\pi}$ a	0.00	0.00	0.00	0.00	1.00	-0.08	0.45	0.00	0.00
C' $\mathrm{bkg}_{\pi\pi}$ b	0.00	0.00	0.00	0.00	-0.08	1.00	-0.17	0.00	0.00
C' $\mathrm{bkg}_{\pi\pi}$ w	0.00	0.00	0.00	0.00	0.45	-0.17	1.00	0.00	0.00
ΔE $\mathrm{bkg}_{\pi\pi}$ c	0.02	-0.05	-0.05	0.24	0.00	0.00	0.00	1.00	-0.17
yield $\mathrm{bkg}_{\pi\pi}$	0.02	0.06	0.06	-0.39	0.00	0.00	0.00	-0.17	1.00

List of Figures

List of Tables

Acronyms

basf2 Belle II Analysis Framework.

BDT Boosted Decision Tree.

CF Cross-Feed.

CKM Cabibbo-Kobayashi-Maskawa.

MC Monte Carlo.

NR Normalized Residuals.

PDF Probability Density Function.

PXD Pixel Vertex Detector.

SM Standard Model.

SVD Silicon Vertex Detector.

Bibliography

[1] A. D. Sakharov. "Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe". In: *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967), pp. 32–35. DOI: 10.1070/PU1991v034n05ABEH002497.

[2] M. B. Gavela et al. "Standard Model CP-Violation and Baryon Asymmetry". In: *Modern Physics Letters A* 09.09 (1994), pp. 795–809. DOI: 10.1142/S0217732394000629.

[3] J. Charles et al. "Future Sensitivity to New Physics in B_d, B_s, and K Mixings". In: *Physical Review D* 89.3 (2014). DOI: 10.1103/PhysRevD.89.033016.

[4] M. Thomson. *Modern Particle Physics*. Cambridge University Press, 2013. ISBN: 978-1-107-03426-6.

[5] D. J. Griffiths. *Introduction to Elementary Particles*. Second, Revised Edition. Wiley-VCH, 2008. ISBN: 978-3-527-40601-2.

[6] M. Robinson et al. "A Simple Introduction to Particle Physics". In: *arXiv (preprint)* (2008). DOI: 10.48550/arXiv.0810.3328.

[7] V. Devanathan. "How Do the Particles Acquire Mass?" In: *Journal of Chennai Academy of Sciences* 1 (2019), pp. 1–20. URL: https://www.tnasc.com/wp-content/uploads/2019/11/JOUR1-1.pdf.

[8] Wikipedia. *Standard Model*. Wikipedia, The Free Encyclopedia. URL: https://en.wikipedia.org/wiki/Standard_Model (visited on 07/27/2023).

[9] E. Noether. "Invariante Variationsprobleme". In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1918 (1918), pp. 235–257. URL: http://eudml.org/doc/59024.

[10] H. D. Politzer. "Reliable Perturbative Results for Strong Interactions?" In: *Physical Review Letters* 30.26 (1973), pp. 1346–1349. DOI: 10.1103/PhysRevLett.30.1346.

[11] D. J. Gross and F. Wilczek. "Ultraviolet Behavior of Non-Abelian Gauge Theories". In: *Physical Review Letters* 30.26 (1973), pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343.

[12] S. L. Glashow. "Partial-Symmetries of Weak Interactions". In: *Nuclear Physics* 22.4 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[13] S. Weinberg. "A Model of Leptons". In: *Physical Review Letters* 19.21 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[14] A. Pich. "The Standard Model of Electroweak Interactions". In: *arXiv (preprint)* (2012). DOI: 10.48550/arXiv.1201.0537.

[15] J. Horejsi. "Fundamentals of Electroweak Theory". In: *arXiv (preprint)* (2022). DOI: 10.48550/arXiv.2210.04526. arXiv: 2210.04526.

[16] ATLAS Collaboration. "Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020.

[17] CMS Collaboration. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021.

[18] H. W. Hamber. *Quantum Gravitation: The Feynman Path Integral Approach*. Berlin: Springer, 2009. ISBN: 978-3-540-85292-6.

[19] S. Dimopoulos, S. Raby, and F. Wilczek. "Supersymmetry and the Scale of Unification". In: *Physical Review D* 24.6 (1981), pp. 1681–1683. DOI: 10.1103/PhysRevD.24.1681.

[20] R. L. Workman et al. "Review of Particle Physics". In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[21] I. J. R. Aitchison. *Supersymmetry in Particle Physics: An Elementary Introduction*. Cambridge, New York: Cambridge University Press, 2007. ISBN: 978-0-521-88023-7.

[22] V. Trimble. "Existence and Nature of Dark Matter in the Universe". In: *Annual Review of Astronomy and Astrophysics* 25.1 (1987), pp. 425–472. DOI: 10.1146/annurev.aa.25.090187.002233.

[23] J. H. Christenson et al. "Evidence for the 2π Decay of the K_2^0 Meson". In: *Physical Review Letters* 13.4 (1964), pp. 138–140. DOI: 10.1103/PhysRevLett.13.138.

[24] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay". In: *Physical Review* 105.4 (1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413.

[25] T. D. Lee and C. N. Yang. "Question of Parity Conservation in Weak Interactions". In: *Physical Review* 104.1 (1956), pp. 254–258. DOI: 10.1103/PhysRev.104.254.

[26] Belle Collaboration. "Observation of Large CP Violation in the Neutral B Meson System". In: *Physical Review Letters* 87.9 (2001), p. 091802. DOI: 10.1103/PhysRevLett.87.091802.

[27] BABAR Collaboration. "Observation of CP Violation in the B^0 Meson System". In: *Physical Review Letters* 87.9 (2001), p. 091801. DOI: 10.1103/PhysRevLett.87.091801.

[28] LHCb Collaboration. "Observation of CP Violation in Charm Decays". In: *Physical Review Letters* 122.21 (2019), p. 211803. DOI: 10.1103/PhysRevLett.122.211803.

[29] W. Pauli. "Exclusion Principle, Lorentz Group and Reflection of Space-Time and Charge". In: *Wolfgang Pauli: Das Gewissen der Physik*. Wiesbaden: Vieweg+Teubner Verlag, 1988, pp. 459–479. DOI: 10.1007/978-3-322-90270-2_41.

[30] CPLEAR Collaboration. "First Direct Observation of Time-Reversal Non-Invariance in the Neutral-Kaon System". In: *Physics Letters B* 444.1 (1998), pp. 43–51. DOI: 10.1016/S0370-2693(98)01356-2.

[31] M. Kobayashi and T. Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". In: *Progress of Theoretical Physics* 49.2 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652.

[32] N. Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Physical Review Letters* 10.12 (1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531.

[33] L. Wolfenstein. "Parametrization of the Kobayashi-Maskawa Matrix". In: *Physical Review Letters* 51.21 (1983), pp. 1945–1947. DOI: 10.1103/PhysRevLett.51.1945.

[34] U. Nierste. "Three Lectures on Meson Mixing and CKM Phenomenology". In: *arXiv (preprint)* (2009). URL: http://arxiv.org/abs/0904.1869.

[35] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". In: *Physical Review D* 2.7 (1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285.

[36] M. Beneke, G. Buchalla, and I. Dunietz. "Mixing-Induced CP Asymmetries in Inclusive B Decays". In: *Physics Letters B* 393.1-2 (1997), pp. 132–142. DOI: 10.1016/S0370-2693(96)01648-6.

[37] R. Fleischer. "B Physics and CP Violation". In: *Heavy Quark Physics*. Vol. 647. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 42–77. DOI: 10.1007/978-3-540-40975-5_2.

[38] M. Gronau and D. London. "Isospin Analysis of CP Asymmetries in B Decays". In: *Physical Review Letters* 65.27 (1990), pp. 3381–3384. DOI: 10.1103/PhysRevLett.65.3381.

[39] K. A. Prothmann. "Measurement of branching ratios and CP asymmetries for the decays $B^0 \to \pi^+\pi^-$, $B^0 \to K^+\pi^-$, $B^0 \to K^+K^-$". PhD thesis. Ludwig-Maximilians-Universität München, 2013. URL: https://edoc.ub.uni-muenchen.de/15477/.

[40] F. Abudinén. "Development of a B^0 flavor tagger and performance study of a novel time-dependent CP analysis of the decay $B^0 \to \pi^0\pi^0$ at Belle II". PhD thesis. Ludwig-Maximilians-Universität München, 2018. URL: https://edoc.ub.uni-muenchen.de/23003/.

[41] M. Gronau and J. L. Rosner. "Combining CP Asymmetries in $B \to K\pi$ Decays". In: *Physical Review D* 59.11 (1999), p. 113002. DOI: 10.1103/PhysRevD.59.113002.

[42] H.-n. Li, S. Mishima, and A. I. Sanda. "Resolution to the $B \to \pi K$ Puzzle". In: *Physical Review D* 72.11 (2005), p. 114005. DOI: 10.1103/PhysRevD.72.114005.

[43] S. Baek et al. "$B \to \pi K$ Puzzle and New Physics". In: *Physical Review D* 71.5 (2005), p. 057502. DOI: 10.1103/PhysRevD.71.057502.

[44] A. J. Buras et al. "New aspects of $B \to \pi\pi, \pi K$ and their implications for rare decays". In: *The European Physical Journal C* 45 (2006), pp. 701–710. DOI: 10.1140/epjc/s2005-02469-4.

[45] R. Fleischer, S. Recksiegel, and F. Schwab. "On Puzzles and Non-Puzzles in $B \to \pi\pi, \pi K$ Decays". In: *The European Physical Journal C* 51.1 (2007), pp. 55–61. DOI: 10.1140/epjc/s10052-007-0277-8.

[46] N. B. Beaudry et al. "The $B \to K\pi$ Puzzle Revisited". In: *Journal of High Energy Physics* 1 (2018), p. 74. DOI: 10.1007/JHEP01(2018)074.

[47] M. Gronau. "A Precise Sum Rule among Four $B \to K\pi$ CP Asymmetries". In: *Physics Letters B* 627.1-4 (2005), pp. 82–88. DOI: 10.1016/j.physletb.2005.09.014.

[48] G. Buchalla, A. J. Buras, and M. E. Lautenbacher. "Weak Decays beyond Leading Logarithms". In: *Reviews of Modern Physics* 68.4 (1996), pp. 1125–1244. DOI: 10.1103/RevModPhys.68.1125.

[49] M. Gronau et al. "Decays of B Mesons to Two Light Pseudoscalars". In: *Physical Review D* 50.7 (1994), pp. 4529–4543. DOI: 10.1103/PhysRevD.50.4529.

[50] H. J. Lipkin et al. "Penguin Trapping with Isospin Analysis and CP Asymmetries in B Decays". In: *Physical Review D* 44.5 (1991), pp. 1454–1460. DOI: 10.1103/PhysRevD.44.1454.

[51] F. Abudinén et al. "B-Flavor Tagging at Belle II". In: *The European Physical Journal C* 82.4 (2022), p. 283. DOI: 10.1140/epjc/s10052-022-10180-9.

[52] Y. Funakoshi et al. "The SuperKEKB Has Broken the World Record of the Luminosity". In: *Proceedings of the 13th International Particle Accelerator Conference*. 13th Int. Particle Acc. Conf. Vol. IPAC2022. 2022. DOI: 10.18429/JACOW-IPAC2022-MOPLXGD1.

[53] T. Abe et al. "Belle II Technical Design Report". In: *arXiv (preprint)* (2010). DOI: 10.48550/arXiv.1011.0352. arXiv: 1011.0352.

[54] SuperB Collaboration. "SuperB: A High-Luminosity Asymmetric e^+e^- Super Flavor Factory. Conceptual Design Report". In: *arXiv (preprint)* (2007). DOI: 10.48550/arXiv.0709.0451. arXiv: 0709.0451.

[55] E. Kou et al. "The Belle II Physics Book". In: *Progress of Theoretical and Experimental Physics* 2019.12 (2019), p. 123C01. DOI: 10.1093/ptep/ptz106.

[56] K. Akai et al. "SuperKEKB Collider". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 907 (2018), pp. 188–199. DOI: 10.1016/j.nima.2018.08.017.

[57] T. Kuhr et al. "The Belle II Core Software". In: *Computing and Software for Big Science* 3.1 (2018), p. 1. DOI: 10.1007/s41781-018-0017-9.

[58] T. Abe et al. "Achievements of KEKB". In: *Progress of Theoretical and Experimental Physics* 2013.3 (2013), 03A001. DOI: 10.1093/ptep/pts102.

[59] K. Oide. "Super B Factories". In: O. Brüning and S. Myers. *Challenges and Goals for Accelerators in the XXI Century*. WORLD SCIENTIFIC, 2016, pp. 343–349. DOI: 10.1142/9789814436403_0020.

[60] L. Andricek et al. "Initial Requirements for an Upgraded VXD in Belle II". In: *Internal Belle II Note*. BELLE2-NOTE-TE-2019-011 (2019).

[61] B. Wang et al. "Operational Experience of the Belle II Pixel Detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1032 (2022), p. 166631. DOI: 10.1016/j.nima.2022.166631.

[62] J. Kemmer and G. Lutz. "New Detector Concepts". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 253.3 (1987), pp. 365–377. DOI: 10.1016/0168-9002(87)90518-3.

[63] L. Zani et al. "The Silicon Vertex Detector of the Belle II Experiment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1038 (2022), p. 166952. DOI: 10.1016/j.nima.2022.166952.

[64] S. Skambraks et al. "A Z-Vertex Trigger for Belle II". In: *IEEE Transactions on Nuclear Science* 62.4 (2015), pp. 1732–1740. DOI: 10.1109/TNS.2015.2439617.

[65] I. Frank and Ig. Tamm. "Coherent Visible Radiation of Fast Electrons Passing Through Matter". In: *Selected Papers*. Berlin, Heidelberg: Springer, 1991, pp. 29–35. DOI: 10.1007/978-3-642-74626-0_2.

[66] M. Starič. "Pattern Recognition for the Time-of-Propagation Counter". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. Proceedings of the Seventh International Workshop on Ring Imaging Cherenkov Detectors 639.1 (2011), pp. 252–255. DOI: 10.1016/j.nima.2010.09.176.

[67] T. Iijima et al. "A Novel Type of Proximity Focusing RICH Counter with Multiple Refractive Index Aerogel Radiator". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 548.3 (2005), pp. 383–390. DOI: 10.1016/j.nima.2005.05.030.

[68] Belle Collaboration. "The Belle Detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. Detectors for Asymmetric B-factories 479.1 (2002), pp. 117–232. DOI: 10.1016/S0168-9002(01)02013-7.

[69] D. Besson and T. Skwarnicki. "Upsilon spectroscopy: Transitions in the bottomonium system". In: *Annual Review of Nuclear and Particle Science* 43.1 (1993), pp. 333–378. DOI: 10.1146/annurev.ns.43.120193.002001.

[70] C. Cecchi et al. "B counting measurement of the Run 1 Belle II data sample". In: *Internal Belle II Note*. BELLE2-NOTE-PH-2023-011 (2023).

[71] S. Jadach, B. F. L. Ward, and Z. Wąs. "The precision Monte Carlo event generator KK for two-fermion final states in e^+e^- collisions". In: *Computer Physics Communications* 130.3 (2000), pp. 260–325. DOI: 10.1016/S0010-4655(00)00048-5.

[72] T. Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Computer Physics Communications* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024.

[73] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1 (2001), pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4.

[74] E. Barberio, B. van Eijk, and Z. Wąs. "Photos – a universal Monte Carlo for QED radiative corrections in decays". In: *Computer Physics Communications* 66.1 (1991), pp. 115–128. DOI: 10.1016/0010-4655(91)90012-A.

[75] S. Agostinelli et al. "Geant4 – a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[76] G. C. Fox and S. Wolfram. "Event shapes in e^+e^- annihilation". In: *Nuclear Physics B* 149 (3 1979), pp. 413–496. DOI: 10.1016/0550-3213(79)90003-8.

[77] W. Waltenberger et al. "The RAVE/VERTIGO vertex reconstruction toolkit and framework". In: *Journal of Physics: Conference Series* 119 (2008), p. 032037. DOI: 10.1088/1742-6596/119/3/032037.

[78] F. Abudinén et al. "B-flavor tagging at Belle II". In: *The European Physical Journal C* 82.4 (2022). DOI: 10.1140/epjc/s10052-022-10180-9.

[79] T. Keck. "FastBDT: A Speed-Optimized and Cache-Friendly Implementation of Stochastic Gradient-Boosted Decision Trees for Multivariate Classification". In: *arXiv (preprint)* (2016). DOI: 10.48550/arXiv.1609.06119.

[80] CLEO Collaboration. "Search for exclusive charmless hadronic B decays". In: *Physical Review D* 53 (1996), pp. 1039–1050. DOI: 10.1103/PhysRevD.53.1039.

[81] A. Hoecker et al. "TMVA - Toolkit for Multivariate Data Analysis". In: *arXiv (preprint)* (2009). DOI: 10.48550/arXiv.physics/0703039.

[82] Belle II Collaboration. *Systematic Corrections Framework*. URL: https://syscorrfw.readthedocs.io/en/latest/ (visited on 12/01/2022).

[83] T. Skwarnicki. "A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances". PhD thesis. Institute of Nuclear Physics, Cracow and DESY, Hamburg, 1986. URL: https://inspirehep.net/files/e31108fa63ba2754e30038cccdf9e3a7.

[84] B. Efron and R. Tibshirani. *An Introduction to the Bootstrap*. New York: Chapman & Hall, 1994. ISBN: 978-1-4899-4541-9.

[85] M. Pivk and F. Le Diberder. "sPlot: a statistical tool to unfold data distributions". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 555.1-2 (2004), pp. 356–369. DOI: 10.1016/j.nima.2005.08.106.

[86] A. Glazov, P. Rados, and A. Rostomyan. "Measurement of the track reconstruction efficiency and fake rate with $e^+e^- \to \tau^+\tau^-$ events". In: *Internal Belle II Note*. BELLE2-NOTE-PH-2020-006 (2020).

[87] C. Cecchi et al. "B counting measurement in Moriond 2022/ICHEP 2022 Belle II dataset". In: *Internal Belle II Note*. BELLE2-NOTE-PH-2022-007 (2022).

[88] Belle Collaboration. "Measurement of the B^+/B^0 production ratio in e^+e^- collisions at the $\Upsilon(4S)$ resonance using $B \to J/\psi(\ell\ell)K$ decays at Belle". In: *Physical Review D* 107 (3 2023), p. L031102. DOI: 10.1103/PhysRevD.107.L031102.

[89] M. Dorigo, D. Ghosh, and M. Mantovano. "Measurement of instrumental asymmetries of kaons and pions". In: *Internal Belle II Note*. BELLE2-NOTE-TE-2022-014 (2022).

[90] Belle II Collaboration. "Measurements of branching fractions and direct CP asymmetries in $B^0 \to K^+\pi^-$, $B^+ \to K^0_S\pi^+$ and $B^0 \to \pi^+\pi^-$ using 2019 and 2020 data". In: *arXiv (preprint)* (2021). DOI: 10.48550/ARXIV.2106.03766.

[91] Belle Collaboration. "Measurements of branching fractions and direct CP asymmetries for $B \to K\pi$, $B \to \pi\pi$ and $B \to KK$ decays". In: *Physical Review D* 87.3 (2013), p. 031103. DOI: 10.1103/PhysRevD.87.031103.

[92] BABAR Collaboration. "Improved Measurements of the Branching Fractions for $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$, and a search for $B^0 \to K^+K^-$". In: *Physical Review D* 75.1 (2007), p. 012008. DOI: 10.1103/PhysRevD.75.012008.

[93] BABAR Collaboration. "Measurement of CP Asymmetries and Branching Fractions in Charmless Two-Body B-Meson Decays to Pions and Kaons". In: *Physical Review D* 87.5 (2013), p. 052009. DOI: 10.1103/PhysRevD.87.052009.

[94] T. E. Browder et al. "New physics at a Super Flavor Factory". In: *Reviews of Modern Physics* 81.4 (2009), pp. 1887–1941. DOI: 10.1103/revmodphys.81.1887.

[95] T. Gershon and A. Soni. "Null tests of the Standard Model at an International SuperBFactory". In: *Journal of Physics G: Nuclear and Particle Physics* 34.3 (2007), pp. 479–492. DOI: 10.1088/0954-3899/34/3/006.

[96] G. Bell et al. "Two-loop current–current operator contribution to the non-leptonic QCD penguin amplitude". In: *Physics Letters B* 750 (2015), pp. 348–355. DOI: 10.1016/j.physletb.2015.09.037.

[97] Belle II Collaboration. "Measurement of branching fractions and direct CP asymmetries for $B \to K\pi$ and $B \to \pi\pi$ decays at Belle II". In: *arXiv (preprint)* (2023). DOI: 10.48550/arXiv.2310.06381. arXiv: 2310.06381.

[98] D. Weyland. "Continuum Suppression with Deep Learning techniques for the Belle II Experiment". Master's thesis. Karlsruhe Institute of Technology, 2017. URL: https://publish.etp.kit.edu/record/21416.

Acknowledgments

But it's becoming increasingly obvious.
I can deny it no longer.
I am small. — Plankton

Next to the hundreds of collaborators within Belle II, without whom this work would not have been possible, a select few have played an indispensable role in supporting this endeavor and my personal journey over these years.

First, I want to thank Professor Allen Caldwell for enabling me to pursue a PhD under his supervision. This appreciation extends to Dr. Hans-Günther Moser for his role in making my PhD journey possible. Hans-Günther, I am grateful for the trust you placed in me, despite my initial lack of experience in high-energy physics. Your unwavering support, especially during the more challenging times, has been invaluable. Thank you!

I also want to express my gratitude to the entire rest of the group at the Max-Planck Institute. Professor Christian Kiesling, your enthusiasm for particle physics, even in the face of analyzing ntuples with ROOT on a virtual machine, is truly inspiring. Thank you, Felix, for allowing me to be the co-founder of our much cherished kiosk. I am telling you, instant ramen would have changed everything for us. Markus, Oskar, Justin, Thibaud, thank you for making data analysis feel a bit more like a team effort. Angie, Elia, thank you for choosing that one particular restaurant in Tokyo, where eating slugs is not optional but mandatory. Endrico, Thomas, Philipp, Fernando, thanks for always offering your help without ever asking anything in return — it's my favorite kind of help.

To my Italian colleagues, Riccardo, Sebastiano, Michele, and Diego, thank you for your guidance during my early days as a clueless PhD student. Without you, I would have never wandered the Roman suburbs at 4 in the morning. Elia, I am grateful to call you my friend — sorry for never replying to your texts.

Finally, my heartfelt gratitude goes to my dear friends and loved ones, who probably still have no idea what my thesis is about but provided their relentless support anyway. Ekaterina, thank you for enduring, encouraging and taking care of me! Maxi, Svenja, Jan, Oma, Opa, Mama, Papa, Worte könnten nicht verpacken, wie dankbar ich für euch bin!