Università degli Studi di Trieste

Dipartimento di Fisica

XXXV Ciclo del Dottorato di Ricerca in Fisica

Measurement of the properties of $B^+ \to \rho^+ \rho^0$ decays at Belle II

Settore scientifico-disciplinare: FIS/04

Candidato:
Riccardo Manfredi

Coordinatore:
Prof. Francesco Longo

Supervisore di tesi:
Dott. Diego Tonelli

ANNO ACCADEMICO 2021–2022

Abstract

This experimental particle-physics thesis reports measurements of properties of $B^+ \to \rho^+\rho^0$ decays for the determination of the quark-mixing matrix angle $\alpha = \arg(-(V_{td}V_{tb}^*)/(V_{ud}V_{ub}^*))$, a fundamental parameter of quark dynamics. The analysis uses the full data set of electron-positron collisions at the $\Upsilon(4S)$ resonance produced by the energy-asymmetric SuperKEKB collider and collected by the Belle II experiment until June 2022. Belle II is an hermetic solenoidal magnetic spectrometer, surrounded by particle-identification detectors, calorimeter, and muon detectors, designed to reconstruct the decay products of 10 GeV electron-positron collisions. The sample corresponds to 387 million pairs of bottom-antibottom mesons.

The analysis is developed using simulated and control-data samples. Only when all procedures are established, the analysis is applied to the sample of signal candidates in data. Data are enriched in $B^+ \to \rho^+\rho^0$ events by a multivariate statistical-learning selection that suppresses the most prominent source of background by minimizing the expected statistical variance of the quantities of interest. A multidimensional fit of sample composition statistically identifies the signal and determines the angular distributions of its decay products. Special care is devoted to keep the systematic uncertainties under control: the main contribution comes from mismodelings of the angular distributions in signal-rich regions. The central values are hidden as the analysis is still under internal collaboration review. The fit results, combined with acceptance and efficiency corrections determined from simulation, allow for determining the $B^+ \to \rho^+\rho^0$ branching fraction,

$$\mathcal{B}(B^+ \to \rho^+\rho^0) = ([\mathrm{xx}]^{+1.40}_{-1.39}\ (\mathrm{stat}) \pm [1.39]\ (\mathrm{syst})) \times 10^{-6},$$

the fraction of longitudinally polarized $B^+ \to \rho^+\rho^0$ decays,

$$f_L = [\mathrm{xx}]^{+0.017}_{-0.015}\ (\mathrm{stat}) \pm [0.008]\ (\mathrm{syst}),$$

and the CP-violating charge-asymmetry,

$$A_{CP} = [\mathrm{xx}] \pm 0.062\ (\mathrm{stat}) \pm [0.041]\ (\mathrm{syst}).$$

The statistical uncertainties are competitive with the current world-best results. The systematic uncertainties are reported in brackets as they are still preliminary.

Contents

Introduction

The Standard Model of elementary particles and their interactions (SM) is the currently accepted theory of particle physics. It is widely recognized as the ultimate success of the reductionist paradigm for describing microphysics at its most fundamental level. By means of about twenty parameters, the Standard Model describes accurately thousands of measurements involving processes mediated by the electromagnetic, weak, and strong interactions that span more than ten orders of magnitude in energy.

However, theoretical considerations and, possibly, experimental inconsistencies support the general belief that the Standard Model might still be an *effective theory* — a theory valid at the energies probed so far, that is incorporated in a yet-unknown and more general theory. Completing the Standard Model is the principal goal of today's particle physics.

Direct approaches, which broadly consist in searching for decay products of non-SM particles produced on mass-shell in high-energy collisions, have been traditionally fruitful. However, their current reach is limited by the collision energy of today's accelerators and by the large investments needed to further it in future. Complementary approaches consist in comparing precise measurements in lower-energy processes where virtual non-SM particles could contribute, with equally precise predictions. The reach of such *indirect* approaches is not constrained by collision energy, but rather by the precision attainable, both in measurements and predictions.

The Belle II experiment is an international collaboration of about 1000 physicists that *indirectly* tests the Standard Model by studying millions of decays of τ leptons (heaviest partners of the electron) and mesons containing the quarks b and c (heavier and longer-lived partners of the fundamental constituents of nuclear matter) produced in electron-positron collisions at energies around $10.58\,\mathrm{GeV}$. Since the start of physics data taking in 2019, Belle II collected samples corresponding to $362\,\mathrm{fb}^{-1}$ of integrated luminosity, or, equivalently, 387 million pairs of quantum-correlated bottom-antibottom mesons (B mesons, bound states of a b quark and a lighter quark).

This work focuses on the experimental determination of α (also indicated as ϕ_2), a fundamental parameter of quark dynamics that limits the current precision of tests of the SM description of charge-parity violation in the weak interactions. The parameter $\alpha = \arg(-(V_{td}V_{tb}^*)/(V_{ud}V_{ub}^*))$ is a combination of couplings between up-type (charge $+2e/3$) and down-type (charge $-e/3$) quarks. Its determination requires the combination of measurements from families of charmless B decays such as $B \to \rho\rho$ and $B \to \pi\pi$ decays. The presence of neutral pions in many of such channels renders Belle II significantly competitive for contributing to the determination of α with respect to concurring experiments. Our measurement uses the decay properties of $B^+ \to \rho^+\rho^0$ decays. Decays $B^+ \to \rho^+(\to \pi^+\pi^0)\rho^0(\to \pi^+\pi^-)$ are reconstructed in the full Belle II data set currently available. The composition of the resulting sample is fit to measure the branching fraction, charge-parity-violating decay-rate asymmetry, and the fraction of decays where both ρ mesons have zero spin projections along their momenta.

The chief challenge is to improve the initial 10^{-6} signal-to-background ratio with a selection sufficiently discriminating to isolate an abundant, low-background signal without introducing excessive correlations between the observables used in the multidimensional fit that determines the parameters of interest. Another key point of this work is the detailed study of the mismodeling between data and simulation that appear in signal-rich regions of the angular distributions, used in the fit to identify the signal polarization.

This thesis is structured as follows. Chapter 1 introduces the flavor sector of the Standard Model; Chapter 2 describes the Belle II experiment; Chapter 3 outlines the relevant experimental features of the $B^+ \to \rho^+\rho^0$ measurement and provides an outline of the general principles of B meson reconstruction; Chapter 4, where the description of my direct, original contributions begins, reports on the reconstruction and selection of the $B^+ \to \rho^+\rho^0$ decay channel; Chapter 5 describes the fit of sample composition, the core of this work; Chapter 6 documents the analysis validation and consistency checks on control data; Chapter 7 reports the results of the fit of $B^+ \to \rho^+\rho^0$ data; Chapter 8 discusses sources of systematic effects and their associated uncertainties; Chapter 9 summarizes the final results of the measurement and implications; a final summary concludes the document.

Charge-conjugate processes are implied throughout the document unless specified otherwise. Generic particle symbols (B, π, ...) indicate indistinctly charged or neutral particles.

Chapter 1

Flavor physics to overcome the Standard Model

This is a concise introduction to the weak interactions of quarks and how they are incorporated in the Standard Model of particle physics. Emphasis is on the role of such interactions in searches for as-yet unknown particles that may complete the Standard Model at high energies. Finally, I specialize the discussion to charmless decays of bottom mesons and the angle α, which are the subject of the measurement described in this thesis.

1.1 The Standard Model of particle physics

The Standard Model (SM) is an effective quantum field theory that describes three of the four fundamental interactions in nature (gravity is not included) [1–6].

The quantum field theory framework results from the unification of quantum mechanics with special relativity and offers the most fundamental description of nature known to date.

A field is a set of values, associated to certain physical properties, assigned to every point in space and time. Quantum fields are fields that pervade the whole spacetime and obey the rules of quantum mechanics. If a quantum field is modified by an appropriate perturbation, the resulting oscillatory states, called field excitations, carry more energy than the resting state and are what we call 'particles'. For instance, the electron is the massive excitation of the corresponding electron field. The quantized nature of the description implies that only certain perturbations that satisfy precise energetic conditions are capable of generating field excitations. In other words, it is not possible, for example, to generate a wave in the electron field that corresponds to half an electron, with half a unit of electric charge.

Quantum fields interact with each other. The Standard Model is the theory that describes their dynamics at energy scales relevant for the subatomic world. Particles and their interactions are described in a Lagrangian formalism, in which every combination of fields and interaction operators that is not forbidden by the symmetries of the dynamics is, in principle, included. Local gauge symmetry, i.e., the invariance of the Lagrangian under space-time-dependent transformations applied to the phases of fields, is the key overarching concept. Interaction terms appear in the free-field Lagrangian after requiring it to be invariant under local gauge symmetries.

The Standard Model is based on the symmetry group

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y,$$

where $SU(3)_C$ is the standard unitary group that describes the strong interactions (quantum chromodynamics, QCD), and C stands for the color charge; $SU(2)_L \otimes U(1)_Y$ is the product of groups that describe the combination of the weak and electromagnetic interactions, with L standing for *left* and Y standing for *hypercharge.*[1]

Spin-1 particles called *gauge bosons* mediate the interactions. Strong interactions are mediated by eight massless particles corresponding to the $SU(3)_C$ generators, called *gluons*: they carry a charge that can be of three kinds, called *color.* Weak interactions are mediated by two charged massive bosons, $W^{\pm}$, and a neutral massive boson, Z^0. Electromagnetic interactions occur between particles carrying electric charge and are mediated by a neutral massless boson, the photon γ. The physical electroweak bosons ($W^{\pm}$, Z^0, γ) arise from the following linear combinations of $SU(2)_L \otimes U(1)_Y$ generators:

$$W^{\pm} = \frac{1}{\sqrt{2}}(W_1 \mp iW_2) \quad \text{and} \quad \begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix},$$

where θ_W is a free parameter, called *Weinberg angle.* The $W^{\pm}$ mass depends on the Z mass via θ_W. Particles acquire mass via the interaction with the Higgs field, which is mediated by a spin-0 particle, the Higgs boson.

Matter particles correspond to excitations of spin-$\frac{1}{2}$ fields and are called *fermions.* Their masses are free parameters. Each fermion is also associated with an anti-particle that has the same mass and opposite internal quantum numbers. Fermions are further classified into two classes, quarks, that are the fundamental constituents of nuclear matter, and leptons, each organized in three weak-isospin doublets.

- Quark doublets are composed each of an up-type quark, with charge $\frac{2}{3}e$, and a down-type quarks, with charge $-\frac{1}{3}e$,

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

  They couple with both the strong and electroweak interactions. Each quark has color and a ‘flavor’ quantum number, which comes in six varieties and is conserved in the electromagnetic and strong interactions, but not in the weak interactions. Due to color confinement [7] free quarks are not observable. They are only observed in their colorless bound states, which include mesons, typically composed of a quark and an anti-quark, and baryons, composed of three quarks. Baryons are assigned a quantum number, called baryon number, found to be conserved even if no symmetry of the Lagrangian requires that.

- Lepton doublets are composed each by a massless neutral neutrino and a massive particle with electric charge $-e$;

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

  They couple only with the electroweak interaction. Each lepton has a lepton family quantum number; their sum in a process, called global lepton number, is found to be conserved in all interactions, although no symmetry of the dynamics prescribes that; individual lepton numbers are not conserved in some processes, notably neutrino oscillations.

Figure 1.1 shows a scheme of the Standard Model particles and their interactions.

---

[1] Only particles with *left* chirality are influenced by the weak interaction.

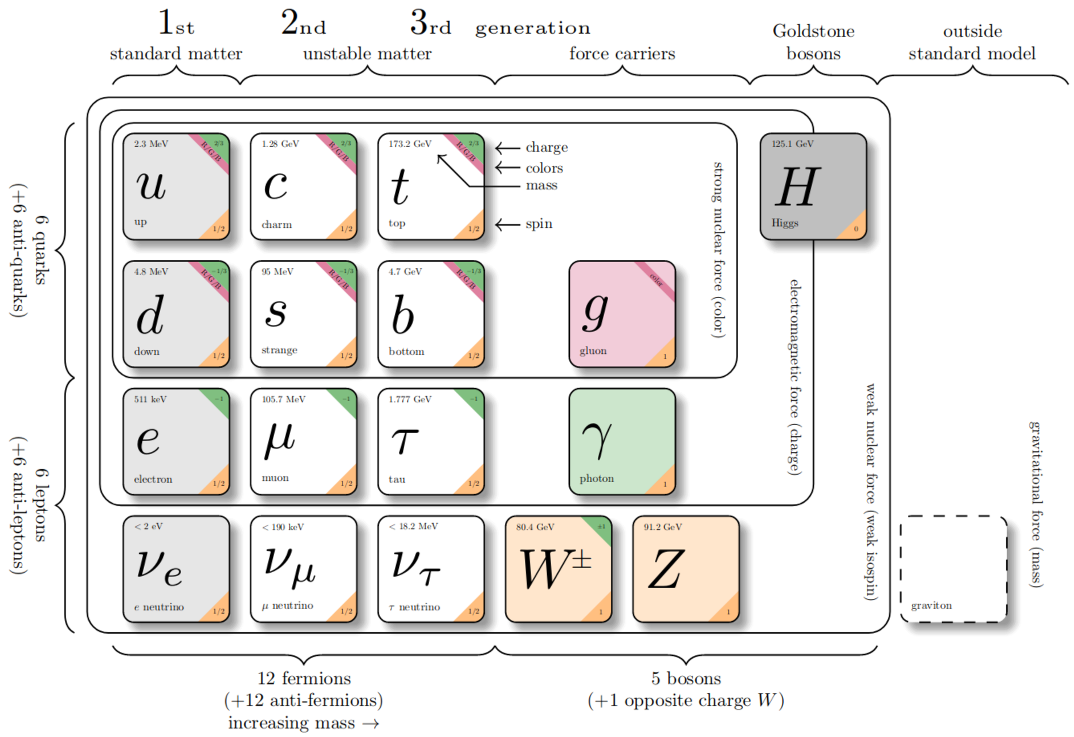


Figure 1.1: Scheme of particles and interactions in the Standard Model.

In addition to gauge symmetry, discrete symmetries are important too in constraining the dynamics. Parity (P) is a transformation that inverts all spatial coordinates; charge conjugation (C) is the exchange of every particle with its own antiparticle; and time reversal (T) inverts the time axis. The product of these three discrete symmetry transformations is found to be conserved in all interactions, as prescribed by foundational axioms of field theory, but they are not conserved individually [8, 9]. Parity symmetry is maximally violated in the weak interactions, while the combined CP symmetry is violated in the weak interactions at the 0.1% level. In principle, the strong interaction too could violate CP symmetry, but no experimental evidence of that has ever been observed. The existence of as-yet unobserved particles (axions) has been postulated to account for that.

1.2 Where do we stand?

The Standard Model was completed in the 1970's and has been successfully tested since, in thousands of measurements whose fractional precisions reach one part per trillion [10]. However, observations and theoretical considerations suggest that the Standard Model is likely to be an effective theory, valid at the eV–TeV energies probed so far, that should be completed by a more general full theory valid over a broader range of high energies. Open questions that support this interpretation include the lack of an explanation for a dynamical origin for the observed asymmetry between matter and antimatter in the universe, the strikingly large differences observed between fermion masses, the possible instability of the Higgs vacuum, the conceptual and technical difficulties in achieving a description of gravity consistent with quantum mechanics, or the postulated large amounts of non-interacting matter (dark matter), introduced to justify cosmological observations.

Extending the Standard Model to higher energy-scales is the main goal of today's particle physics, in an attempt at addressing these and other open issues. Current strategies to extend the Standard Model can broadly be classified into two synergic approaches.

The energy-frontier, *direct* approach aims at using high-energy collisions to produce on-shell particles (that is, particles satisfying the energy-momentum conservation in the production process) not included in the SM, and detect directly their decay products, thus gaining direct evidence of their existence.[2] Historically this offered striking experimental evidence of new phenomena, when energetically accessible, but its reach is limited by the maximum energy available at colliders.

The intensity-frontier, *indirect* approach broadly consists in searching for significant differences between precise measurement and equally precise SM predictions in lower-energy processes sensitive to non-SM contributions. A simplified, semi-intuitive conceptual representation of the subtending idea is that exchanges of virtual (off-mass-shell) particles of arbitrary high mass, including those not described in the Standard Model, occur in the amplitude, thus altering the amplitudes in an observable manner. The presence of virtual particles, which may imply a temporary non-conservation of energy, is allowed by Heisenberg's uncertainty principle $\Delta E \Delta t > \frac{\hbar}{2}$. Experimental evidence is typically harder to establish, but the reach is not bounded by the maximum collision energy reachable by experiments. A large portion of the effort in this approach is centered on the weak-interactions of quarks (so called 'flavor physics').

1.3 Flavor physics in the Standard Model

Although technically flavor physics includes also lepton flavor, I restrict the scope by referring solely to the quark interactions here.

The role of flavor in shaping the Standard Model has been central since the early days of particle physics. However, its prominence in determining the theory can perhaps be tracked down to the early 1960's with the apparent inconsistency between weak coupling constants measured in muon decay, neutron decay, and strange-particle decays. Such inconsistency was first addressed by Gell-Mann and Levy [11] and then Cabibbo [12], who postulated differing mass (d) and weak (d') eigenstates for down-type quarks. This was achieved by introducing a mixing angle (θ_C) between the s-quark and d-quark, the only two down-type quarks known at the time. While Cabibbo's theory addressed economically the difference of weak coupling constants, it also predicted a rate for the $K^0_{\mathrm{L}} \to \mu^+\mu^-$ and other kaon decays inconsistent with the experimental exclusion limits at the time. Glashow, Iliopoulos, and Maiani addressed the conundrum by postulating the existence of a fourth quark (c) of $2\,\mathrm{GeV}/c^2$ mass, whose contribution in the $K^0_{\mathrm{L}} \to \mu^+\mu^-$ decay amplitude would cancel the u-quark contribution, suppressing the branching fraction down to values consistent with experimental limits [13]. The charm quark was then discovered four years after the prediction, showing the compelling power of the indirect approach. In addition, in 1973 when only three quarks were known, Kobayashi and Maskawa generalized Cabibbo's theory from a four-quark model to a six-quark model to accommodate the phenomenon of CP violation observed in 1964 [14]. They introduced a complex unitary matrix to describe the relations between mass and weak interaction eigenstates of quarks as seen by $W^{\pm}$ bosons. This is known as the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix or V_{CKM}. The $N \times N$ CKM matrix has $(N-1)^2$ free parameters [15], where N is the number of quarks families. If $N = 2$, the only free parameter is the Cabibbo angle θ_C, whereas if

---

[2] *Mass shell* is jargon for mass hyperboloid, which identifies the hyperboloid in energy–momentum space describing the solutions to the mass-energy equivalence equation $E^2 = (pc)^2 + m^2c^4$. A particle *on-mass-shell* satisfies this relation.

$N = 3$, the free parameters are three Euler angles (θ_{12}, θ_{13}, and θ_{23}) and a complex phase (δ), which allows for *CP*-violating couplings. The matrix is written as

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

where primed quark symbols indicate the weak-interaction eigenstates and the unprimed symbols the mass eigenstates. The V_{ij} matrix element encapsulates the coupling between an up-type i and down-type j quarks. It is most conveniently written in the so-called *Wolfenstein parametrization* [16], an expansion in the small parameter $\lambda = \sin\theta_C \approx 0.23$ that makes explicit the observed hierarchy between its elements,

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

where

$$\lambda = \frac{V_{us}}{\sqrt{V_{ud}^2 + V_{us}^2}} \quad A\lambda^2 = \lambda\frac{V_{cb}}{V_{us}} \quad A\lambda^3(\rho + i\eta) = V_{ub}^* .$$

The parameter λ expresses the mixing between the first and second quark generations, A and ρ are real parameters, and η is a complex phase that introduces *CP* violation. The unitarity condition $V_{\text{CKM}}V_{\text{CKM}}^\dagger = \mathbb{1}$ yields nine relations,

$$\begin{array}{ccc} |V_{ud}|^2 + |V_{cd}|^2 + |V_{td}|^2 = 1 & V_{us}^*V_{ud} + V_{cs}^*V_{cd} + V_{ts}^*V_{td} = 0 & V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* = 0 , \\ |V_{us}|^2 + |V_{cs}|^2 + |V_{ts}|^2 = 1 & V_{ub}^*V_{ud} + V_{cb}^*V_{cd} + V_{tb}^*V_{td} = 0 & V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* = 0 , \\ |V_{ub}|^2 + |V_{cb}|^2 + |V_{tb}|^2 = 1 & V_{ub}^*V_{us} + V_{cb}^*V_{cs} + V_{tb}^*V_{ts} = 0 & V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* = 0 , \end{array}$$

which are sums of three complex numbers each. The six equations summing to zero prompt a convenient geometric representation in terms of so-called *unitarity triangles* in the complex plane. A *CP* conserving theory would yield null-area triangles or, equivalently, a vanishing Jarlskog invariant $J = \Im(V_{us}V_{cb}V_{ub}^*V_{cs}^*)$ [17–19]. All elements of the second equation in the second row have similar magnitudes, yielding a notable triangle referred to as 'the Unitarity Triangle', shown in Figure 1.2. Conventionally, side sizes are normalized to the length of the base, and the three angles are labelled α or ϕ_2, β or ϕ_1, and γ or ϕ_3.

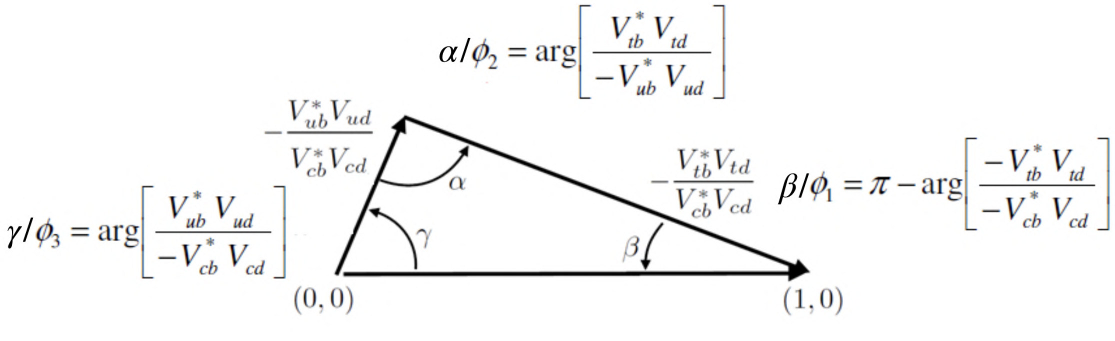


Figure 1.2: Graphical representation of the Unitarity Triangle.

In addition to *CP* violation, the flavor-mixing phenomenon, which involves flavored neutral mesons $|M\rangle$, enriches significantly the phenomenology. Flavor quantum numbers are conserved in strong interactions and thus the flavor eigenstates are eigenstates of strong interactions. Weak interactions do not conserve flavor allowing $|M\rangle$ to undergo a transition into $|\overline{M}\rangle$ (or vice versa), which changes flavor by two units. Because the full Hamiltonian contains strong and weak interactions, its eigenstates (which are the particles we observe, with definite masses and lifetimes) are linear superpositions of flavor eigenstates $|M\rangle$ and $|\overline{M}\rangle$,

$$\begin{aligned} |M_+(t)\rangle &\equiv p|M(t)\rangle + q|\bar{M}(t)\rangle, \\ |M_-(t)\rangle &\equiv p|M(t)\rangle - q|\bar{M}(t)\rangle. \end{aligned} \tag{1.1}$$

This phenomenon is called flavor mixing and it induces flavor oscillations in the time evolution of neutral flavored mesons. As an example, Fig. 1.3 shows the leading-order Feynman diagrams contributing to neutral $B^0 - \overline{B}^0$ mixing. Flavor oscillations generate additional time-evolution paths to the simple decay, which interfere with the decay thus enriching the dynamics and our opportunities to study it.

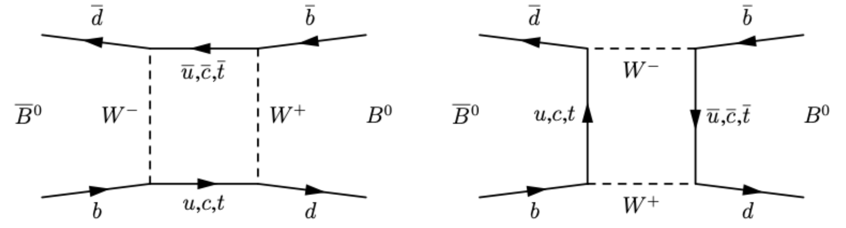


Figure 1.3: Leading order Feynman diagrams contributing to neutral $B^0 - \overline{B}^0$ mixing.

1.4 Flavor physics to overcome the Standard Model

Many physicists find the current understanding of flavor dynamics and *CP* violation unsatisfactory. The observed hierarchies between quark masses and couplings seem too regular to be accidental and the abundance of free parameters (six quark masses and four couplings) suggests the possibility of a deeper, more fundamental theory based on a possibly reduced set of parameters. In addition, while the CKM mechanism offers a framework to include *CP* violation in the SM, it does not really enlighten the origin for such a singular phenomenon. But even in the absence of a deeper understanding of the origin of *CP* violation, naturalness arguments indicate that most generic extensions of the SM would involve additional sources of *CP* violation. These and other considerations support the notion that a more detailed and complete study of the phenomenology of quarks dynamics and *CP* violation may reveal useful information to guide searches for Standard Model extensions.

The abundance and diversity of experimentally accessible processes to measure redundantly a reduced set of parameters makes indirect searches in the flavor sector a powerful option for exploring non-SM dynamics. In fact, even if no deviations from the Standard Model are found, the resulting stringent constraints on SM extensions are expected to remain useful in informing future searches.

The two classes of flavor-physics processes most promising for probing contributions of non-SM particles are *flavor-changing-neutral-currents* and *CP-violating* processes.

1.4.1 Flavor-changing neutral currents

Flavor-changing neutral currents (FCNC) are processes in which quark flavor changes in the transition but quark electric charge does not. The processes are suppressed in the Standard Model, because they occur only through second-order amplitudes involving the internal exchange of $W^{\pm}$ bosons ('loop amplitudes'), as shown in Fig. 1.4.

Such amplitudes are naturally sensitive to non-SM contributions, since any particle with proper quantum numbers and nearly arbitrary mass can replace the SM-quark closed-line in these diagrams thus altering the rate. FCNC are therefore powerful in identifying contributions from non-SM particles if rate enhancements or suppressions with respect to Standard Model expectations are observed.

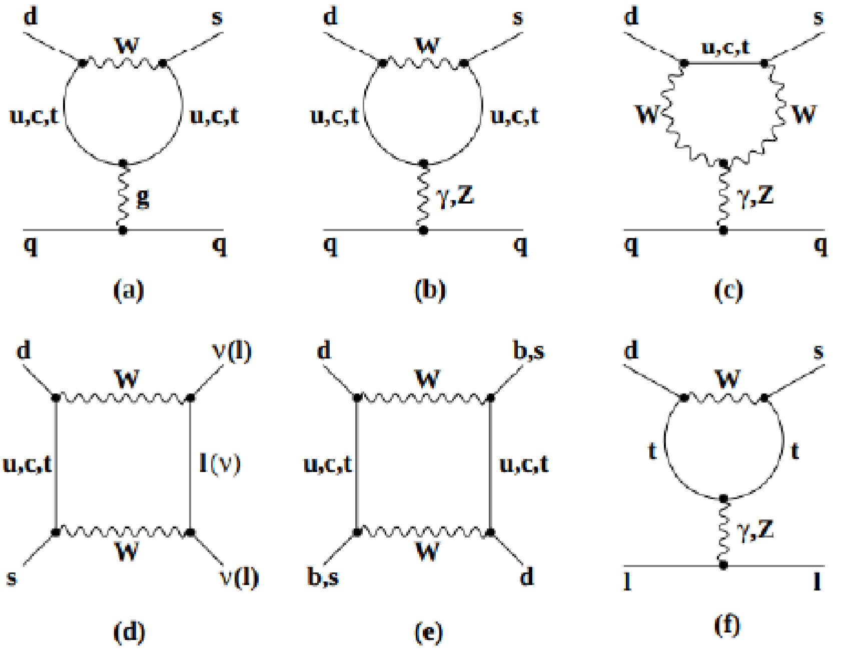


Figure 1.4: Examples of leading FCNC diagrams.

1.4.2 Violation of charge-parity symmetry

In addition to rate alterations, the phenomenon of *CP* violation offer additional avenues to uncover or characterize possible non-SM contributions. Alterations of the *CP*-violating phases with respect to those predicted by the SM are generically expected in a broad class of SM-extensions. Observing experimental evidence of those phases offers further opportunities to explore the dynamics even if total rates are unaffected.

Depending on how the *CP*-violating complex phase enters the dynamical evolution of a particle, *CP* violation can be classified into three distinct phenomenologies. In the most complete case, one considers the amplitudes of the transitions of a flavored meson $|M\rangle$ and of its antiparticle $|\bar{M}\rangle$, into a final state $|f\rangle$ and into the *CP* conjugate $|\bar{f}\rangle$, respectively

$$\begin{aligned} A_f &= \langle f|\mathcal{H}_{\text{eff}}|M\rangle = A\left(|M\rangle \to |f\rangle\right), & \bar{A}_f &= \langle f|\mathcal{H}_{\text{eff}}|\bar{M}\rangle = A\left(|\bar{M}\rangle \to |f\rangle\right), \\ A_{\bar{f}} &= \langle \bar{f}|\mathcal{H}_{\text{eff}}|M\rangle = A\left(|M\rangle \to |\bar{f}\rangle\right), & \bar{A}_{\bar{f}} &= \langle \bar{f}|\mathcal{H}_{\text{eff}}|\bar{M}\rangle = A\left(|\bar{M}\rangle \to |\bar{f}\rangle\right), \end{aligned}$$

with the effective Hamiltonian operator $\mathcal{H}_{\text{eff}}$.

a) *CP* violation in decay, or direct *CP* violation, occurs if

$$\left|\frac{A_f}{\bar{A}_{\bar{f}}}\right| \neq 1 \text{ or } \left|\frac{\bar{A}_f}{A_{\bar{f}}}\right| \neq 1, \tag{1.2}$$

thus appearing experimentally as a difference between the $|\bar{M}\rangle \to |\bar{f}\rangle$ and $|M\rangle \to |f\rangle$ decay rates. This is the only *CP* phenomenology possible for both neutral and charged mesons, and for baryons. Since in general, different intermediate states contribute to a decay, the total decay amplitudes A_f and $\bar{A}_{\bar{f}}$ are written as sums of the individual contributions

$$A_f = \sum_i |A_i| e^{i(\delta_i + \phi_i)}, \qquad \bar{A}_{\bar{f}} = \sum_i |A_i| e^{i(\delta_i - \phi_i)},$$

where symbols δ_i indicate *CP*-conserving phases, and ϕ_i are *CP*-violating phases associated with the elements of the CKM matrix that appear in a given amplitude. As *CP*-conserving couplings are real, the corresponding phases remain invariant under *CP* transformation. Since a *CP* transformation turns CKM coefficients into their complex conjugate, the *CP*-violating phases flip sign. The direct *CP* violation condition in Eq. (1.2) is satisfied if

$$|A_f|^2 - |\bar{A}_{\bar{f}}|^2 = -2 \sum_{i,j} |A_i||A_j| \sin(\delta_i - \delta_j) \sin(\phi_i - \phi_j) \neq 0.$$

Thus, *CP* violation in decay can occur only if at least two amplitudes with different *CP*-conserving and *CP*-violating phases contribute to a decay process. The squared absolute magnitudes of the total decay amplitudes $|A_f|^2$ and $|\bar{A}_{\bar{f}}|^2$ are accessible experimentally since they are proportional to the total decay rates. The individual amplitudes A_i are often difficult to compute theoretically as they typically involve contributions from strong-interaction amplitudes at low energy. These are non-perturbative and therefore hard to calculate leading to large uncertainties. Thus, observables that depend only on the weak phases, such as the CKM angles, allow to test SM predictions in a more reliable way.

b) *CP* violation in mixing or indirect *CP* violation implies,

$$\left|\frac{q}{p}\right| = \left|\frac{1-\varepsilon}{1+\varepsilon}\right| \neq 1 \;\; \Rightarrow \;\; |\epsilon| \neq 0, \tag{1.3}$$

where q and p are introduced in Eq. (1.1). In this case, *CP* violation generates a difference between the flavor-oscillation rates $|M\rangle \to |\bar{M}\rangle$ and $|\bar{M}\rangle \to |M\rangle$, which can be observed as a charge-dependent asymmetry in the yields of charged leptons from semileptonic decays of the oscillating B^0 mesons.

c) *CP* violation by interference of mixing and decay is observed when the neutral mesons $|M\rangle$ and $|\bar{M}\rangle$ can decay into a common final state $|f\rangle$, preferentially a pure *CP* eigenstate

$$CP|f_{CP}\rangle = \pm|f_{CP}\rangle.$$

Even if *CP* is conserved in mixing and in decay separately, i.e., if $|\bar{A}_{f_{CP}}/A_{f_{CP}}| = |q/p| = 1$, the combination of the decay and mixing phases can generate a total phase difference and thus an interference between these two processes, generating consequently a violation of *CP* symmetry. Introducing a new complex quantity λ_{CP}, the condition for this kind of *CP* violation can be written as

$$\mathrm{Im}(\lambda_{CP}) \neq 0, \quad \text{where} \quad \lambda_{CP} \equiv \frac{q}{p} \cdot \frac{\bar{A}_{f_{CP}}}{A_{f_{CP}}} = \left|\frac{q}{p}\right| \cdot \left|\frac{\bar{A}_{f_{CP}}}{A_{f_{CP}}}\right| e^{-i(\phi_M + \phi_D)}.$$

1.5 Current experimental status

Measurements of parameters associated with quark-flavor physics have been performed in many dedicated, or general-purpose, experiments in the last three decades, including CLEO, CPLEAR, NA32, NA48, NA62, KTeV, SLD, OPAL, L3, ALEPH, DELPHI, BaBar, Belle, CDF, CDFII, LHCb, Atlas, and CMS [20].

The current status of constraints on sides and angles of the Unitarity Triangle is shown in Fig. 1.5 [21]. Measurements of $\sin 2\beta$ reached a precision of 2.5%, mainly due to the availability of large samples of $B^0 \to J/\psi K^0$ decays in e^+e^- and $p\overline{p}$ collisions, while the angle α is known down to a 4% precision from $B \to hh$ decays (where B is charged or neutral, and h represents a charged or neutral π or ρ) in e^+e^- and pp collisions. The angle γ is measured with 5% precision using combinations of several measurements involving $B \to DK$ decays (B, D, K charged or neutral) reconstructed in e^+e^- and pp collisions. Discrepancies in the determinations of $|V_{cb}|$ and $|V_{ub}|$ are found between values measured using different analyses of semileptonic decays, mainly performed in e^+e^- collisions.

The decay width difference of the $B_s^0 - \overline{B_s}^0$ system is determined with 6% precision in pp collisions, while measurements are not yet precise enough to discern the expected non-zero value for the $B^0 - \overline{B}^0$ system. Mass differences in both systems are known with better than 1% precision from pp and $p\overline{p}$ collisions. In addition, many other measurements in charm and kaon physics contribute that are not straightforwardly represented in the Unitarity Triangle.

The resulting global picture is that the CKM interpretation of quark-flavor phenomenology is the dominant mechanism at play in the dynamics.

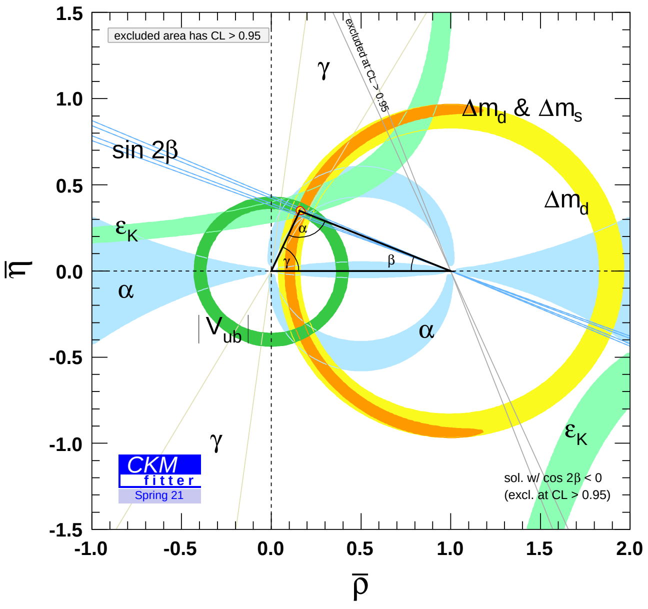


Figure 1.5: Current constrains on sides and angles of the Unitarity Triangle.

Recent direct searches for non-SM physics, mainly in pp collisions at the Large Hadron Collider (LHC) at CERN, excluded large portions of the parameter space for several proposed SM extensions, but showed no conclusive evidence of non-SM physics so far. Since plans for a higher-energy collider in the near future are still fluid, the systematic study of flavor physics emerges as a promising program to search for non-SM in the next decade. Despite the first-order consistency of the experimental flavor picture with the CKM theory, possible deviations of up to 10–15% are still unconstrained, especially those associated to loop-mediated processes, leaving sufficient room for non-SM physics. It is especially promising that most of the relevant measurements are currently dominated by statistical uncertainties, offering therefore fruitful opportunities for the two experiments that will contribute the most in the next decade, LHCb and Belle II.

1.6 The angle α

With an uncertainty of about 4%, the CKM angle $\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right)$ is currently among the major limiting factors on the global consistency check of the Unitarity Triangle.

Any B decay induced by a $b \to u\bar{u}d$ transition (decays of charged and neutral B mesons to two, three or four pions) is sensitive to α. Useful decays with three and four pions in the final state occur through intermediate resonances, for example $B \to \rho\pi \to 3\pi$, $B \to \rho\rho \to 4\pi$ and $B \to a_1\pi \to 4\pi$. Figure 1.6 shows the leading-order tree and QCD-penguin Feynman diagrams, together with the color-suppressed EW penguin diagrams, contributing to the decays $B^{i+j} \to h_1^i h_2^j$ of charged and neutral B mesons. The final state $h_1^i h_2^j$ indicates a pair of unflavoured light mesons such that h is π or ρ and i, j is $-$, 0, $+$. A more complete overview of higher-order amplitudes is in Refs. [22, 23].

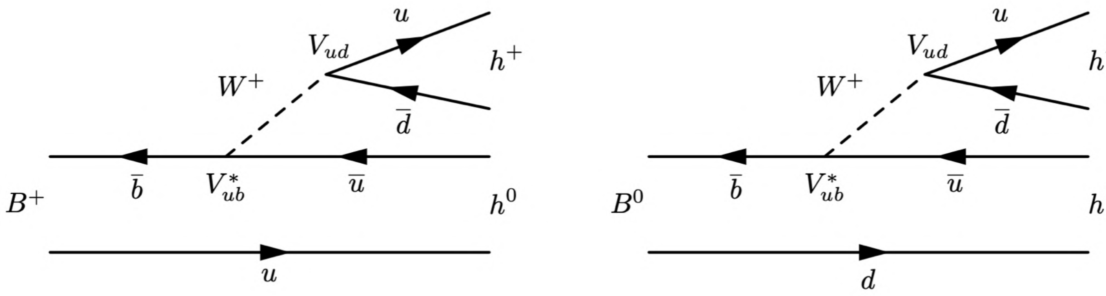
Considering, as an example, the decays of neutral B^0 mesons only, and assuming that only tree-level diagrams contribute (Fig. 1.6a, 1.6b), one obtains

$$\lambda_{hh} = \left(\frac{q}{p}\right)_{B^0}\left(\frac{\bar{A}^{hh}}{A^{hh}}\right) = \frac{V_{tb}^*V_{td}}{V_{tb}V_{td}^*}\frac{V_{ud}^*V_{ub}}{V_{ud}V_{ub}^*} = e^{2i\alpha}. \tag{1.4}$$

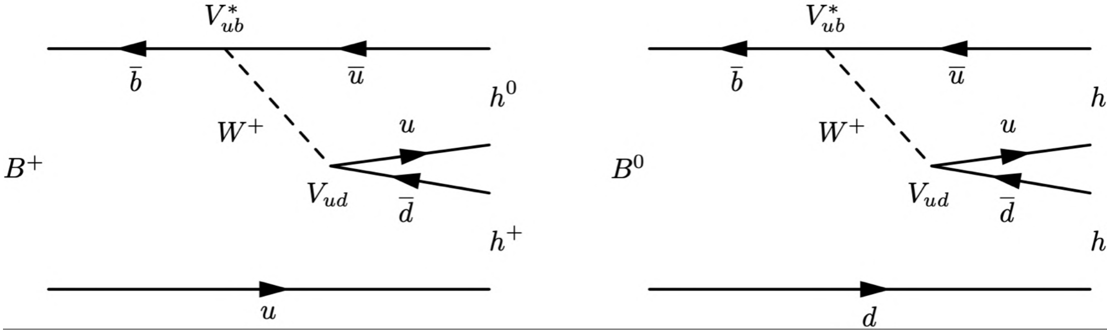
where the mixing and decay amplitudes are transparently shown. The $\frac{V_{ud}^*V_{ub}}{V_{ud}V_{ub}^*}$ term corresponds to the tree level decay amplitude, while $\frac{V_{tb}^*V_{td}}{V_{tb}V_{td}^*}$ term corresponds to the mixing amplitude. In the mixing amplitude, we consider only box diagrams with an internal top quark neglecting those with a charm and up quark, as the mixing amplitude goes with m_{up}^2 (where m_{up} is the mass of the up-type quark contributing to the loop).
Hence, if penguin amplitudes were nonexistent, a measurement of decay-rate asymmetry between flavor-tagged $B^0 \to hh$ decays would offer direct access to α through the corresponding CP-violating coefficients of the cosine and sine terms of the time evolution, $A_{CP} = 0$, $S_{CP} = \eta_{CP}\sin(2\alpha)$. The observed value $A_{CP}^{\rho^+\rho^-} = 0.00 \pm 0.09$ [10] for the decay $B^0 \to \rho^+\rho^-$ shows that the approximation of tree-amplitude dominance might be reasonable to a 10% uncertainty. If large penguin amplitudes would contribute, then the direct CP asymmetry would sense it unless the difference in CP-conserving phases is too small. However, the measured value $A_{CP}^{\pi^+\pi^-} = 0.32 \pm 0.04$ for the decay $B^0 \to \pi^+\pi^-$ differs considerably from zero, revealing that penguin contributions are not negligible in general, and especially so for decays $B^0 \to h^0h^0$, where the leading-order tree contribution is color suppressed.

Penguin contributions have to be considered. Since the fine-structure constant α is smaller than the strong coupling coefficient α_{s}, EW penguin amplitudes are expected to be of $\mathcal{O}(0.1)$ of QCD penguin amplitudes [24, 25] and can be neglected to a good approximation. The $B \to h_1^i h_2^j$ decay amplitude, where CKM coefficients are factorized, reads [26]

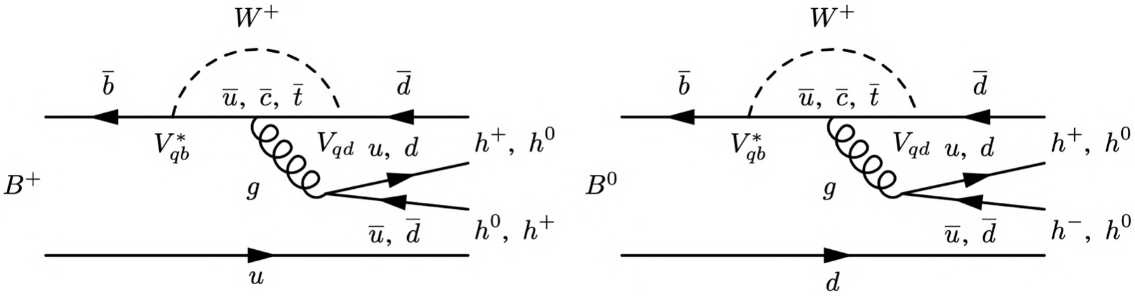
$$A^{ij} = \langle h_1^i h_2^j|\mathcal{H}_{\text{eff}}|B\rangle = V_{ud}V_{ub}^*(T_u^{ij} + P_u^{ij}) + V_{cd}V_{cb}^*P_c^{ij} + V_{td}V_{tb}^*P_t^{ij},$$



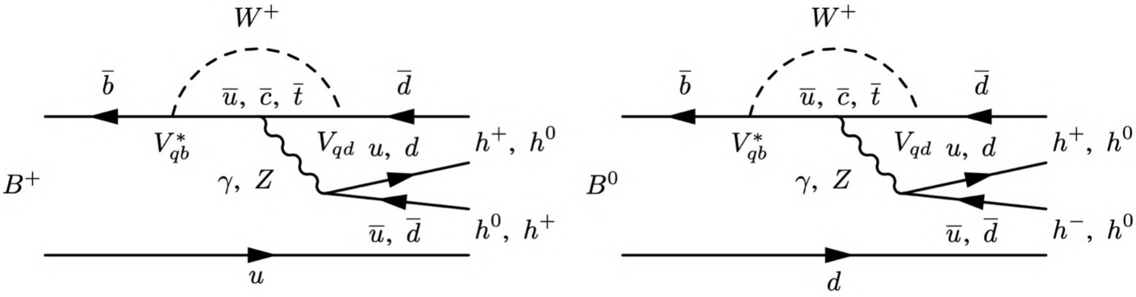
(a) Color-favored tree diagrams T.



(b) Color-suppressed tree diagrams C.



(c) QCD-penguin diagram P.



(d) Color-suppressed EW penguin diagrams P_{EW}^{C}.

Figure 1.6: Dominant tree-level and QCD-penguin Feynman diagrams together with the color-suppressed EW-penguin diagrams contributing to $B^+ \to h^- h^0$ (left) and to $B^0 \to h^+ h^-,\ h^0 h^0$ (right). The generic shorthand h indicates a pion or ρ meson.

where $\mathcal{H}_{\text{eff}}$ is the effective Hamiltonian describing the transition, T_u^{ij} is the hadronic tree amplitude, and P_u^{ij}, P_c^{ij} and P_t^{ij} are the hadronic QCD-penguin amplitudes with quarks u, c, and t in the W loop. Using the unitarity relation of the CKM matrix given in Sec. 1.3, the decay amplitude A^{ij} is rewritten as

$$A^{ij} = V_{ud}V_{ub}^*(T_u^{ij} + P_u^{ij} - P_c^{ij}) + V_{td}V_{tb}^*(P_t^{ij} - P_c^{ij}).$$

By defining the tree and penguin amplitudes as $T_{uc}^{ij} = T_u^{ij} + P_u^{ij} - P_c^{ij}$ and $P_{tc}^{ij} = P_t^{ij} - P_c^{ij}$, one obtains

$$A^{ij} = V_{ud}V_{ub}^* T_{uc}^{ij} + V_{td}V_{tb}^* P_{tc}^{ij}.$$

Including the magnitudes squared of the CKM products in the amplitudes, i.e., using $T_{uc}^{ij} = T^{ij}/|V_{ud}V_{ub}^*|^2$ and $P_{tc}^{ij} = -P^{ij}/|V_{td}V_{tb}^*|^2$, the CP-violating quantity λ_{hh} becomes

$$\begin{aligned}\lambda_{hh} = \frac{q}{p}\cdot\frac{\bar{A}^{ij}}{A^{ij}} &= \frac{V_{tb}^*V_{td}}{V_{tb}V_{td}^*}\frac{\frac{1}{V_{ud}V_{ub}^*}\bar{T}^{ij} - \frac{1}{V_{td}V_{tb}^*}\bar{P}^{ij}}{\frac{1}{V_{ud}^*V_{ub}}T^{ij} - \frac{1}{V_{td}^*V_{tb}}P^{ij}} = \frac{V_{tb}^*V_{td}}{V_{tb}V_{td}^*}\frac{V_{cd}^*V_{cb}}{V_{cd}V_{cb}^*}\frac{\frac{V_{cd}V_{cb}^*}{V_{ud}V_{ub}^*}\bar{T}^{ij} - \frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\bar{P}^{ij}}{\frac{V_{cd}^*V_{cb}}{V_{ud}^*V_{ub}}T^{ij} - \frac{V_{cd}^*V_{cb}}{V_{td}^*V_{tb}}P^{ij}}\\ &= e^{-2i\beta}\frac{e^{-i\gamma}\bar{T}^{ij} - e^{i\beta}\bar{P}^{ij}}{e^{i\gamma}T^{ij} - e^{-i\beta}P^{ij}}.\end{aligned}$$

The amplitudes T^{ij} and P^{ij} can be written in terms of magnitude and hadronic phase, leading to

$$A^{ij} = r^{i\gamma}e^{i\delta_T}|T^{ij}| - e^{-i\beta}e^{i\delta_P}|P^{ij}|.$$

CP invariance of strong interactions requires the hadronic amplitudes T^{ij} and P^{ij} to be invariant under CP transformations. Thus, a CP transformation on A^{ij} applies a complex conjugation on weak phases only. Rotating consistently all CP-transformed amplitudes $\bar{A}^{ij}$ in order to absorb the mixing phase, i.e., $\tilde{A}^{ij} = e^{-2i\beta}\bar{A}^{ij}$, the ratio between amplitudes $\tilde{A}^{ij}$ and A^{ij} yields

$$\frac{\tilde{A}^{ij}}{A^{ij}} = \lambda_{hh} = \frac{e^{-i(\beta+\gamma)}e^{i(\delta_T-\delta_P)}|T^{ij}| - |P^{ij}|}{e^{i(\beta+\gamma)}e^{i(\delta_T-\delta_P)}|T^{ij}| - |P^{ij}|} = \frac{e^{i(\delta+\alpha)}|T^{ij}| + |P^{ij}|}{e^{i(\delta-\alpha)}|T^{ij}| + |P^{ij}|}, \tag{1.5}$$

where $\delta = \delta_T - \delta_P$ and $\alpha - \pi = -\beta - \gamma$ (see Fig. 1.2). If penguin contributions are negligible, i.e., $|P^{ij}| \approx 0$, one obtains the Eq. (1.4) result for λ_{hh}. If penguin contributions are non-negligible, a direct determination of α is not possible, and an effective angle α_{eff}, also referred to as penguin *polluted* angle,

$$\frac{\tilde{A}^{ij}}{A^{ij}} = \lambda_{hh} = \left|\frac{\bar{A}^{ij}}{A^{ij}}\right| e^{2i\alpha_{\text{eff}}}, \tag{1.6}$$

is observed.

In 1990, Michael Gronau and David London proposed to determine the size of penguin pollution from data using known symmetry relations among the hadronic amplitudes [27], based on a $SU(2)$ isospin analysis of $B \to \pi\pi$ decays. A later review clarified that the method applies to $B \to \rho\rho$ and $B \to \rho\pi$ too [28].

A further approach to determine the penguin contribution is to exploit the (approximate) $SU(3)$ flavor symmetry between the decays $B \to K\pi$, $B \to KK$, and $B \to \pi\pi$ [29, 30], or, similarly, for the decays $B \to K^*\rho$ and $B \to \rho\rho$ [31]. For the $B \to \rho\pi$ system, isospin symmetry in a Dalitz plot analysis, and flavour symmetry $SU(3)$ have been applied to determine α [32, 33]. A treatment based on $SU(3)$ was proposed also for the $B \to a_1(1260)\pi$ system [34]. Other approaches consider the extraction of α from

$B \to a_0(980)\pi \to \eta\pi\pi$, $B \to a_0(980)\rho \to \eta\pi\pi\pi$, and other B decays to resonances with isospin $I = 1$ [35]. However, the uncertainties associated with the assumption of $SU(3)$ flavor symmetry are large and poorly known, making those methods not relevant for CKM parameters estimation.

1.6.1 Current status

Currently, the precision in the determination of $\alpha = (86.4^{+4.3}_{-4.0})^\circ$ is dominated by the $B \to \rho\rho$ decays, and to a lesser extent, by the $B \to \pi\pi$ system [26].

In the following, the determination of α based on an isospin analyses of the systems $B \to \pi\pi$ and $B \to \rho\rho$ is outlined.

1.6.2 Isospin analysis

The π and ρ mesons are, respectively, the lightest pseudoscalar (spin-0) and vector (spin-1) mesons. Both particles form $SU(2)$ isospin triplets ($I = 1$):

- $\pi(u\overline{d})$, $\pi^0(\frac{u\overline{u}-d\overline{d}}{\sqrt{2}})$, $\pi^-(\overline{u}d)$ with $I_3 = -1, 0, +1$, and
- $\rho^+(u\overline{d})$, $\rho^0(\frac{u\overline{u}-d\overline{d}}{\sqrt{2}})$, $\rho^-(\overline{u}d)$ with $I_3 = -1, 0, +1$.

The validity of the Gronau-London isospin analysis relies on the near exact conservation of isospin symmetry during the hadronization in $B \to \pi\pi$ and $B \to \rho\rho$ decays. Thus, relations between decay amplitudes derived from the isospin symmetry are used to separate the effects of the tree and the QCD penguin contributions to obtain the unpolluted value of α [27].

Given a final state $h_1^i h_2^j$ with two identical mesons h_1, $h_2 = h = \pi$ or ρ, the decay amplitude A^{+0}, A^{+-} and A^{00} is written as

$$A^{ij} \equiv \langle h^i h^j | \mathcal{H}_{\text{eff}} | B^{i+j} \rangle. \tag{1.7}$$

Since the final-state pions and ρ mesons are bosons, the total wave function of the final state $h_1^i h_2^j$ must be symmetric under particle exchange. For $i \neq j$, the symmetrized final states are

$$|h^i h^j\rangle = \sqrt{\frac{1}{2}} \left(|h_1^i h_2^j\rangle + |h_1^j h_2^i\rangle \right).$$

Since B mesons are spin-0 particles, the total angular momentum of the final state is $J = 0$. For a single π or ρ the isospin is 1. Spin sum rules dictate that the total final-state isospin I_{f} for a pair $\pi\pi$ or $\rho\rho$ can be 0, 1, or 2. However, due to Bose statistics only final states with $I_{\text{f}} = 0$ or 2 are allowed. A final state $I_{\text{f}} = 1$ would be antisymmetric as prescribed by the symmetry $(-1)^{J+I}$ under particle exchange [28]. This is exact for final-state particles with equal masses. In the case of ρ mesons, which have a significant width, the possible mass difference between the final-state particles could generate a final state with $I_{\text{f}} = 1$ [36]. Nevertheless, lack of a specific enhancement of the $I_{\text{f}} = 1$ amplitude shows that the results for α are insensitive to the ρ width [37]. Determination of any possible $I_{\text{f}} = 1$ contribution requires more data and is neglected in this discussion. The three relevant final states for

$I_{\mathrm{f}} = 0$ or 2 are

$$|h^+h^0\rangle = \sqrt{\frac{1}{2}}\left(|h_1^+h_2^0\rangle + |h_1^0h_2^+\rangle\right) = |2,1\rangle, \tag{1.8}$$

$$|h^+h^-\rangle = \sqrt{\frac{1}{2}}\left(|h_1^+h_2^-\rangle + |h_1^+h_2^-\rangle\right) = \sqrt{\frac{1}{3}}|2,0\rangle + \sqrt{\frac{2}{3}}|0,0\rangle, \tag{1.9}$$

$$|h^0h^0\rangle = \sqrt{\frac{2}{3}}|2,0\rangle - \sqrt{\frac{1}{3}}|0,0\rangle. \tag{1.10}$$

The decay amplitudes in Eq. (1.7) can be factorized in the weak decay $b \to u\overline{u}d$ corresponding to an isospin transition ΔI, and the hadronization into two light mesons [26]. Following Ref. [28], the Hamiltonian for the quark transition $\overline{b} \to \overline{u}u\overline{d}$, in terms of $A_{\Delta I}$ amplitudes, is of the form

$$\mathcal{H}_{\text{eff}} = A_{\frac{3}{2}}\left|\frac{3}{2},+\frac{1}{2}\right\rangle + A_{\frac{1}{2}}\left|\frac{1}{2},+\frac{1}{2}\right\rangle. \tag{1.11}$$

Application of the operator to the initial states B^+ and B^0 yields

$$\mathcal{H}_{\text{eff}}|B^+\rangle = \mathcal{H}_{\text{eff}}|\frac{1}{2},+\frac{1}{2}\rangle = \sqrt{\frac{3}{4}}A_{\frac{3}{2}}|2,1\rangle + (A_{\frac{1}{2}} - \frac{1}{2}A_{\frac{3}{2}})|1,1\rangle, \tag{1.12}$$

$$\mathcal{H}_{\text{eff}}|B^0\rangle = \mathcal{H}_{\text{eff}}|\frac{1}{2},-\frac{1}{2}\rangle = \sqrt{\frac{1}{2}}A_{\frac{3}{2}}|2,0\rangle + \sqrt{\frac{1}{2}}(A_{\frac{1}{2}} + A_{\frac{3}{2}})|1,0\rangle + \sqrt{\frac{1}{2}}A_{\frac{1}{2}}|0,0\rangle. \tag{1.13}$$

The equations above correspond to four-quark states $|\overline{u}u\overline{d},q\rangle$ with $q = u$ or d, while the states in Eq. (1.8)–(1.10) are two-meson states. The transition between the two involves hadronization and other rescattering effects.

Following the Wigner-Eckhart theorem, these amplitudes can be expressed in terms of reduced matrix elements $A_{\Delta I,I_{\mathrm{f}}}$, where ΔI is the isospin shift and I_{f} is the final-state isospin. Hadronization and strong rescattering effects that are absent in the amplitudes $A_{\Delta I}$ are included into the amplitudes $A_{\Delta I,I_{\mathrm{f}}}$. The projection of the states $\mathcal{H}_{\text{eff}}|B^{i+j}\rangle$ onto the final states yields

$$\overset{(-)}{A}{}^{+0} = \langle h^{+(-)}h^0|\mathcal{H}_{\text{eff}}|B^{+(-)}\rangle = \sqrt{\frac{3}{4}}\overset{(-)}{A}_{\frac{3}{2},2}, \tag{1.14}$$

$$\overset{(-)}{A}{}^{+-} = \langle h^+h^-|\mathcal{H}_{\text{eff}}|\overset{(-)}{B^0}\rangle = \sqrt{\frac{1}{6}}\overset{(-)}{A}_{\frac{3}{2},2} + \sqrt{\frac{1}{3}}\overset{(-)}{A}_{\frac{1}{2},0}, \tag{1.15}$$

$$\overset{(-)}{A}{}^{00} = \langle h^0h^0|\mathcal{H}_{\text{eff}}|\overset{(-)}{B^0}\rangle = \sqrt{\frac{1}{3}}\overset{(-)}{A}_{\frac{3}{2},2} - \sqrt{\frac{1}{6}}\overset{(-)}{A}_{\frac{1}{2},0}. \tag{1.16}$$

The CP-conjugated amplitudes $\bar{A}_{\Delta I,I_{\mathrm{f}}}$ carry strong phases identical to those of $A_{\Delta I,I_{\mathrm{f}}}$, but opposite weak phases. Eqs. (1.14–1.16) yields two relations among the hadronic amplitudes

$$\begin{aligned} A^{+0} - A^{00} &= \sqrt{\frac{1}{2}}A^{+-}, \\ \bar{A}^{+0} - \bar{A}^{00} &= \sqrt{\frac{1}{2}}\bar{A}^{+-}, \end{aligned} \tag{1.17}$$

which allow extracting the penguin pollution. These are referred to as isospin triangles since they can be represented as triangles in the complex space (Fig. 1.7). The triangle relations hold also for consistently rotated amplitudes. I use the convention $\tilde{A}^{ij} = e^{-2i\beta}\bar{A}^{ij}$. Strong penguin amplitudes can lead only to $\Delta I = \frac{1}{2}$ transitions. Since the amplitudes in Eq. (1.14) lack $\Delta I = \frac{1}{2}$ components, the decay $B^{\pm} \to h^{\pm}h^{0}$ occurs purely as a tree amplitude. Using Eq. (1.5), one obtains

$$\frac{\tilde{A}^{+0}}{A^{+0}} = \frac{e^{i(\delta+\alpha)}}{e^{i(\delta-\alpha)}} = e^{2i\alpha} \quad \Rightarrow \quad A_{CP}^{h^+h^-} = 0. \tag{1.18}$$

In this case, the imaginary component of the ratio cannot be measured since no mixing-induced CP parameter is present. However, the measured CP parameters $A_{CP}^{\pi^+\pi^0} = 0.03 \pm 0.04$ and $A_{CP}^{\rho^+\rho^0} = -0.05 \pm 0.05$ agree with zero [10]. Therefore, $A_{CP}^{h^+h^-}$ is not taken into account for the determination of α. Using Eq. (1.6) for the penguin polluted decay $B^0 \to h^+h^-$, the CKM triangle relations and the relation in Eq. (1.18), yields all ingredients to determine α through isospin symmetry. By convention, the A^{+-} amplitude is chosen to be real; yielding [38],

$$\begin{aligned}
A^{+0} &= |A^{+0}|e^{i(\delta-\alpha)}, & \tilde{A}^{+0} &= |A^{+0}|e^{i(\delta+\alpha)}, \\
A^{+-} &= |A^{+-}|, & \tilde{A}^{+-} &= |\bar{A}^{+-}|e^{2i\alpha_{\text{eff}}}, \\
A^{00} &= A^{+0} - \frac{A^{+-}}{\sqrt{2}} = e^{i(\delta-\alpha)}\left(|A^{+0}| - \frac{|A^{+-}|}{\sqrt{2}}e^{-i(\delta-\alpha)}\right), \\
\tilde{A}^{00} &= \tilde{A}^{+0} - \frac{\tilde{A}^{+-}}{\sqrt{2}} = e^{i(\delta+\alpha)}\left(|A^{+0}| - \frac{|\bar{A}^{+-}|}{\sqrt{2}}e^{-i(\delta+\alpha-2\alpha_{\text{eff}})}\right).
\end{aligned} \tag{1.19}$$

For the branching fractions $\mathcal{B}^{ij}$ and the CP parameters A_{CP}^{ij} and S_{CP}^{ij},

$$\begin{aligned}
\mathcal{B}^{+0} &= \tau_{B^+}|A^{+0}|^2, & \mathcal{B}^{+-} &= \tau_{B^0}\frac{|A^{+-}|^2 + |\bar{A}^{+-}|^2}{2}, \\
\mathcal{B}^{00} &= \tau_{B^0}\left(|A^{+0}|^2 + \frac{|A^{+-}|^2 + |\bar{A}^{+-}|^2}{4} - \frac{|A^{+0}|}{\sqrt{2}}\left(|A^{+-}|c + |\bar{A}^{+-}|\bar{c}\right)\right), \\
A_{CP}^{+-} &= \frac{|\bar{A}^{+-}|^2 - |A^{+-}|^2}{|\bar{A}^{+-}|^2 + |A^{+-}|^2}, & S_{CP}^{+-} &= \frac{2|\bar{A}^{+-}||A^{+-}|\sin(2\alpha_{\text{eff}})}{|\bar{A}^{+-}|^2 + |A^{+-}|^2}, \\
A_{CP}^{00} &= \frac{|\bar{A}^{+-}|^2 - |A^{+-}|^2 - 2\sqrt{2}|A^{+0}|\left(|\bar{A}^{+-}|\bar{c} - |A^{+-}|c\right)}{4|A^{+0}|^2 + |\bar{A}^{+-}|^2 + |A^{+-}|^2 - 2\sqrt{2}|A^{+0}|\left(|\bar{A}^{+-}|\bar{c} + |A^{+-}|c\right)},
\end{aligned} \tag{1.20}$$

where $c = \cos(\alpha - \delta)$ and $\bar{c} = \cos(\alpha + \delta - 2\alpha_{\text{eff}})$. The lifetimes of charged and neutral B mesons, τ_{B^+} and τ_{B^0}, differ and are included in the branching fractions. If all branching fractions and CP-violation parameters in Eq. (1.20) are measured, the system has six linear independent equations with six real positive variables: $|A^{+0}|$, $|A^{+-}|$, $|\bar{A}^{+-}|$, δ, α_{eff}, and α. The value of α is determined up to an eight-fold ambiguity in the range $[0, \pi]$ because each isospin triangle can have two possible orientations, leading to a four-fold trigonometric ambiguity

$$(\alpha, \delta) \leftrightarrow (\delta, \alpha),\ (2\alpha_{\text{eff}} - \alpha, 2\alpha_{\text{eff}} - \delta),\ (2\alpha_{\text{eff}} - \delta, 2\alpha_{\text{eff}} - \alpha). \tag{1.21}$$

An additional symmetry exists involving also α_{eff} that prescribes the functions c, $\bar{c}$, and S_{CP}^{+-} to be invariant under the reflection

$$(\alpha_{\text{eff}}, \alpha, \delta) \leftrightarrow \left(\frac{\pi}{2} - \alpha_{\text{eff}},\ \frac{\pi}{2} - \alpha,\ \frac{\pi}{2} - \delta\right). \tag{1.22}$$

An illustration of the eight-fold α ambiguity is shown in Figure 1.7.

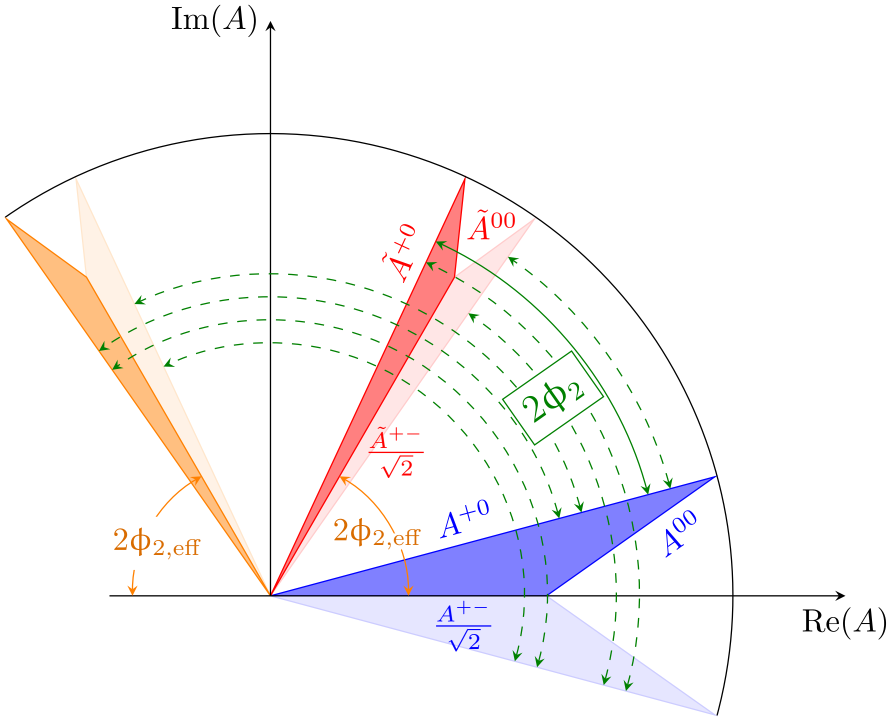


Figure 1.7: Geometrical representation of the isospin triangular relations (Eq. (1.17)) in the complex plane of $B^{i+j} \to h^i h^j$ amplitudes. The blue and the red shaded areas correspond to the isospin triangles. The angle between the CP conjugate charged amplitudes A^{+-} and $\tilde{A}^{+-}$ corresponds to twice the weak phase α_{eff} (orange solid lines). The angle between the CP conjugate charged amplitudes A^{+0} and $\tilde{A}^{+0}$ corresponds to twice the CKM angle α (green solid line). The other triangles with lighter shade represent the mirror solutions allowed by the discrete ambiguities in Eq. (1.20), with the corresponding values for α represented by the green dashed lines.

1.6.2.1 The S_{CP}^{00} constraint

An additional constraint can partially lift the eight-fold ambiguity in the determination of α, the mixing-induced CP-violation parameter of the decay $B^0 \to h^0 h^0$,

$$S_{CP}^{00} = \frac{4|A^{+0}|^2 \sin(2\alpha) + 2|A^{+-}||\bar{A}^{+-}|\sin(2\alpha_{\text{eff}}) - 2\sqrt{2}|A^{+0}|\left(|\bar{A}^{+-}|\bar{s} + |A^{+-}|s\right)}{4|A^{+0}|^2 + |\bar{A}^{+-}|^2 + |A^{+-}|^2 - 2\sqrt{2}|A^{+0}|\left(|\bar{A}^{+-}|\bar{c} + |A^{+-}|c\right)}, \quad (1.23)$$

where $s = \sin(\alpha+\delta)$ and $\bar{s} = \sin(\alpha - \delta + 2\alpha_{\text{eff}})$ are not invariant under the transformations of (α, δ) in Eq. (1.21), thus fixing the orientation of each isospin triangle.

No measurement of S_{CP}^{00} has yet been reported for the decay $B^0 \to \pi^0\pi^0$. This is a formidable challenge because the time-dependent analysis of $B^0 \to \pi^0\pi^0$ requires a precise reconstruction of the B^0-decay vertex, which cannot be achieved in the dominant four-photons final state.

For the decay $B^0 \to \rho^0\rho^0$, a single measurement of $S_{CP}^{\rho^0\rho^0}$ is available by the BaBar experiment with large uncertainties $S_{CP}^{\rho^0\rho^0} = 0.3 \pm 0.7 \pm 0.2$ [39]. However, for $B \to \rho\rho$, the sides A^{00} and $\tilde{A}^{00}$ of the isospin triangles are much smaller than the other sides, and the triangles are squashed into lines. The eight-fold ambiguity is therefore already reduced by a factor four.

1.6.2.2 Comments on $B \to \pi\pi$ and $B \to \rho\rho$ phenomenology

A major difference between the $\pi\pi$ and the $\rho\rho$ final states is that $\rho\rho$ dynamics is further enriched by angular momentum degrees of freedom. The CP eigenvalue of the final state $h^i h^j$ is given by $\eta_{CP} = (-1)^L$. In the case of $\pi\pi$, the CP eigenvalue is even: the total spin S is zero, and the orbital angular momentum L has to be zero to conserve the initial $J = L + S = 0$. The pair of spin-1 ρ mesons can have spin configurations $S = 0$, 1, and 2. The conservation of $J = 0$ imposes the orbital angular momentum L to be equal and oppositely aligned to S, leading to $L = 0$, 1 or 2; where $L = 0$ and 2 correspond to CP-even, and $L = 1$ to CP-odd final states. Thus, independent isospin analyses have to be performed for the CP-even and the CP-odd modes, requiring the branching fractions and the CP asymmetries in Eq. (1.20)–(1.23) to be measured in both cases.

Angular analyses of the decays $B^0 \to \rho^+\rho^-$ and $B^+ \to \rho^+\rho^0$ showed that the fraction f_{L} of decays leading to longitudinally polarized mesons $\rho_{\mathrm{L}}\rho_{\mathrm{L}}$, i.e., with $L = 0$, dominates over a negligible fraction of events leading to final states with $L \neq 0$ [39–42]. The current world-average values are $f_{\mathrm{L},\rho^+\rho^0} = 0.950 \pm 0.016$ and $f_{\mathrm{L},\rho^+\rho^-} = 0.990^{+0.021}_{-0.019}$ [10]. Consequently, the isospin analysis of $B \to \rho\rho$ is performed only for decays leading to longitudinally polarized ρ_{L} mesons.

Chapter 2

The Belle II experiment at the SuperKEKB collider

The data used in this work were collected by the Belle II experiment. This chapter outlines the Belle II detector at the SuperKEKB accelerator, with the subdetectors more relevant for the reconstruction of $B^+ \to \rho^+\rho^0$ decays.

2.1 The SuperKEKB collider

SuperKEKB is an electron-positron (e^+e^-) energy-asymmetric collider, designed to produce more than 600 $B\overline{B}$ pairs per second via decays of $\Upsilon(4S)$ mesons produced at threshold [43] ($B^0\overline{B}^0$ and B^+B^- in approximately equal proportions). Such colliders are called 'B-factories', and were proposed in the 1990's for the dedicated exploration of CP-violation in B mesons. The main goal of B-factories is to produce low-background quantum-correlated $B\overline{B}$ pairs at high rates.

Intense beams of electrons and positrons are brought to collision at the energy corresponding to the $\Upsilon(4S)$ meson mass, 10.58 GeV, which is just above the $B\overline{B}$ production kinematic threshold. The great majority of collisions yields electroweak processes ($e^+e^- \to e^+e^-$, $e^+e^- \to \gamma\gamma$, etc.) that are scarcely interesting and straightforwardly discarded using global event quantities (see Fig. 2.1). More interesting for flavor physics are the collisions that produce hadrons (hadronic events). In these, the finely tuned collision energy is key. The production of $\Upsilon(4S)$ mesons, which decay in $B\overline{B}$ pairs 96% of the times with little available energy to produce additional particles, suppresses backgrounds from competing nonresonant hadron production. In addition, colliding beams of point-like particles allow for knowing precisely the collision energy, which sets stringent constraints on the collision's kinematic properties, thus offering means of further background suppression. Since bottom mesons are produced in a strong-interaction decay, flavor is conserved, and the null net bottom content of the initial state implies production of a flavorless $B\overline{B}$ pair. Even though B^0 and $\overline{B}^0$ undergo flavor oscillations before decaying, their time-evolution is quantum-correlated in such a way that no B^0B^0 or $\overline{B}^0\overline{B}^0$ pairs are present at any time. Angular-momentum conservation implies that the decay of a spin-1 particle in two spin-0 particles yields total angular momentum $J = 1$. Because the simultaneous presence of two identical particles in an antisymmetric state would violate Bose statistics, the system evolves coherently as an oscillating $B^0\overline{B}^0$ particle-antiparticle pair until either one decays. This allows identification of the bottom (or antibottom) content of one meson at the time of decay of the other, if the latter decays in a final state accessible only by either bottom or antibottom states. This important capability is called 'flavor tagging' and allows measurements of flavor-dependent decay rates, as needed in many determinations of CP-violating quantities.

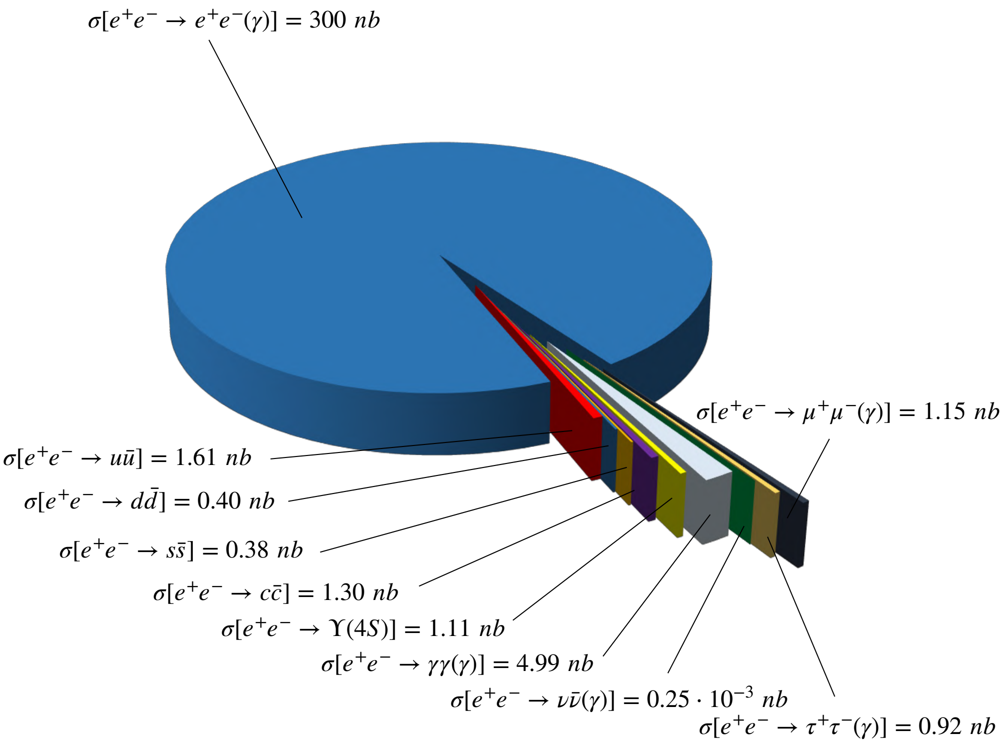


Figure 2.1: Cross sections of the main final states produced in e^+e^- collision at the $\Upsilon(4\mathrm{S})$ center-of-mass energy.

Figure 2.2 shows the hadron-production cross-section in e^+e^- collisions as functions of the final-state mass. The various peaks are radial excitations of the Υ meson overlapping the nearly uniform background at about 4 nb which represents the so-called continuum of lighter-quark pair-production, $e^+e^- \to q\overline{q}$, where q identifies a u, d, c, or s quark.

Because the $\Upsilon(4\mathrm{S})$ mesons are produced at threshold, they would be nearly at rest in the laboratory frame in an energy-symmetric collider. The resulting B mesons too would be produced with low momentum ($\approx 10\,\mathrm{MeV}/c$) in the laboratory, because of the $21\,\mathrm{MeV}/c^2$ difference between the $\Upsilon(4\mathrm{S})$ mass and the $B\overline{B}$ pair mass. With such low momenta they would only travel approximately $1\,\mu$m before decaying rendering the $10\ \mu$m typical spatial resolution of vertex detectors insufficient to separate B-decay vertices and enable the study of the decay-time evolution. Asymmetric beam energies are used to circumvent this limitation. By boosting the collision center-of-mass along the beam in the laboratory frame, they achieve B-decay vertex separations resolvable with current vertex detectors [44]. SuperKEKB (Fig. 2.3) implements a 7–4 GeV energy-asymmetric double-ring design, which achieves a vertex displacement of about $130\,\mu$m.

Electrons are produced in a thermionic gun with a barium-impregnated tungsten cathode, then accelerated to 7 GeV with a linear accelerator (linac) and injected in the high-energy ring (HER). Positrons are produced by colliding electrons on a tungsten target, then isolated by a magnetic field, accelerated to 4 GeV with the linac and injected in the low-energy ring (LER).

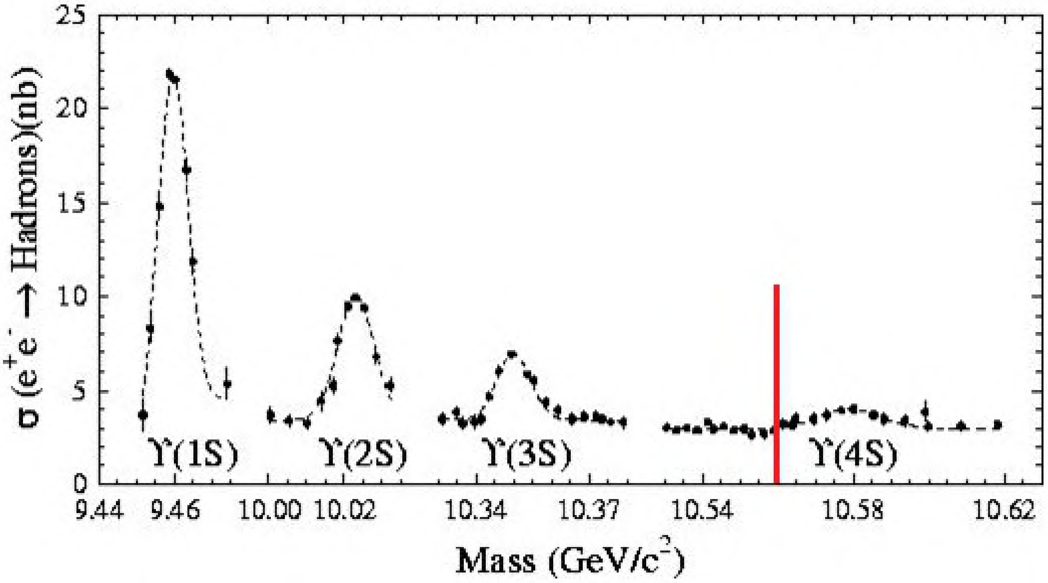


Figure 2.2: Hadron production cross section from e^+e^- collisions as a function of the final-state mass. The vertical red line indicates the $B\overline{B}$ production threshold.

The electrons and positrons continuously collide at a single interaction point, around which the Belle II detector is installed. To achieve high luminosities, a nano-beam, large crossing-angle collision scheme is implemented [45]. This is an innovative configuration based on keeping small horizontal and vertical emittance and large crossing angle, as shown in Figure 2.4. This is obtained with a final-focus superconducting-quadrupole-magnet system (QCS), made of magnets, corrector coils, and compensation solenoids; a QCS magnet is installed at each longitudinal end of the interaction region. Conceptually the nano-beam scheme mimics a collision with many short micro-bunches, allowing great advantages in luminosity with respect to previous standard schemes. The reduction of the luminous volume size to about 5% with respect to the predecessor KEKB, combined with doubling of beam currents, is expected to yield a factor 40 gain in intensity.

The performance of the SuperKEKB collider is characterized in terms of the luminosity $\mathcal{L}$, which is a measure of collision intensity. The rate of any given process

$$\text{rate [events s}^{-1}] = \mathcal{L}\ [\text{cm}^{-2}\ \text{s}^{-1}] \times \sigma\ [\text{cm}^2],$$

is the product of its cross-section, σ, and the instantaneous luminosity $\mathcal{L}$,

$$\mathcal{L} = \frac{\gamma_\pm}{2er_e}\left(1+\frac{\sigma_y^*}{\sigma_x^*}\right)\frac{I_\pm \xi_{y\pm}}{\beta_{y\pm}^*}\cdot\frac{R_\mathcal{L}}{R_{\xi_y}},$$

where γ is the relativistic Lorentz factor, e is the absolute value of the electron charge, r_e is the classical radius of electron, σ_x^* and σ_y^* are the widths of the bunch at IP on the transvere plane, I is the current of the beam, β_y^* is the vertical betatron function at the IP, ξ_y is the vertical beam-beam parameter, $R_\mathcal{L}$ and R_{ξ_y} are the reduction factors of luminosity and the vertical beam-beam parameter due to non-vanishing crossing angle. The ratio of these reduction factors is close to unity, while the design values for the other parameters are reported in Table 2.1.

The integral of instantaneous luminosity over time T, called integrated luminosity,

$$\mathcal{L}_{int} = \int_0^T \mathcal{L}(t')dt'$$

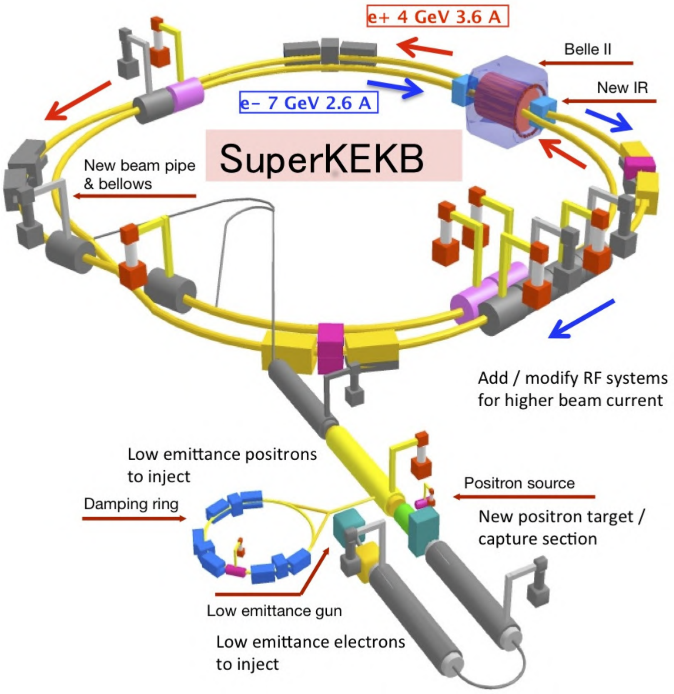


Figure 2.3: Illustration of the SuperKEKB collider.

	Design	Achieved
Energy [GeV]	4.0/7.0	4.0/7.0
ξ_y	0.090/0.088	0.0407/0.0279
β_y^* [mm]	0.27/0.41	1.0/1.0
I [A]	3.6/2.62	1.321/1.099

Table 2.1: Design and achieved values for SuperKEKB fundamental parameters (LER/HER).

is a direct measure of the number of produced events of interest $N = \mathcal{L}_{int} \times \sigma$. Physics data-taking started in March 2019, and Belle II to date has integrated 424 ± 3 fb^{-1} of luminosity. In 2022, SuperKEKB also broke the instantaneous-luminosity world record, achieving $4.7{\times}10^{34}$ cm^{-2}s^{-1}. In spite of these achievements, a number of technological and scientific challenges have significantly reduced SuperKEKB performance compared to design. A number of issues associated with beam injection, collimation, and the capability to reduce the transverse dimensions of the beams without generating uncontrollable beam backgrounds limited the capability to deliver the expected samples of data in its first three years. Consolidation, improvement and development work is ongoing to overcome these difficulties.

2.2 The Belle II detector

Belle II (Fig. 2.5) is a large-solid-angle, approximately cylindrical multi-purpose magnetic spectrometer surrounded by a calorimeter and particle-identification systems, installed around the SuperKEKB interaction point. It is designed to determine energy, momentum,

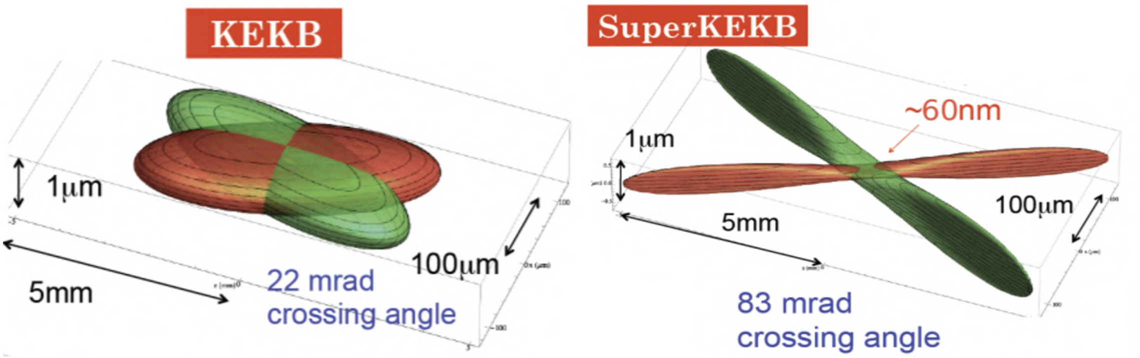


Figure 2.4: Two-dimensional sketch of the nano-beam mechanism implemented in SuperKEKB (right) compared with the previous KEKB collision scheme (left).

and identity a broad range of particles produced in 10.58 GeV e^+e^- collisions. Belle II is approximately a cylinder of about 7 m in length and 7 m in diameter, composed of several subsystems, each dedicated to a specific aspect of event reconstruction. Starting from the interaction point, an emerging particle would traverse the beryllium pipe, a two-layer silicon-pixel vertex-detector (PXD), a four-layer silicon-strip vertex-detector (SVD), a central wire drift-chamber (CDC), aerogel threshold forward Cherenkov counter (ARICH), time-of-propagation Cherenkov counters (TOP), an array of CsI(Tl) crystals (ECL), a superconducting solenoidal magnet, and multiple layers of resistive plate counters (KLM). The main technological specifications of the Belle II subsystems are summarized in the Table 2.2.

The main experimental strengths of the Belle II detector are a highly hermetic coverage; robust and precise reconstruction of charged-particle trajectories (tracks), which provides precisely reconstructed decay-vertices and high momentum resolution; high-efficiency charged-particle identification and neutral-particle reconstruction. A detailed description of Belle II and its performance is in Ref. [46]. In the following, I concentrate on the tracking, electromagnetic calorimeter, and particle identification systems, which are the aspects of the detector more relevant for the analysis reported in this thesis.

2.2.1 Coordinates

Belle II employs a right-handed Cartesian coordinate system with origin in the interaction point. The z-axis is parallel to the electron-beam direction, which is parallel to the magnetic field within the solenoid; the y-axis points vertically upward, and the x-axis is horizontal and pointing outward of the accelerator tunnel. The polar angle, θ, is referred to the positive z-axis. The azimuthal angle, ϕ, is referred to the positive x-axis in the xy-plane. The radius, $r = \sqrt{x^2 + y^2}$, is defined in a cylindrical coordinate system and measured from the origin in the xy-plane. Throughout this work, *longitudinal* means parallel to the electron beam direction (to the z-axis), and *transverse* means perpendicular to the electron beam direction, i.e., in the xy-plane.

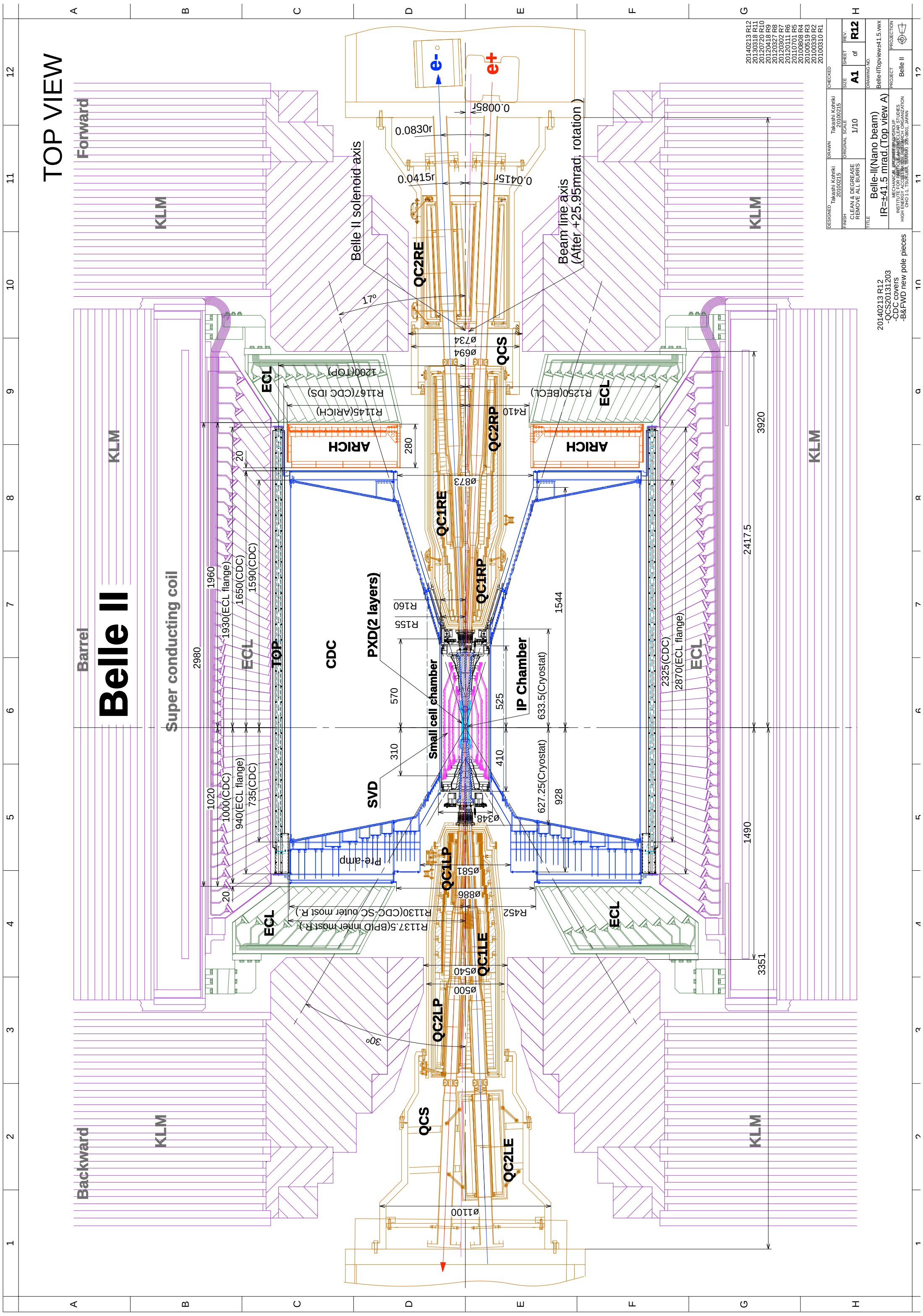


Figure 2.5: Top view of Belle II, the beam pipe at IP and final-focus magnets, all subdetectors are highlighted.

Purpose	Name	Component	Configuration	Readout channels	θ coverage
Beam pipe	Beryllium		Cylindrical, inner radius 10 mm, 10 μm Au, 0.6 mm Be, 1 mm paraffin, 0.4 mm Be		
Tracking	PXD	Silicon pixel (DEPFET)	Sensor size: 15×(L1 136, L2 170) mm^2, Pixel size: 50×(L1a 50, L1b 60, L2a 75, L2b 85) μm^2; two layers at radii: 14, 22 mm	10^6	[17°;150°]
	SVD	Silicon strip	Rectangular and trapezoidal, strip pitch: 50(p)/160(n) - 75(p)/240(n) μm, with one floating intermediate strip; four layers at radii: 38, 80, 115, 140 mm	2.45×10^5	[17°;150°]
	CDC	Drift chamber with He-C_2H_6 gas	14336 wires in 56 layers, inner radius of 160mm outer radius of 1130 mm	1.4×10^5	[17°;150°]
Particle ID	TOP	RICH with quartz radiator	16 segments in ϕ at $r \approx 120$ cm, 275 cm long, 2cm thick quartz bars with 4×4 channel MCP PMTs	8×10^3	[31°;128°]
	ARICH	RICH with aerogel radiator	2×2 cm thick focusing radiators with different n, HAPD photodetectors	7.8×10^4	[14°;30°]
Calorimetry	ECL	CsI(Tl)	Barrel: $r = 125 - 162$cm, end-cap: $z = -102 - +196$cm	6624 (Barrel), 1152 (FWD), 960 (BWD)	[12.4°;31.4°], [32.2°;128.7°], [130.7°;155.1°]
Muon ID	KLM	barrel:RPCs and scintillator strips	2 layers with scintillator strips and 12 layers with 2 RPCs	θ 1.6×10^4, ϕ 1.6×10^4	[40°;129°]
	KLM	end-cap: scintillator strips	12 layers of (7-10)×40 mm^2 strips	1.7×10^4	[25°;40°], [129°;155°]

Table 2.2: Summary of the Belle II components and specifications.

2.2.2 Beam pipe

The beam pipe is a vacuum enclosure where the beams circulate inside the detector. Multiple Coulomb scattering of the charged particles in the beam-pipe wall would spoil the vertex-position resolution; this dictates a thin beam-pipe wall made of a low-Z material. Moreover, since the vertex resolution is roughly inversely proportional to the distance between the interaction point and the first track sampling, the beam pipe has to be narrow and as close as possible to the interaction point. Beam-induced backgrounds and heating of the pipe wall complicates such positioning. Hence, the beam pipe should be constantly cooled and shielded from the vertex detector. The Belle II detector beam pipe is made of two beryllium cylinders, 0.6 mm thick at radius of 10 mm and 0.4 mm thick at radius of 12 mm, respectively. A 1.0 mm gap between the inner and outer walls of the pipe is filled with helium for cooling. The beam pipe is coated with a 10 μm gold sheet that absorbs low-energy photons, which could damage the silicon detector.

2.2.3 Tracking system

Ability to reconstruct efficiently charged-particle trajectories (tracks) and measure precisely their momenta are essential for Belle II physics since most decays of interest, as the $B^+ \to \rho^+\rho^0$ decay that is the object of this thesis, involve charged particles in the final state. At Belle II, tracking is achieved through an integrated system consisting of up to six layers of silicon and a drift chamber, all immersed in a super-conducting axial magnetic field.

2.2.3.1 Magnet

A 1.5 T axial magnetic field is maintained in a cylindrical volume of 3.4 m diameter and 4.4 m length. The field is oriented along the z-direction and provided by an aluminum-stabilized superconducting solenoid made of NbTi/Cu alloy. The main parameters of the coil are summarized in Table 2.3. The solenoid surrounds all the sub-detectors up to the KLM. The iron yoke of the detector serves as the return path of magnetic flux.

2.2.3.2 Silicon-pixel vertexing detector

The innermost detector is a pixel vertex detector (PXD). Its goal is to sample the trajectories of final-state charged particles in the vicinity of the decay position (vertex) of their long lived ancestors, so that the decay point can be inferred by extrapolation inward.

PXD sensors are based on depleted field-effect transistor technology [47]. They are made of p-channel MOSFET integrated on a silicon substrate, which is fully depleted by applying an appropriate voltage. Incident particles generate electron-hole pairs in the depleted region, and thus induce a current passing through the MOSFET. Sensors are 75 μm thick, which allows on-pixel integration of most of the electronics.

The PXD has two layers at 14 mm and 22 mm radius, respectively, and a full length of 174 mm at the radius of the outer layer. It comprises around 8 million pixels, $50 \times (50 - 55)\mu\text{m}^2$ (inner layer) and $50\times(70-85)\mu\text{m}^2$ (outer layer) each. The polar acceptance ranges from 17° to 150°. The design impact-parameter resolution is 12 μm, achieved by weighting the charge deposited in neighbouring pixels. To simplify pattern recognition, tracks are first reconstructed in the outer tracking volume, where lower occupancy aids track finding, and extrapolated to the PXD radius, to define regions of interest around their expected intersection points. If a firing pixel is found inside this region, it is kept in the pattern

Items	Parameters
Cryostat	
Radius: outer/inner	2.00 m/1.70 m
Central field	1.5 T
Total weight	23t
Effective cold mass	$\approx$ 6t
Length	4.41 m
Coil	
Effective radius	1.8 m
Length	3.92 m
Conductor dimensions	3×33 mm^2
Superconductor	NbTi/Cu
Stabilizer	99.99% aluminium
Nominal current	4400 A
Inductance	3.6 H
Stored energy	35 MJ
Typical charging time	0.5 h
Liquid helium cryogenics	Forced flow two phase
Cool down time	$<$ 6 days
Quench recovery time	$<$ 1 day

Table 2.3: Main parameters of the solenoid coil.

recognition algorithm, otherwise it is discarded. For the data used in this thesis, the full first pixel layer is used, along with 1/6 of the second layer.

2.2.3.3 Silicon-strip vertexing detector

Around the PXD is SVD [48], a silicon detector aimed at reconstructing decay vertices and low-momentum charged-particle tracks at high resolution.

SVD uses a double-sided silicon strip technology. Each sensor is made of a silicon n-doped bulk with an highly p-doped implant on one side. A voltage is applied to enhance the depletion region at the p-n junction, and removes intrinsic charge-carriers from the region. Traversing charged particles ionize the silicon, freeing electron-hole pairs that drift due to the electric field, inducing a signal in highly granular strip electrodes implanted at both ends of the depletion region. The fine segmentation of SVD sensors reduces latency, in order to deal with the high rates.

SVD is structured into four concentric layers at radii of 39, 80, 104 and 135 mm, composed by, respectively, 7, 10, 12, and 16 independently-readout modules called ladders, arranged in a cylindrical geometry. As shown in Figure 2.6, SVD has a polar-asymmetric geometry that mirrors the asymmetry in particle density resulting from the center-of-mass boost. The polar acceptance ranges from 17° to 150°.

Sensors are 300 μm thick, and the separation between adjacent strips ($\mathrm{d_{pitch}}$) ranges from 50 μm to 240 μm. Hence, the spatial resolution $\mathrm{d_{pitch}}/\sqrt{12}$ varies with the polar angle. Since the charge associated with an incident particle is usually distributed among several strips, position resolution is improved by interpolation.

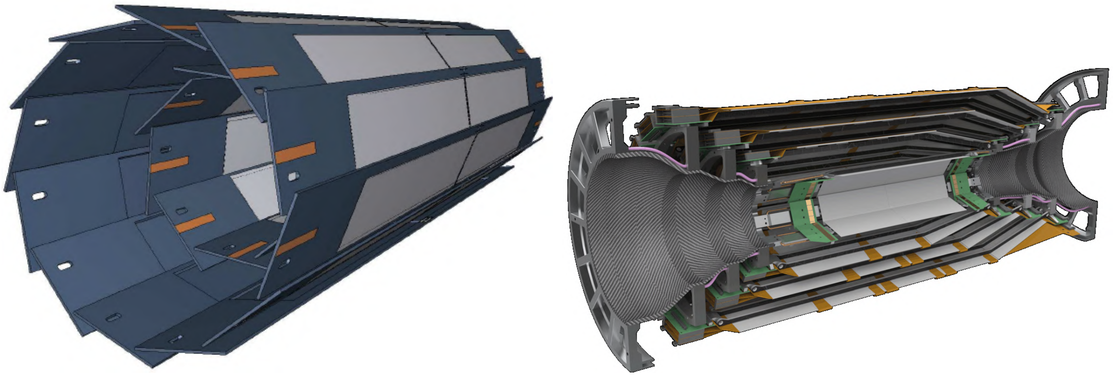


Figure 2.6: (Left) scheme of the PXD detector. (Right) exploded view of a SVD detector half.

2.2.3.4 Central drift chamber

The CDC [49] is a drift chamber. It samples charged-particle trajectories at large radii, thus providing accurate measurements of momentum and electric charge, trigger signals for events containing charged particles, and information on identification of charged-particle species by measuring their specific-ionization energy-loss (dE/dx).

When a charged particle traverses the CDC volume, it ionizes the gas, freeing electrons and positive ions from gas atoms. An applied electric field then moves these charges until they approach the sense wires, where high field gradients cause an abrupt acceleration with secondary ionizations that induce an electric signal whose time is digitized. The particle trajectory is inferred from the time between the collision and the signal.

The CDC inner radius is 16 cm and outer radius is 113 cm. The chamber is composed of 14336 30-μm-diameter sense wires, divided in 56 layers, immersed in a gaseous mixture of 50% He and 50% C_2H_6, while 42240 126-μm-diameter aluminum wires shape the electric field. The azimuthal acceptance ranges from 17° to 180°.

The spatial resolution is about 100 μm and the dE/dx resolution is 11.9% for an incident angle of 90°. Figure 2.7 shows a sliced view of the CDC and the possible wires configurations.

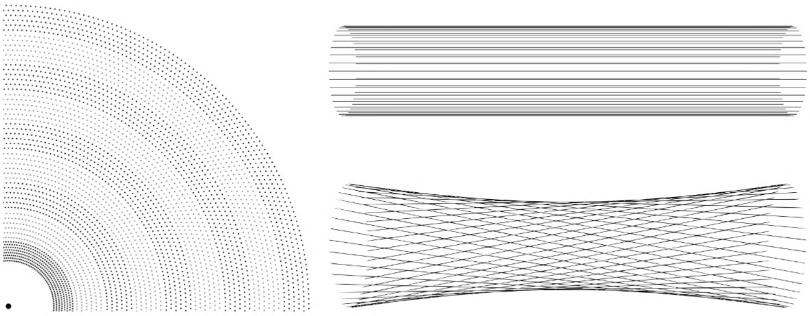


Figure 2.7: A quadrant of a slice of the transverse projection of the drift chamber (left); wire orientation for axial (top right) and stereo (bottom right) layers. The skew is exaggerated for visualization purposes

2.2.3.5 Tracking performance

Charged track reconstruction mainly relies on the CDC information. Momentum resolution is improved by combining CDC tracks with PXD and SVD hit information, especially for low-momentum particles. The transverse momentum resolution is $\sigma(p_T)/p_T = 0.0011 p_T[\mathrm{GeV}/c] \oplus 0.0025/\beta$as is in Fig. 2.8. The momentum- and angle-dependent impact parameter resolutions are $\sigma_{xy} = 10 \oplus 25/(p\beta \sin^{3/2}\theta)\mu\mathrm{m}$
$\sigma_z = 15 \oplus 27/(p\beta \sin^{5/2}\theta)\mu\mathrm{m}$.

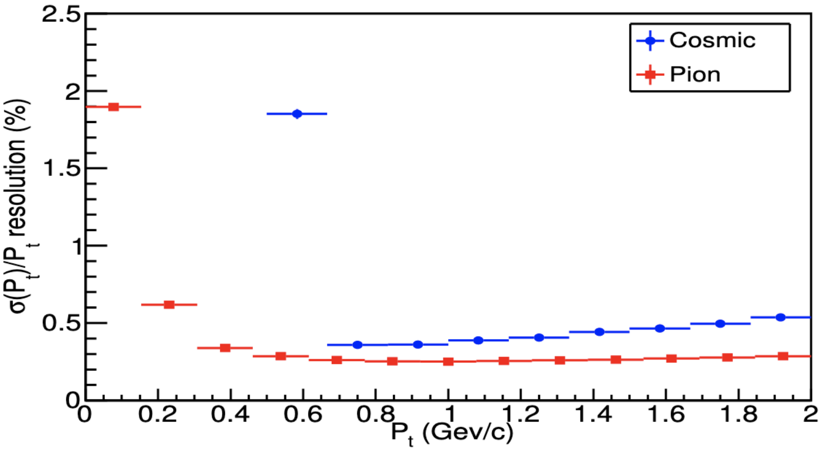


Figure 2.8: Transverse momentum resolution for collision and cosmic ray data.

2.2.4 Electromagnetic calorimeter

The ECL [50] measures the energy of photons and electrons. High energy photons and electrons entering the calorimeter initiate an electromagnetic shower through bremsstrahlung and electron-positron pair production. The energy is mostly converted to photons, which are collected by the photodiodes. In contrast to hadrons, which pass through the calorimeter with minimal energy loss, photons and electrons dissipate their entire energy.

The configuration, mechanical structure, and crystals of Belle II ECL are those of the Belle's calorimeter. The readout electronic boards have been completely upgraded to handle SuperKEKB's higher luminosity. The layout is shown in Figure 2.9. The ECL consists of three parts: the barrel, the forward endcap, and the backward endcap section. The barrel section is 3.0 m long with 1.25 m of inner radius; the endcaps are located at $z = +2.0$ m (forward) and -1.0 m (backward) from the interaction point. Table 2.4 summarizes the geometrical parameters of each section.

Item	θ coverage	θ segmentation	ϕ segmentation	Number of crystals
Forward endcap	12.4°–31.4°	13	48–144	1152
Barrel	32.2°–128.7°	46	144	6624
Backward endcap	130.7°–155.1°	10	64–144	960

Table 2.4: ECL parameters.

High momentum π^0 detection requires good separation of two nearby photons and a precise determination of the opening angle. This requires a segmented calorimeter structure. The ECL consists of a highly segmented array of 8736 cesium iodide crystals doped

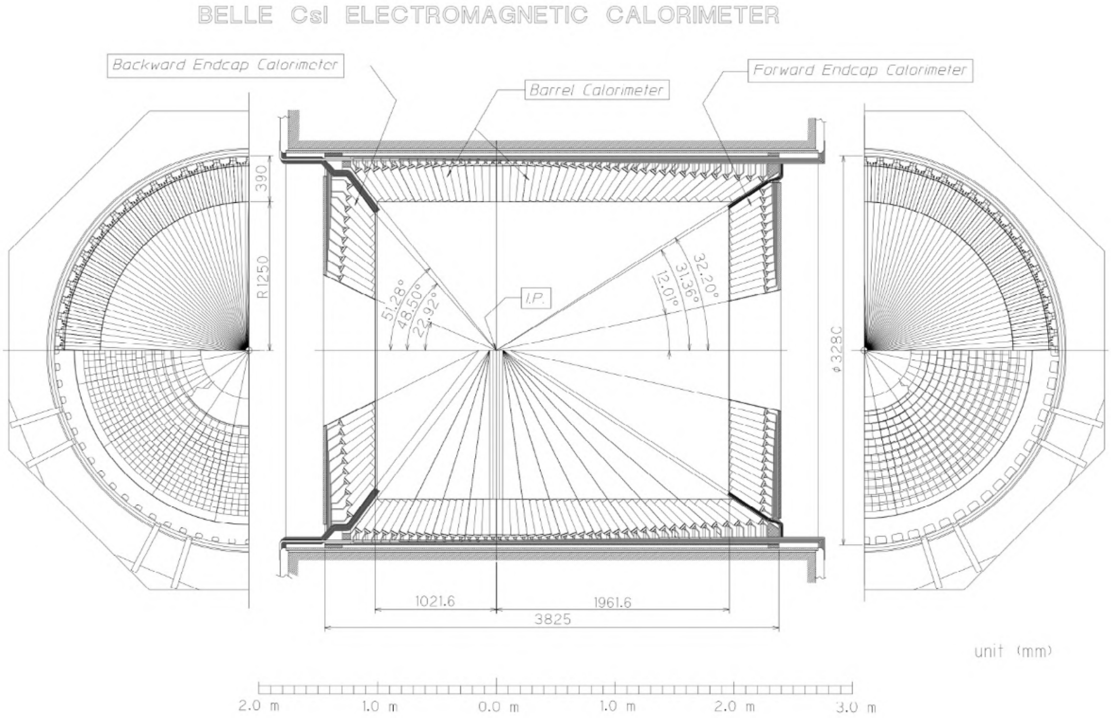


Figure 2.9: ECL layout.

with thallium (CsI(Tl)). Thallium shifts the energy of the excitation light into the visible spectrum. The light is detected by a independent pair of silicon PIN photodiodes and charge-sensitive preamplifiers installed at the outer end of each crystal.

Each crystal is arranged so as to point projectively to the nominal interaction point. The crystals are inclined to prevent photons from escaping through gaps between crystals by about 1.3° in the θ and ϕ direction in the barrel section, about 1.5° and about 4° in the θ direction in the forward and backward sections.

The 30 cm crystal length (corresponding to $16.2X_0$) is chosen to minimize the fluctuations of shower leakages out of the outermost end of the crystals that would spoil energy resolution. The crystals were designed in such a way that a photon injected at the center of the crystal would deposit 80% of its energy in the crystal on average. A typical crystal in the barrel section has 55×55 mm^2 front face and 65×65 mm^2 rear face; the dimensions of the crystals in the endcap sections vary from 44.5 to 70.8 mm and from 54 to 82 mm for front and rear face, respectively. The energy resolution ranges from $\sigma_E/E = 4\%$ at 100 MeV to 1.6% at 8 GeV. The resolution on the reconstructed π^0 mass is of $4.5\,\mathrm{MeV}/c^2$.

The ECL also uses Bhabha scattering to measure luminosity. Because the Bhabha cross section is predicted with high accuracy in QED, a precise inference of luminosity is achieved from the measured rate of Bhabha events in a volume of known acceptance.

2.2.5 Particle identification

Belle II combines measurements of time-of-propagation, Cherenkov radiation, and ionization energy loss in the tracker and drift chamber to identify charged particles.

2.2.5.1 Time-of-propagation detector

The TOP detector [51] measures the time of propagation of the Cherenkov photons emitted from charged particles passing through its quartz bars and internally reflected within a radiator. It is made of 16 quartz bars mounted at 1.2 m from the IP. Each bar has three main components (Fig. 2.10): a long bar acts as Cherenkov radiator, where photons are generated and propagated; a focusing mirror is mounted at the forward end; and a prism mounted at the backward end collects photons and guides them to a photomultiplier (PMT). The polar coverage ranges from 31° to 128°. On average, photons originated from slower particles take more time to reach the PMT, because of the inverse proportionality between β and $\cos\theta_C$. The time resolution is about 100 ps, allowing separation of pions from kaons at $0.4-4\,\mathrm{GeV}/c$ momenta with kaon identification efficiency of 85% and pion misidentification rate of 10% (Fig. 2.11).

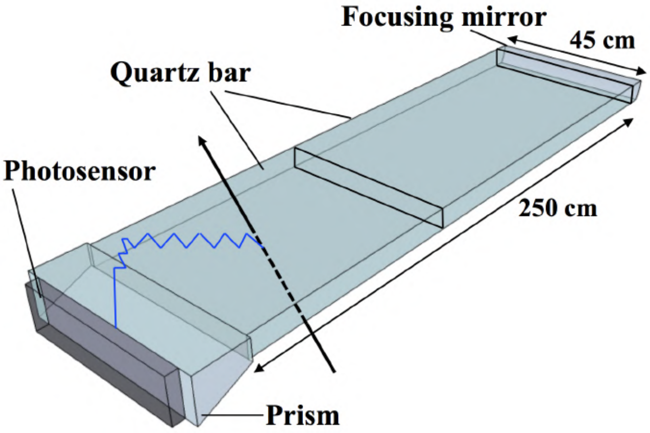


Figure 2.10: Scheme of a TOP bar. A charged particle crossing the radiator and emitting Cherenkov photons, which are collected at the PMT, is also represented.

2.2.5.2 Aerogel ring-imaging Cherenkov counter

The ARICH detector [52] identifies charged particles by measuring the Cherenkov ring produced when passing through a radiator. It consists of 420 modules for photon detection in seven layers extending from 0.56 to 1.14 m radius, and 248 aerogel tiles installed on the detector endcaps. The aerogel radiator produces Cherenkov photons when traversed by charged particles in a certain momentum range. Next to the radiator is an expansion volume where photons are propagated, to form rings on position-sensitive photodiodes. Photocathodes are then used to convert photons into photoelectrons and generate electric signals. Two adjacent radiators with different refraction indexes are used to generate enough photons for achieving sufficient resolution, as shown in Figure 2.12. The ARICH separates pions from kaons across all their momentum spectrum and discriminates also pions, electrons, and muons below 1 GeV/c with 4σ separation or more.

2.2.6 Trigger and data acquisition system

The e^+e^- collisions at the energy corresponding to the $\Upsilon(4S)$ mass produce a variety of processes. As the events of interest are only a fraction of the total cross section, an online event-selection system (trigger) is used to distinguish them from background in real time, and to feed them to data acquisition system (DAQ). The physics processes of interest

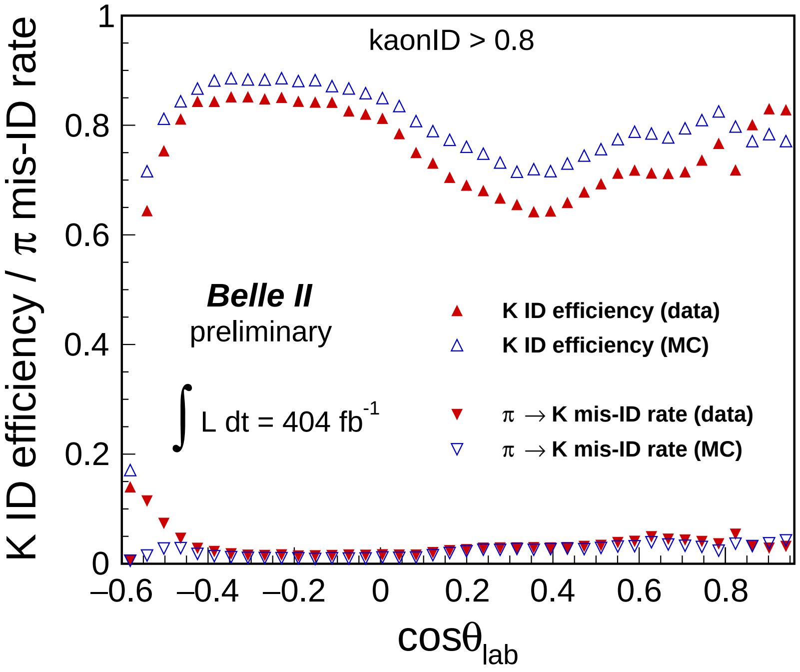


Figure 2.11: Kaon identification efficiency and pion fake rate as functions of momentum from the combination of all Belle II PID detectors. Up-pointing markers show kaon efficiency and down-pointing markers show pion fake rate.

include hadronic, Bhabha, μ/τ-pair and two photon events. Preferably discarded events include beam-related background resulting from synchrotron radiation, scattering of the beams on the residual gas, interactions in the beam pipe, and cosmic ray events.

The Belle II trigger is organized in a two-level logic, with a level 1 (L1) hardware trigger followed by the software-based high level trigger (HLT).

The L1 trigger, designed for a maximum rate of 30 kHz, uses input from four subdetectors: (i) the CDC, that provides track information on the r-ϕ plane and in z to suppress tracks not originating from the interaction point; (ii) the ECL, that gives information on total energy deposit and cluster multiplicity; (iii) the TOP, that provides timing and hit topology information; and (iv) the KLM, that gives high-efficiency trigger for muons. These are used to achieve a low-level reconstruction that is fed to the global decision logic (GDL), which sends the proper trigger signal if the events passes the selection requirements. The L1 logic is implemented using field programmable gate arrays that allow to have a fixed latency of 5 µs, with an uncertainty on the trigger timing (jitter) of ∼10 ns.

Expected cross sections and trigger accept rates for physics processes of interest at the design instantaneous luminosity of $8{\times}10^{35}\text{cm}^{-2}\text{s}^{-1}$ are given in Table 2.5.

Events selected by the L1 trigger are input to the HLT, that makes a decision using information from all the subdetectors except for PXD. The online software reconstruction is similar to that used in offline. A first selection, performed after the first step of the reconstruction and aimed at discarding about half of the events, is based on requirements on track multiplicity, vertex position and total ECL energy deposit. After the remaining steps of the standard reconstruction are completed, further physics-level selection are performed. After this stage, the number of events is reduced to about 1/8.

During this procedure, PXD data for events that pass the L1 selection are stored in the dedicated online data reduction system. Once an event passes the selection, HLT

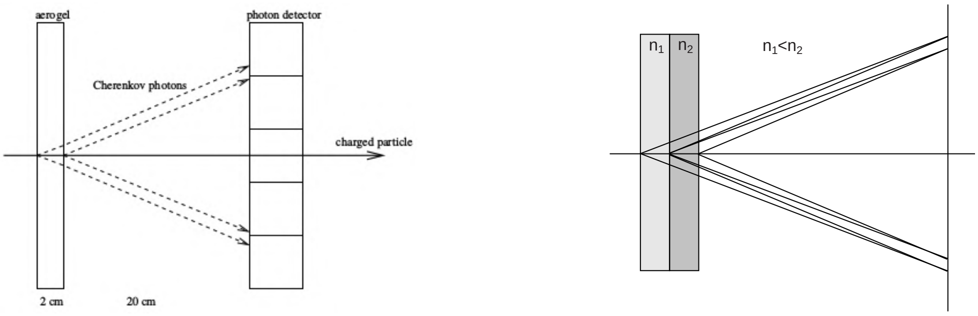


Figure 2.12: Scheme of the ARICH detector.

Process	σ [nb]	Rate [Hz]
$e^+e^- \to \Upsilon(4S)$	1.2	960
$e^+e^- \to q\overline{q}(q = u, d, s, c)$	2.8	2200
$e^+e^- \to \mu^+\mu^-$	0.8	640
$e^+e^- \to \tau^+\tau^-$	0.8	640
$e^+e^- \to e^+e^-$ (Bhabha scattering) $\theta_{lab} > 17°$	44	350*
$e^+e^- \to \gamma\gamma$ $\theta_{lab} > 17°$	2.4	19*
Two photon events ($\theta_{lab} > 17°$ & $p_T \geq 0.1$ GeV/c)	≈ 80	≈ 1500

Table 2.5: Expected cross sections and trigger rates of various physics processes at $8{\times}10^{35}\text{cm}^{-2}\text{s}^{-1}$ luminosity [46]. Bhabha and $\gamma\gamma$ rates (*) are artificially reduced by a factor of 100.

extrapolates the tracks found by CDC and SVD to the PXD layers, defining regions of interest (ROIs). These are passed to the data reduction system, and only hits matching with a ROI are transmitted to the DAQ system. This keeps the PXD data size to about 100 kB/event.

Fully reconstructed events are stored in DST files. The size of a DST of a typical hadronic event is 100 kB. The large amount of information stored in DST files are reduced into mini-DST to isolate subsets of events of physics processes of interest like hadronic events. The size of a mini-DST of an hadronic event is around 40 kB. I analyze the data passed through the "hlt_hadron" skim, which imposes a set of loose requirements to isolate hadronic meson events. The mini-DST files are produced through the Belle II software [53].

Chapter 3

Experimental considerations

This chapter discusses the experimental features of $B^+ \to \rho^+\rho^0$ decays and their analysis in a B factory, along with a concise overview of the experimental status of $B \to \rho\rho$ measurements.

3.1 Generalities on $B \to \rho\rho$ decays

The $B^+ \to \rho^+\rho^0$ decay involves the decay of spin-0 particle into a pair of spin-1 particles. Because the total angular momentum is conserved, the orientations of the ρ spins, combined with their orbital angular momentum should sum to zero. In other words, $\vec{0} = \vec{1} + \vec{1} + \vec{L}$, which yields three possible orbital angular momentum states for the ρ pair, $\vec{L} = 0, 1, 2$. $B^+ \to \rho^+\rho^0$ decays are therefore classified into two categories, "longitudinally" and "transversely" polarized. Longitudinally polarized decays are B^+ decays proceeding through the $L = 0$ state, where spin projections of both ρ mesons align along their momenta and helicity λ (i.e., spin projection along momentum direction) equals zero. Transversely polarized decays proceed through the $L = 1, 2$ states, implying that ρ meson helicities equal ±1.

The polarization states are distinguished experimentally from each other through the decay angles of the pions with respect to each other and the ρ mesons. The three angles that describe $B^+ \to \rho^+\rho^0$ decays in the helicity basis are shown in Fig. 3.1. The helicity axes are defined as the negative of the momentum direction of the B^+ in the rest frame of the ρ meson. Hence, the ρ^+ and ρ^0 helicity axes are back to back, pointing along their decay axis outward from the decay position of the B^+. Two ρ helicity angles, θ_{ρ^+} and θ_{ρ^0}, are defined as angles between the momenta of the charged pions and the helicity axes in the rest frame of the ρ mesons. The third angle ϕ is an azimuthal angle between two ρ decay planes.

The angular distribution of the π^+ and $\pi^{0(-)}$ from the $\rho^{+(0)}$ decay is trivial, since the final state particles are spinless. In general, the angular dependence for $B^+ \to \rho^+\rho^0$ decays is expressed in terms of spherical harmonics as [54, 55]

$$\frac{d^3\Gamma}{d\cos\theta_{\rho^0}\, d\cos\theta_{\rho^+}\, d\phi} \propto \left| \sum_{\lambda=-1}^{1} H_\lambda \times Y_1^{\lambda}(\theta_{\rho^+}, \phi) \times Y_1^{-\lambda}(\pi - \theta_{\rho^0}, 0) \right|^2, \tag{3.1}$$

where H_λ ($\lambda = 0, \pm1$) indicates the helicity amplitudes associated with longitudinally, or transversely, polarized decays. By substituting $Y_1^{\pm1}(\theta,\phi) \to \mp\frac{1}{2}\sqrt{\frac{3}{2\pi}}\sin\theta e^{\pm i\phi}$, $Y_1^0(\theta,\phi) \to \frac{1}{2}\sqrt{\frac{3}{\pi}}\cos\theta$ and expanding the right-hand side of Eq. (3.1), we obtain the full angular distribution of $B^+ \to \rho^+\rho^0$ decays

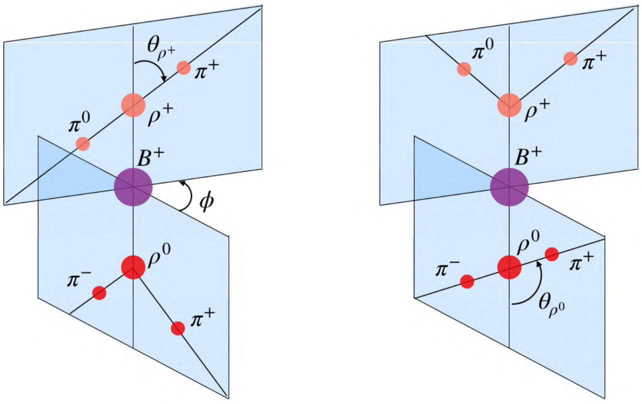


Figure 3.1: Geometric definition of the helicity basis angles.

$$\frac{d^3\Gamma}{d\cos\theta_{\rho^0} d\cos\theta_{\rho^+} d\phi} = \frac{9}{16\pi}\Bigg[\frac{1}{2}\sin^2\theta_{\rho^+}\sin^2\theta_{\rho^0}(|H_{+1}|^2+|H_{-1}|^2)+2\cos^2\theta_{\rho^+}\cos^2\theta_{\rho^0}|H_0|^2+$$
$$\sin^2\theta_{\rho^+}\sin^2\theta_{\rho^0}[\cos 2\phi \mathrm{Re}(H_{+1}H^*_{-1})-\sin 2\phi \mathrm{Im}(H_{+1}H^*_{-1})]-$$
$$\frac{1}{2}\sin 2\theta_{\rho^+}\sin 2\theta_{\rho^0}[\cos\phi \mathrm{Re}(H_{+1}H^*_0+H_{-1}H^*_0)-\sin\phi \mathrm{Im}(H_{+1}H^*_0-H_{-1}H^*_0)]\Bigg]. \quad (3.2)$$

Integration over the angles θ_{ρ^+}, θ_{ρ^0}, and ϕ yields the integrated width $\Gamma = |H_0|^2 + |H_+|^2 + |H_-|^2$. The ratio

$$f_L = \frac{\Gamma_L}{\Gamma} = \frac{|H_0|^2}{|H_0|^2 + |H_+|^2 + |H_-|^2} \quad (3.3)$$

defines the fraction of longitudinally polarized $B^+ \to \rho^+\rho^0$ decays that is a fundamental input to the $B \to \rho\rho$ isospin analysis (Sec. 1.6.2.2). By integrating Eq. (3.2) over the angle ϕ (assuming the azimuthal acceptance is uniform) we obtain

$$\frac{d^2\Gamma}{d\cos\theta_{\rho^0} d\cos\theta_{\rho^+}} = \frac{9}{16}|H|^2\Big[4f_L\cos^2\theta_{\rho^0}\cos^2\theta_{\rho^+} + (1-f_L)\sin^2\theta_{\rho^0}\sin^2\theta_{\rho^+}\Big], \quad (3.4)$$

where $|H|^2 \equiv |H_0|^2 + |H_+|^2 + |H_-|^2$.

Equation (3.4) shows that f_L can be extracted from a one-parameter fit to the distributions of $\cos\theta_{\rho^0}$ and $\cos\theta_{\rho^+}$. This is exemplified by the two-dimensional angular distributions associated with some specific f_L values shown in Fig. 3.2.

3.2 Experimental status

Measurements of $B \to \rho\rho$ decay properties have been reported by previous experiments. The first observation of this decay was reported in 2003 by the Belle experiment [41]. Then the Belle and BaBar experiments studied them in detail. The most recent results are summarized in Table 3.1. The Belle measurements in $B^0 \to \rho^+\rho^-$ [40] and $B^0 \to \rho^0\rho^0$ [56] decays are based on the full data set, corresponding to 772×10^6 $B\overline{B}$ pairs, while the $B^+ \to \rho^+\rho^0$ measurement is only based on 85×10^6 $B\overline{B}$ pairs, just about 10% of the

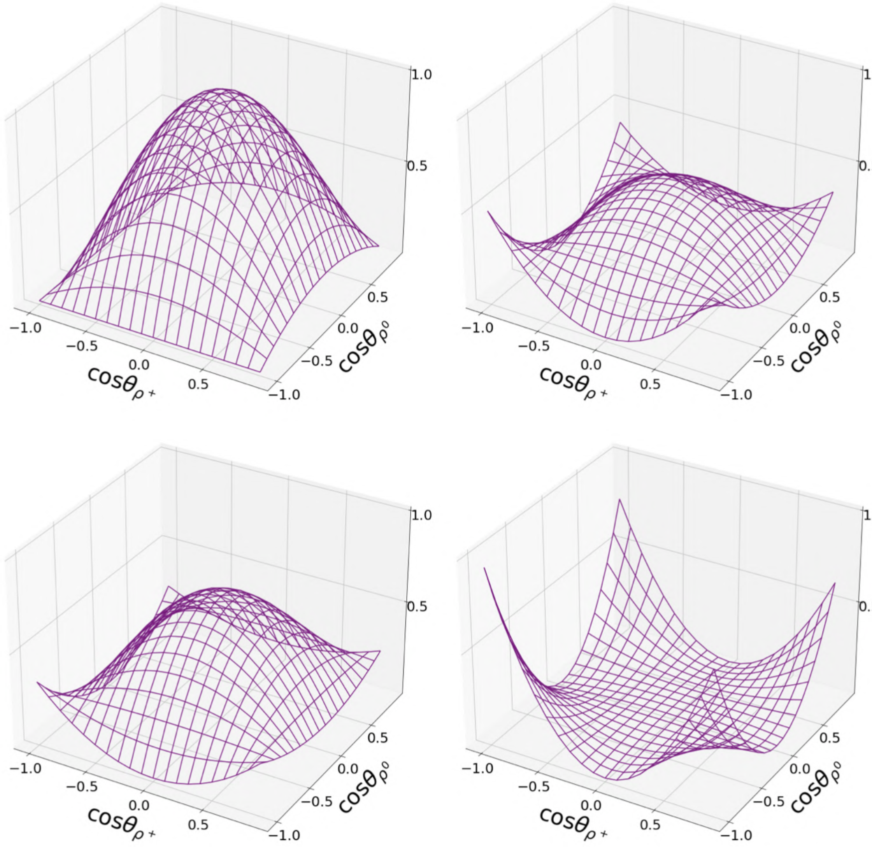


Figure 3.2: Theoretical two-dimensional angular distributions for $f_L = 0$, 1/3, 1/2, and 0.95, respectively.

sample. The BaBar measurements on the $B^0 \to \rho^0\rho^0$ [39] and $B^+ \to \rho^+\rho^0$ [42] channels are based on the full data set corresponding to $465{\times}10^6$ $B\overline{B}$ pairs; and the $B^0 \to \rho^+\rho^-$ measurement [57] is based on $383.6{\times}10^6$ $B\overline{B}$ pairs. The LHCb experiment reported a Dalitz-plot analysis of the $B^0 \to \rho^0\rho^0$ decay with a sample corresponding to 3 fb^{-1} of integrated luminosity [58].

3.3 *B* decay reconstruction at Belle II

Before delving into the details of my work, it helps introducing the general aspects of B decay reconstruction at Belle II.

Figure 3.3 shows a sketch of a $B^+ \to \rho^+\rho^0$ decay. Confined bunches of electrons and positrons are brought to collision in the interaction point (IP). Various final states are produced, with proportions given by the associated cross sections. In 0.4% of collisions, a $\Upsilon(4S)$ meson is produced, and 96% of them decay into a pair of B mesons, labeled as signal B-meson, B_{rec}, and partner B-meson, B_{tag}, in Fig. 3.3. These events produce about ten tracks in acceptance on average, which are used by the trigger to identify *hadronic* events, as opposed to QED events, and save them onto permanent memory.

The 1.5 ps B meson lifetime, combined with a typical momentum of 1.5 GeV/c in the laboratory frame, results in a flight length of about 130 μm. This allows reconstructing the decay position for background discrimination and for measurements of quantities depending

	$\mathrm{N_{sig}}$	$\mathcal{B} \times 10^{-6}$	f_L	$\mathcal{S}_{\mathrm{CP}}$	$\mathcal{A}_{\mathrm{CP}}$
Belle					
$\rho^+\rho^0$	58.7±13.2	$31.7 \pm 7.1^{+3.8}_{-6.7}$	$0.95 \pm 0.11 \pm 0.02$	-	$0.00 \pm 0.22 \pm 0.03$
$\rho^+\rho^-$	1754±94	$28.3 \pm 1.5 \pm 1.5$	$0.988 \pm 0.012 \pm 0.023$	$-0.13 \pm 0.15 \pm 0.05$	$0.00 \pm 0.10 \pm 0.06$
$\rho^0\rho^0$	166±59	$1.02 \pm 0.30 \pm 0.15$	$0.21^{+0.18}_{-0.22} \pm 0.15$	-	-
BaBar					
$\rho^+\rho^0$	1122±63	$23.7 \pm 1.4 \pm 1.4$	$0.950 \pm 0.015 \pm 0.006$	-	$-0.054 \pm 0.055 \pm 0.010$
$\rho^+\rho^-$	729±60	$25.5 \pm 2.1^{+3.6}_{-3.9}$	$0.992 \pm 0.024^{+0.026}_{-0.013}$	$-0.17 \pm 0.20^{+0.05}_{-0.06}$	$-0.01 \pm 0.15 \pm 0.06$
$\rho^0\rho^0$	99^{+35}_{-34}	$0.92 \pm 0.32 \pm 0.14$	$0.75 \pm^{+0.11}_{-0.14} \pm 0.04$	$0.3^{+0.7}_{-0.2} \pm 0.04$	$0.2 \pm 0.8 \pm 0.3$
LHCb					
$\rho^0\rho^0$	634±28	$0.94 \pm 0.17 \pm 0.09 \pm 0.06$	$0.745^{+0.048}_{-0.058} \pm 0.034$	-	-

Table 3.1: Current status of measurements on $B \to \rho\rho$ decays. The first uncertainties are statistical and the second are systematic.

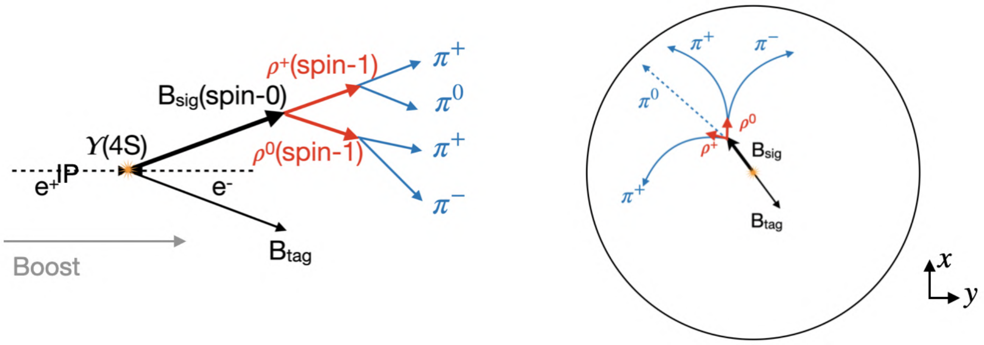


Figure 3.3: Sketch of a $B^+ \to \rho^+(\to \pi^+\pi^0)\rho^0(\to \pi^+\pi^-)$ decay in the (left) longitudinal and (right) transverse views.

on decay time. In this example relevant for my work, the B meson decays into a charged and a neutral ρ meson, each promptly decaying into pairs of pions. Final-state pions have typical momenta of 0.3–3 GeV/c in the laboratory frame. The trajectories of charged pions are curved by the Lorentz force produced by the axial magnetic field that pervades the tracking volume. They are reconstructed by a geometric fit of the observed pattern of measurement points, called *hits*, excited in the active layers of the tracking detectors. Knowledge of trajectories, magnetic field, and detector material allows for determining the momentum and charge of charged final-state particles. The kinematic properties of the neutral pion are reconstructed by combining the four-momenta of the two photons from its decay. These are identified and reconstructed by using information on their energy depositions in the electromagnetic calorimeter, where signals from adjacent *hit cells* are topologically grouped in *clusters*. The trajectory is inferred from the position of the cluster in the calorimeter and the reconstructed production space-point.

Reconstructed final-state particles are then used to determine the kinematic properties of the signal candidate of interest B^+_{rec}. A kinematic fit combines the four momenta of the final-state particles, along with the space-point where they originate by using momentum and energy conservation.

In the analysis, I use the `treeFitter` [59] [60] algorithm to reconstruct the decay-vertex position of $B^\pm$ candidates using only charged tracks and constraining the production vertex with the event-dependent interaction-point profile. In addition, I constrain the diphoton mass to the known π^0 mass (`massKFit` [61]). Every combination that meets the reconstruction quality-requirements is a possible decay candidate. However, decay candidates can originate from genuine signal events or background events. Background events can in turn be associated with two broadly defined sources,

Continuum, these are light mesons, such as pions, originated from light quark-antiquark pairs produced in the e^+e^- collision, that form random combinations accidentally meeting the reconstruction and selection requirements;

B **decays** (other than signal), these are candidates where one or more final-state particles are misidentified or when a different decay yields the same final states.

Various quantities are available to discriminate signal from backgrounds.

3.3.1 Kinematic fit variables

Information on reconstructed vertices is used to distinguish between continuum and $B\overline{B}$ events through two powerful observables,

Vertex displacement. In continuum events, all tracks typically originate from the interaction point, unlike in $B\overline{B}$ events, where the vertices of two B mesons are displaced due to the lifetime of B mesons combined with their momenta. Thus, the displacement ΔZ between the decay position of the signal B vertex and that of the other B along the beam axis offers good discrimination of $B\overline{B}$ events from continuum [62]. The uncertainty on ΔZ, $\delta\Delta Z$, is also used. Figure 3.4 shows the distributions of the ΔZ and $\delta\Delta Z$ variables in simulation. The signal peak in the ΔZ distribution is broader than the peak for continuum events. As it is more straightforward to reconstruct the signal decay vertex due to its displacement from the interaction point, the uncertainty of ΔZ is larger for continuum events.

χ^2 probability of the vertex fit. Frequently, multiple decay candidates corresponding to various combinations of particles that meet the quality and reconstruction requirements are reconstructed in an event. Each candidate is associated to a reconstructed decay vertex. The corresponding vertex fit probability is used to determine a p-value. This may be used to choose the decay candidate most likely to correspond to the genuine signal, as accidental combinations of random charged particles are more likely to yield worse fit qualities. Figure 3.5 shows the distribution of the vertex p-value. Worse fit qualities are obtained for continuum events as expected.

In addition to these basic vertex variables, Belle II exploits the available event information to construct discriminating variables at particle-, candidate-, and event-level to separate signal from background.

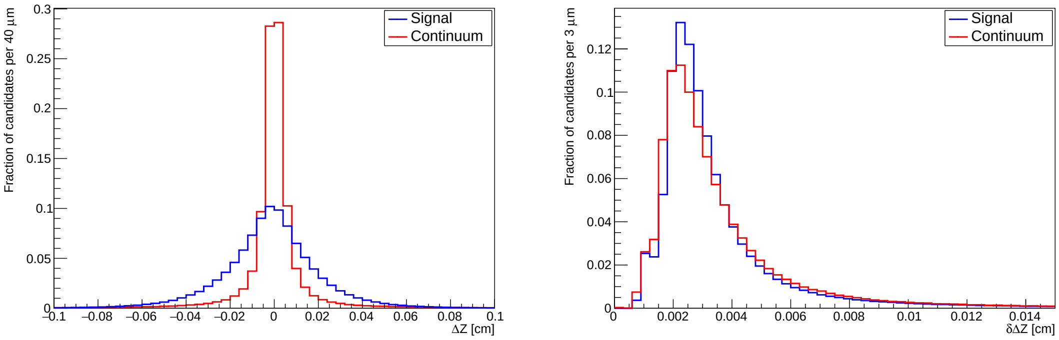


Figure 3.4: Distribution of (left) ΔZ and (right) $\delta\Delta Z$ in simulated signal $B^+ \to \rho^+\rho^0$ and continuum events. Distributions normalized to unity.

3.3.2 Particle-level variables

Several discriminating variables are specific to each reconstructed final-state particle. Relevant examples include

Hit multiplicity. Information related to the number of samplings a charged particle trajectory undergoes in the tracking detectors (*hits*) is used to select the quality of the tracks. Not only the resolution of high-level physics quantities, such as particle

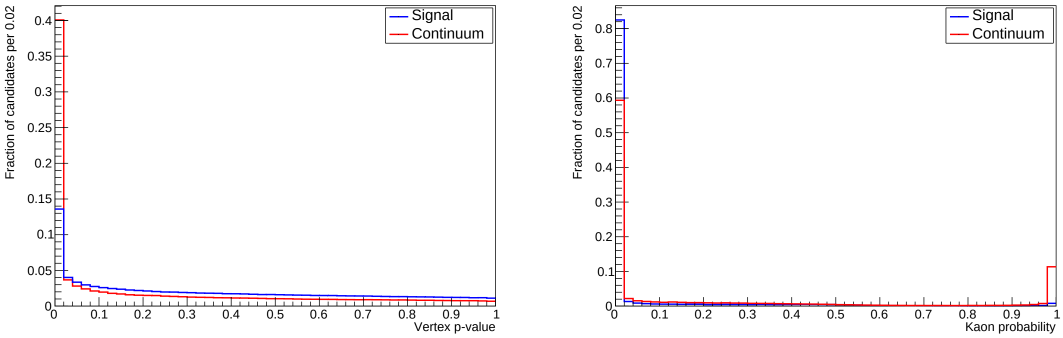


Figure 3.5: Distribution of (left) vertex quality p-value and (right) probability of pion to be reconstructed as kaon in simulated signal $B^+ \to \rho^+\rho^0$ and continuum events. Distributions normalized to unity.

momenta, depends on the number of hits, but also spurious tracks from accidental combinations of unrelated hits are typically suppressed by hit-multiplicity requirements;

Track displacement. Interactions between beam particles within the same bunch, or with residual gas, may result in interactions of the beam halo with the SuperKEKB or Belle II infrastructure, yielding intense showers of secondary particles that illuminate the detector. These are referred to as *beam backgrounds*. Measures of track displacement from the IP are effective to suppress such backgrounds, because beam-background tracks do not usually point back to the interaction point. The quantities typically used are the transverse (dr) and longitudinal (dz) distances of a track from the IP;

Particle identification (PID) information. Several detectors provide track-specific information associated with the identity of the corresponding charged particle. This information expresses the probability of observing the detected PID signal assuming true a mass hypothesis out of two possibilities (kaon or pion, pion or electron, and pion or muon) and enhances discrimination against background, especially from candidates from misreconstructed B decays.

3.3.3 Candidate-level variables

A class of higher-level discriminating features is associated with the reconstructed decay candidates. Relevant examples include

Beam-constrained masses and energies. Using distinctive kinematic information about the signal is a common approach to suppress background in many experimental environments. A widely used and effective discriminator is the candidate's invariant mass, since fully reconstructed genuine signal events cluster at a specific mass value and background shows typically broader distributions.

The peculiar kinematic environment of B-factory colliders provides additional constraints that further background separation. The $\Upsilon(4S)$ is produced almost at threshold and decays in two same-mass particles, B and $\overline{B}$. If the B meson is correctly reconstructed, the energy of its decay products equals half of the collision energy in the center-of-mass frame. This is optimally exploited by two variables specific of B-Factories (all quantities in the $\Upsilon(4S)$ frame),

- beam-energy-constrained mass M_{bc}, defined as

  $$M_{\mathrm{bc}} = \sqrt{s/4 - |\vec{p}_B^{\,*}|^2}, \tag{3.5}$$

  where $\vec{p}_B^{\,*}$ is the momentum of the B meson reconstructed from the momenta of its decay products, and s is the squared collision energy. While conceptually similar to the B invariant mass, M_{bc} is known with significantly more precision since the beam-energy spread is significantly smaller than the uncertainty on the reconstructed B-meson energy. In the M_{bc} distribution, all B decays reconstructed using all their decay products peak at the B meson mass regardless if the decay products are correctly or incorrectly identified, while non-B events have a smooth distribution. Both distributions drops to zero at the kinematic limit of half the collision energy. This makes M_{bc} powerful to separate B-events from $q\overline{q}$ continuum events.
- energy difference $\Delta E = E_B^* - \sqrt{s}/2$. This is the difference between the reconstructed B-candidate energy and half of the collision energy, which is known with high precision. If the B meson is correctly reconstructed, the energy of the decay products equals approximately half of the collision energy. Therefore, B signals peak at zero in the ΔE distribution, while continuum background follows a smooth distribution. In addition to discriminating against continuum, ΔE suppresses background from misidentified B decays. If a B final-state particle is misidentified as another, its reconstructed energy, and consequently that of the B candidate, will depart from its true energy because of the mismatch in mass, resulting in a ΔE shift.

Figure 3.6 shows a sketch of the ΔE and M_{bc} distributions for signal decays, continuum background events, and candidates reconstructed in other $B\overline{B}$ events, demonstrating the significant discrimination power of these variables.

3.3.4 Event-level variables

At a further level of abstraction, global-event information is used too in signal-from-background discrimination.

Hadronic e^+e^- cross-sections are dominated by non-B events, so-called continuum background, consisting in production of light $q\overline{q}$ pairs (where q is u, d, c, s) that mostly yield pions and kaons. Because of the kinematic features associated with at-threshold $B\overline{B}$ production, variables capable to capture the 'shape' of the event, that is, the spatial and phase-space distributions of final-state particles, offer powerful discrimination of $B\overline{B}$ events from continuum.

Figure 3.7 shows a sketch of the event shapes of a continuum and a $B\overline{B}$ event. In a $B\overline{B}$ event, both B mesons are nearly at rest in the Υ(4S) frame. The B decay products are therefore emitted isotropically in that frame, unlike light quarks which are produced with a comparatively large initial momentum due to their small mass compared to the total collision energy. This results in a collimated fragmentation into two narrow back-to-back jets of light hadrons. Hence, the spatial and energy-momentum distributions of $B\overline{B}$ decay products are approximately spherical, compared to pencil-like shapes for continuum.

Information based on these distributions is therefore useful to discriminate signal events from continuum and it is exploited through several variables. Relevant examples include

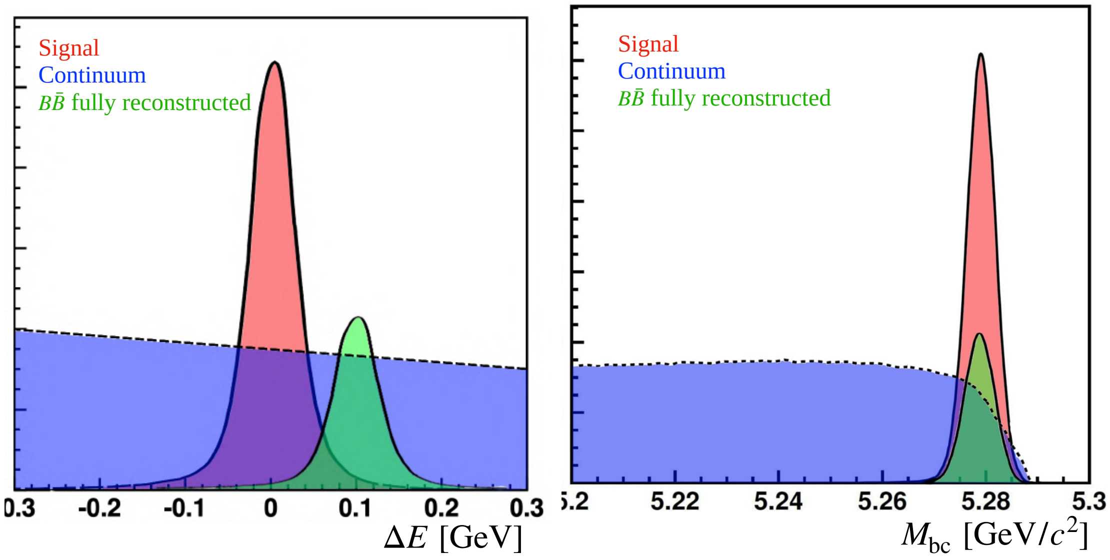


Figure 3.6: Sketched distributions of (left) ΔE and (right) $M_{\rm bc}$ for signal, continuum and fully reconstructed candidates in $B\overline{B}$ events.

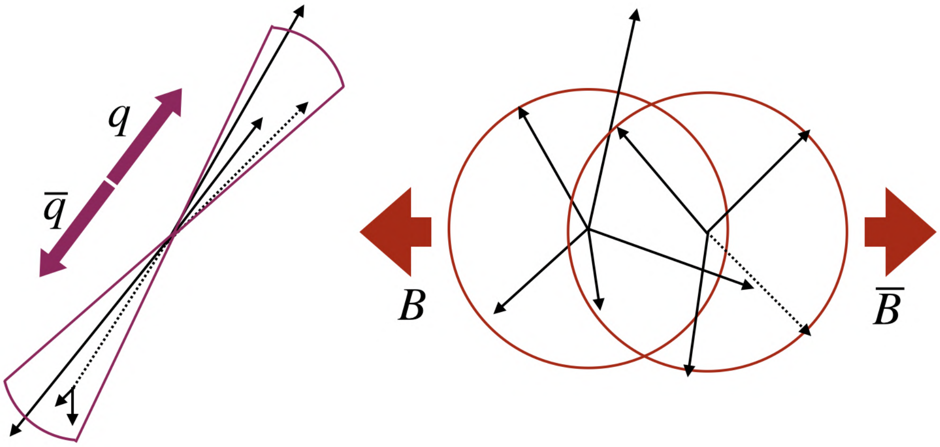


Figure 3.7: Event-shape sketch for a (left) pencil-like continuum and (right) sphere-like $B\overline{B}$ event.

***B* meson direction.** The momenta of spin-0 B mesons produced in a spin-1 $\Upsilon(4S)$ meson decay have a $\sin^2\theta^*$ angular distribution, where θ^* is the angle between the B-candidate momentum and the beam axis in the $\Upsilon(4S)$ frame. The momenta of spin-1/2 quarks, and their resulting jets follow a $1{+}\cos^2\theta^*$ angular distribution. The variable $|\cos\theta^*|$ allows therefore to distinguish signal B decays and background candidates reconstructed from continuum events (Fig. 3.10).

Thrust. For a collection of N momenta $\mathbf{p}_i$ ($i = 1, ..., N$), thrust T is defined as

$$T = \frac{\sum_{i=1}^{N} |\mathbf{T}\cdot\mathbf{p}_i|}{\sum_{i=1}^{N} |\mathbf{p}_i|}, \tag{3.6}$$

where $\mathbf{T}$ is the unit vector that maximize the total momentum projection, and therefore, the longitudinal projection of particle's momenta [63]. In a continuum event, particles typically have a larger longitudinal momentum component, due to momentum conservation, as they are produced in jets induced by initial quarks, and therefore show values of T closer to unity. The thrust is calculated for the momenta of particles produced in the signal B meson decay, T_{sig}, and for the momenta of all the other particles in the event, T_{tag} (Fig. 3.8, left panel). Continuum events tend to cluster at higher values of the T_{tag} distribution, unlike signal.

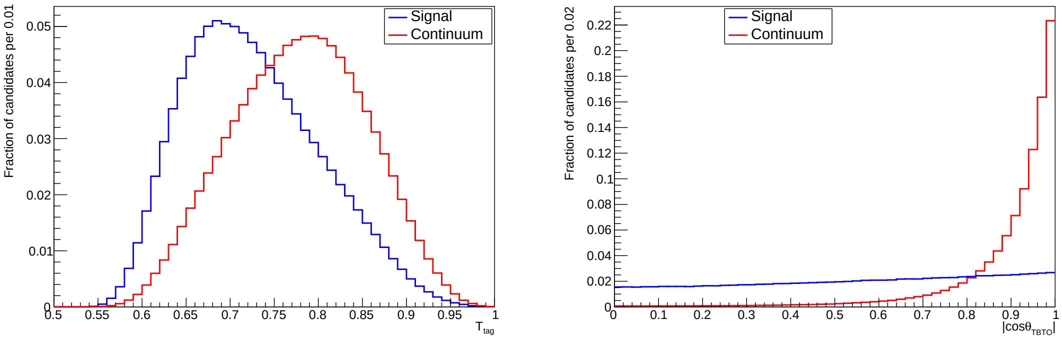


Figure 3.8: Distribution of (left) T_{tag} and (right) $|\cos\theta_{\text{TBTO}}|$ in simulated signal and continuum events. Distributions normalized to unity.

Thrust angles. The cosine of the angles between the $\mathbf{T}_{\text{sig}}$ and $\mathbf{T}_{\text{tag}}$ axes, $|\cos\theta_{\text{TBTO}}|$, and between the $\mathbf{T}_{\text{sig}}$ and the z axis, $|\cos\theta_{\text{TBz}}|$, are also used to discriminate between $B\overline{B}$ and continuum events offering two of the most powerful discriminators. Figure 3.8 (right) shows $|\cos\theta_{\text{TBTO}}|$ distributions for simulated signal and continuum events. Since the momenta of the decay products of B^+ and B^- are isotropically distributed, $\mathbf{T}_{\text{sig}}$ and $\mathbf{T}_{\text{tag}}$ are randomly distributed, $|\cos\theta_{\text{TBTO}}|$ and $|\cos\theta_{\text{TBz}}|$ follow uniform distributions. For $q\overline{q}$ events, particle momenta are collimated, resulting in strongly directional $\mathbf{T}_{\text{sig}}$ and $\mathbf{T}_{\text{tag}}$ vectors and yielding peaking $|\cos\theta_{\text{TBTO}}|$ and $|\cos\theta_{\text{TBz}}|$ distributions.

CLEO cones. The CLEO collaboration introduced nine variables capable of offering additional discrimination against continuum [64]. "CLEO cones" are a refinement of the concept of thrust: they are based on the sum of the absolute values of the momenta of all particles within angular sectors around the thrust axis, in intervals of 10°. This results in nine concentric cones (Fig. 3.9, left panel). The event is “folded” such that the particle content of two cones of same aperture but pointing in opposite directions are combined. Since the flight directions of the B-mesons’ decay products are less correlated than those of continuum events, CLEO-cone distributions allow for continuum discrimination (Fig. 3.9, right panel).

Fox-Wolfram moments. Given a total number of N particles in an event, with momenta $\mathbf{p}_i^*$ in the $\Upsilon(4S)$ frame, the ℓth order Fox-Wolfram moment H_ℓ is defined as

$$H_\ell = \sum_{i,j}^{N} \frac{|\mathbf{p}_i^*| \cdot |\mathbf{p}_j^*|}{s} \cdot \mathrm{P}_\ell(\cos\theta_{i,j}^*), \tag{3.7}$$

where $\theta_{i,j}^*$ is the angle between $\mathbf{p}_i^*$ and $\mathbf{p}_j^*$, $\sqrt{s}$ is the total energy in the $\Upsilon(4S)$ frame, and P_ℓ is the ℓth order Legendre polynomial [65]. The most discriminating example

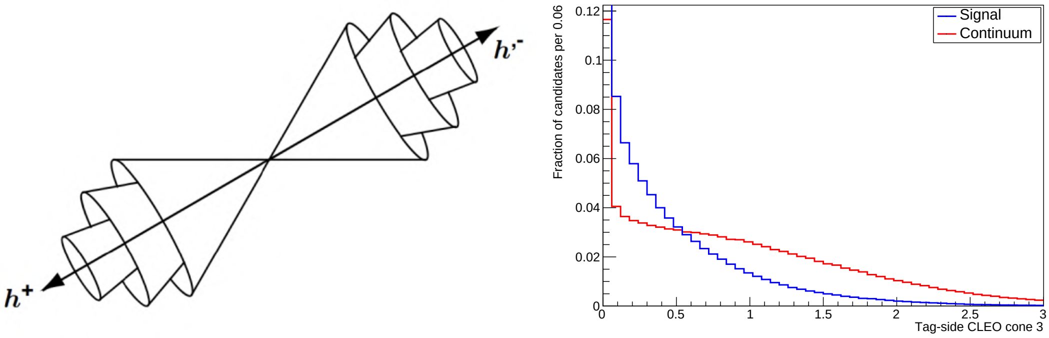


Figure 3.9: (Left) sketch of the first three CLEO cones. (Right) distribution of the third CLEO cone for simulated signal and continuum events. Distributions normalized to unity.

is the normalized second Fox-Wolfram moment $R_2 = H_2/H_0$, which captures the "shape" of spatial distribution and energy flow. Events with collimated jets, and therefore more likely to originate from continuum, show values of R_2 close to 0.5, while $B\overline{B}$ events cluster at lower R_2 values (right panel of Fig. 3.10).

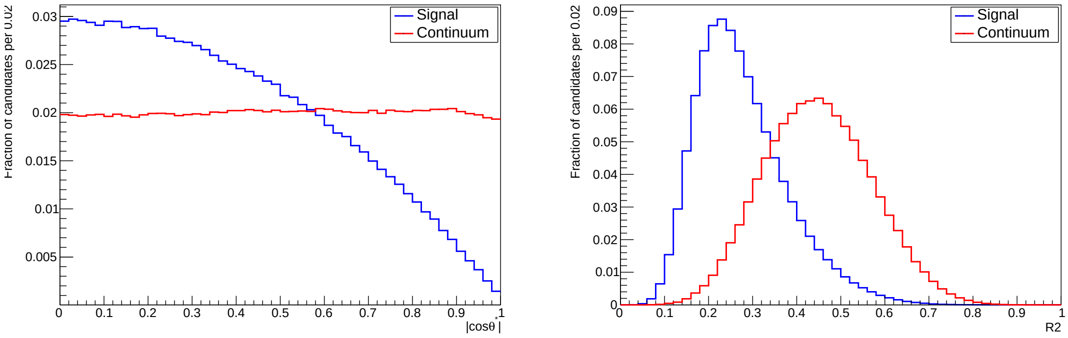


Figure 3.10: Distribution of (left) $|\cos\theta^*|$ and (right) R_2 for simulated signal and continuum events. Distributions normalized to unity.

Kakuno-Super-Fox-Wolfram moments. The discrimination power provided by the Fox-Wolfram moments deteriorates when particles are not reconstructed because of the presence of neutrinos, finite acceptances, or inefficiencies. To correct for this limitation, the Belle collaboration developed a variant called Kakuno-Super-Fox-Wolfram moments (KSFW) [66]. All reconstructed particles associated with the signal B candidate (denoted as s for "signal") are treated separately from those in the rest-of-event (denoted as o for "others"). The $H^{so}_{x\ell}$ KSFW moments are sorted into three categories depending on whether the particle is charged ($x = c$), neutral ($x = n$), or missing ($x = m$). For even ℓ,

$$H^{so}_{x\ell} = \sum_i \sum_{j_x} |\mathbf{p}^*_{j_x}| \cdot \mathrm{P}_\ell(\cos\theta^*_{i,j_x}), \tag{3.8}$$

where i runs over the signal B decay products, and j_x over all other particles belonging to generic category x; $\mathbf{p}^*_{j_x}$ is the momentum of the particle j_x in the $\Upsilon(4S)$ frame; and $\mathrm{P}_\ell(\cos\theta^*_{i,j_x})$ is the ℓth order Legendre polynomial of the cosine of the angle between the particles i and j_x in the $\Upsilon(4S)$ frame. For odd ℓ one obtains $H^{so}_{n\ell} = H^{so}_{m\ell} = 0$ and

$$H^{so}_{c\ell} = \sum_{i} \sum_{j_x} q_i \cdot q_{j_x} \cdot |\mathbf{p}^*_{j_x}| \cdot \mathrm{P}_\ell(\cos\theta^*_{i,j_x}), \tag{3.9}$$

where q_i and q_{j_x} are the charges of the particles i and j_x. There are eleven $H^{so}_{x\ell}$ moments in total, two for $\ell = 1, 3$ and nine (3×3) for $\ell = 0, 2, 4$.

Other five KSFW moments are associated with the rest-of-event particles,

$$H^{oo}_\ell = \begin{cases} \sum_j \sum_k |\mathbf{p}^*_j| \cdot |\mathbf{p}^*_k| \cdot \mathrm{P}_\ell(\cos\theta^*_{j,k}) & (\ell = \text{even}) \\ \sum_j \sum_k q_j \cdot q_k \cdot |\mathbf{p}^*_j| \cdot |\mathbf{p}^*_k| \cdot \mathrm{P}_\ell(\cos\theta^*_{j,k}) & (\ell = \text{odd}), \end{cases} \tag{3.10}$$

where j and k extend over all rest-of-event particles.

To eliminate the dependence on ΔE, the $H^{so}_{x\ell}$ moments are normalized to H^{max}_0, and the H^{oo}_ℓ moments to $(H^{max}_0)^2$, where $H^{max}_0 = 2(\sqrt{s} - \Delta E)$. Figure 3.11 shows the distribution of the KSFW moments for $\ell = 2$, which offer the higher discrimination power.

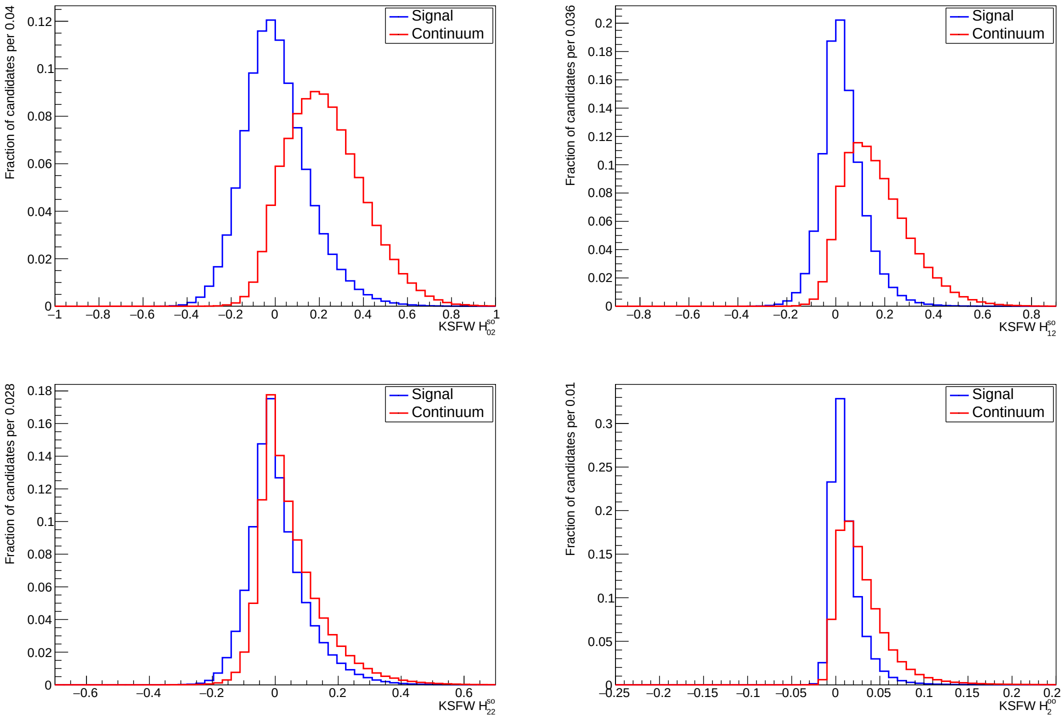


Figure 3.11: Distributions of the $\ell = 2$ KSFW moments for simulated signal and continuum events. Distributions normalized to unity.

Flavor variables. Many variables used by algorithms dedicated to inferring the quark content of the partner B meson (flavor tagging) offer additional discriminating information. If such algorithms indicate that a partner B meson is reconstructed in an event, that information obviously correlates with the presence of a B meson and therefore helps suppressing continuum. The flavor of the B meson accompanying the signal candidate is identified via properties of its decay products, charged leptons, kaons, and pions. Flavor tagging algorithms yield the flavor of the tagged meson,

$q = \pm 1$, and a flavor-tagging quality factor ($r = 1 - 2w = 0 - -1$), proportional to the probability that the identification is correct, where w is the mistagging probability, $r = 0$ means no flavor discrimination, and $r = 1$ means unambiguous flavor assignment. Hence, continuum events are expected to cluster close to zero of $q \cdot r$ distribution, while B events peak at $q \cdot r = \pm 1$ (Fig. 3.12, top left panel).

Additional discriminating flavor-related variables include (Fig. 3.12),

|qpKinLepton| – the probability of the tag-side track with highest probability for being a primary lepton,

|qpKaon| – the probability of the tag-side track with highest probability for being a $b \to c \to s$ kaon,

|qpSlowPion| – the probability of the tag-side track with highest probability for being a low-momentum pion from the decay of a primary D^*.

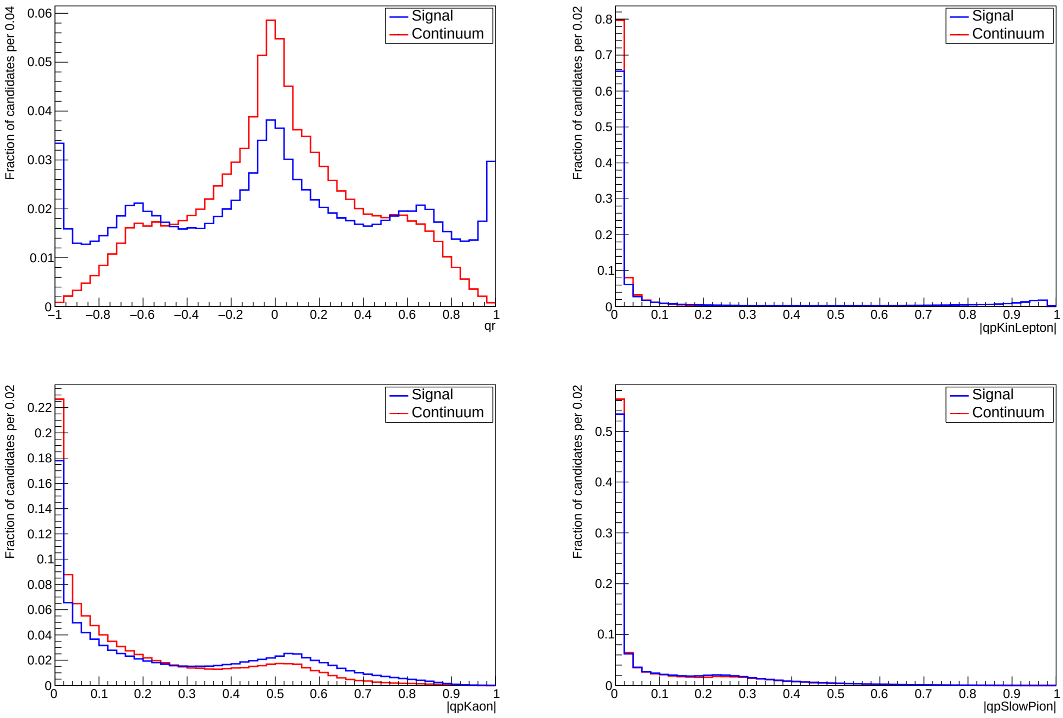


Figure 3.12: Distributions of (top left) qr, (top right) |qpKinLepton|, (bottom left) |qpKaon|, and (bottom right) |qpSlowPion| in simulated signal and continuum events. Distributions normalized to unity.

3.4 Analysis overview

This work aims at measuring the branching fraction, *CP* asymmetry, and fraction of longitudinally polarized $B^+ \to \rho^+\rho^0$ decays, which are the parameters that mostly limit the 4% precision on the current determination of α CKM angle. This is a "blind" analysis; I develop the whole analysis using only simulated and control-data samples without looking into the signal-enriched region in data until all procedures are established and final. Such approach prevents from tuning the measurement procedure towards the expected or desired result, reducing the chances for bias.

A priori, the principal challenge is to overcome the initial 10^{-6} signal-to-background ratio with a selection sufficiently discriminating to isolate an abundant, low-background signal without introducing intractable correlations in the multidimensional fit that determines the parameters of interest. First, I restrict the sample to hadronic events by applying baseline requirements determined from simulation or control-data samples. Then, I design a multivariate statistical-learning method that combines non-linearly about 20 kinematic, decay-time, final-state topology, and flavor-tagger variables to suppress the dominant continuum background. Also, I address remaining backgrounds from bottom meson production, which are less pervasive in rate than continuum, but similarly harmful as their experimental signatures mirror signal. I optimize the selection by maximizing the expected statistical precision on the physics parameters of interest.

After the selection, continuum and B background events still dominate the sample. I perform a multidimensional fit that allows determining sample composition. The main challenge is to develop a model that properly accounts for the multidimensional dependences existing between fit observables. The fit development is the core of this analysis. The outcome is a robust model as shown by tests of the estimator properties using simulated data under various experimental conditions.

A crucial point is to understand and correct possible discrepancies that can arise between the simulation and the data. To study these effects, I implement the key steps of the analysis on abundant control samples that share relevant features with the signal decay, and perform additional studies in signal-empty control regions ("sidebands") of the signal sample.

Finally, I apply simulation-based efficiency and acceptance corrections to the fit results to determine the final $B^+ \to \rho^+\rho^0$ results. Significant effort is devoted keeping the impact of systematic uncertainties under control.

Once every step is validated, I apply the analysis to the full Belle II data sample available as of this writing.

Chapter 4

Signal sample selection

This chapter describes signal candidate reconstruction, where I select and combine tracks and photons applying requirements to enrich the sample in signal $B^+ \to \rho^+\rho^0$ events, and background suppression, which aims at improving purity to extract a visible $B^+ \to \rho^+\rho^0$ signal.

4.1 Data sample

4.1.1 Experimental data

The Belle II experiment started the physics data taking in March 2019, accumulating electron-positron collision data corresponding to an integrated luminosity of 424 ± 3 fb^{-1}. I use the full data set collected at the energy of the $\Upsilon(4S)$ resonance mass up to July 2022, corresponding to 362 ± 2 fb^{-1}, to reconstruct the signal and control modes. I also use data collected at a collision energy 60 MeV below the $\Upsilon(4S)$ mass (called henceforth "off-resonance" data), and corresponding to 42.3 ± 0.3 fb^{-1} of integrated luminosity, to characterize and study continuum background.

4.1.2 Simulated data

I use simulated data to optimize the event selection, to estimate acceptances and efficiency for reconstructing and selecting signal and control modes, and to study and model the relevant distributions used in the sample composition fit.

Simulated samples are based on the Monte Carlo approach. Monte Carlo samples are produced using event generators, which are computer programs that use pseudorandom number generators to produce sets of four-vectors reproducing final states of e^+e^- collisions according to theoretical models of particle kinematics and interactions. Generated data are then subjected to detector simulation, where models of the detector geometry and material are interfaced with models of interactions of particles with matter and signal formation to reproduce the expected values of the quantities observed in the detector. The resulting simulated data contain information about reconstructed particles and about the generated *true* particles. By matching these, we understand whether the observed particles are reconstructed properly, or what are the most frequent misreconstruction occurrences, and what are backgrounds. This truth-matching procedure is useful to optimize selection requirements, calculate signal efficiency, classify sample components, and for many consistency checks.

Figure 4.1 shows a sketch of the generation sequence for an hadronic event in Belle II. The properties of virtual photons, created in the electron-positron annihilation, and its subsequent splitting into a quark-antiquark pair, which in turn produces the observed hadrons, is simulated by `Pythia` [67]. Finally, the decay of the heavy hadron (top right corner in Fig. 4.1) is generated according to a certain `EvtGen` model [68]. The photon emission by final-state charged particles is generated by `PHOTOS` [69].

For signal studies, I simulate exclusive samples of $e^+e^- \to B^+B^-$ events, where one of the B mesons is forced to decay to the $\rho^\pm\rho^0$ final state using the `EvtGen` generator. `EvtGen` simulates the time evolution and known decays of B-mesons according to the relevant experimental and theoretical knowledge. The decay model is `SVV_HELAMP`, which uses the helicity amplitude formalism to describe the decay of a pseudoscalar into two vector particles. Subsequent ρ decays are modeled by the `VSS` model that describes the decay of a vector to two pseudoscalar particles. These models account for the spin configuration of the process, which is essential to simulate accurately the relevant angular distributions used in the analysis. In addition to `EvtGen`, the `PHOTOS` generator is used to account for photon emission by charged final-state particles, known as final state radiation (FSR). Such photons can distort the shapes of relevant distributions in data. The sample is then fed through the standard Belle II detector simulation, based on the CERN `GEANT4` package [70], which simulates interaction with matter and signal formation yielding simulated data in the same format as experimental data. To account for possible charge-dependent instrumental asymmetries, I generate B^+ and B^- signal decays separately. I also produce separately longitudinally polarized and transversely polarized $B^+ \to \rho^+\rho^0$ decays. In total, I generate 10×10^6 $B^+ \to \rho^+\rho^0$ decays for both polarizations, which are about a factor 1000 more than expected in data.

For continuum background studies, I use centrally produced simulated samples corresponding to almost three times the size of the full Belle II data sample available to this writing. These samples include $e^+e^- \to u\bar{u}$, $e^+e^- \to d\bar{d}$, $e^+e^- \to s\bar{s}$, and $e^+e^- \to c\bar{c}$ events generated using `KKMC` [71] interfaced with `Pythia`, and `PHOTOS`.

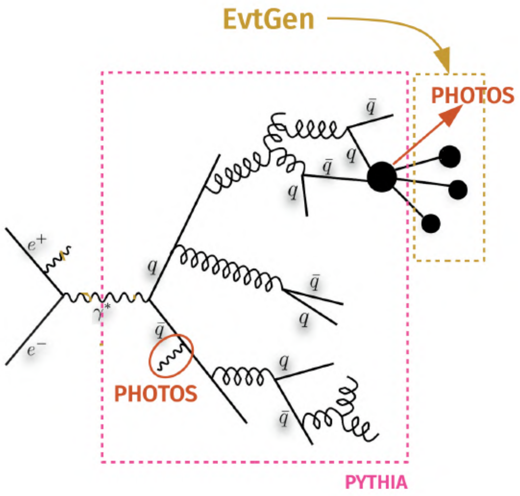


Figure 4.1: Sketch of the various portions of each simulation package models of an hadronic event in Belle II.

For $B\overline{B}$ background studies, I use centrally simulated $e^+e^- \to B^+B^-$ and $e^+e^- \to B^0\overline{B}^0$ events generated with `EvtGen`, where B mesons undergo all their allowed decays. The relative proportions among the B decay modes are based on known values or upper limits for measured/observed decays [10], and on arbitrary or educated guesses for unobserved decays.

Exclusive simulated samples each dedicated to a specific decay mode are also produced for peaking background and signal decays. Decay modes, number of generated events, and generator models used are summarized in Table 4.1. We assume that a_1 meson decays only in two-body states, $\rho\pi$ or $f_0\pi$.

In addition to full-fledged simulated samples, I also use simplified simulated samples ("toy") based on pseudorandom extraction of simulated values of data based on the likelihood function of the fit of sample composition. These are used mostly for prototyping the analysis and assessing systematic uncertainties.

Decay channel	f_L	N ($\times 10^6$)	`EvtGen` model
$B^+ \to \rho^+(\to \pi^+\pi^0)\rho^0(\to \pi^+\pi^-)$	1	10	`PHOTOS SVV_HELAMP`
$B^+ \to \rho^+(\to \pi^+\pi^0)\rho^0(\to \pi^+\pi^-)$	0	10	`PHOTOS SVV_HELAMP`
$B^+ \to \rho^+(\to \pi^+\pi^0)\rho^0(\to \pi^+\pi^-)$	0.95	10	`PHOTOS SVV_HELAMP`
$B^+ \to \rho^+ f_0$		2	`PHOTOS SVS`
$B^+ \to a_1^+\pi^0$		2	`PHOTOS SVS`
$B^0 \to a_1^{\pm}\pi^+$		2	`PHOTOS SVS`
$B^+ \to a_1^0\pi^+$		2	`PHOTOS SVS`
$B^+ \to \rho^0\pi^+\pi^0$		2	`PHOTOS PHSP`
$B^+ \to \rho^+\pi^+\pi^-$		2	`PHOTOS PHSP`
$B^+ \to \rho^-\pi^+\pi^+$		2	`PHOTOS PHSP`
$B^+ \to f_0\pi^+\pi^0$		2	`PHOTOS PHSP`

Table 4.1: Exclusive simulated samples used for signal, control channel, and peaking background studies along with the `EvtGen` decay models used. `PHSP` is a model that averages all spins of particles in the initial and the final state. `SVS` is a model that describes the decay of a pseudoscalar to a vector and pseudoscalar particles. We also use `SVV_HELAMP` model for decays of a pseudoscalar to an axial vector plus a vector.

4.1.3 Basic data structures

As many physics topics are studied, Belle II data are analyzed multiple times by hundreds of collaborators. To ease the analysis of such a large amount of data, various centralized data processing steps are implemented.

Raw data are processed centrally to produce summary data, called mDST, which are reduced in size and contain higher-level information related to primitives for physics analysis, including four-momenta, vertices, particle-identification information, and others. A second centralized step consists in applying loose selection criteria on the mDST data to obtain analysis-specific subsets (*skims*) further reduced in size so that each collaborator can access them. The most relevant of such subsets for our analysis is the `hlt_hadron` skim.

4.1.4 HLT hadron selection

Several processes in e^+e^- collisions at $\sqrt{s} = 10.58$ GeV occur with much higher rates than to $B\overline{B}$ pair production (Table 2.5). To restrict the sample to hadronic events, which are those of interest in this work, QED events, i.e. Bhabha scattering, lepton or photon pair production, and beam gas interaction events where a single electron or positron interacts with residual gas molecules, are suppressed.

This is achieved with a centralized loose selection based on a minimum number of charged particles and calorimeter clusters meeting basic quality criteria [72]. Charged particles are required to have $p_T > 200$ MeV/c and impact parameters $|dz| < 4$ cm and $|dr| < 2$ cm to suppress tracks generated in collisions of beams with residual gas molecules. To make it into the `hlt_hadron` skim an event should meet the following conditions

- have three or more good tracks originated from a primary vertex, to suppress Bhabha, $e^+e^- \to \mu^+\mu^-$, two-photon, and low multiplicity beam-gas events, as all these processes produce two or fewer tracks in the final state.
- the event should **not** pass the HLT selection for Bhabha events, which consists in a set of requirements on the number of tracks identified as $e^\pm$, their momenta, the angle between them and the total energy deposit in the ECL.

After applying these criteria, more than 99% (85%) of all $B\overline{B}$ ($q\overline{q}$) events are retained, with a rejection of more than 99% of non-hadronic events as shown in Table 4.2.

	$B\overline{B}$	$c\overline{c}$	$q\overline{q}$	$\tau^+\tau^-$	$\mu\mu$	ee
ϵ	0.995	0.963	0.798	0.273	0.0096	0.0014

Table 4.2: Efficiencies for various processes selected through the hadronic stream selection.

4.2 Signal candidates reconstruction

The full decay chain of interest is $B^+ \to (\rho^+ \to \pi^+\pi^0(\to \gamma\gamma))(\rho^0 \to \pi^+\pi^-)$. The first major challenge of this analysis is to isolate a low-background signal. The four pions in the final state, which are also the most common hadrons in continuum events, along with lack of narrow resonances in the intermediate state to suppress combinatorial background, call for a dedicated discrimination.

4.2.1 Charged pions

To reconstruct $\pi^\pm$ candidates I use successfully reconstructed tracks with impact parameters $|dz| < 3$ cm and $|dr| < 0.5$ cm to suppress tracks originated from beam-gas events. Tracks are also required to be in the CDC angular acceptance and to have at least 20 hits in the CDC. This selection retains 95% of correctly truth-matched pions in simulated signal decays.

4.2.2 Neutral pions

To reconstruct π^0 candidates, I require photons with energies $E_\gamma > 20$ (22.5) MeV in the barrel (endcap) section of the ECL to reduce beam-induced background and ECL noise, which varies across the polar compartments of the calorimeter. I then require the diphoton mass to meet $115 < m(\gamma\gamma) < 150$ MeV/c^2 corresponding to $\pm 3\sigma$ in resolution and constrain the mass of the resulting candidates to the known value by applying a kinematic fit (`massKFit` [61]). Only candidates for which the `massKFit` has converged are retained. The cosine of the angle between the momentum difference of the two photons in the π^0 frame and the π^0 momentum in the lab frame (`cosHelicityAngleMomentum`) is required to be lower than 0.98. The requirements on the diphoton mass and `cosHelicityAngleMomentum` suppress misreconstructed neutral-pion candidates as shown in Fig. 4.2. True neutral pions tend to cluster in the 105–150 MeV/c^2 mass range and be simmetrically distributed in `cosHelicityAngleMomentum`, unlike misreconstructed pions.

To further suppress photon candidates coming from soft energy deposits of beam background particles in the calorimeter, I increase the threshold on the photon energy and exploit an additional `photonMVA` variable, provided with the calorimeter reconstruction and obtained by combining various information of the calorimeter cluster. This selection retains 81% of correctly truth-matched neutral pions from simulated signal decays.

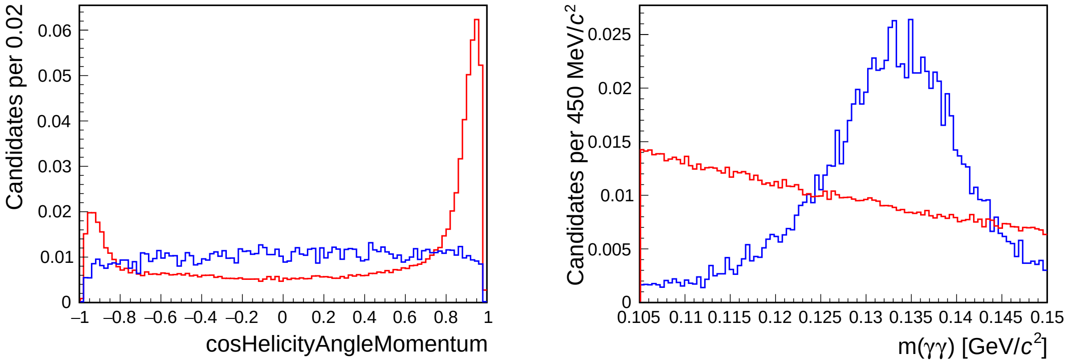


Figure 4.2: Distribution of (left) `cosHelicityAngleMomentum` and (right) diphoton mass in simulated signal data for (blue) truth-matched π^0 candidates and (red) background candidates. Distributions normalized to unity.

4.2.3 ρ mesons

Candidate π^+ and $\pi^-(\pi^0)$ are paired in a kinematic fit to form $\rho^0(\rho^+)$ meson candidates. The resulting invariant mass is required to be consistent with that of a ρ meson by imposing $0.52 < m(\pi^+\pi^{-(0)}) < 1.06$ GeV/c^2 (Fig. 4.3), which retains 97% of all signal ρ mesons according to simulation.

In addition, I require the ρ^+ helicity angle to meet $\cos\theta_{\rho^+} < 0.75$ to compromise between signal efficiency and background contamination from misreconstructed low-energy π^0 candidates. B candidates containing such neutral pions tend to peak at high values of $\cos\theta_{\rho^+}$. By excluding $\cos\theta_{\rho^+} > 0.75$ events I remove a large fraction of continuum background events as shown in Fig. 4.4.

4.2.4 B mesons

B candidates are reconstructed by combining the energy-momentum four-vectors of ρ^+ and ρ^0 mesons into a kinematic fit that constrains three final state tracks to originate from a common primary vertex. I restrict the data sample to the $5.27 < M_{\rm bc} < 5.29$ GeV/c^2 and $-0.15 < \Delta E < 0.15$ GeV ranges (Fig. 4.5). The $M_{\rm bc}$ range narrowly encloses the signal region to reduce background, as we choose not to include $M_{\rm bc}$ among the fit observables because it shows a dependence on ΔE through π^0 reconstruction imperfections. Excluding $M_{\rm bc}$ avoids further sophistication of an already complicated fit model. Monte Carlo simulation shows that this choice sacrifices at worst 5% of signal yield. After the selection, multiple B meson candidates per event are reconstructed in a fraction of events as shown

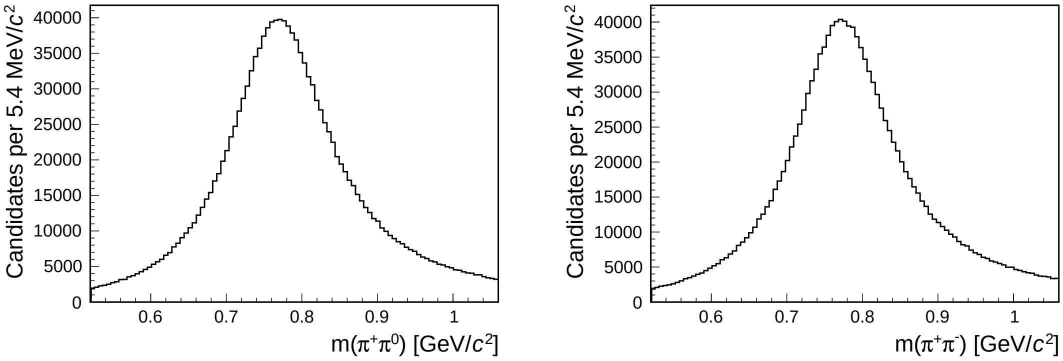


Figure 4.3: Distribution of (left) $\pi^+\pi^0$ and (right) $\pi^+\pi^-$ invariant mass reconstructed in simulated signal decays.

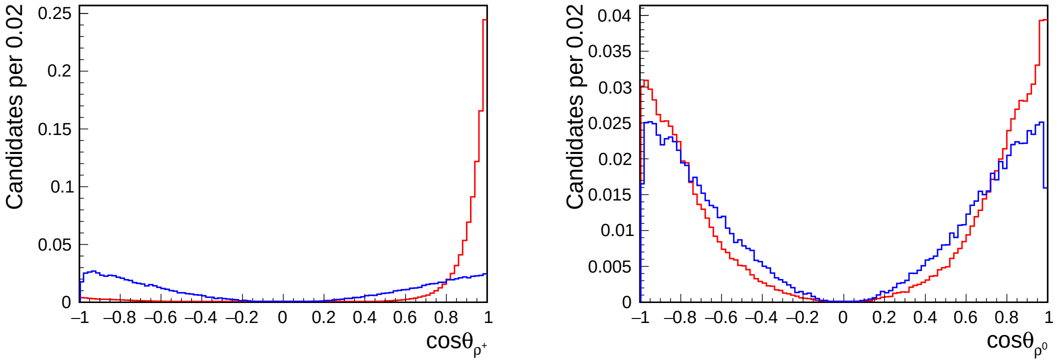


Figure 4.4: Distribution of the cosines of the helicity angles for (left) ρ^+ meson and (right) ρ^0 meson from simulated (blue) signal and (red) background components. Distributions normalized to unity.

in Fig. 4.6. In average, 1.13 B candidates per event are reconstructed in each simulated event.

I restrict the samples to one candidate per event by ranking the candidates according to $\chi^2(\pi^0, \text{vertex}) = \chi^2(\pi^0) + \chi^2(\text{vertex})$, where $\chi^2(\pi^0)$ is the χ^2 probability of the π^0 fit and $\chi^2(\text{vertex})$ is the χ^2 probability of the signal B-vertex fit. Simulation shows that the efficiency of such criterion is 96% (98%) on longitudinally (transversely) polarized signal decays, to be compared with the 93% (96%) figure associated with a random choice.

4.3 Continuum background

The resulting sample after this first selection is still dominated by events from direct production of light-pair of quarks and other B meson decays, as shown by Fig. 4.7 of the ΔE distributions in realistically simulated and data samples resulting from the baseline selection. No visible signal peak is observed at $\Delta E \approx 0$ region. Further background suppression is needed.

At the $\Upsilon(4S)$ resonance, more than 3/4 of all hadronic events come from $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$) processes (see Fig. 2.1). Random combinations of particles produced in such events form large amounts of background $B^+ \to \rho^+\rho^0$ candidates that outnumber the signal decays by orders of magnitude. To discriminate signal $B^+ \to \rho^+\rho^0$ from con-

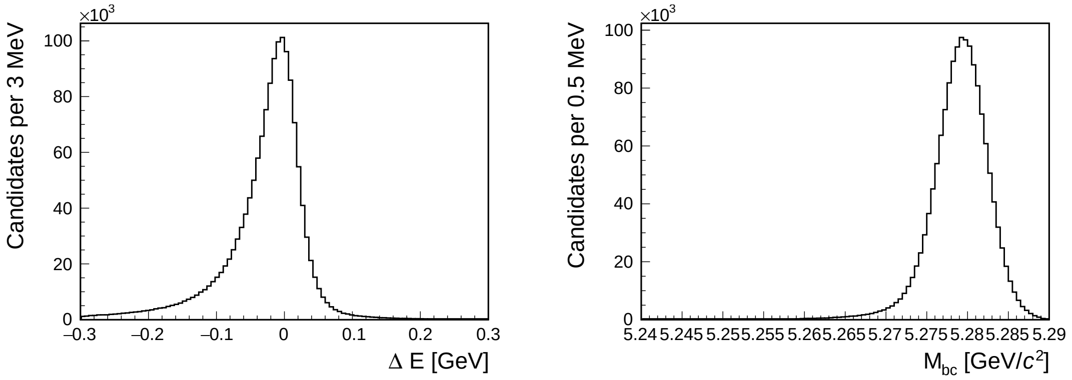


Figure 4.5: Distribution of (left) ΔE and (right) $M_{\rm bc}$ reconstructed in simulated signal decays.

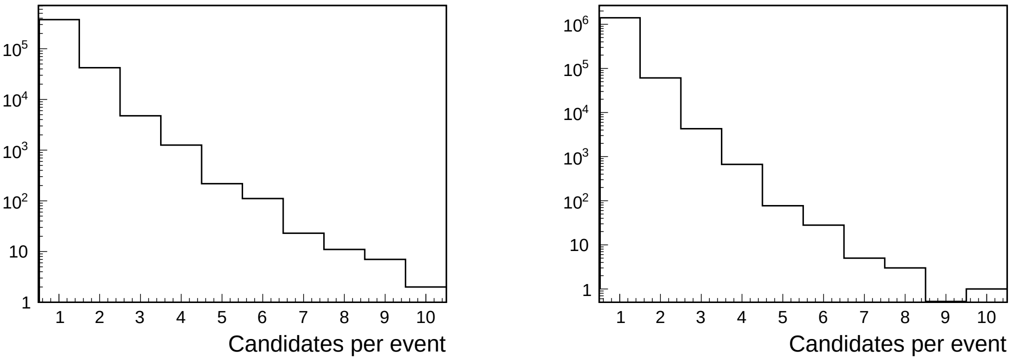


Figure 4.6: Distributions of B candidate multiplicity in simulated (left) longitudinally and (right) transversely polarized signal decays. The vertical axis is log-scaled.

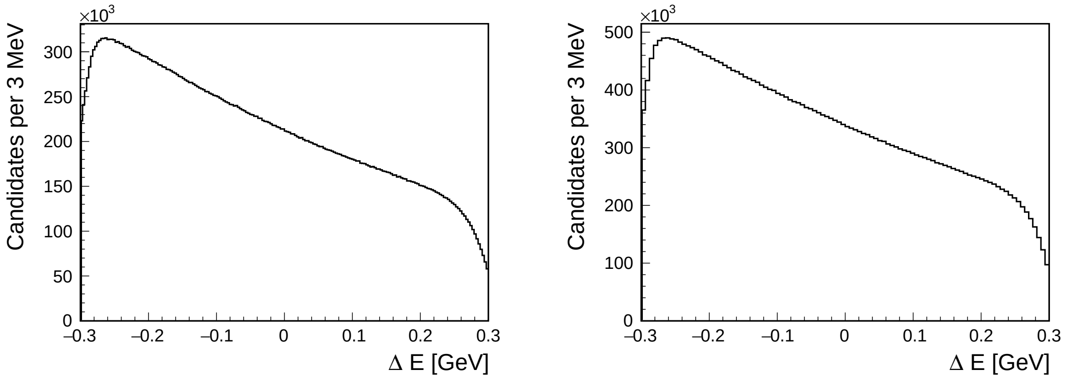


Figure 4.7: Distribution of ΔE in (left) realistically simulated and (right) data sample selected with the baseline suppression.

tinuum, I combine non-linearly various discriminating variables in a boosted decision tree (FBDT) [73], a supervised multivariate classifier. The idea is that nonlinear discriminators may achieve a better signal-to-background separation than sequential application of one-dimensional restrictions ("cuts") as they can capture and exploit statistical differences present in the multidimensional correlations over the variable space.

A multivariate classifier estimates the probability of a datum to belong to a given class;

this probability is inferred from a set of explanatory observables $\mathbf{x} = (x_1, ..., x_n)$. The algorithm operates in two phases. In the fitting phase, the classifier is "trained" using data with known classification (training sample). In this "supervised" phase the internal configuration of the classifying function that maps the inputs into a classification output is adjusted so as to maximize the rate of successful classification. Successful classifications are known because the classification is known for the training data. In the application phase, the resulting classifier is applied to new data-points with unknown classification (testing sample). In this phase, the internal configuration of the classifying function is established from the training and used to classify the test data. I use an FBDT based on a stochastic gradient-boosted decision tree.

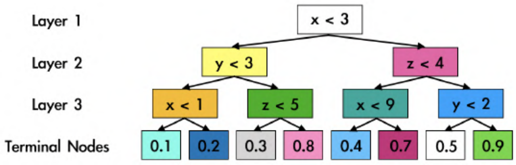


Figure 4.8: Schematic example of a three-layer decision tree. At each node of the tree a binary decision is made until a terminal node is reached. The numbers in the terminal node correspond to a probability of the test data-point to be signal.

The decision tree is a specific type of supervised classifier that approximates the optimal classifying function by applying a set of consecutive binary requirements on each of the given discriminating observables. The maximum number of consecutive requirements is typically configurable and is called the depth of the tree. A schematic example of a decision tree is shown in Fig. 4.8 for a simple case of three discriminating observables and two classes.

The tree requirements are determined during fitting. By varying the requirements at each node, the tree estimates the probability of a data-point to belong to a certain class of events. These predictions are then compared with the known true classification available for the training sample. Only requirements that result in accurate predictions are kept in the node. Hence each requirement maximizes locally the separation between classes of events on the given training sample.

The predictions of a tree with many consecutive requirements (a deep tree) is often driven by the statistical fluctuations of the training data-sample, instead of the genuinely distinctive features. This "over-fitting" reduces the predictive power of the tree.

To reduce overfitting, ensembles of shallow trees are combined into a "boosted" tree. While individually each shallow tree may give inaccurate predictions, combining them sequentially yields a model that is less likely to over-fit and has good classification performance. Boosting proceeds by fitting an initial tree to the data; then a second tree is built targeted at classifying accurately only the events where the first tree performs poorly; and then the sequence is repeated many times. Each successive tree attempts to correct the shortcomings of the combination of previous trees. A common used boosting technique is gradient boosting, which builds consequent trees such that the overall misclassification rate is minimized [74].

The robustness of gradient-boosted decision-trees against over-fitting is improved by using random subsamples of the training data set instead of the full training sample in each boosting step. The strength of possible correlations between trees is reduced thus achieving enhanced discriminating capabilities. This approach is called stochastic gradient boosted decision tree [75] and it is what is used here.

4.3.1 Choice of discriminating observables

I explore a broad set of discriminating variable to choose the FBDT inputs that offer the optimal separation of signal from continuum. The challenge is that variables should not only provide separation, but also not introduce intractable dependences among observables used in the subsequent fit of sample composition. Otherwise, fitting the sample composition could prove prohibitively complicated. Moreover, the chosen discriminating observables need to be well-reproduced by simulation. Otherwise, the classifier trained on the simulated samples will be suboptimal when applied to data.

I focus on the event-level observables described in Section 3.3.4, because they are the most powerful to separate B from continuum background as shown in many previous B-factory analyses. Starting from a total of about 40 observables, I first study statistical dependences between each of the potential discriminating observables and the observables of our fit of sample composition, ΔE, $\cos\theta_{\rho^+}$, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^0}$, and $m(\pi^+\pi^-)$ (see Sec. 5.1). I converge to a subset of 17 discriminating variables that are not very correlated with the relevant fit observables.

To ensure that the chosen observables are well-reproduced in the simulation, I compare their distributions obtained in simulated and collision data. I use the $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decay mode as it has a similar final state as the signal decay and is sufficiently abundant. I reconstruct $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays in on-resonance simulated and collision data applying the same continuum suppression selection as for signal. I compare simulation and data separately for background events in the sideband region $5.211 < M_{\mathrm{bc}} < 5.224$ GeV/c^2 and signal events in the region $M_{\mathrm{bc}} > 5.27$ GeV/c^2 statistically obtained by subtracting the sideband distributions. The sideband region is chosen such that it contains in total as many events as the simulated continuum in the signal region, as shown in Fig. 4.9. Figures 4.10–4.15 show data-simulation distributions of the chosen 17 discriminating observables. As observed in other Belle II analyses, a few discriminating observables are mismodeled in simulation, e.g., $|\cos\theta_{\mathrm{TBTO}}|$ and R_2 in Figs. 4.10 and 4.15 respectively. A brief investigation indicates that such mismodelings might be due to two causes. One is the possible wrongly simulated sample-composition of the partner-B side of the event. The other is an imperfect sideband subtraction, which may occur if the sample composition varies rapidly between the events chosen as background proxy and the events where the signal is. However, the observed mismodeling is unlikely to be disruptive for the rest of the analysis: it might lead, at worst, to a moderately suboptimal final selection, but it does not bias the measurements results, provided that data-driven selection-efficiency validation and correction is applied. In addition, the data-simulation consistency in the FBDT output shows that the final discriminator is largely insensitive to these mismodelings (Fig. 5.3).Hence, we don't investigate these further.

4.3.2 Classifier training and test

I train the FBDT using 5×10^5 truth-matched signal events (with realistic proportions of longitudinally and transversely polarized and of positively and negatively charged B candidates) and 5×10^5 simulated continuum events passing the baseline selection.

A convenient way to estimate the classifier performance is through a receiver operating characteristic (ROC) curve, which represents signal efficiency as a function of background rejection (defined as the complement to the background efficiency). The classifier performance is better when the ROC curve approaches the top right corner of the plot. This is expressed quantitatively by the area under the curve (AUC), which equals one for 100% signal efficiency with 100% background rejection. I test the performance of the FBDT

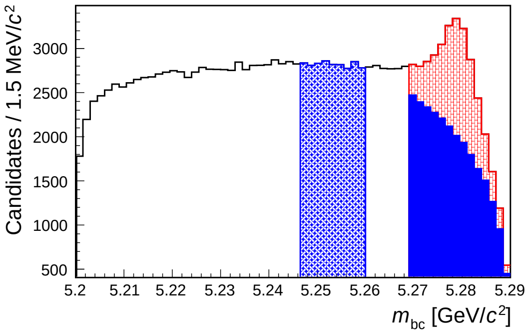


Figure 4.9: Distribution of M_{bc} and definition of the (blue) side-band and (red) signal region for simulated $B^+ \to \overline{D}^0(\to K^+\pi^-\pi^0)\,\pi^+$ decays.

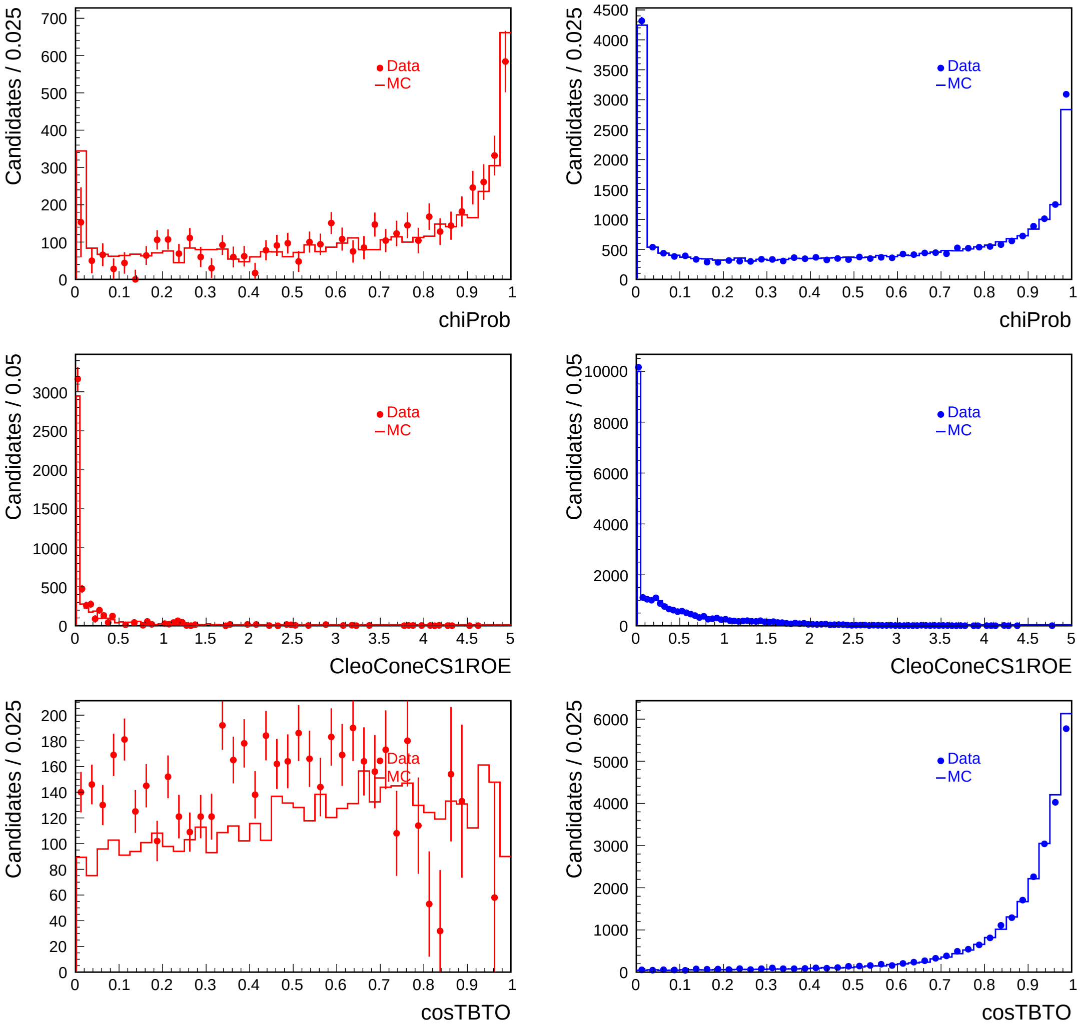


Figure 4.10: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (1/6).

classifier in various configurations and compare it also with existing configurations used by other Belle II analyses. Figure 4.16 shows ROC curves for two configurations: the red

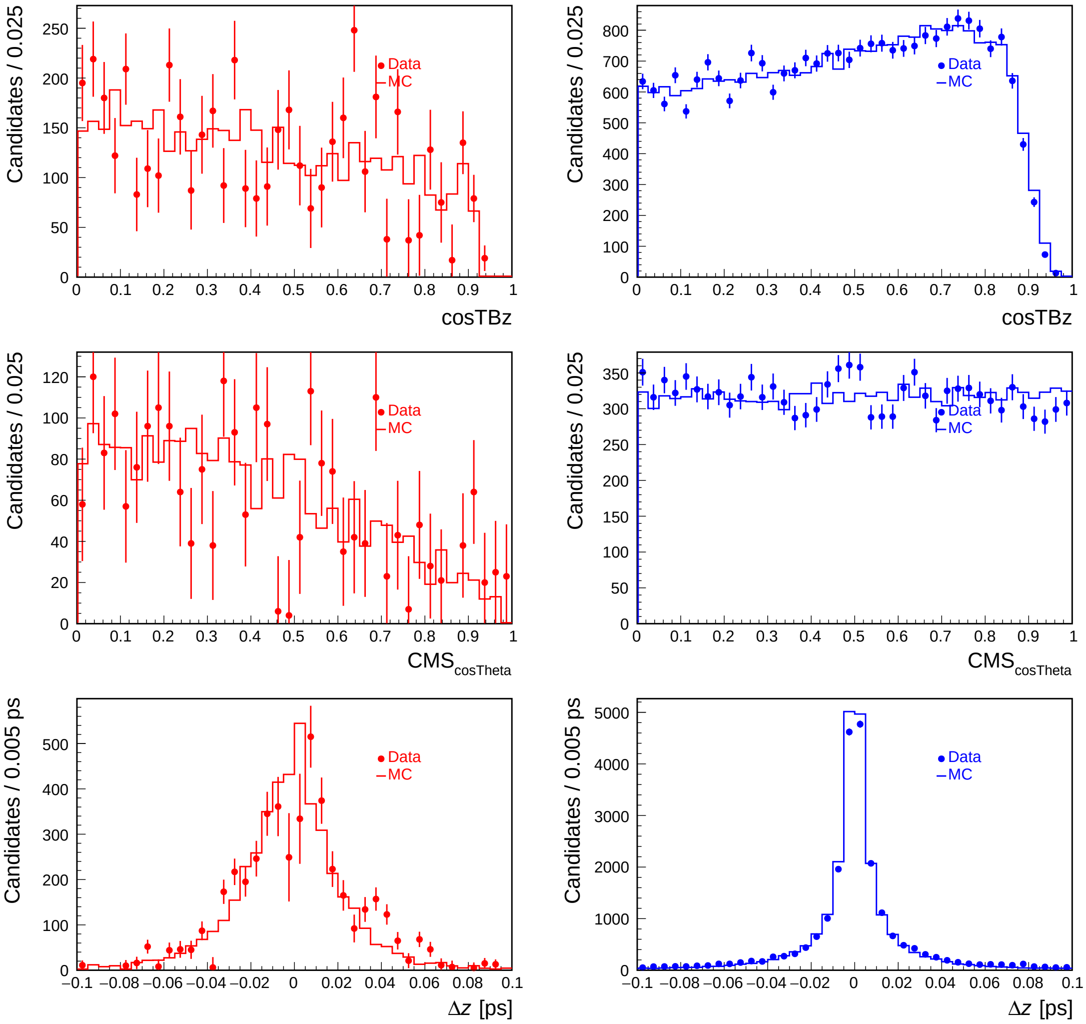


Figure 4.11: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (2/6).

curve corresponds to a standard configuration used in many Belle II analyses of B decays in hadronic final states, while the blue is the output of the classifier that I trained specifically for the $B^+ \to \rho^+\rho^0$ channel, removing from the input set of variables the ones too correlated with the fit observables. The latter has the largest AUC value. I therefore use this configuration throughout the analysis.

I test the FBDT on independent simulated samples with same composition as the training data to ensure that the model is not over-fitted. Figure 4.17 compares distributions of classifier output for training and testing samples. The distributions obtained from training and testing data agree for both signal and background components, showing negligible, if any, overfitting.

4.3.3 Selection optimization

The ROC curve (Fig. 4.16) qualifies the classifier performance but it does not indicate which requirement on the classifier output provides the best selection, as this depends on the physics goal at hand. Since in this analysis the continuum suppression output is also

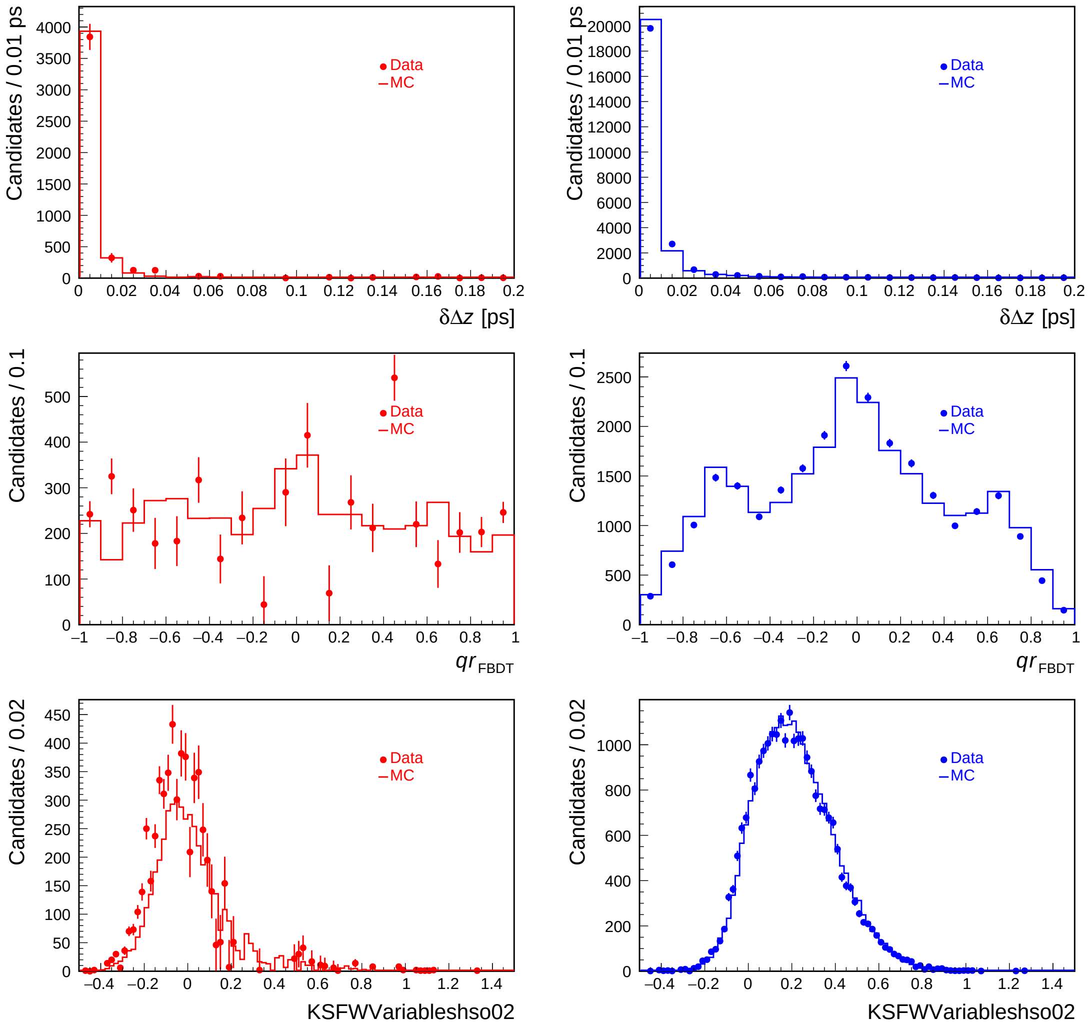


Figure 4.12: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (3/6).

an observable used in the fit of sample composition, the choice of the optimal threshold is not straightforward. Usual figures-of-merit such as $S/\sqrt{S+B}$ may fail to capture the real statistical power of the measurement. I optimize the selection by minimizing the average expected statistical variance of the parameters of interest. These are determined by analyses of simplified simulated samples drawn from probability density functions based on empirical modeling of realistically simulated samples. This is ideally the optimal approach as it enhances the statistical power on the very results one is targeting. However, implementation is challenging since it implies repeating the measurement on many simulated samples selected through different criteria, with several technical implications in terms of remodeling, computing time, and so forth.

In addition to the requirement on the classifier output, I also need to optimize the particle identification requirement associated with the probability for a pion to be identified as a kaon, to reduce backgrounds from other B decays where kaons are misidentified. As particle identification information is used in the flavor-tagging, which is an input to the continuum suppression FBDT, continuum suppression and particle identification can

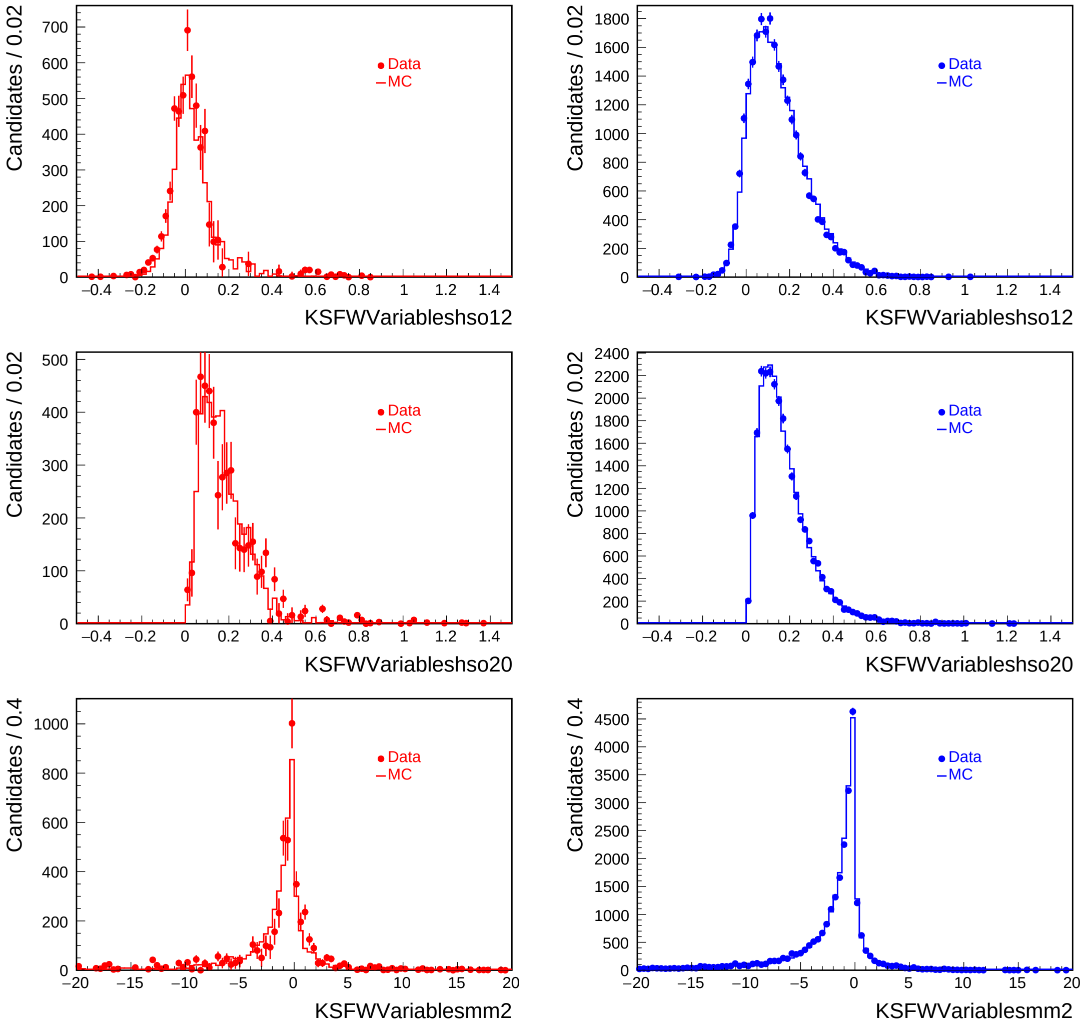


Figure 4.13: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (4/6).

be correlated. Therefore, a simultaneous optimization over the two-dimensional space of continuum suppression and particle identification is needed.

I use a simplified version of the sample composition fit that includes the relevant features for this study but allows a swifter modeling. Details on the fit are discussed in the next chapter. Here I only outline the general aspects relevant to the optimization. The likelihood is the same as used in the full sample-composition fit but the number of components is reduced. I include only the two signals and the combinatorial backgrounds from continuum $B\overline{B}$ events, neglecting (both in generation and fitting) the other peaking backgrounds from $B\overline{B}$. This choice is not expected to impact the optimization as those are B decays that are nearly indistinguishable from the signal for the continuum suppression classifier; longitudinal and transversely polarized signal components are taken in proportions corresponding to $f_L = 0.95$, the current known value. To determine the average expected statistical uncertainty, I study simplified simulated ("toy") samples selected through various continuum suppression and pion identification requirements sampled in a two dimensional grid. For each test selection choice, 1000 toy samples are generated all of the same sample size. In

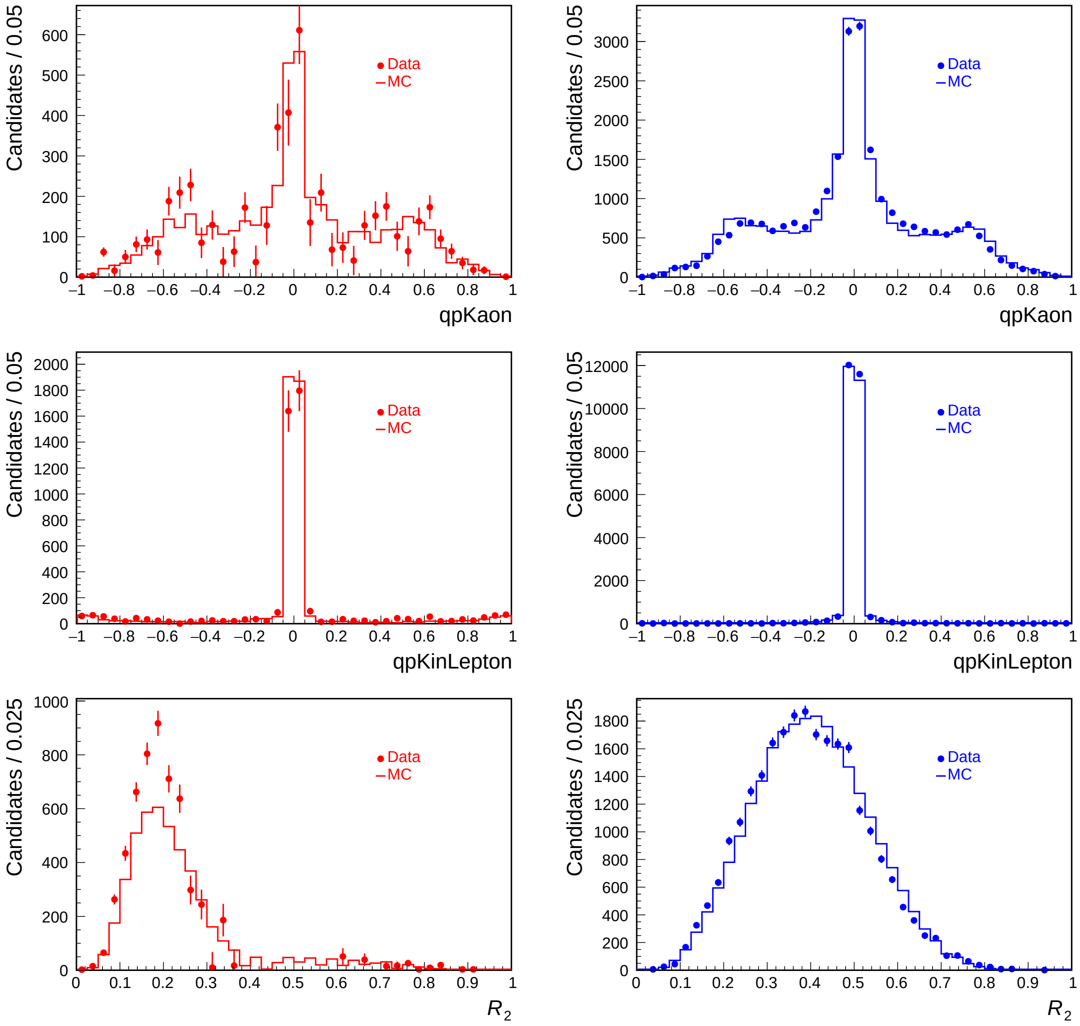


Figure 4.14: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (5/6).

generation, the yields of each component are drawn according to multinomial distributions mean valued at the yields expected in data, except for the continuum component which is generated with the complementary yield to match the predefined sample size. All the distributions are drawn from simulated samples. I fit the composition of toy samples using a model appropriate for each of the selection scenarios. I then study the distributions of residuals $\mathcal{R}(\theta_i) = \hat{\theta}_i - \theta_i$ for the three parameters of interest, where θ_i is the true value and $\hat{\theta}_i$ is the fit estimate.

The results are shown in Fig. 4.18. The resulting optimal values for $\mathcal{B}$ and f_L are an upper bound at 0.95 for the probability of a pion to be identified as a kaon and a lower bound at 0.96 for the continuum suppression output. I choose these requirements even if they are not optimal for A_{CP}: I prefer to give up some statistical precision, but work with a sample that is not dominated by the continuum background, that could have resulted in additional large systematic uncertainties.

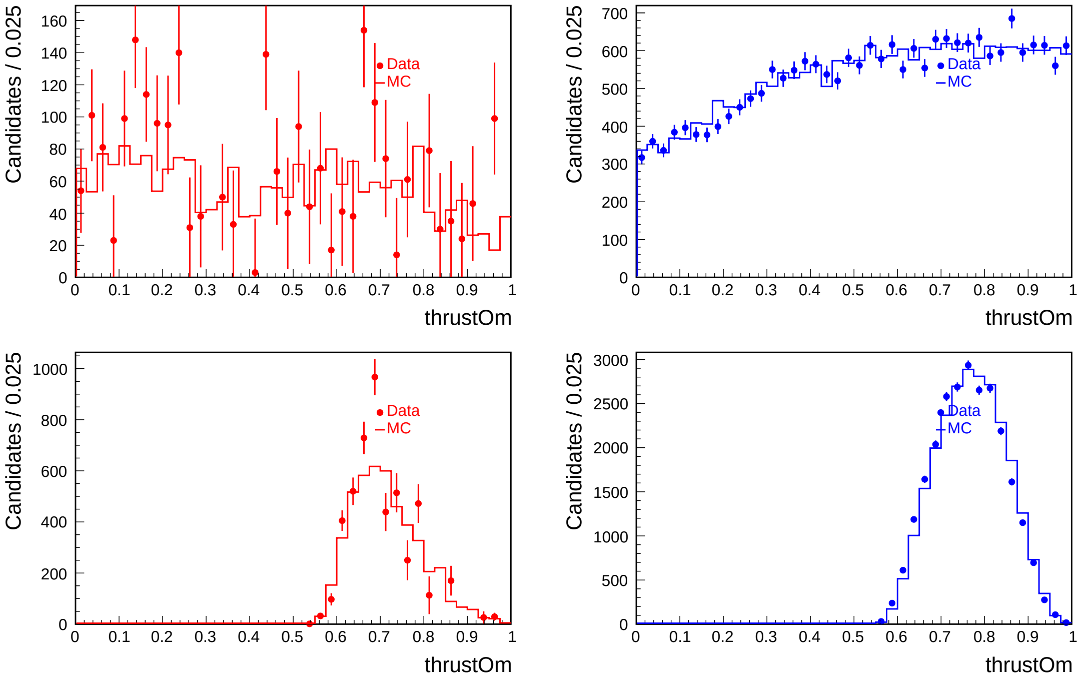


Figure 4.15: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (6/6).

4.4 Self cross-feed background

Generated signal $B^+ \to \rho^+\rho^0$ events are classified into two categories after reconstruction: properly reconstructed ("true") $B^+ \to \rho^+\rho^0$ signal decays and misreconstructed $B^+ \to \rho^+\rho^0$ decays (*self cross-feed*). Simulation shows that the ratio of self cross-feed to signal candidates is about 20%. This is a potential limitation for the analysis reach. Large cross-feed adds complexity to the fit of sample composition. To reduce this complexity, the yield of the self cross-feed component in the fit is constrained to the fraction found in simulated signal events. To further investigate the self cross-feed source, I classify self cross-feed candidates as shown in Fig. 4.19. The composition is dominated by random combination of correctly reconstructed charged particles or photons; in particular, for about half of the self cross-feed component candidates a ρ meson is correctly reconstructed, resulting in peaking structures in some of the fit observables.

4.5 Peaking backgrounds

Charmless B decays with same or similar final states as the $B^+ \to \rho^+(\to \pi^+\pi^0)\rho^0(\to \pi^+\pi^-)$ signal tend to peak in the distributions of some of the observables used in the fit of sample composition. Such peaking backgrounds cannot be vetoed because most of them proceed through broad intermediate resonances that overlap with the signal ρ mass range. Vetoing those would also reduce significantly signal efficiency. Therefore B peaking backgrounds should be included and properly modeled in the fit of sample composition.

I first study $B^+ \to \rho^+\rho^0$ candidates reconstructed in the simulated generic sample and passed through the full selection to single out rare B peaking backgrounds. I also survey previous $B^+ \to \rho^+\rho^0$ analyses by Belle [41] and BaBar [42] to identify possible peaking

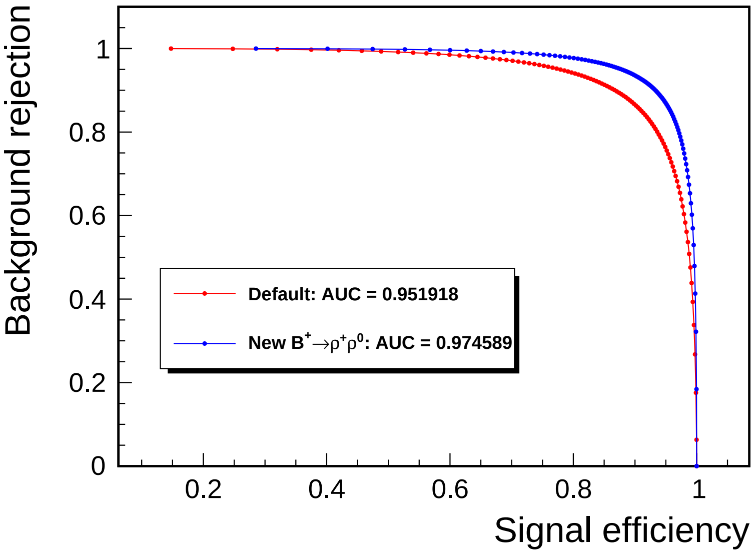


Figure 4.16: Receiver operating characteristic for two configurations of the continuum-suppression classifier: (red) standard classifier used in many hadronic B decay analyses in Belle II; (blue) customed classifier trained for $B^+ \to \rho^+\rho^0$ and with optimized input variables.

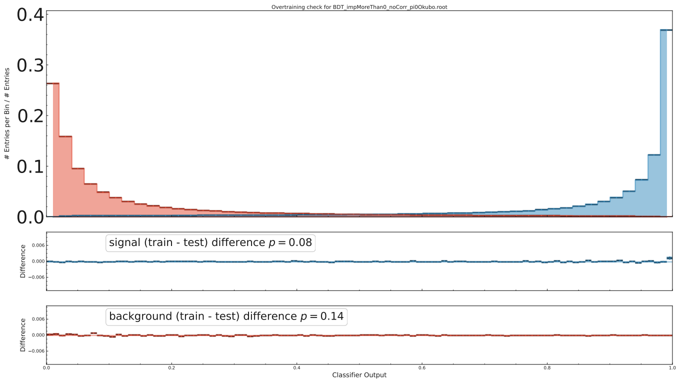


Figure 4.17: Output of the FBDT classifier on the training and testing samples.

backgrounds that may not be included or properly modeled in the simulated samples.

Table 4.3 summarizes the branching fractions and expected yields of rare charmless B decay background modes expected to contribute in the sample. When branching fractions are not known, I conservatively assume them to be 10^{-5}.

4.6 Combinatorial B decay background

The second largest source of background are $B\overline{B}$ events, which yield accidental combinations of B decay products that mimic signal. B decay backgrounds are less pervasive than continuum, but still harmful as they have similar or identical final states, as signal. These backgrounds may both dilute the signal, thus reducing the statistical precision, and mimic the signal, therefore biasing the results.

I study generic $e^+e^- \to B\overline{B}$ simulated samples (see Sec. 4.1.2) to assess the impact

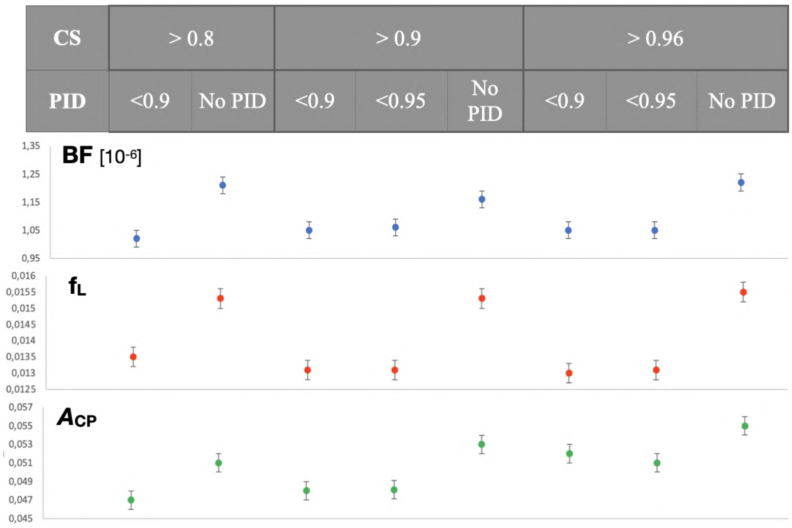


Figure 4.18: Average expected statistical uncertainty on $\mathcal{B}$, f_L and A_{CP} as a function of continuum suppression and pion-identification requirements as resulting from analyses of simplified simulated experiments.

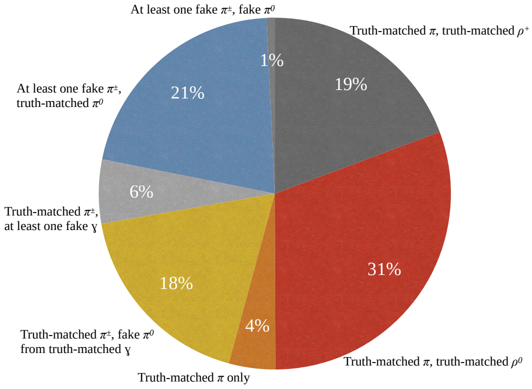


Figure 4.19: Composition of self-cross-feed candidates assuming realistic proportions between longitudinally and transversely polarized events.

of $B\overline{B}$ backgrounds. Favored B decays contribute significantly to the sample composition due to their large branching fractions. The major contributions come from hadronic and semileptonic B decays, such as $B \to D^{(*)}\rho$, $B \to D^{(*)}\pi$, or $B \to D^{(*)}l\overline{\nu}$. Some of these backgrounds can generate ΔE distributions with peaks under, or in proximity, of the $B^+ \to \rho^+\rho^0$ peak, and I try to reduce them studying possible targeted vetoes.

4.6.1 Charm vetoes

To identify and remove $B \to D$ backgrounds, I inspect the distributions of the two- and three-body masses of intermediate states, where tracks are either assigned with the pion or kaon mass hypothesis. Narrow peaks associated with properly reconstructed D decays allow for straightforward vetoes, as shown in Fig. 4.20. Some peaking structures are visible at 1.865 GeV/c^2, the known D^0 mass, in both the two- and three-body masses distributions, generated by $B^+ \to \overline{D}^0(\to \pi^+\pi^-)\pi^+\pi^0$ and $B^+ \to \overline{D}^0(\to \pi^+\pi^-\pi^0)\pi^+$ decays. In addition,

	Channel	$\mathcal{B}\,[10^{-6}]$
Known	$B^+ \to a_1(1260)^+\pi^0$	26 ± 7
	$B^0 \to a_1(1260)^\mp\pi^\pm$	26 ± 5
	$B^+ \to a_1(1260)^0\pi^+$	20.4 ± 5.8
Unknown	$B^+ \to \rho^0\pi^+\pi^0$	10
	$B^+ \to \rho^+\pi^+\pi^-$	10
	$B^+ \to \rho^0\pi^+\pi^+$	10
	$B^+ \to f_0(980)\pi^+\pi^0$	10

Table 4.3: Overview of charmless peaking backgrounds: Channels are grouped into decays with known branching fraction, processes for which an upper limit exists on the branching fraction, and decays with unknown branching fraction. If experimental results are not available, I conservatively assume $\mathcal{B} = 10^{-5}$.

peaks from $B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+\pi^0$ and $B^+ \to \overline{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays where the pion mass is mistakenly assigned to the kaon are visible at 1.790 and 1.775 GeV/c^2 in the two- and three-body masses distributions respectively. Another smaller peak is also visible at 3.105 GeV/c^2 in the two–body mass distribution, and is generated by $J/\Psi \to \mu^+\mu^-$ decays, where muons are mistakenly identified as pions.

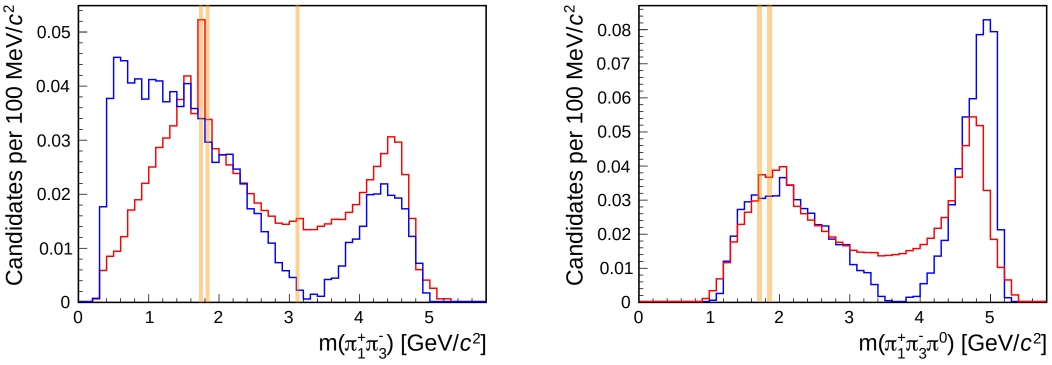


Figure 4.20: Mass distributions for (left) two- and (right) three-body combinations of final-state pions in $B^+ \to \rho^+\rho^0$ decays reconstructed from simulated signal and $B\overline{B}$ samples. The distributions for candidates reconstructed in (red) favored B decay events are compared with the distributions for (blue) truth-matched signal candidates (in same proportions of longitudinal and transverse polarization). The red orange lines show the veto regions.

Object	Requirement
Charged pions	$\|dz\| < 3$ cm
	$\|dr\| < 0.5$ cm
	in CDC acceptance
	number of CDC hits > 20
converged track fit	
	$\mathcal{R}(K, \pi) < 0.95$
Neutral pion	$E_\gamma > 50$ MeV for ECL barrel
	$E_\gamma > 60$ MeV for forward ECL endcap
	$E_\gamma > 100$ MeV for backward ECL endcap
	number of cluster hits per γ > 1.5
	$\|$timing$(\gamma)\|$ < 200ns
	photonMVA(γ) > 0.1
	$105 < m(\gamma\gamma) < 150$ MeV/c^2
	$\|$cos$\Theta_H\|$ < 0.98
	converged vertex fit
Charged ρ meson	$0.52 < m(\pi^+\pi^0) < 1.06$ GeV/c^2
	cos$\theta^{hel}_{\rho^+} < 0.75$
Neutral ρ meson	$0.52 < m(\pi^+\pi^-) < 1.06$ GeV/c^2
B meson	$M_{\rm bc} > 5.27$ GeV/c^2
Best candidate selection	$\chi^2(\pi^0, \text{vertex})$
Continuum suppression	$C_{\rm FBDT} > 0.96$
Charm veto	$1.78 < m(\pi\pi) < 1.80$ GeV/c^2
	$1.855 < m(\pi\pi) < 1.875$ GeV/c^2
	$3.095 < m(\pi\pi) < 3.115$ GeV/c^2
	$1.75 < m(\pi\pi\pi^0) < 1.80$ GeV/c^2
	$1.84 < m(\pi\pi\pi^0) < 1.90$ GeV/c^2

Table 4.4: Summary of selection.

Chapter 5

Determination of the sample composition

The reconstructed sample of signal candidates is expected to contain a prominent fraction of background events. The individual sample contributions need to be discriminated for extracting the desired measurements of the $B^+ \to \rho^+\rho^0$ *properties. In this chapter, I describe how kinematic and continuum-suppression information is combined in a maximum likelihood fit to determine sample composition.*

5.1 The fit

The composition of the $B^+ \to \rho^+\rho^0$ sample is determined statistically with a multivariate maximum likelihood fit of the unbinned distributions of kinematic, angular, and continuum-suppression discriminating observables. I write a likelihood $\mathcal{L}(\vec{\theta}|\vec{x})$, function of the n unknown parameters of the model $\vec{\theta} = (\theta_1, ..., \theta_n)$ at the given set $\vec{x}$ of observed data. For any specific set of values of the unknown parameters $\vec{\theta}$, $\mathcal{L}$ is the joint probability density for obtaining the observed set of values of the discriminating observables in the sample, $\vec{x}$. In fitting, I maximize the likelihood function with respect to the unknown parameters by minimizing twice the negative natural logarithm by imposing

$$-2\frac{\partial(\ln\mathcal{L})}{\partial\vec{\theta}} = 0. \tag{5.1}$$

5.1.1 Minimizer

I use the fitting package `MINUIT` [76], which calculates numerically Eq. (5.1). In a first fit, I use the `MIGRAD` minimization algorithm, which features a variable-metric method with inexact line search, a stable metric updating scheme, and checks for positive-definiteness to identify coarsely the maximum. Then I use `MINOS`, which uses the likelihood ratio to include non-linearities in the uncertainty calculation, therefore providing more accurate estimates.

5.1.2 Fit components

The sample of $B^+ \to \rho^+\rho^0$ candidates resulting from the final selection receives contributions from four main sources, signal, self-cross feed, continuum background, and backgrounds from other $B\overline{B}$ events, as shown in Fig. 5.1. I treat longitudinally and transversely polarized decays separately, as the goal is to measure the fraction of longitudinally polarized

$B^+ \to \rho^+\rho^0$ decays. For each signal component a self-cross feed component is included. Continuum and combinatorial $B\overline{B}$ backgrounds dominate the sample composition. In addition, three components of peaking $B\overline{B}$ backgrounds are included in the fit.

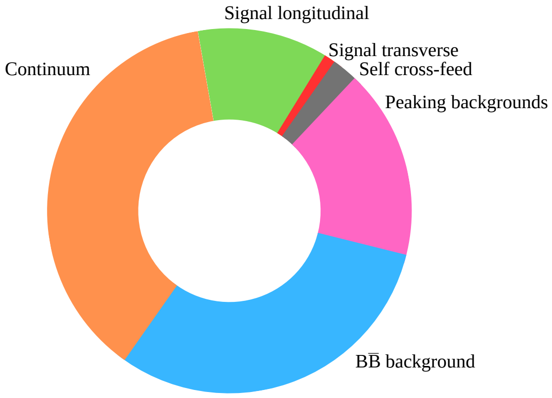


Figure 5.1: Pie-chart illustrating the $B^+ \to \rho^+\rho^0$ sample composition (assuming $f_L = 0.95$) resulting from the final selection as expected by simulation. Unobserved decays are assumed to have $\mathcal{B} = 10^{-5}$.

Truth information in simulation suggests the expected proportions of each component in simulated samples similar in size to the Belle II sample (assuming $f_L = 0.95$), and offers a snapshot of the complexity of the sample composition (Table 5.1).

Component	Expected yield
Longitudinally polarized $B^+ \to \rho^+\rho^0$	700
Transversely polarized $B^+ \to \rho^+\rho^0$	60
Longitudinally polarized self-cross feed	133
Transversely polarized self-cross feed	5
Combinatorial $B\overline{B}$ background	1870
$B \to f_0(980)\pi\pi$	320
$B \to \rho\pi\pi$	440
$B \to a_1(1260)\pi$	260
Continuum background	2270

Table 5.1: Summary of $B^+ \to \rho^+\rho^0$ sample composition expected from simulation.

The expected peaking-background yields are determined according to the branching fractions listed in Table 4.3.

5.1.3 Fit observables

I choose a set of discriminating observables that allow distinguishing signal from background and separating longitudinally and transversely polarized signal decays. The goal is to achieve a trade off between effective signal-to-background separation and reasonable fit complexity both in terms of number of observables and of their multidimensional dependencies.

ΔE is the most powerful observable to separate exclusively reconstructed signal from continuum events (Fig. 5.2 top left). It also has the additional advantage to discriminate signal from other fully reconstructed B decays.

Another powerful discriminating observable against continuum background is the output of the FBDT classifier (Fig. 5.2 top right). The distribution of the classifier output C_{FBDT} has a non-trivial shape, as shown in Figure 5.3. I therefore transform it into $C'_{\text{FBDT}} = \log\frac{C_{\text{FBDT}} - C_{\text{FBDT}_{\text{min}}}}{C_{\text{FBDT}_{\text{max}}} - C_{\text{FBDT}}}$, which is more easily described by an analytical model without degrading discriminating power.

The signal decay includes a pair of intermediate ρ mesons. The invariant masses $m(\pi^+\pi^0)$ and $m(\pi^+\pi^-)$ offer therefore further discriminating power between signal and both continuum and $B\overline{B}$ backgrounds, which have at most one ρ meson in the intermediate state (Fig. 5.2 middle plots). Finally, I use helicity-angle observables, $\cos\theta_{\rho^+}$ and $|\cos\theta_{\rho^0}|$ to separate longitudinally and transversely polarized $B^+ \to \rho^+\rho^0$ decays (Fig. 5.2 bottom plots).

In summary, the fit uses six discriminating observables:

1. ΔE – difference between observed and expected B energy;
2. C'_{FBDT} – log-transformed continuum suppression classifier output;
3. $m(\pi^+\pi^0)$ – invariant mass of charged and neutral pions;
4. $m(\pi^+\pi^-)$ – invariant mass of opposite charge pions;
5. $\cos\theta_{\rho^+}$ – cosine of the helicity angle of ρ^+ meson;
6. $\cos\theta_{\rho^0}$ – cosine of the helicity angle of ρ^0 meson.

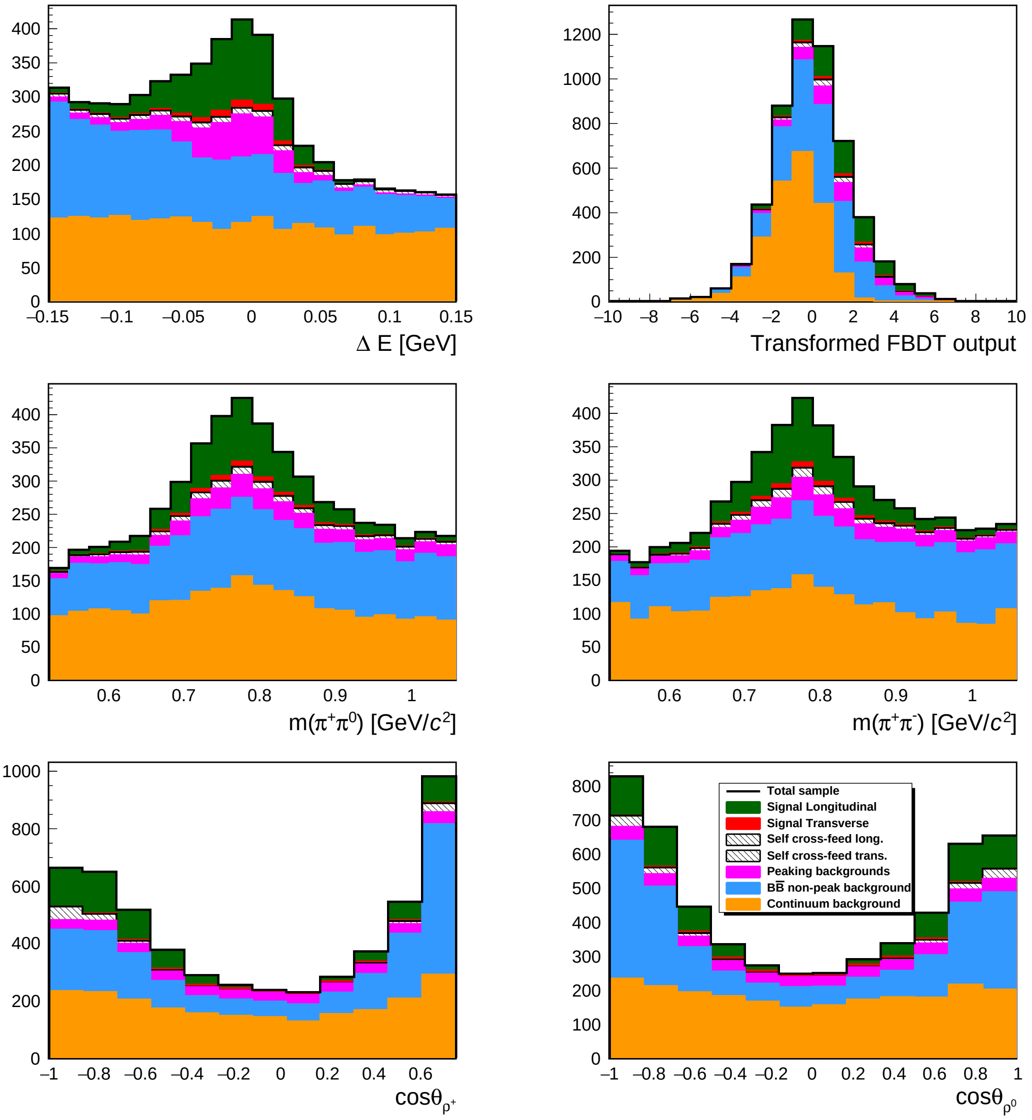


Figure 5.2: Distributions of (top left) ΔE, (top right) transformed continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $m(\pi^+\pi^-)$, (bottom left) $\cos\theta_{\rho^+}$, and (bottom right) $\cos\theta_{\rho^0}$ for simulated $B^+ \to \rho^+\rho^0$ components. Proportions mirror realistic expected yields.

5.1.4 Fit parameters

For convenience, I parametrize the likelihood directly in terms of the parameters of interest, branching fraction and fraction of longitudinally polarized $B^+ \to \rho^+\rho^0$ decays, so that they are determined by the fit, along with their statistical uncertainties including correlations.

The branching fraction $\mathcal{B}$ for a B decay into final state f is expressed as a ratio of the number of observed candidates N corrected by reconstruction and selection efficiency ε and the total number of collected $B\overline{B}$ pairs $N_{B\overline{B}}$,

$$\mathcal{B}(B \to f) = \frac{N}{\varepsilon N_{B\overline{B}}}. \tag{5.2}$$

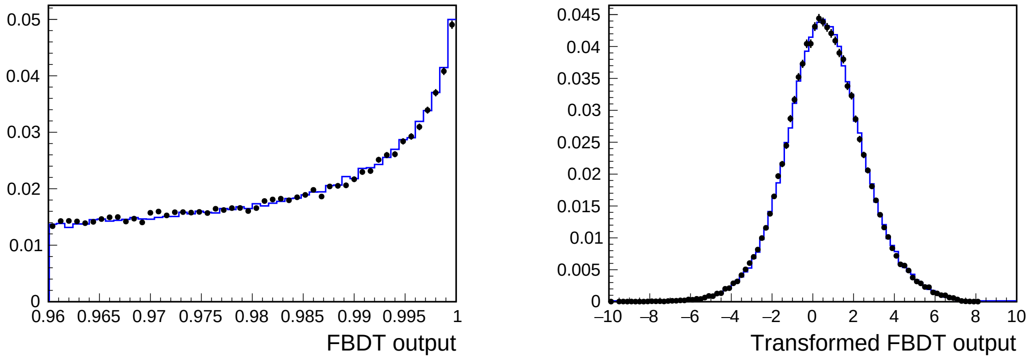


Figure 5.3: Distributions of (left) FBDT classifier output and (right) its log-transformation for $B^+ \to \bar{D}^0(\to K^+\pi^-)\pi^+$ decays reconstructed in (points) collision and (solid) simulated data.

Since I distinguish between polarizations, (Eq. 5.2) transforms into

$$\mathcal{B}(B^+ \to \rho^+\rho^0) = \frac{N_L/\varepsilon_L + N_T/\varepsilon_T}{N_{B\overline{B}}}, \tag{5.3}$$

where N_L (N_T) is the number of longitudinally (transversely) polarized $B^+ \to \rho^+\rho^0$ decays and ε_L (ε_T) is the corresponding selection efficiency.

The fraction of longitudinally polarized B decays f_L is

$$f_L = \frac{N_L/\varepsilon_L}{N_L/\varepsilon_L + N_T/\varepsilon_T}. \tag{5.4}$$

Currently Belle II recorded 387×10^6 $B\overline{B}$ pairs. As the ratio of charged to neutral $B\bar{B}$ pair production is consistent with one, $f_{B^+B^-}/f_{B^0\bar{B}^0} = 1.06 \pm 0.012 \pm 0.019 \pm 0.047$ [77], the total number of B^+B^- pairs used in this analysis is around 199×10^6.

In addition to the parameters of interest, the likelihood depends on a set of nuisance parameters whose values are not necessarily of interest but influence the values of parameters of interest. These are self-cross-feed-to-signal-yield ratios and the yields of the other background components.

The fit has difficulties in distinguishing signal from self-cross-feed candidates because their shapes are similar. I therefore impose a Gaussian constraint on the self-cross-feed-to-signal-yield-ratio based on simulation, assuming that self-cross-feed-to-signal-yield-ratio in data is well reproduced by simulation.

The yields of decay modes for which an external measurement exists are Gaussian constrained to the observed values. The fractions of non-peaking $B\overline{B}$ background and unknown peaking B decays are allowed to float.

5.1.5 Model

Distributions for some of the observables are known based on physics-motivated models. For instance, we know that the signal ΔE distribution is approximately Gaussian since its shape is dominated by a multitude of small resolution effects, or that the invariant dipion masses of ρ mesons are well reproduced by Breit-Wigner functions. For many other observables, excessive complexity prevents from identifying physics-motivated models. In those cases, I empirically find models that are capable to reproduce adequately the distribution shapes.

In addition, when the distributions are too complex to be straightforwardly described by analytical functions, I use binned templates, as is the case for p.d.f. describing angular distributions. For each component, shapes are extracted from the corresponding simulated samples of various sizes between 10^4 and 10^7 events.

5.2 Likelihood modeling

The likelihood function $\mathcal{L}$ is written as a product, over the N events in the sample, of the single-event likelihoods $\mathcal{L}_i$:

$$\mathcal{L}(\vec{\theta}|\vec{x}) \equiv \mathcal{L}(\vec{\theta}) = \prod_{i=1}^{N} \mathcal{L}_i(\vec{\theta}). \tag{5.5}$$

Each single event likelihood has the following expression

$$\begin{aligned}\mathcal{L}_i = N_{(\rho^+\rho^0)_L} p_i^{(\rho^+\rho^0)_L} &+ N_{(\rho^+\rho^0)_T} p_i^{(\rho^+\rho^0)_T} \\ &+ r_{\mathrm{sxf}\ (\rho^+\rho^0)_L} N_{(\rho^+\rho^0)_L} p_i^{\mathrm{sxf}\ (\rho^+\rho^0)_L} \\ &+ r_{\mathrm{sxf}\ (\rho^+\rho^0)_T} N_{(\rho^+\rho^0)_T} p_i^{\mathrm{sxf}\ (\rho^+\rho^0)_T} \\ + N_{B\overline{B}} {p_i}^{B\bar{B}} &+ N_{f_0\pi\pi} p_i^{f_0\pi\pi} + N_{a_1\pi} p_i^{a_1\pi} + \\ N_{\rho\pi\pi} p_i^{\rho\pi\pi} &+ (1 - N_{\mathrm{sum}}) p_i^{\mathrm{cont}},\end{aligned} \tag{5.6}$$

where p_i^y indicates the six-dimensional probability density function of the generic component y (cont means continuum), and N_{sum} is the sum of all independent yields.

I extract central values and uncertainties of branching ratios from a non-extended unbinned fraction fit. I then add in quadrature a Poisson $\sqrt{N}$ factor to the `MINOS` uncertainty of the branching fraction and all yields for estimating properly the uncertainty from the fluctuation of the total number of events.

5.2.1 Identification of dependencies between fit observables

Special care is needed in writing a multivariate likelihood, since potentially large biases may arise if the densities of dependent observables are factorized improperly. To determine the proper probability density functions, dependencies between the fit observables need to be identified and modeled.

For each component I identify the major dependencies by studying distributions of each observable conditional on each of the others. This is achieved by inspecting the distributions of one variable for non-overlapping narrow intervals ("slices") of the others and checking whether the distribution of the probe observable depends on the range chosen in the other or not. For instance, Fig. 5.4 shows how the distribution of ΔE in slices of $\cos\theta_{\rho^+}$ is obtained. The shape of the ΔE distribution changes as a function of $\cos\theta_{\rho^+}$, indicating a dependence between these two observables.

In principle one should study distributions of each variable sliced simultaneously in all five remaining dimensions. This approach is complicated by the reduction in statistical information that become critical if simulated samples are partitioned across too many dimensions. However, a coarse study of the full-fledged partitioning shows that all major features are mostly captured by one-dimensional slicings, where each observable is shown in slices of another, integrated over the remaining four observables. I therefore show here distributions obtained by one-dimensional slicing only.

Figure 5.5, as an example, shows such distributions for all the combinations of pairs of fit observables for the longitudinally polarized signal component. The largest dependencies are found between the ΔE, $m(\pi^+\pi^0)$, and cosθ_{ρ^+} observables. This dependence, which is probably due to the photon energy measurement, is also found in the transversely polarized signal and both self-cross-feed components. A complete set of plots illustrating dependencies is in Appendix A. Tables 5.2–5.3 show matrices that summarize observed dependencies between fit observables in the simulated signal sample. For self-cross-feed components, a dependence between $m(\pi^+\pi^-)$ and cosθ_{ρ^0} is also observed. All observed dependencies between fit observables in the simulated self-cross feed sample are summarized in Tables 5.4–5.5.

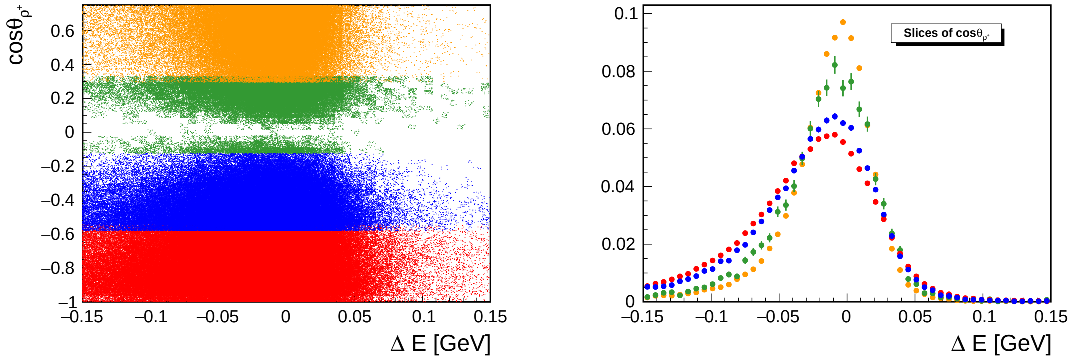


Figure 5.4: Example of dependencies between fit observables: distributions of (left) cosθ_{ρ^+} as a function of ΔE and (right) ΔE in "slices" of cosθ_{ρ^+} for simulated longitudinally polarized signal decays. Different colors indicate different cosθ_{ρ^+} slices.

	ΔE	C'_{FBDT}	$m(\pi^+\pi^0)$	$m(\pi^+\pi^-)$	$\cos\theta_{\rho^+}$	$\cos\theta_{\rho^0}$
ΔE	■		●		■	
C'_{FBDT}		■				
$m(\pi^+\pi^0)$			■		●	
$m(\pi^+\pi^-)$				■		
$\cos\theta_{\rho^+}$					■	
$\cos\theta_{\rho^0}$						■

Table 5.2: Matrix representing dependencies between fit observables in the simulated sample of longitudinally polarized signal $B^+ \to \rho^+\rho^0$ decays. Squares indicate strong dependencies, circles indicate mild dependencies.

	ΔE	C'_{FBDT}	$m(\pi^+\pi^0)$	$m(\pi^+\pi^-)$	$\cos\theta_{\rho^+}$	$\cos\theta_{\rho^0}$
ΔE	■		●		■	
C'_{FBDT}		■				
$m(\pi^+\pi^0)$			■		●	
$m(\pi^+\pi^-)$				■		
$\cos\theta_{\rho^+}$					■	
$\cos\theta_{\rho^0}$						■

Table 5.3: Matrix representing dependencies between fit observables in the simulated sample of transversely polarized signal $B^+ \to \rho^+\rho^0$ decays. Squares indicate strong dependencies, circle indicate mild dependencies.

	ΔE	C'_{FBDT}	$m(\pi^+\pi^0)$	$m(\pi^+\pi^-)$	$\cos\theta_{\rho^+}$	$\cos\theta_{\rho^0}$
ΔE	■		●		■	
C'_{FBDT}		■				
$m(\pi^+\pi^0)$			■	●	■	
$m(\pi^+\pi^-)$				■		■
$\cos\theta_{\rho^+}$					■	
$\cos\theta_{\rho^0}$						■

Table 5.4: Matrix representing dependencies between fit observables in the simulated sample of longitudinally polarized self-cross feed $B^+ \to \rho^+\rho^0$ candidates. Squares indicate strong dependencies, circles indicate mild dependencies.

	ΔE	C'_{FBDT}	$m(\pi^+\pi^0)$	$m(\pi^+\pi^-)$	$\cos\theta_{\rho^+}$	$\cos\theta_{\rho^0}$
ΔE	■		■		■	
C'_{FBDT}		■				
$m(\pi^+\pi^0)$			■		■	
$m(\pi^+\pi^-)$				■		■
$\cos\theta_{\rho^+}$					■	
$\cos\theta_{\rho^0}$						■

Table 5.5: Matrix representing dependencies between fit observables in the simulated sample of transversely polarized self-cross feed $B^+ \to \rho^+\rho^0$ candidates. Squares indicate strong dependencies, circles indicate mild dependencies.

For continuum and non-peaking $B\overline{B}$ backgrounds, there are no visible dependencies among variables with the available size of simulated samples. Since this is about three times the expected yields in data, I neglected dependencies among the fit observables for these components. I also inspect "sliced" distributions for the three peaking background modes. The same dependence due to the photon energy measurement, as for signal, is evident for all peaking background modes (Figs A.6–A.8). Similar dependencies as for self cross-feed are observed for the $B \to a_1\pi$ decay mode. All observed dependencies between fit observables in the peaking background samples are summarized in Tables 5.6–5.5.

	ΔE	C'_{FBDT}	$m(\pi^+\pi^0)$	$m(\pi^+\pi^-)$	$\cos\theta_{\rho^+}$	$\cos\theta_{\rho^0}$
ΔE	■		●		■	
C'_{FBDT}		■				
$m(\pi^+\pi^0)$			■		■	
$m(\pi^+\pi^-)$				■		
$\cos\theta_{\rho^+}$					■	
$\cos\theta_{\rho^0}$						■

Table 5.6: Matrix representing dependencies between fit observables in the simulated sample of $B^+ \to f_0(980)\pi\pi$ events. Squares indicate strong dependencies, circles indicate mild dependencies.

	ΔE	C'_{FBDT}	$m(\pi^+\pi^0)$	$m(\pi^+\pi^-)$	$\cos\theta_{\rho^+}$	$\cos\theta_{\rho^0}$
ΔE	■				■	
C'_{FBDT}		■				
$m(\pi^+\pi^0)$			■	●		
$m(\pi^+\pi^-)$				■		
$\cos\theta_{\rho^+}$					■	●
$\cos\theta_{\rho^0}$						■

Table 5.7: Matrix representing dependencies between fit observables in the simulated sample of $B \to \rho\pi\pi$ events. Squares indicate strong dependencies, circles indicate mild dependencies.

	ΔE	C'_{FBDT}	$m(\pi^+\pi^0)$	$m(\pi^+\pi^-)$	$\cos\theta_{\rho^+}$	$\cos\theta_{\rho^0}$
ΔE	■				●	●
C'_{FBDT}		■				
$m(\pi^+\pi^0)$			■		■	●
$m(\pi^+\pi^-)$				■	●	■
$\cos\theta_{\rho^+}$					■	●
$\cos\theta_{\rho^0}$						■

Table 5.8: Matrix representing dependencies between fit observables in the simulated sample of $B \to a_1(1260)\pi$ background events. Squares indicate strong dependencies, circles indicate mild dependencies.

This study allows identifying the factorization of multidimensional p.d.f.'s that approximates better the real dependencies present in the sample. Not all dependencies can be exactly reproduced in the model, since the finite size of the simulated samples used to identify them imposes a limit on the sensitivity to the effects probed. In addition, mirroring in detail all observed dependencies in the fit model is not straightforward since using joint distributions of more than three-dimension poses increasingly hard technical challenges. Hence a systematic work of trial and error is required to converge to the fit model that captures sufficiently well the dependencies of the sample, as demonstrated by low, if any, biases in the estimates, while permitting a reasonable technical handling. Documenting all the failed attempts is beyond the scope of this thesis, but converging to an adequate model has been an essential and challenging part of this work. The model is described in the following sections.

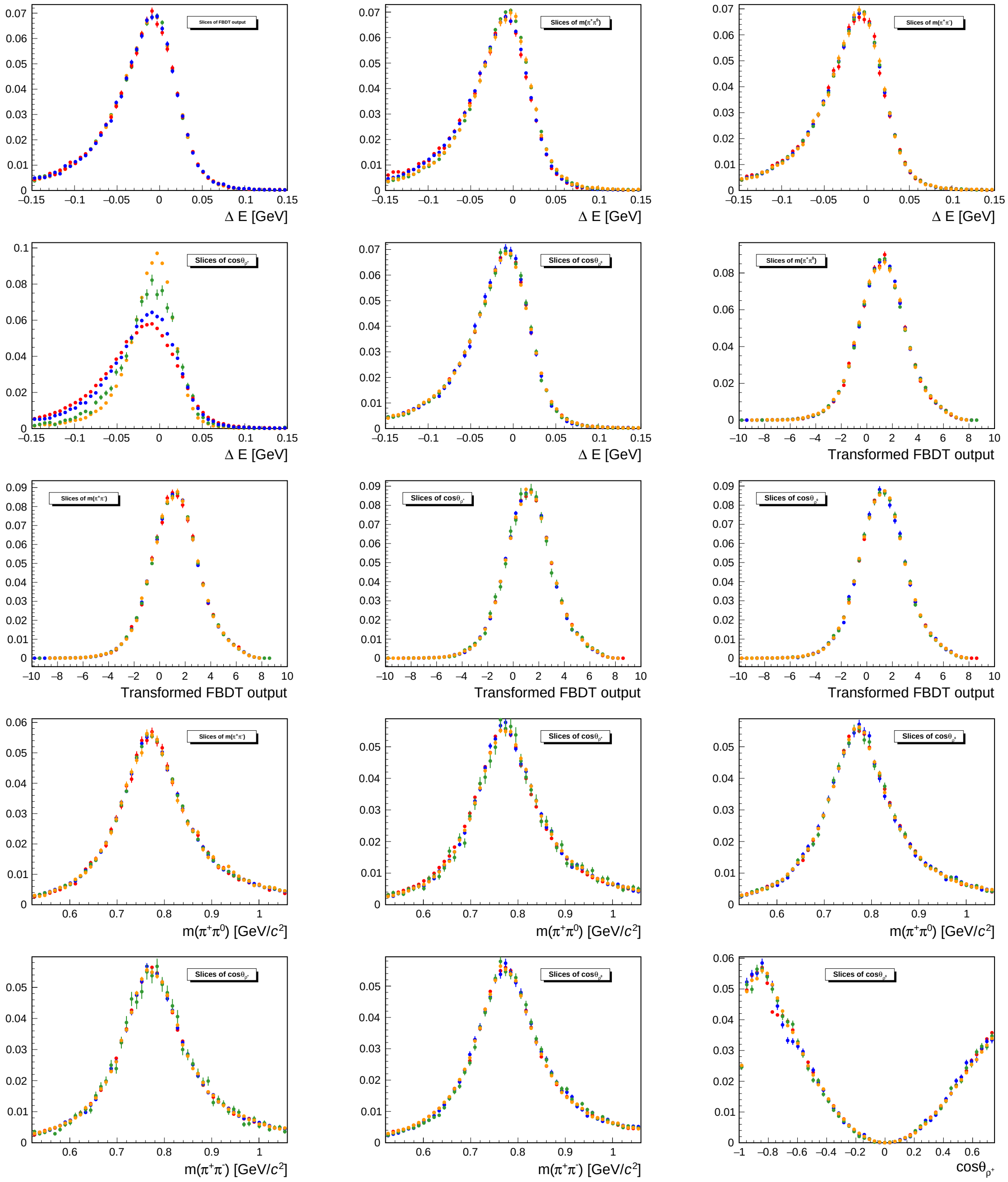


Figure 5.5: Example of dependence studies: distributions of fit observables in slices of other fit observables for simulated longitudinally polarized signal decays. Distributions normalized to unity. The full set of distributions in Appendix A.

5.2.2 Signal p.d.f.

The probability density function that describes each of the signal components is factorized as the product of five terms,

$$p_j = p_j(\Delta E|\mathrm{cos}\theta_{\rho^+})p_j(C'_{\mathrm{FBDT}})p_j(m(\pi^+\pi^0))p_j(m(\pi^+\pi^-))p_j(\mathrm{cos}\theta_{\rho^+})p_j(\mathrm{cos}\theta_{\rho^0}), \qquad (5.7)$$

where index j corresponds to the longitudinally or transversely polarized signal. $p(\Delta E|\mathrm{cos}\theta_{\rho^+})$ is a conditional p.d.f. that accounts for the dependence between the two observables: ΔE is modeled empirically as the sum of three Gaussian functions,

$$p(\Delta E) \propto f_1 \frac{e^{-\frac{1}{2}(\frac{\Delta E-\mu_1}{\sigma_1})}}{\sigma_1\sqrt{2\pi}} + f_2 \frac{e^{-\frac{1}{2}(\frac{\Delta E-\mu_2}{\sigma_2})}}{\sigma_2\sqrt{2\pi}} + (1-f_1-f_2)\frac{e^{-\frac{1}{2}(\frac{\Delta E-\mu_3}{\sigma_3})}}{\sigma_3\sqrt{2\pi}} \qquad (5.8)$$

where the weight f_i, mean μ_i, and width σ_i of each Gaussian are determined independently from a fit in simulation in two different bins of $\mathrm{cos}\theta_{\rho^+}$; $p(\mathrm{cos}\theta_{\rho^+})$ is a one-dimensional histogram template. Figure 5.6 shows that the conditional p.d.f. model chosen to describe ΔE and $\mathrm{cos}\theta_{\rho^+}$ is adequate. $p(C'_{\mathrm{FBDT}})$ is the p.d.f. that describes the transformed continuum suppression output ("continuum suppression term"), and is described with the sum of two Gaussian functions

$$p(\Delta E) \propto f \frac{e^{-\frac{1}{2}(\frac{\Delta E-\mu_1}{\sigma_1})}}{\sigma_1\sqrt{2\pi}} + (1-f)\frac{e^{-\frac{1}{2}(\frac{\Delta E-\mu_2}{\sigma_2})}}{\sigma_2\sqrt{2\pi}}; \qquad (5.9)$$

figure 5.7 shows that this functional form is adequate. The p.d.f. of the ρ^+ and ρ^0 invariant mass terms are modeled using the physics-motivated Breit-Wigner function,

$$p(m(\pi^+\pi^{0,-})) \propto \frac{1}{(m(\pi^+\pi^{0,-}) - m_{\rho^{+,0}})^2 + \left(\frac{\Gamma_{\rho^{+,0}}}{2}\right)^2}, \qquad (5.10)$$

where $m_{\rho^{+,0}}$ and $\Gamma_{\rho^{+,0}}$ are determined from a fit in simulation. Figure 5.8 shows the result of the modeling on the simulation. The mismodelings shown in the pulls are accounted for in the systematic uncertainty, and are not expected to have significant impact on the final result.

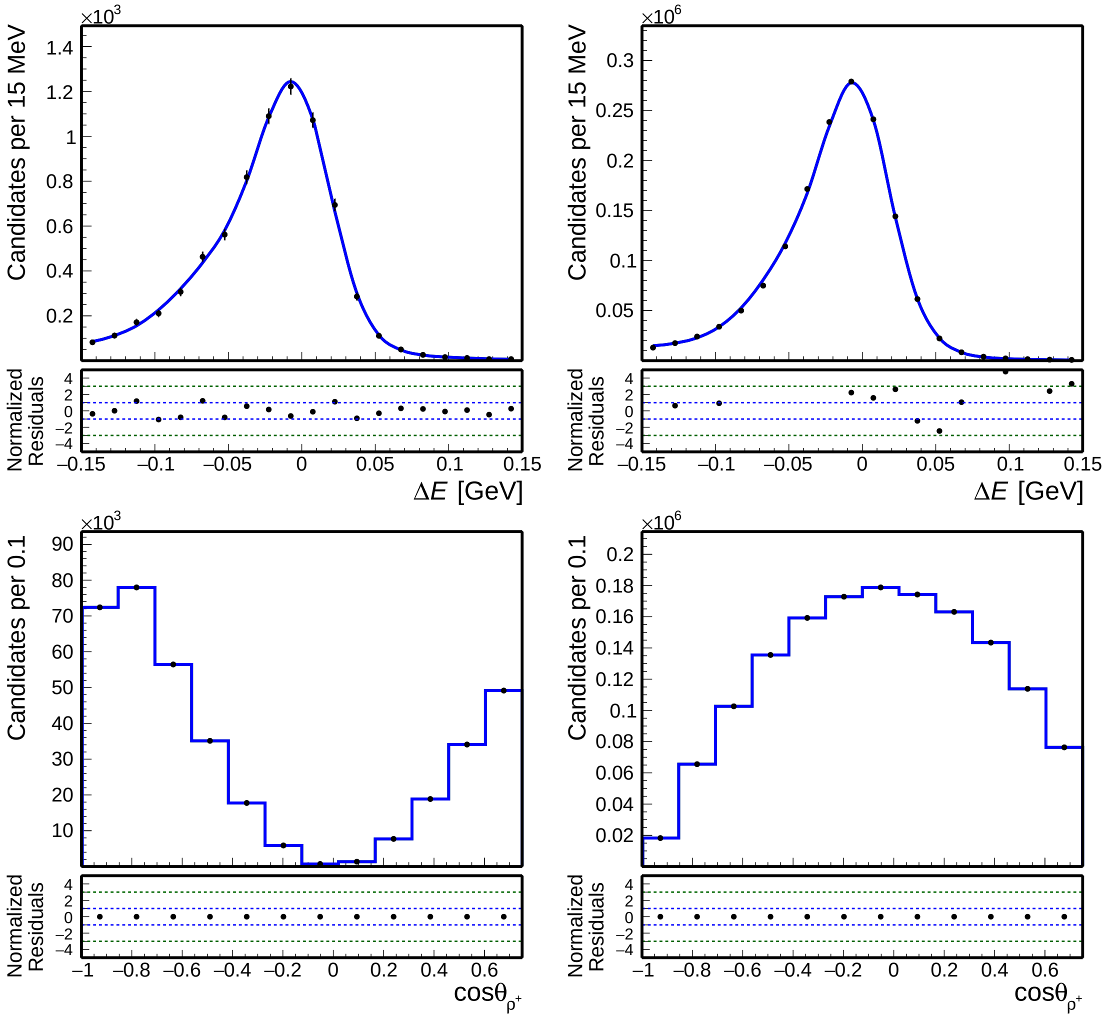


Figure 5.6: Example of fit modeling: distributions of (top) ΔE and (bottom) $\cos\theta_{\rho^+}$ for simulated longitudinally polarized (left) and (right) transversely polarized signal. The corresponding model shapes are overlaid (solid). Full set of distributions in Appendix A.

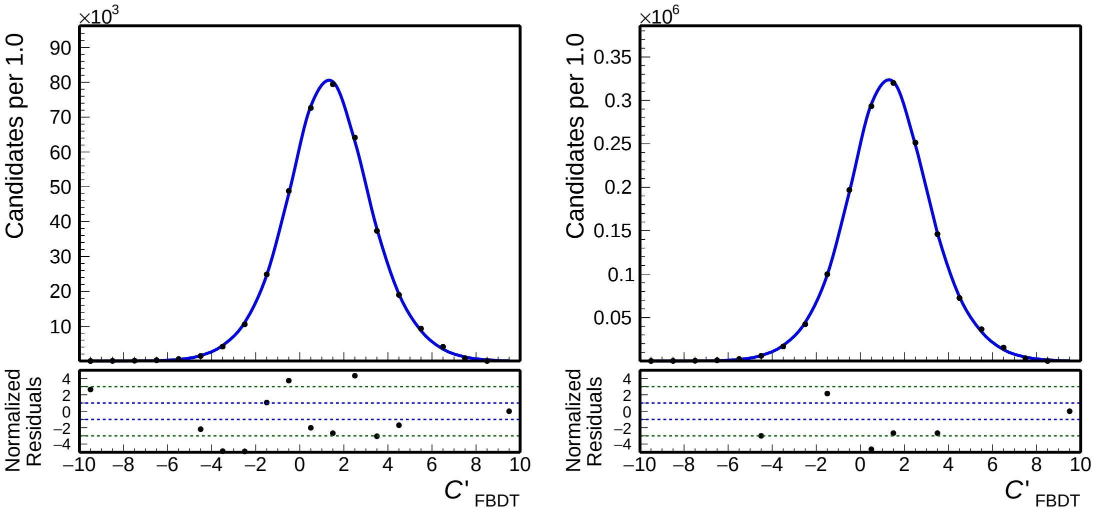


Figure 5.7: Example of fit modeling: distributions of continuum suppression output for (left) simulated longitudinally polarized and (right) transversely polarized signal. The corresponding model shapes are overlaid (solid). Full set of distributions in Appendix A.

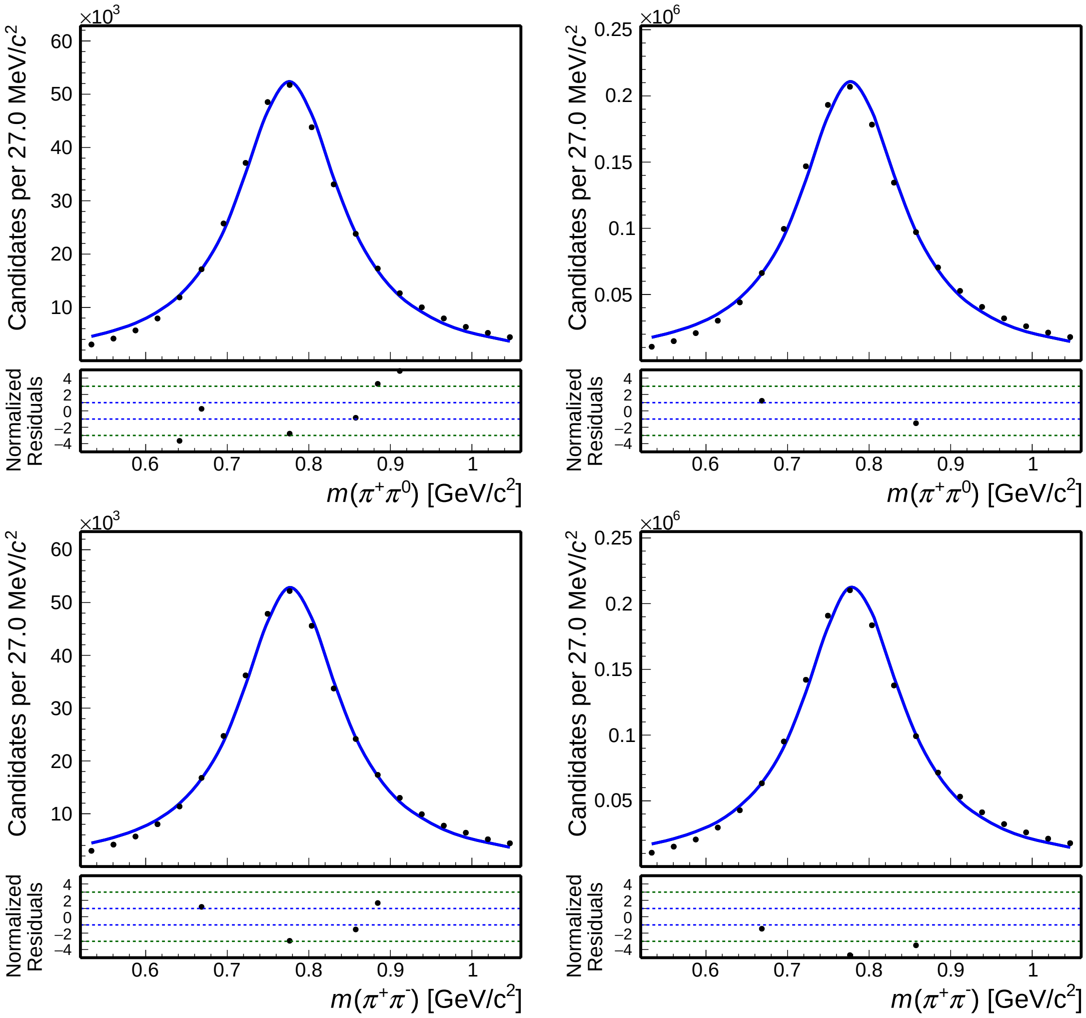


Figure 5.8: Distributions of (top) $m(\pi^+\pi^0)$ and (bottom) $m(\pi^+\pi^-)$ for simulated (left) longitudinally polarized and (right) transversely polarized signal. The corresponding model shapes are overlaid (solid). Full set of distributions in Appendix A.

5.2.3 Self-cross feed p.d.f.

Probability density functions for longitudinally and transversely polarized self-cross feed candidates have different structures.

The probability density function that describes longitudinally polarized self-cross feed is factorized as the product of three terms,

$$p_j = p_j(\Delta E|\cos\theta_{\rho^+})p_j(C'_{\rm FBDT})p_j(m(\pi^+\pi^0)|\cos\theta_{\rho^+})p_j(m(\pi^+\pi^-)|\cos\theta_{\rho^0})p_j(\cos\theta_{\rho^+})p_j(\cos\theta_{\rho^0}), \tag{5.11}$$

where $p(x_i|\cos\theta_{\rho^{+,0}})$ are conditional p.d.f. that describe the x_i observable in two different bins of the helicity angle $\cos\theta_{\rho^{+,0}}$. ΔE and $C'_{\rm FBDT}$ are described with a sum of three Gaussian functions (Eq. (5.8) and a sum of two Gaussian functions (Eq. (5.9) respectively, while the invariant mass distributions are modeled with the sum of a Breit-Wigner (Eq. (5.10) and a quadratic function. In Figure A.9 the corresponding one-dimensional projections of the model show that this is adequate to describe the longitudinal self cross-feed component.

Given the expected low statistics, the probability density function that describes transversely polarized self-cross feed is factorized as the product of six one-dimensional terms,

$$p = p(\Delta E)p(C'_{\mathrm{FBDT}})p(m(\pi^+\pi^0))p(m(\pi^+\pi^-))p(\cos\theta_{\rho^+})p(\cos\theta_{\rho^0}), \tag{5.12}$$

where the helicity angles are modeled with histogram templates, while the other observables are described with analytical functions. Figure A.10 (top panels) shows that sums of three Gaussian functions (Eq. (5.8) are adequate to describe the ΔE and continuum suppression terms. The p.d.f. of the $\rho^{+,0}$ mass terms are described with sums of a Breit-Wigner function (Eq. (5.10) and a quadratic function, shown to be adequate in Figure A.10 (bottom panels).

5.2.4 Peaking background p.d.f.

The probability density functions for the peaking backgrounds have similar structures to the signal and self cross-feed components. Continuum suppression is described with the sum of two Gaussian functions (Eq. 5.9), while helicity angles are described with one-dimensional histogram templates. The ΔE and ρ^+ mass distributions are all modeled with p.d.f. conditional to $\cos\theta_{\rho^+}$, with a sum of three Gaussian functions (Eq. 5.8) or a sum of Breit-Wigner (Eq. 5.10) and quadratic shapes respectively. For the $B \to \rho\pi\pi$ component, the ρ^0 mass distribution is modeled with an independent one-dimensional function, a sum of Breit-Wigner (Eq. 5.10) and quadratic shape. For the $B \to a_1\pi$ component, the ρ^0 mass shape is again modeled with a sum of Breit-Wigner (Eq. 5.10) and quadratic shape, but conditional to the $\cos\theta_{\rho^0}$ distribution. For the $B \to f_0\pi\pi$ component, the ρ^0 mass distribution is modeled with a Crystal Ball function peaking at the f_0 mass,

$$p(m(\pi^+\pi^-)) \propto \begin{cases} \exp(-\frac{(x-\bar{x})^2}{2\sigma^2}), & \text{for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A \cdot \left(B - \frac{x-\bar{x}}{\sigma}\right)^{-n}, & \text{for } \frac{x-\bar{x}}{\sigma} \leq -\alpha \end{cases}, \tag{5.13}$$

where α, n, $\bar{x}$ and σ are shape parameters, and A and B are other functions of these parameters.

5.2.5 $B\bar{B}$ and continuum background p.d.f.

Both the large non-peaking background components are described with a model factorized into six one-dimensional terms,

$$p = p(\Delta E)p(C'_{\mathrm{FBDT}})p(m(\pi^+\pi^0))p(m(\pi^+\pi^-))p(\cos\theta_{\rho^+})p(\cos\theta_{\rho^0}), \tag{5.14}$$

where the helicity angles are described with histogram templates. The ΔE terms are modeled with a second-order polynomial function,

$$p(\Delta E) \propto a\Delta E^2 + b\Delta E + 1, \tag{5.15}$$

where a and b are determined from a modeling fit in simulation. The continuum suppression is modeled in both components with the sum of two Gaussian functions (Eq. (5.9). The $\rho^{+,0}$ mass distributions are modeled with sums of a Gaussian function and a quadratic function for the $B\bar{B}$ background, and with sums of a Breit-Wigner (Eq. (5.10) and a quadratic function for the continuum.

5.2.6 P.d.f. summary

I summarize the model choices in Table 5.9 not to overload the text with model details. One-dimensional projections of all the fit models are in Appendix A. The resulting single-event likelihood consists of a combination of various binned templates and analytical functions. The shape parameters are fixed from simulation calibrated on control samples. In the fit, I set a Gaussian constraint on the self-cross-feed-to-signal-yield-ratio observed in simulation. The yield of the $B \to a_1\pi$, for which an external measurement exists, is also Gaussian constrained to the observed values. The yields of unobserved modes are determined freely by the fit. The model neglects the interference between B decays into the signal-like four-pion-final-state, for which a systematic uncertainty is assigned (see Sec. 8.9).

Component	ΔE	C'_{FBDT}	$m(\pi^+\pi^0)$	$m(\pi^+\pi^-)$	cosθ_{ρ^+}	cosθ_{ρ^0}
$B^+ \to \rho^+\rho^0$ ($f_L = 1$)	3G$\vert$cosθ_{ρ^+}	2G	BW	BW	1D hist	1D hist
$B^+ \to \rho^+\rho^0$ ($f_L = 0$)	3G$\vert$cosθ_{ρ^+}	2G	BW	BW	1D hist	1D hist
$B^+ \to \rho^+\rho^0$ ($f_L = 1$) scf	3G$\vert$cosθ_{ρ^+}	2G	BW+Pol2$\vert$cosθ_{ρ^+}	BW+Pol2$\vert$cosθ_{ρ^0}	1D hist	1D hist
$B^+ \to \rho^+\rho^0$ ($f_L = 0$) scf	3G$\vert$cosθ_{ρ^0}	2G	BW+Pol2	BW+Pol2	1D hist	1D hist
$B\bar{B}$	Pol2	2G	G+Pol2	G+Pol2	1D hist	1D hist
$B^+ \to f_0\pi^+\pi$	3G$\vert$cosθ_{ρ^+}	2G	BW+Pol2$\vert$cosθ_{ρ^+}	CB+Pol2	1D hist	1D hist
$B^+ \to \rho\pi\pi$	3G$\vert$cosθ_{ρ^+}	2G	BW+Pol2	BW+Pol2	1D hist	1D hist
$B^+ \to a_1^+\pi$	3G$\vert$cosθ_{ρ^+}	2G	BW+Pol2$\vert$cosθ_{ρ^+}	BW+Pol2$\vert$cosθ_{ρ^0}	1D hist	1D hist
Continuum	Pol2	2G	BW+Pol2	BW+Pol2	1D hist	1D hist

Table 5.9: Model summary. The shorthand "1D hist" means one-dimensional histogram template, "3G (2G)" means sum of three (two) Gaussian functions, "BW" indicates a Breit-Wigner, "Poln" indicates a polynomial of n-th degree, and "CB" indicates the Crystall Ball shape.

5.3 Estimator properties

Maximum likelihood estimates can show statistical bias and non Gaussian dispersions because of multiple causes. The most common is that the model is too simplified and fails to capture some of the relevant data features. Both phenomena may lead to unsatisfactory results. To investigate the presence and entity of these phenomena, I study the distributions of estimators on simulated data. I use the pull value of each fit parameter θ_i, defined as the uncertainty-weighted residual,

$$\text{Pull}(\theta_i) = \frac{\hat{\theta}_i - \theta_i}{\hat{\sigma}_{\hat{\theta}_i}}, \tag{5.16}$$

where θ is the known true value of the parameter, $\hat{\theta}_i$ is the estimate, and $\hat{\sigma}_{\hat{\theta}_i}$ is the estimate of the uncertainty. Distributions of residuals are studied over ensembles of simplified simulated experiments ("toys") that simulate the experimental circumstances of the fit on data. The samples are obtained by mixing simulated signal, self-cross-feed, continuum and $B\overline{B}$ background events in realistic proportions and sizes.

I generate 1000 toy samples by drawing the yield of each component according to a Poisson distribution centered at the true values expected from simulation or previous measurements. Then, I fit the composition of each toy as if they were data, and construct the pull-value distributions of the fit results.

Ideally, for an unbiased and asymptotic maximum likelihood estimator, the pull distribution is a Gaussian distribution mean-valued at zero with unit variance. A mean differing from zero corresponds to a bias in units of statistical uncertainty, while a variance smaller (larger) than one corresponds to an over-(under) estimation of the statistical uncertainty.

Figure 5.9 shows pull distributions for the $\mathcal{B}(B^+ \to \rho^+\rho^0)$ and f_L parameters. The widths of the distributions are compatible with one showing no significant over-(under-) estimation of the uncertainty. The means are compatible with zero, showing no intrinsic biases in the fit procedure.

Figure 5.10 shows residual distributions for the $\mathcal{B}(B^+ \to \rho^+\rho^0)$ and f_L parameters. The widths of the distributions correspond to the expected average statistical uncertainties. By comparing these results with the current world-best, I expect to significantly improve the precision of $\mathcal{B}(B^+ \to \rho^+\rho^0)$.

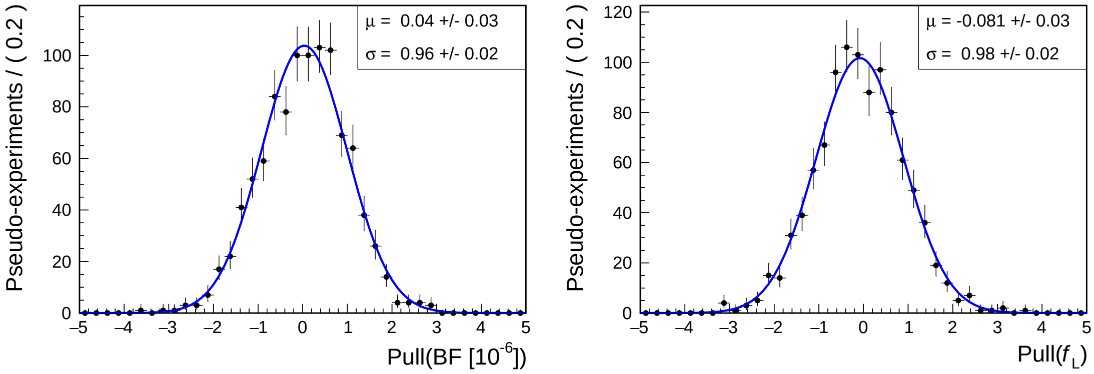


Figure 5.9: Pull distributions of (left) branching fraction and (right) fraction of longitudinally polarized $B^+ \to \rho^+\rho^0$ decays. Results of fits to Gaussian functions are overlaid (solid).

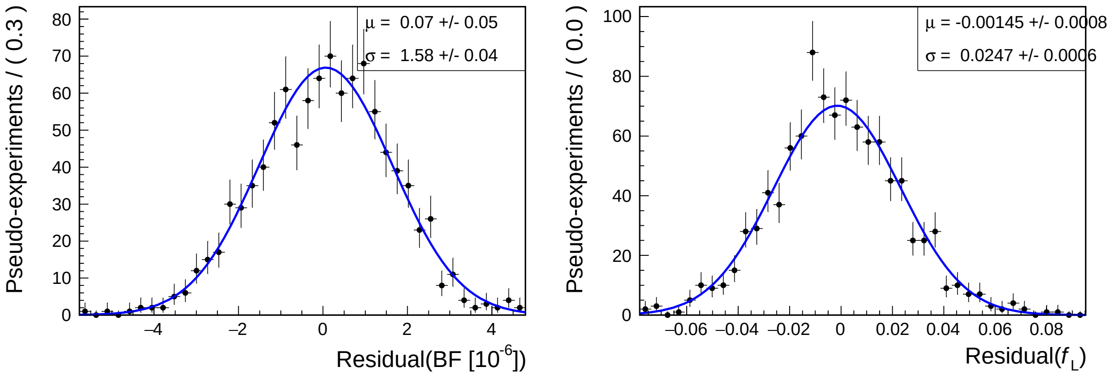


Figure 5.10: Residual distributions of (left) branching fraction and (right) fraction of longitudinally polarized $B^+ \to \rho^+\rho^0$ decays. Results of fits to Gaussian functions are overlaid (solid).

5.4 Charge-asymmetry fit

I modify the likelihood given in Eq. (5.2) to extract the charge-asymmetry $A_{\rm CP}$. The likelihood is a sum of two functions describing samples of positively and negatively charged candidates,

$$\mathcal{L}_i = \mathcal{L}_i^+ + \mathcal{L}_i^-, \tag{5.17}$$

where $\mathcal{L}_i^+$ and $\mathcal{L}_i^-$ are given by

$$\begin{aligned}\mathcal{L}_i^+ = N^+_{(\rho^+\rho^0)_L} p_i^{(\rho^+\rho^0)_L} &+ N^+_{(\rho^+\rho^0)_T} p_i^{(\rho^+\rho^0)_T} \\ +r_{\text{sxf }(\rho^+\rho^0)_L} N^+_{(\rho^+\rho^0)_L} & p_i^{\text{sxf }(\rho^+\rho^0)_L} \\ +r_{\text{sxf }(\rho^+\rho^0)_T} N^+_{(\rho^+\rho^0)_T} & p_i^{\text{sxf }(\rho^+\rho^0)_T} \\ +N^+_{B\overline{B}} p_i^{B\bar{B}} + N^+_{f_0\pi\pi} p_i^{f_0\pi\pi} &+ N^+_{a_1\pi} p_i^{a_1\pi} + \\ N^+_{\rho\pi\pi} p_i^{\rho\pi\pi} &+ (1 - N^+_{\rm sum}) p_i^{\rm cont},\end{aligned} \tag{5.18}$$

and

$$\begin{aligned}\mathcal{L}_i^- = N^-_{(\rho^+\rho^0)_L} p_i^{(\rho^+\rho^0)_L} &+ N^-_{(\rho^+\rho^0)_T} p_i^{(\rho^+\rho^0)_T} \\ +r_{\text{sxf }(\rho^+\rho^0)_L} N^-_{(\rho^+\rho^0)_L} & p_i^{\text{sxf }(\rho^+\rho^0)_L} \\ +r_{\text{sxf }(\rho^+\rho^0)_T} N^-_{(\rho^+\rho^0)_T} & p_i^{\text{sxf }(\rho^+\rho^0)_T} \\ +N^-_{B\overline{B}} p_i^{B\bar{B}} + N^-_{f_0\pi\pi} p_i^{f_0\pi\pi} &+ N^-_{a_1\pi} p_i^{a_1\pi} + \\ N^-_{\rho\pi\pi} p_i^{\rho\pi\pi} &+ (1 - N^-_{\rm sum}) p_i^{\rm cont},\end{aligned} \tag{5.19}$$

where N^+ (N^-) is the number of positively (negatively) charged candidates in the sample, $f^+_{\rm sig}$ is the fraction of positively charged $B^+ \to \rho^+\rho^0$ decays, $f^+_{B\overline{B}}$ is the fraction of positively charged $B\overline{B}$ background candidates and is fixed to 0.5, and $f^+_{\rm sum}$ ($f^-_{\rm sum}$) is the sum of all independent fractions corresponding to positively (negatively) charged candidates in the sample. In addition to signal and self cross-feed components, I also account for yields CP-asymmetries for the peaking background components. I assume the combinatorial $B\overline{B}$ background has no CP-asymmetry.

I verify that probability density functions used in Section 5.1 are adequate to describe distributions of both positively and negatively charged candidates. For the helicity angles distributions, I use different charge-specific histogram templates.

Figures 5.11–5.12 show example projections obtained from a fit to a realistic simulated sample. A signal peak is observed at $\Delta E \approx 0$, overlapping smooth continuum and non-peaking $B\overline{B}$ background distributions. In the continuum-suppression-output distribution, continuum events tend to peak at $C'_{\rm FBDT} \approx 0$, while the peak corresponding to $B\overline{B}$ events (including signal and background) is shifted toward $C'_{\rm FBDT} \approx 1$. Two prominent ρ-meson peaks are observed in the $m(\pi^+\pi^-)$ and $m(\pi^+\pi^0)$ distributions, as expected. Longitudinally polarized signal decays tend to cluster at the edges of the $\cos\theta_{\rho^+}$ distribution, while transversely polarized signal decays populate the central part of the spectrum. Fit projections show that the fit reproduces adequately all the observed distributions and provides confidence on the accuracy of our model.

Finally, I study the A_{CP} estimator properties as described in Section 5.3. The pull distribution shows that the estimator is unbiased and the statistical uncertainty is correctly

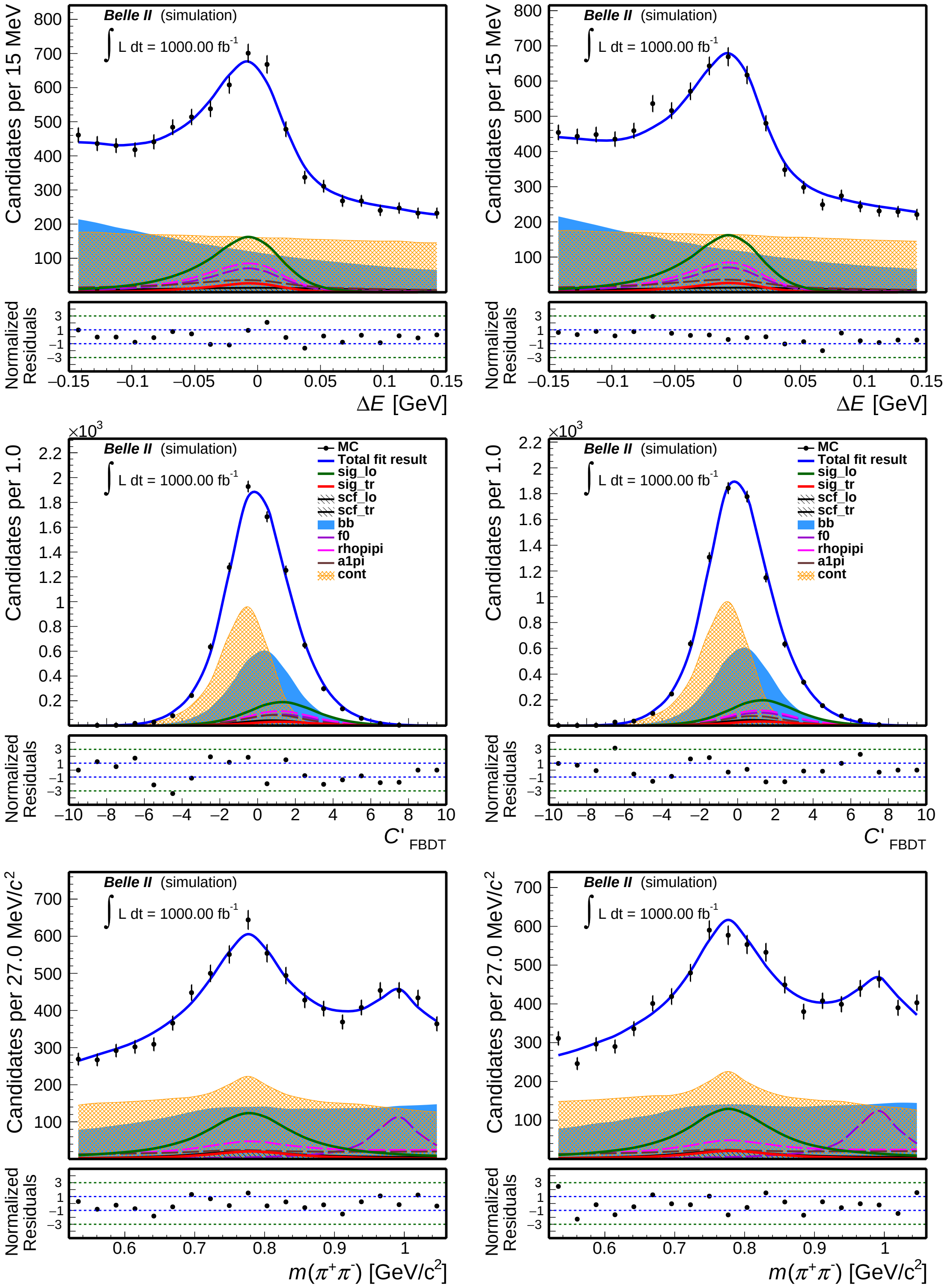


Figure 5.11: Example of fit projections: distributions of (top) ΔE, (middle) continuum suppression output, and (bottom) $m(\pi^+\pi^0)$ for a realistic simulated Belle II sample of (left) negatively and (right) positively charged $B^+ \to \rho^+\rho^0$ candidates with fit projections overlaid.

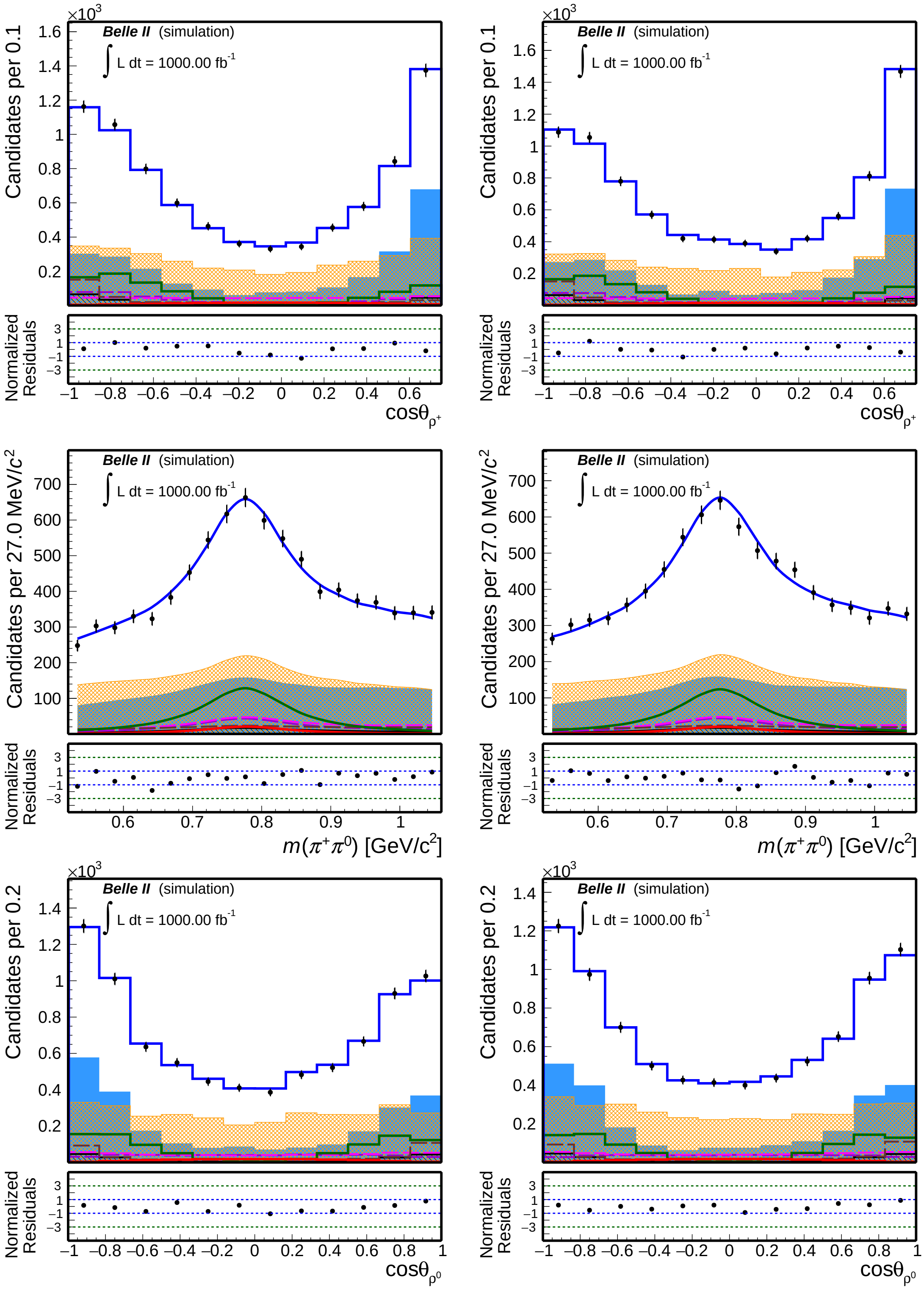


Figure 5.12: Example of fit projections: distributions of (top) cosθ_{ρ^+}, (middle) $m(\pi^+\pi^-)$, and (bottom)|cosθ_{ρ^0}| for a realistic simulated Belle II sample of (left) negatively and (right) positively charged $B^+ \to \rho^+\rho^0$ candidates with fit projections are overlaid.

estimated (Fig. 5.13, left plot). By inspecting the residual distribution (Fig. 5.13, right plot), I expect to obtain results for A_{CP} of $B^+ \to \rho^+\rho^0$ decay competitive with the current word's best measurements.

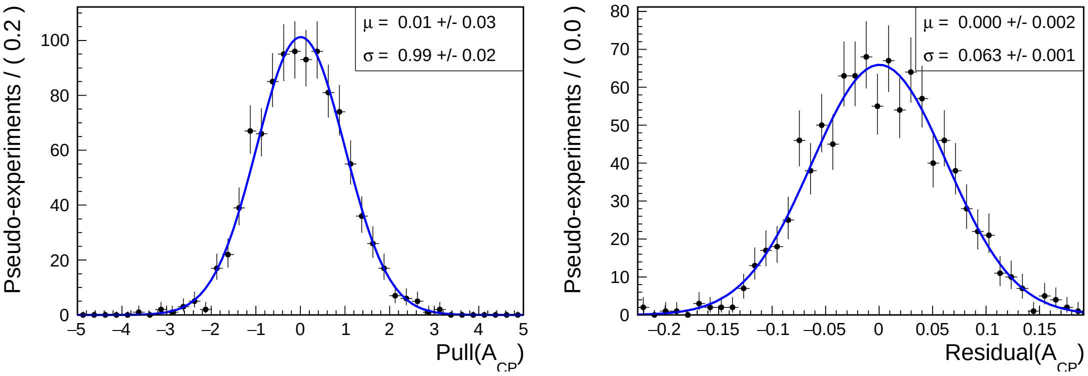


Figure 5.13: Distributions of (left) pull and (right) residual for the charge-parity asymmetry of $B^+ \to \rho^+\rho^0$ decays. Results of fit to Gaussian functions are overlaid (solid).

Chapter 6

Analysis validation on data

Data and simulations are known to feature differences that have the potential of spoiling the measurement results. This chapter discusses the validation of the analysis on data to account for such differences.

6.1 Introduction

The sample-composition fit is developed based on realistically simulated samples. However, potentially harmful discrepancies are known to exist between Belle II data and simulation. In order to ensure a reliable estimation of the desired parameters in data, it is essential to identify and correct such discrepancies. A variety of control samples in data are used to this end. I apply minimum variations of the $B^+ \to \rho^+\rho^0$ selection and fit to an abundant control channel that shares relevant features with the signal decay, in search for possible undetected issues or unmodeled effects. In addition, I identify appropriate control regions in the signal-candidate sample ("sidebands") to further compare simulated and observed data. Finally, projections and uncertainties (but not central values) of fits of a $B^+ \to \rho^+\rho^0$ sample in half of the data set are inspected to assess realistically the effects of various analysis choices.

6.2 Validation of signal modeling for ρ^+, ΔE, and C'_{FBDT}

I validate on data the signal fit model, because the shapes determined on simulation might differ in data. Known minor miscalibrations in the electromagnetic-calorimeter energy-reconstruction result in variations of the ΔE and $m(\pi^+\pi^0)$ shapes, which are correlated with the reconstructed π^0 energy. I extract empiric correction factors by performing a fit of the abundant $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+(\to \pi^+\pi^0)$ control channel, which features signal-sample-like topology and final states, reconstructed in Belle II data and simulation. A total of 2.5×10^4 signal events are expected in data, which is a factor of approximately 35 larger than our expected signal yield. I apply the full selection and fit to the $B^+ \to \bar{D}^0\rho^+$ channel, adapted to the unavoidable differences in topology and final states. This sample serves the main purpose of validating or calibrating the treatment of the ΔE, C'_{FBDT}, $m(\pi^+\pi^0)$ and $\cos\theta_{\rho^+}$ signal shapes.

6.2.1 Selection

The selection is the same as for the signal, except for the following variations due to the differences between the two decays:

- I invert the particle-identification selection for the kaon with respect to the other charged pions to suppress combinatorial background from pions;
- I lift charm vetoes, since the $B^+ \to \bar{D}^0\rho^+$ contains a charm meson instead of a ρ^0;
- I require the $K^+\pi^-$ mass to be in the $1.7 < m(K^+\pi^-) < 2.0$ GeV/c^2 range. This loose criterion is not optimized for reducing background. On the contrary, its purpose is to accept a fraction of continuum background so as to mimic the $B^+ \to \rho^+\rho^0$ conditions.

I choose one candidate per event with the same strategy as for signal.

6.2.2 $B^+ \to \bar{D}^0\rho^+$ likelihood modeling

The $B^+ \to \bar{D}^0\rho^+$ composition is statistically determined with a multivariate maximum likelihood fit to the unbinned distributions of discriminating observables as much similar as possible to the fit of the signal sample. The single event likelihood is

$$\mathcal{L}_i = N_{\text{sig}}p_i^{\text{sig}} + N_{\text{sxf}}p_i^{\text{sxf}} + N_{B\bar{B}}p_i^{B\bar{B}} + (N - N_{\text{sig}} - N_{\text{sxf}} - N_{B\bar{B}})p_i^{\text{cont}}, \tag{6.1}$$

where p^{sig} describes the probability density function of the observables for the signal events, p^{sxf} for the self cross-feed events, $p^{B\bar{B}}$ for the non-peaking $B\bar{B}$ background events, and p^{cont} for the continuum events; N_j corresponds to the yield of generic component j, to be determined by the fit, and N is the total number of events. Contributions from rare B decays are considered negligible.

The shapes for each component are determined from fits to realistically simulated samples. All background shapes are obtained from the same generic simulated sample as in $B^+ \to \rho^+\rho^0$ decays.

The single event likelihood for signal and self cross-feed is factorized, as for the $B^+ \to \rho^+\rho^0$ signal case, as

$$\mathcal{L}_i = p_i(\Delta E|\text{cos}\theta_{\rho^+})p_i(C'_{\text{FBDT}})p_i(m(\pi^+\pi^0))p_i(\text{cos}\theta_{\rho^+}), \tag{6.2}$$

where the ΔE distribution is described with a sum of three Gaussians (5.8) conditional to the $\text{cos}\theta_{\rho^+}$ value. The continuum suppression is described with the sum of two Gaussians (5.9), and $p_i(\text{cos}\theta_{\rho^+})$ is an histogram template. The $m(\pi^+\pi^0)$ invariant mass is described with a Breit-Wigner (5.10) for the signal, and with the sum of a Breit-Wigner and a quadratic function for self cross-feed. As shown in Fig. 6.1, the model is adequate for ΔE, C'_{FBDT} and $\text{cos}\theta_{\rho^+}$. The $m(\pi^+\pi^0)$ shows mismodeling between the simulation and the fit results; however, this effect is negligible for the scope of this study.

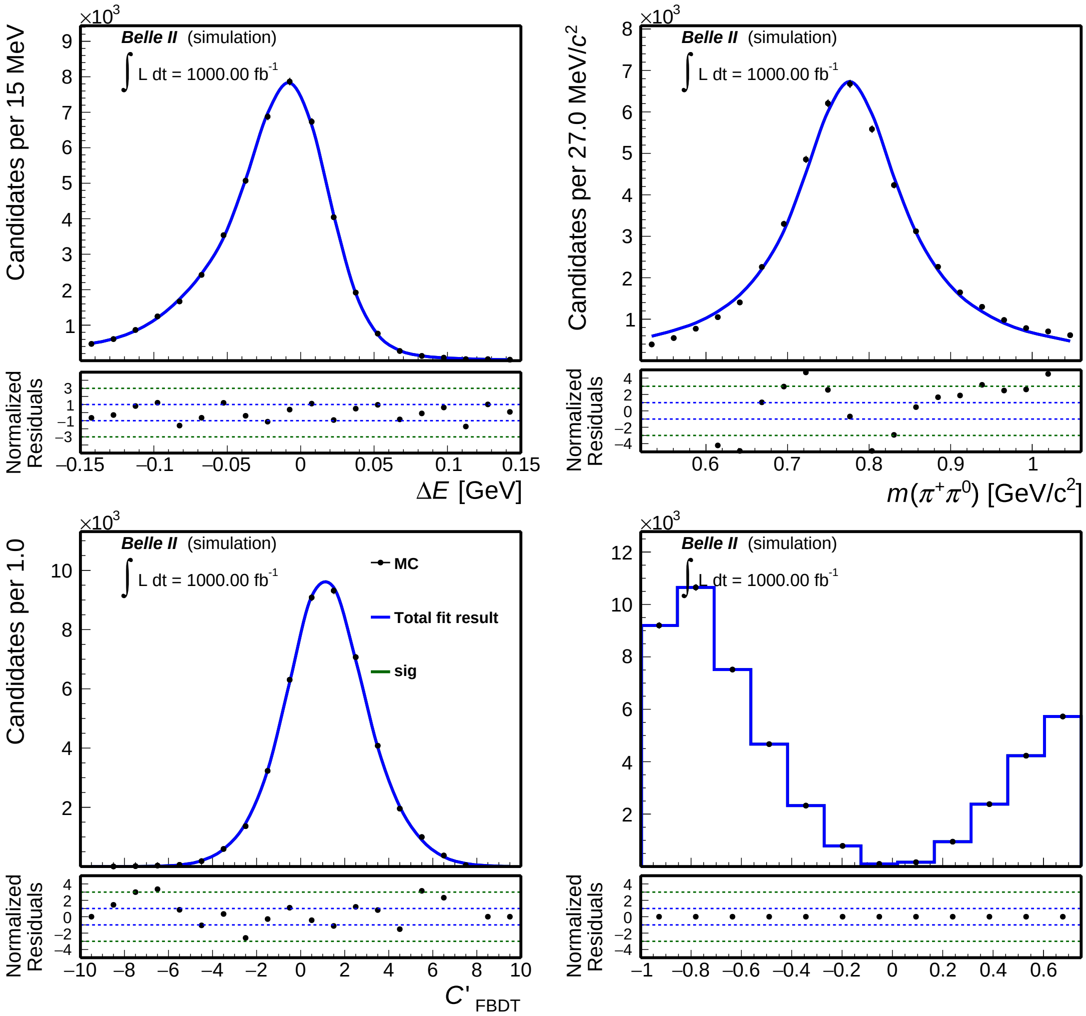


Figure 6.1: Control channel fit modeling: distributions of (top left) ΔE, (top right) $m(\pi^+\pi^0)$, (bottom left) continuum suppression, and (bottom right) $\cos\theta_{\rho^+}$ for simulated $B^+ \to \bar{D}^0\rho^+$ signal. The corresponding model shapes are overlaid (solid).

For candidates reconstructed from $B\overline{B}$ or continuum background events, the single-event likelihood is factorized as

$$\mathcal{L}_i = p_i(\Delta E)p_i(m(\pi^+\pi^0))p_i(\cos\theta_{\rho^+})p_i(C'_{\mathrm{FBDT}}). \tag{6.3}$$

For both components, the ΔE and C'_{FBDT} distributions are described with the sum of three and two Gaussians respectively. The invariant mass $m(\pi^+\pi^0)$ is described by the sum of a Breit-Wigner (Gaussian) distribution with a quadratic function for the $B\overline{B}$ (continuum) background component. For both components, the helicity angle is described with an histogram template. Analytical functions and binned templates used to describe the single-event $B^+ \to \bar{D}^0\rho^+$ likelihood are summarized in Table 6.1.

6.2.3 $B^+ \to \bar{D}^0\rho^+$ results

The likelihood of the $B^+ \to \bar{D}^0\rho^+$ channel is

$$\mathcal{L} = \prod_{i=1}^{n}\sum_{j} N_j\mathcal{L}_i(\Delta E, C'_{\mathrm{FBDT}}, m(\pi^+\pi^0), \cos\theta_{\rho^+}). \tag{6.4}$$

Component	ΔE	$C'_{\rm FBDT}$	$m(\pi^+\pi^0)$	cosθ_{ρ^+}
$B^+ \to \bar{D}^0\rho^+$	3G|cosθ_{ρ^+}	2G	BW	1D hist
$B^+ \to \bar{D}^0\rho^+$ scf	3G|cosθ_{ρ^+}	2G	BW	1D hist
$B\bar{B}$	Exp	2G	BW+Pol2	1D hist
Continuum	Exp	2G	G+Pol2	1D hist

Table 6.1: $B^+ \to \bar{D}^0\rho^+$ model summary. The shorthand "1D hist" means one-dimensional histogram template, "3G (2G)" means sum of three (two) Gaussian functions, "BW" indicates a Breit-Wigner, "Poln" indicates a polynomial of n-th degree.

I allow for additional model flexibility through calibration factors $\mu_{(\Delta E, m(\pi^+\pi^0))}$ and $\sigma_{(\Delta E, m(\pi^+\pi^0))}$ for all peaking components (Gaussians or Breit-Wigner) to enable the fit to capture differences between data and simulation in ΔE and $m(\pi^+\pi^0)$ resolutions and global scales; the factors for each fit observables are common to all components and incorporated in the models by transforming each mean and width as $\mu \longrightarrow \mu + \mu_C(\Delta E, m(\pi^+\pi^0))$ and $\sigma \longrightarrow \sigma \cdot \sigma_C(\Delta E, m(\pi^+\pi^0))$. These factors are meant to account empirically for the known mismodeling of the electromagnetic-calorimeter energy

I perform a four-dimensional fit to the $B^+ \to \bar{D}^0\rho^+$ candidates reconstructed from the same Belle II data set used for the signal-sample fit. I include in the fit shift and scale factors for the ΔE and $m(\pi^+\pi^0)$ signal shapes, while for $C_{\rm FBDT}$ and cosθ_{ρ^+} signal shapes are well reproduced in the simulation and do not need additional calibration. Figure 6.2 shows $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ distributions with fit projections overlaid. A prominent signal peak is observed at $\Delta E \approx 0$. In the continuum suppression output distribution, events tend to peak at $C'_{\rm FBDT} \approx 1$, as the continuum background in the $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ sample is small. In the dipion mass distribution, a clear ρ^+ meson peak is visible. The distribution of the cosine of ρ^+ helicity angle has the structure expected from simulation. The fit projections show imperfect modeling, especially in the angular distribution. I expect any possible impact to be largely reduced in the $B^+ \to \rho^+\rho^0$ channel, given the smaller sample-size. However, the angular mismodeling has the potential of biasing the polarization results and as such deserves further studies. These are discussed in Sec. 6.3.1 on samples where it is more straightforward to separate mismodeling of signal from mismodeling of $B\bar{B}$ backgrounds. The resulting calibration factors, $\mu_{\Delta E} = -3.4 \pm 0.4$ MeV and $\mu_{m(\pi^+\pi^0)} = -29 \pm 4$ MeV$/c^2$, are not expected to depend critically on the observed mismodeling and are subsequently used in the model of the $B^+ \to \rho^+\rho^0$ fit. Data suggest $\sigma_{(\Delta E,\ m(\pi^+\pi^0))}$ values compatible with unity, i.e., no smearing of the peaking distributions is needed. I keep the fit results on the sample composition hidden to avoid influencing other analyses currently ongoing in the Belle II collaboration targeting this channel. This seems appropriate as the branching-fraction reference for this decay dates back to 1994, and is based on a CLEO measurements on a small sample [78].

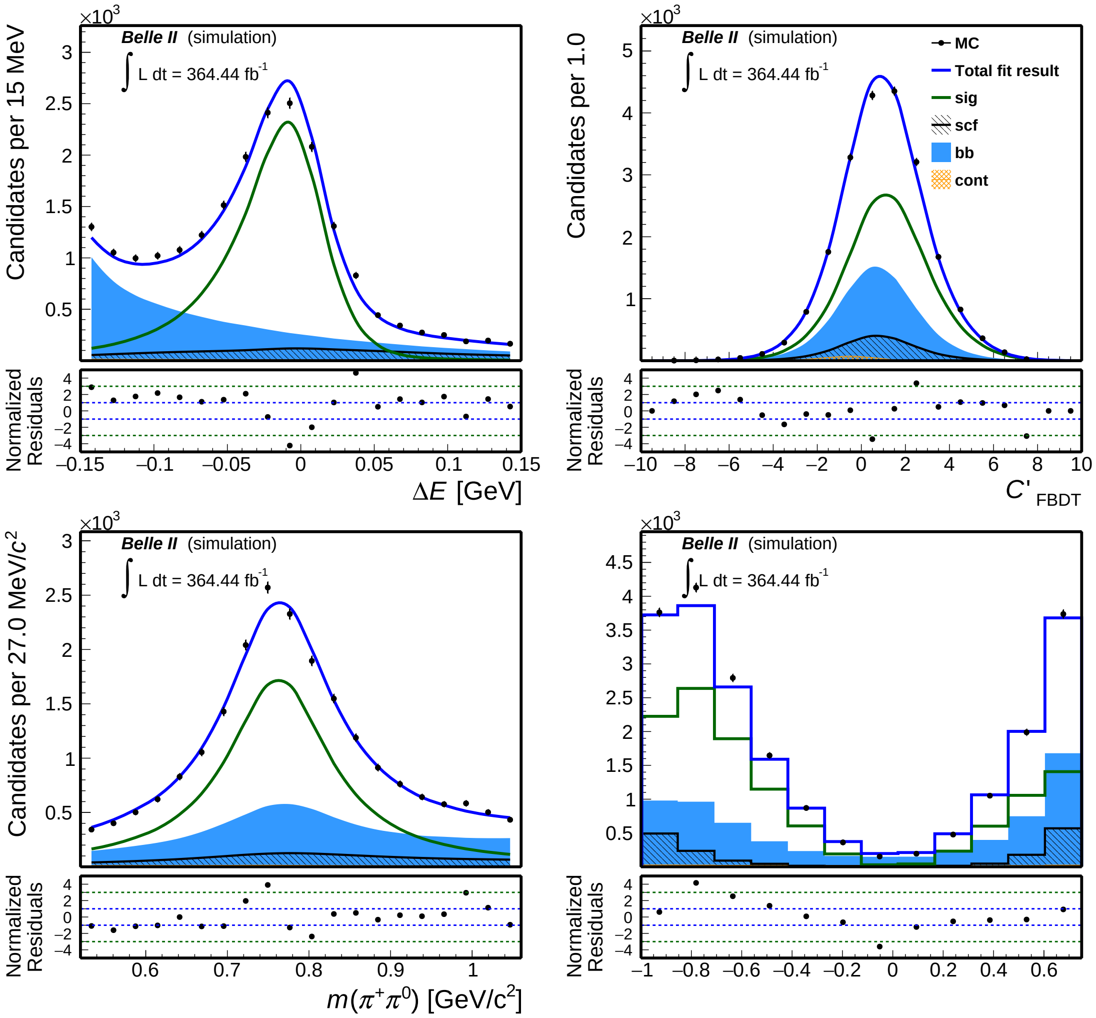


Figure 6.2: Distributions of (top left) ΔE, (top right) transformed continuum suppression output, (bottom left) $m(\pi^+\pi^0)$, and (bottom right) $\cos\theta_{\rho^+}$ for $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ decays reconstructed in the Belle II data set. Fit projections are overlaid (blue solid line).

6.3 Validation of background modeling

After calibrating the peaking signal shapes, I validate the modeling of the dominant background components, which come from continuum and $B\overline{B}$ events. I restrict the analysis to signal-free control regions in $B^+ \to \rho^+\rho^0$ dominated by these two components, and compare with simulation scaled to the appropriate luminosity. Signal decays are expected to cluster in the $-0.15 < \Delta E < 0.1$ GeV and $5.27 < M_{\rm bc} < 5.29$ GeV$/c^2$ regions. Hence, I require $\Delta E > 0.1\,$GeV or $\Delta E < -0.15\,$GeV and $M_{\rm bc} < 5.27\,$GeV$/c^2$, where continuum and $B\overline{B}$ background dominate and neither signals nor peaking backgrounds are expected to contribute significantly. The two-dimensional distribution of ΔE and $M_{\rm bc}$ of candidates reconstructed in the data is shown in Figure 6.3.

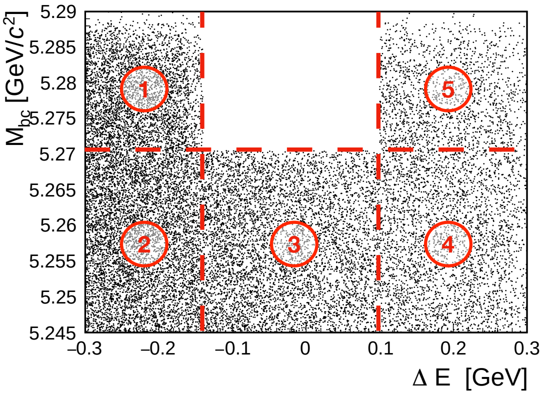


Figure 6.3: Distribution of $M_{\rm bc}$ as a function of ΔE for events reconstructed in collision data, where the signal region is excluded. The red dashed lines and numbers indicate the five sideband subregions referred to in text.

Figure 6.4 shows the comparison between data and simulation for the fit observables in the whole sideband. Significant discrepancies are visible in both helicity-angle distributions. Data seems to cluster more frequently toward the edges on the distributions with respect to simulation. This is particularly insidious, since the edges of the angular distributions are also the regions where the majority of the signal is expected and where most sensitivity to its polarization resides. Since helicity angles correlate strongly with the momenta of the ρ mesons' decay products, I compare the distributions of these (Fig. 6.5) and of relevant basic observables associated with neutral pions (Fig. 6.6) reconstructed in data and simulations. In an attempt at making the comparison more transparent in distinguishing instrumental from dynamics effects, the momenta of the ρ^0's charged pions momenta are also displayed after classifying them separately by magnitude. As expected, discrepancies are visible in the momenta of charged particles, especially at low momentum, and in the π^0 momentum. These two effects must be considered jointly, given the correlation between the momenta of final-state particles from ρ decays. The significant mismodeling observed in the momenta of the charged pions from the ρ^0 decay indicates that at least a significant cause of the π^0 momentum mismodeling is its correlation with the charged-pion decay-product. Hence, I first focus on instrumental effects associated with the reconstruction of charged particles. Possible causes of the observed data-simulation mismodelings include

- a common mismodeling of track acceptance in simulation, which would affect all the reconstructed candidates, also in the signal region;
- an angular shape mismodeling of at least one of the two sample components in the

simulation;

- different $B\overline{B}$-continuum proportions in simulation and data;
- a combination of some or all the previous effects.

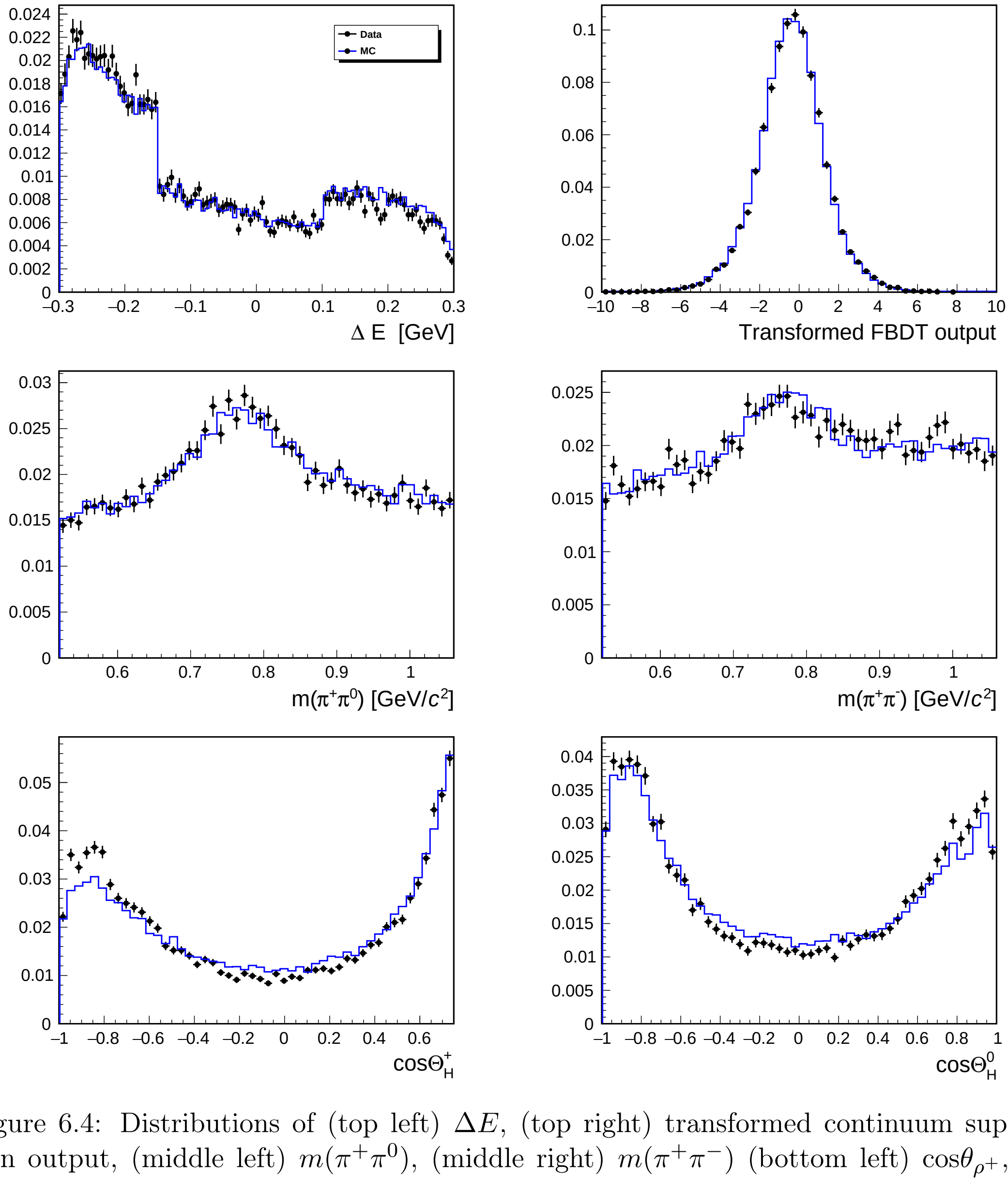


Figure 6.4: Distributions of (top left) ΔE, (top right) transformed continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $m(\pi^+\pi^-)$ (bottom left) $\cos\theta_{\rho^+}$, and (bottom right) $\cos\theta_{\rho^0}$ from $B^+ \to \rho^+\rho^0$ candidates reconstructed in (black dots) collision and (blue line) simulated data outside the signal box ($M_{bc} > 5.27$ GeV/c^2 and $-0.15 < \Delta E < 0.10$ GeV).

I further divide the sideband region into five independent regions (Fig. 6.3) that are expected to have differing proportions of continuum and $B\overline{B}$ background to gain additional indications about specific sources of discrepancies, if any. Simulation shows that in region 2 and region 5 the $B\overline{B}$-to-continuum ratio is 9%; region 1 is the most abundant in $B\overline{B}$ events with $B\overline{B}$-to-continuum ratio 45%; $B\overline{B}$-to-continuum ratio is 32% and 18% in region

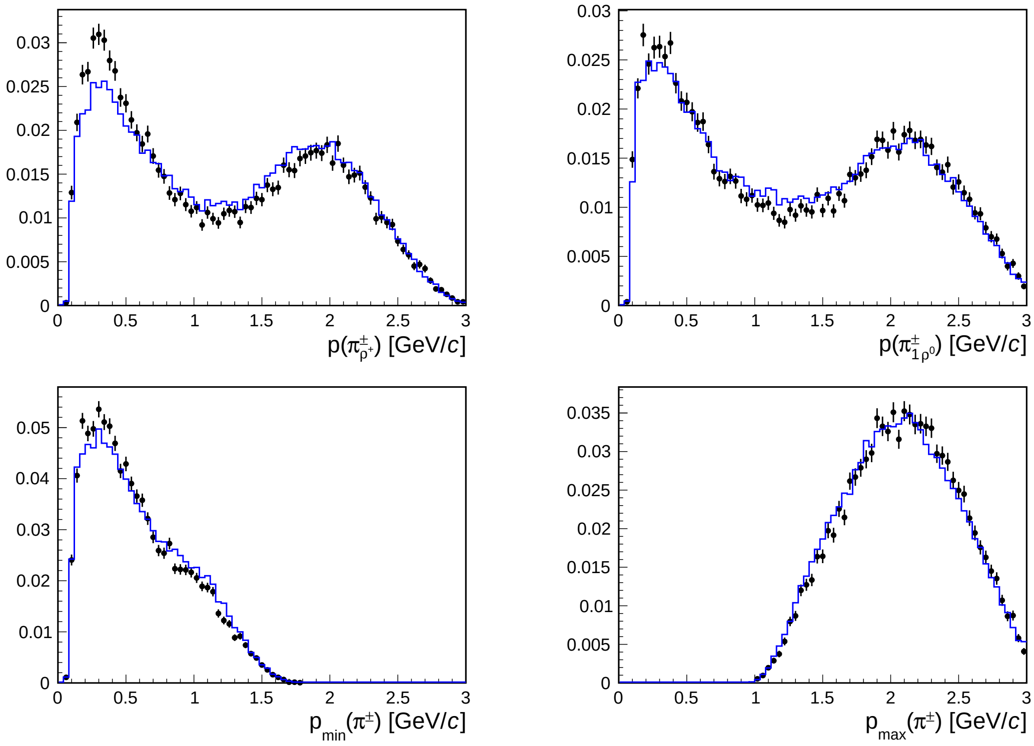


Figure 6.5: (Top) momentum distributions of (left) π from the ρ^+ and (right) π's from the ρ^0 reconstructed in (black dots) collision and (blue line) simulated data outside the signal box ($M_{bc} > 5.27$ GeV$/c^2$ and $-0.15 < \Delta E < 0.10$ GeV).
(Bottom) momentum distributions of the π from the ρ^0 with the (left) lower and (right) higher momentum, reconstructed in (black dots) collision and (blue line) simulated data outside the signal box ($M_{bc} > 5.27$ GeV$/c^2$ and $-0.15 < \Delta E < 0.10$ GeV).

3 and region 4, respectively. Figures 6.7–6.8 show that similar patterns of data-simulation mismodeling are present in all sidebands.

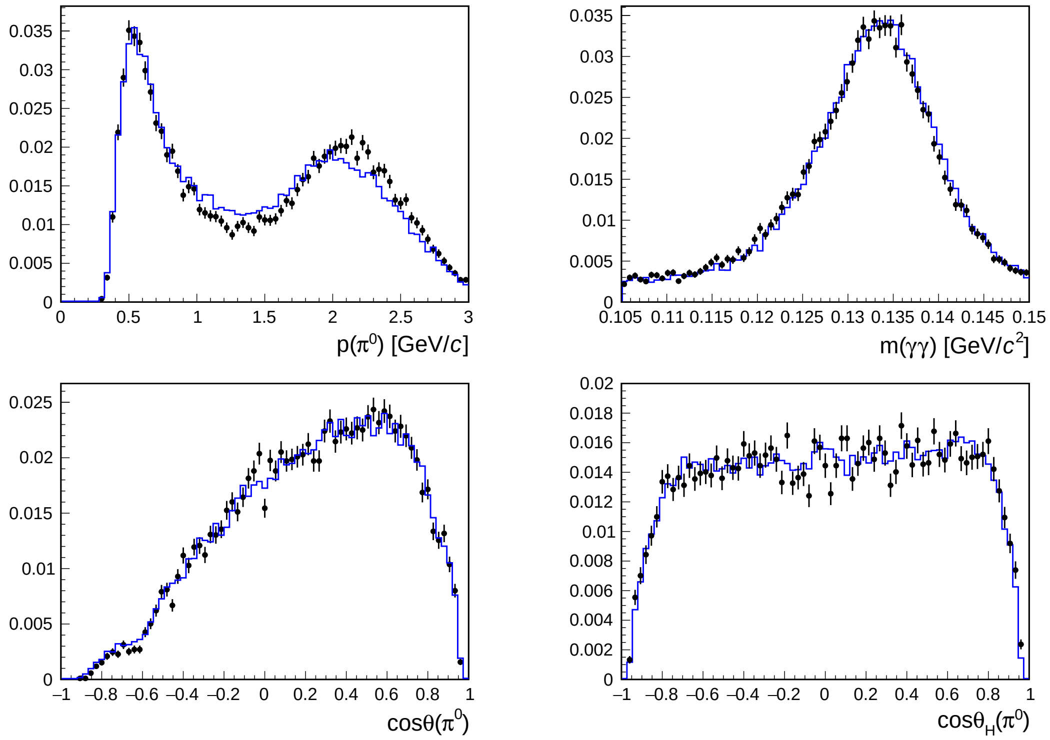


Figure 6.6: Distributions of (top left) momentum, (top right) diphoton mass, (bottom left) cosine of the polar angle, and (bottom right) cosine of the helicity angle of neutral pion candidates reconstructed in (black dots) collision and (blue line) simulated data outside the signal box ($M_{bc} > 5.27$ GeV/c^2 and $-0.15 < \Delta E < 0.10$ GeV).

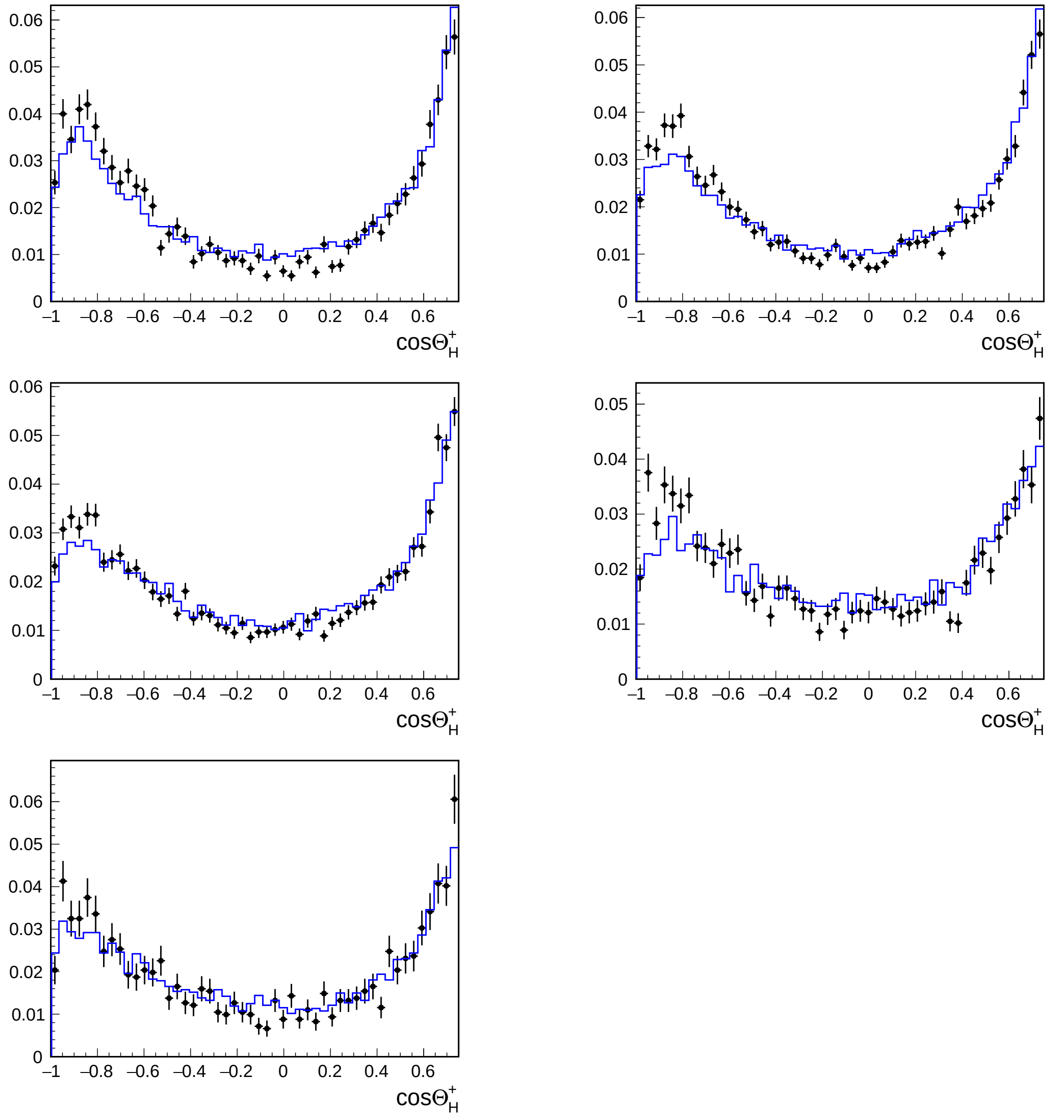


Figure 6.7: Distributions of cosθ_{ρ^+} from $B^+ \to \rho^+\rho^0$ candidates reconstructed in (solid) simulation and in (dots) data for the (top left) region 1, (top right) region 2, (middle left) region 3, (middle right) region 4, and (bottom left) region 5 of sidebands defined in Figure 6.3.

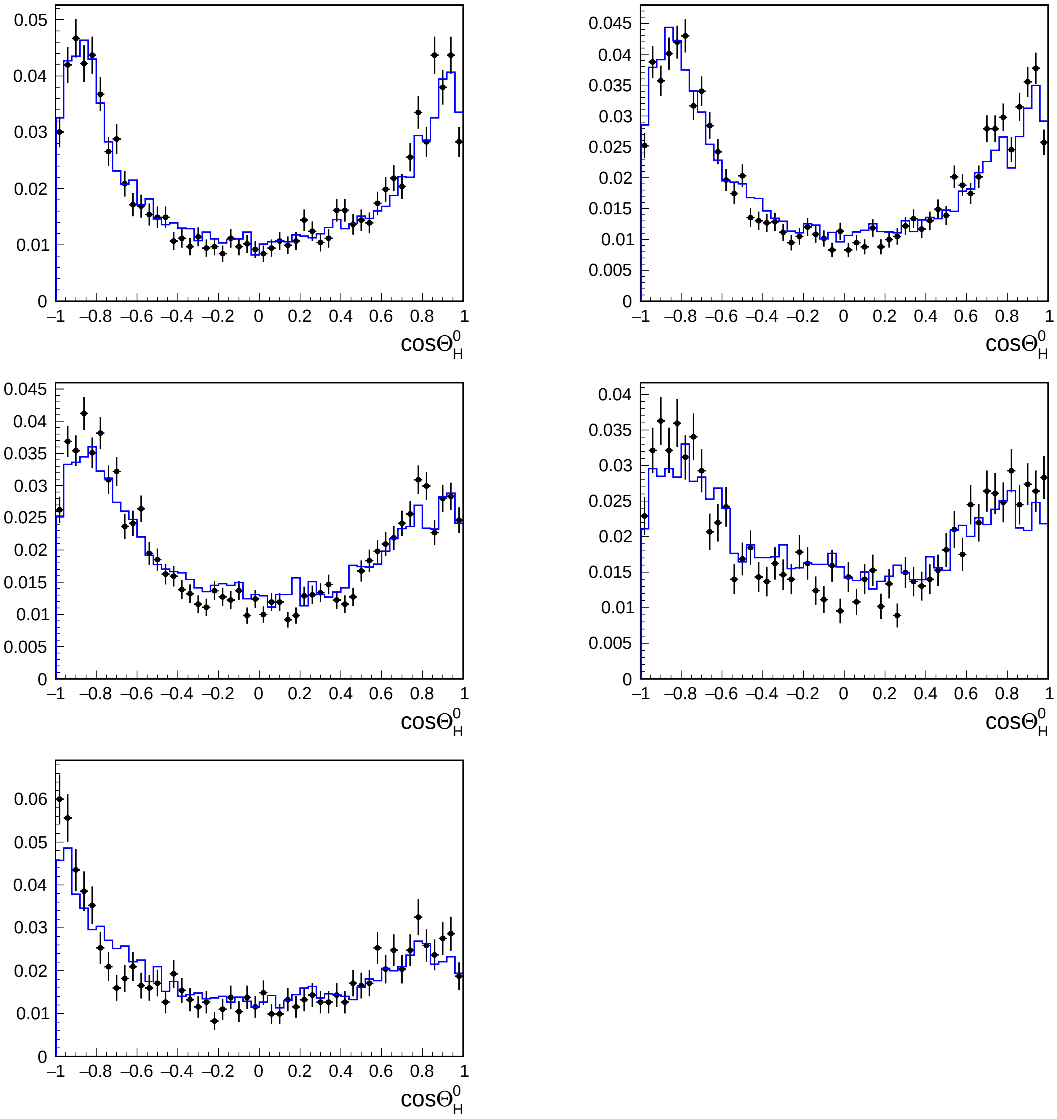


Figure 6.8: Distributions of $\cos\theta_{\rho^0}$ from $B^+ \to \rho^+\rho^0$ candidates reconstructed in (solid) simulation and in (dots) data for the (top left) region 1, (top right) region 2, (middle left) region 3, (middle right) region 4, and (bottom left) region 5 of sidebands defined in Figure 6.3.

6.3.1 Track-acceptance mismodeling

To investigate possible mismodeling of tracking efficiencies or acceptances, I use samples of tracks from signal-rich samples of abundant control B decays. The choice of pure signal samples is dictated by the need of making tests independent of uncertainties on sample composition, which requires suppressing backgrounds to a negligible level. In addition to the $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+(\to \pi^+\pi^0)$ channel, discussed in Sec 6.2, I reconstruct the following channels

> $B^+ \to \bar{D}^0(\to K^+\pi^-)\pi^+$. This channel allows for validating track efficiencies and acceptances in two different momentum ranges accessible through the two final-state pions, and in the simplest kinematic configuration of a two-body decay.
>
> $B^0 \to D^{*-}(\to \bar{D}^0(\to K^+\pi^-)\pi^-)\pi^+$. The low-energy pion from the D^* allows for probing the tracking efficiency in the low-momentum region, and the presence of the spin-1 D^* meson allows for checking proper modeling of decays with richer angular configurations. In addition, a selection on the $D^* - D^0$ mass difference helps to further reduce the background contamination.

I reconstruct all control channels with the same baseline track-quality, continuum suppression, and PID requirements used for the $B^+ \to \rho^+\rho^0$ signal. To obtain signal-pure samples, I impose restrictive requirements in ΔE, M_{bc}, and in the $m(K\pi)$ mass. For the $B^0 \to D^{*-}(\to \bar{D}^0(\to K^+\pi^-)\pi^-)\pi^+$ channel, I also restrict the $D^* - D^0$ mass difference ($\Delta m_{D^*-D^0}$), which has a narrow peak around 145 MeV/c^2 for signal candidates. Table 6.2 summarizes the selection on the control channels and the fraction of signal candidates (purity) of each sample, as determined from the simulation.

Channel	$B^+ \to \bar{D}^0\rho^+$	$B^+ \to \bar{D}^0\pi^+$	$B^0 \to D^{*-}(\to \bar{D}^0\pi^-)\pi^+$
ΔE [MeV]	[−40; +20]	[−20; +20]	[−20; +20]
M_{bc} [GeV/c^2]	[5.277; 5.283]	[5.27; 5.3]	[5.27; 5.3]
$m(K\pi)$ [GeV/c^2]	[1.858; 1.872]	[1.858; 1.872]	[1.858; 1.872]
$\Delta m_{D^*-D^0}$ [MeV/c^2]	n/a	n/a	[144; 147]
Purity	0.94	0.99	0.97

Table 6.2: Summary of the additional signal-enriching selections on the control channels used to isolate pure signal samples.

Comparison between (normalized) pion-momentum distributions in data and simulation for all the control channels are shown in Figures 6.9 — 6.11. For the $B^+ \to \bar{D}^0\rho^+$ channel, the cosine of the helicity angle of the ρ^+ is also shown. All the distributions show good agreement. This indicates that no significant global tracking efficiency or acceptance mismodeling exists in the momentum range $0.1 < p_{\pi^\pm} < 3.2$ GeV/c^2 for pions selected with the $B^+ \to \rho^+\rho^0$ selection, ruling this out as a cause for the angular mismodeling.

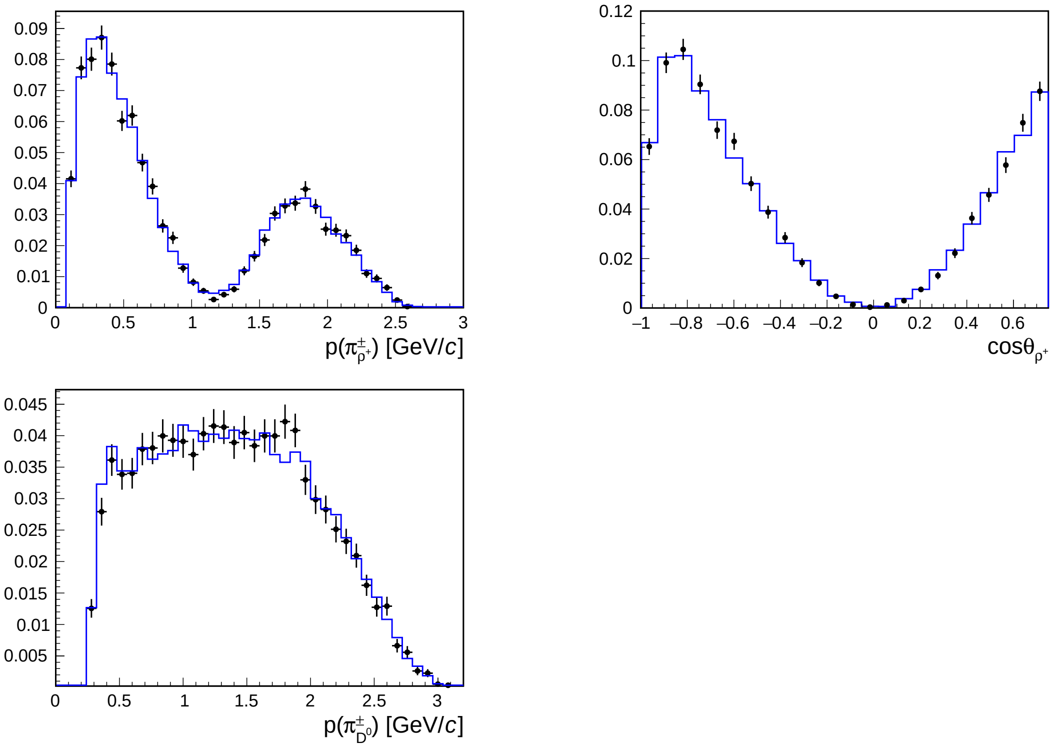


Figure 6.9: Distributions the (top left) momentum of pions coming from the ρ^+ decay, (top right) cosθ_{ρ^+}, and (bottom left) momentum of pions coming from the D^0 decay in $B^+ \to \bar{D}^0\rho^+$ candidates reconstructed in (solid line) simulation and (points) in the Belle II data set. Distributions are normalized to the same area.

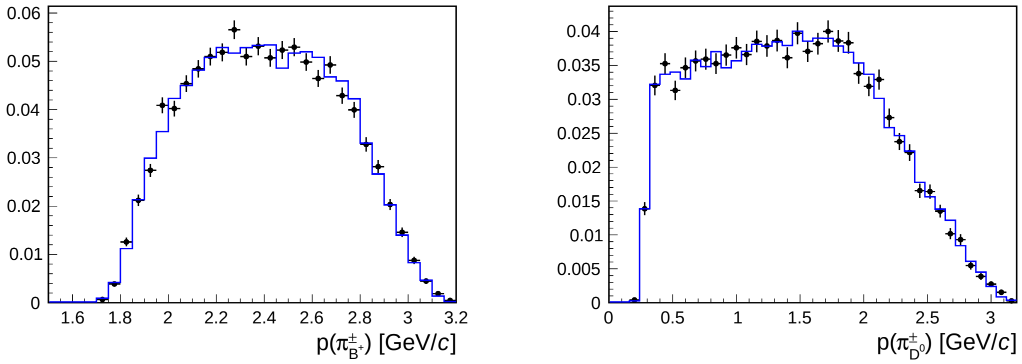


Figure 6.10: Momentum distributions for pions coming from the (left) B^+ decay and (right) D^0 decay in $B^+ \to \bar{D}^0\pi^+$ candidates reconstructed in (solid line) simulation and (points) in the Belle II data set. Distributions are normalized to the same area.

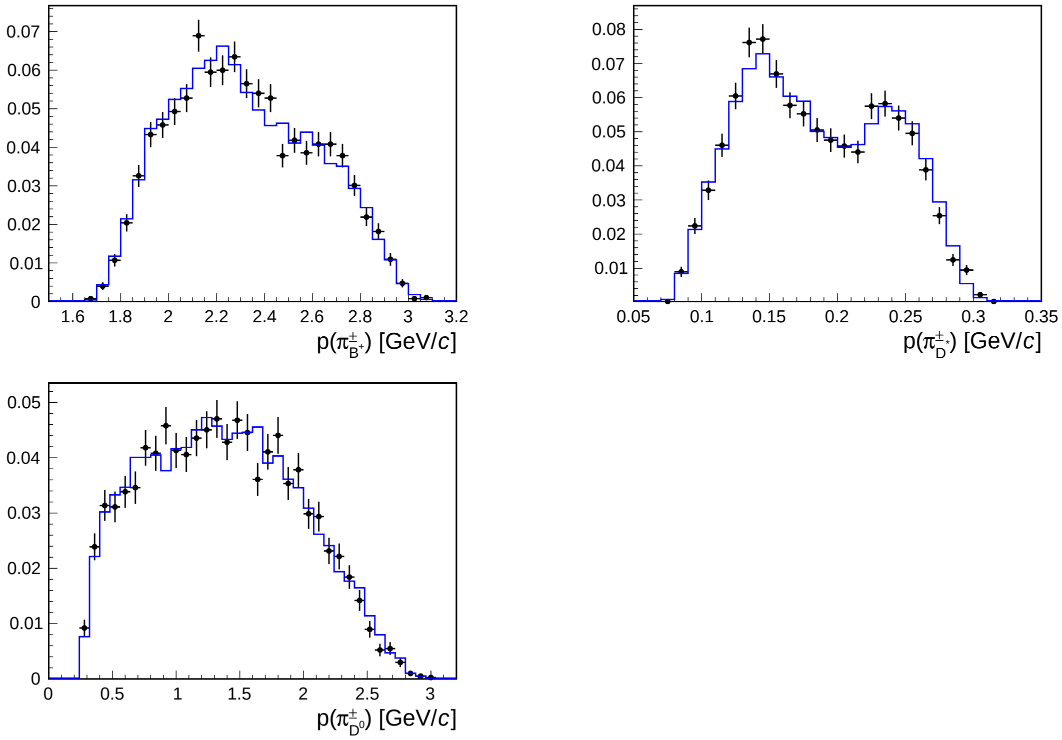


Figure 6.11: Momentum distributions for pions coming from the (top left) B^0 decay, (top right) D^* decay, and (bottom left) D^0 decay in $B^0 \to D^{*-}(\to \bar{D}^0\pi^-)\pi^+$ candidates reconstructed in (solid line) simulation and (points) in the Belle II data set. Distributions are normalized to the same area.

6.3.2 Shape mismodeling of the continuum or $B\overline{B}$ components

The angular mismodeling observed in the sidebands might be due to a mismodeling in the shape of at least one of the two dominant components. Data taken at collision energies inferior (about 60 MeV) to the $\Upsilon(4S)$ resonance, so-called off-resonance data, offer a reliable proxy for pure continuum. However, the size of the off-resonance sample is small, as only 10%–15% of data-taking time is devoted to that. I partially mitigate this sample-size limitation by loosening the requirement on the continuum suppression output, passing from a 0.96 threshold (used in the $B^+ \to \rho^+\rho^0$ analysis) to 0.9, which increases the off-resonance sample size available for my tests by a factor of 4.6. It is important to ensure that the added events do not show different angular features, which would spoil the objective of the test. Simulation shows that the distributions of momenta of the final-state particles and the helicity angles of the ρ mesons do not change significantlyas a function of the continuum suppression output (Fig. 6.12), thus validating this extension.

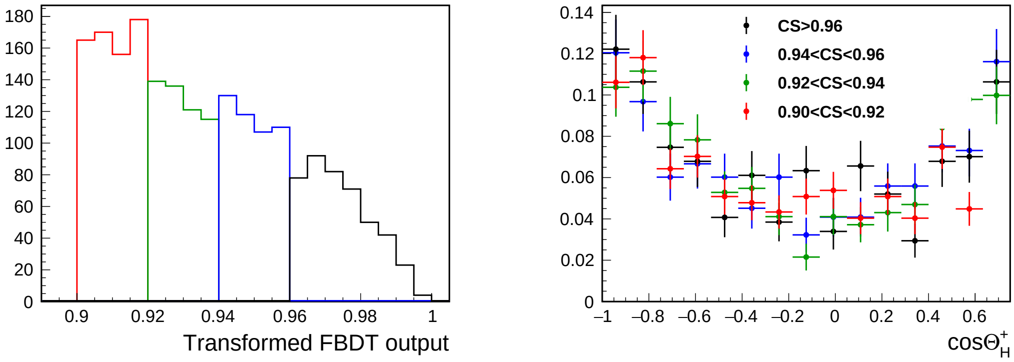


Figure 6.12: Distributions of (left) continuum suppression output and (right) cosθ_{ρ^+} in different ranges of continuum suppression output. Helicity angles distributions are normalized to the same area. The cosθ_{ρ^+} are not reported for simplicity as they show even better agreement.

Figure 6.13 shows comparisons of the helicity-angle distributions between off-resonance data and continuum simulation normalized to the same area. Large discrepancies are visible, and similar to the ones observed in the signal sidebands (Fig. 6.4). Such discrepancies are confirmed by similar effects at low values of pion momenta, shown in Fig. 6.14. The mismodeling of momenta for continuum final-state-tracks might be one culprit of the observed data-simulation mismodeling in the sidebands.

A natural option to avoid the mismodeling in the continuum helicity-angle distributions is to use directly the enriched off-resonance sample to prepare the helicity-angle templates used in the $B^+ \to \rho^+\rho^0$ fit. This solution trades off the benefit of reducing significantly a modeling systematic uncertainty (data-simulation differences) that is harder to assess with the penalty of increasing another modeling systematic uncertainty (off-resonance sample size) whose origin is statistical in nature and therefore more straightforward to assess.

A first validation of this alternate continuum modeling based on off-resonance data is achieved by fitting the data sidebands. The strategy is the same as for the signal $B^+ \to \rho^+\rho^0$ fit, but the sample composition is restricted to $B\overline{B}$ and continuum components only. In addition, ΔE is dropped from the set of fit observables, to avoid difficulties associated with modeling the cusp due to the ΔE-M_{bc} signal-box exclusion. All shapes are fixed from simulation: C'_{FBDT} and $m(\pi\pi)$ are described with analytical shapes, while the helicity angles are modeled with histogram templates from the off-resonance sample.

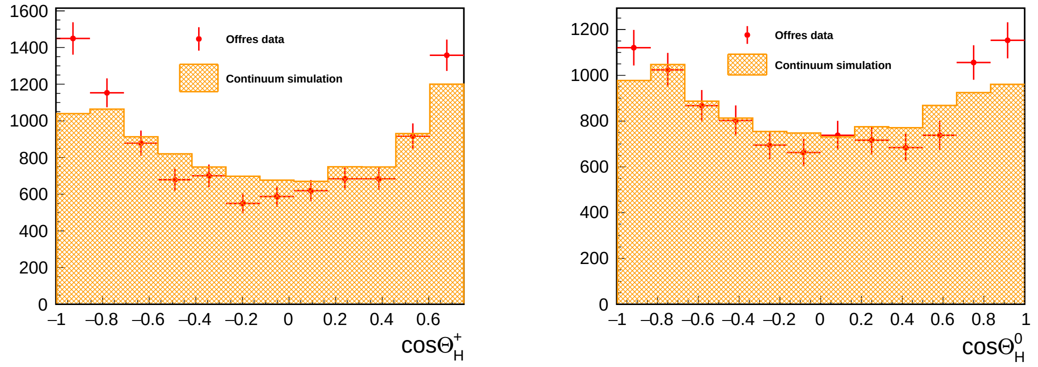


Figure 6.13: Distributions of (left) $\cos\theta_{\rho^+}$, and (right) $\cos\theta_{\rho^0}$ reconstructed in (solid) continuum simulation and (points) off-resonance data, enriched with an enlarged continuum suppression output window. Distributions are normalized to the same area.

Since the off-resonance sample size is smaller than the size of the fitted sample, continuum templates have significant statistical fluctuations in their bin content, which are included in the fit. The results of the $B^+ \to \rho^+\rho^-$ sideband fit that employs a continuum model taken from off-resonance data are shown in Table 6.3, compared with the values expected from simulation. Fit projections are shown in Figure 6.15. The fit describes the data reasonably. Indications of moderate mismodeling affecting the $m(\pi^+\pi^0)$ and helicity-angle distributions are seen. However, agreement with data is largely improved with respect to the a fit where all shapes are extracted from simulation (Fig. 6.16). The resulting sample composition disagrees with expectation from simulation, indicating that either the sample composition is wrongly described in the simulation, or the $B\overline{B}$ shape is wrong, or both.

Parameter	Fit result	MC value
$A_{CP}^{B\overline{B}}$	-0.069 ± 0.015	-0.009
$B^0 \to \bar{D}^0\pi^+\pi^-$ non resonant	0.514 ± 0.007	0.4597
Other backgrounds	0.322	

Table 6.3: Results of the fit of the data sideband, compared with the true values of the simulation.

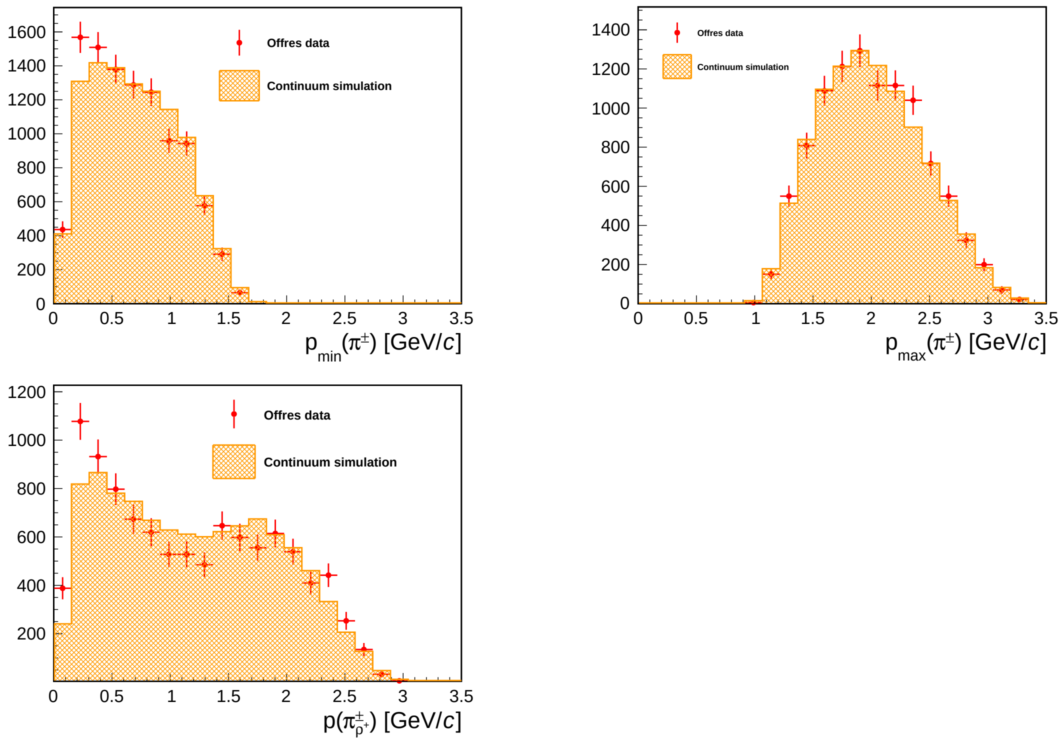


Figure 6.14: Momentum distributions of (top left) π from the ρ^0 with lower momentum, (top right) π from the ρ^0 with higher momentum, and (bottom left) π from the ρ^+ reconstructed in (solid) continuum simulation and (points) off-resonance data, enriched with the enlarged continuum suppression output window. Distributions are normalized to the same area.

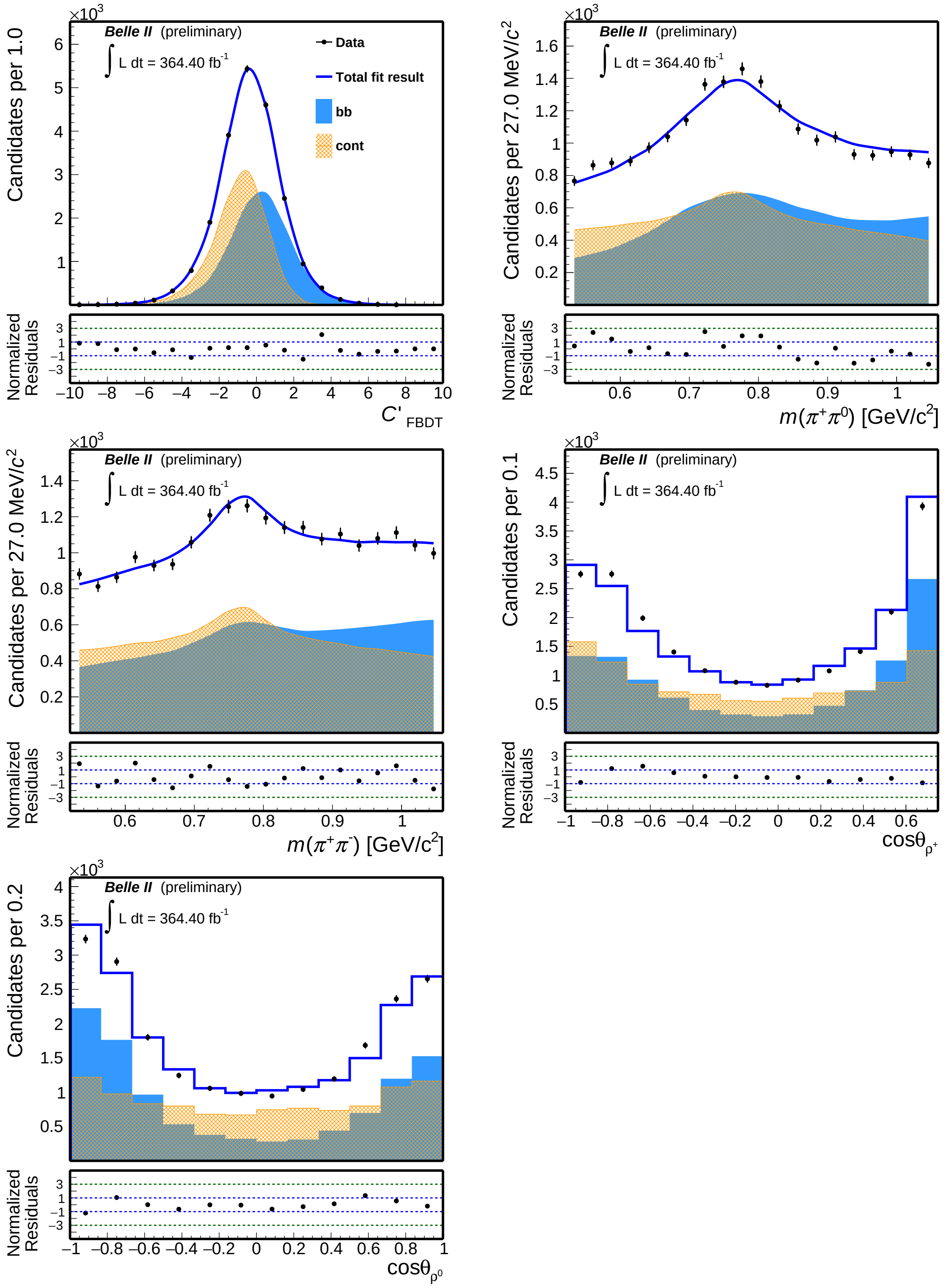


Figure 6.15: Distributions of (top left) transformed continuum suppression output, (top right) $m(\pi^+\pi^0)$, (middle left) $m(\pi^+\pi^-)$, (middle right) $\cos\theta_{\rho^+}$, and (bottom left) $\cos\theta_{\rho^0}$ for $B^- \to \rho^-\rho^0$ candidates reconstructed in the Belle II data sidebands ($M_{bc} > 5.27$ GeV$/c^2$ and $-0.15 < \Delta E < 0.10$ GeV). Projections of a fit in which the continuum model is adjusted according to the off-resonance studies described in text are overlaid.

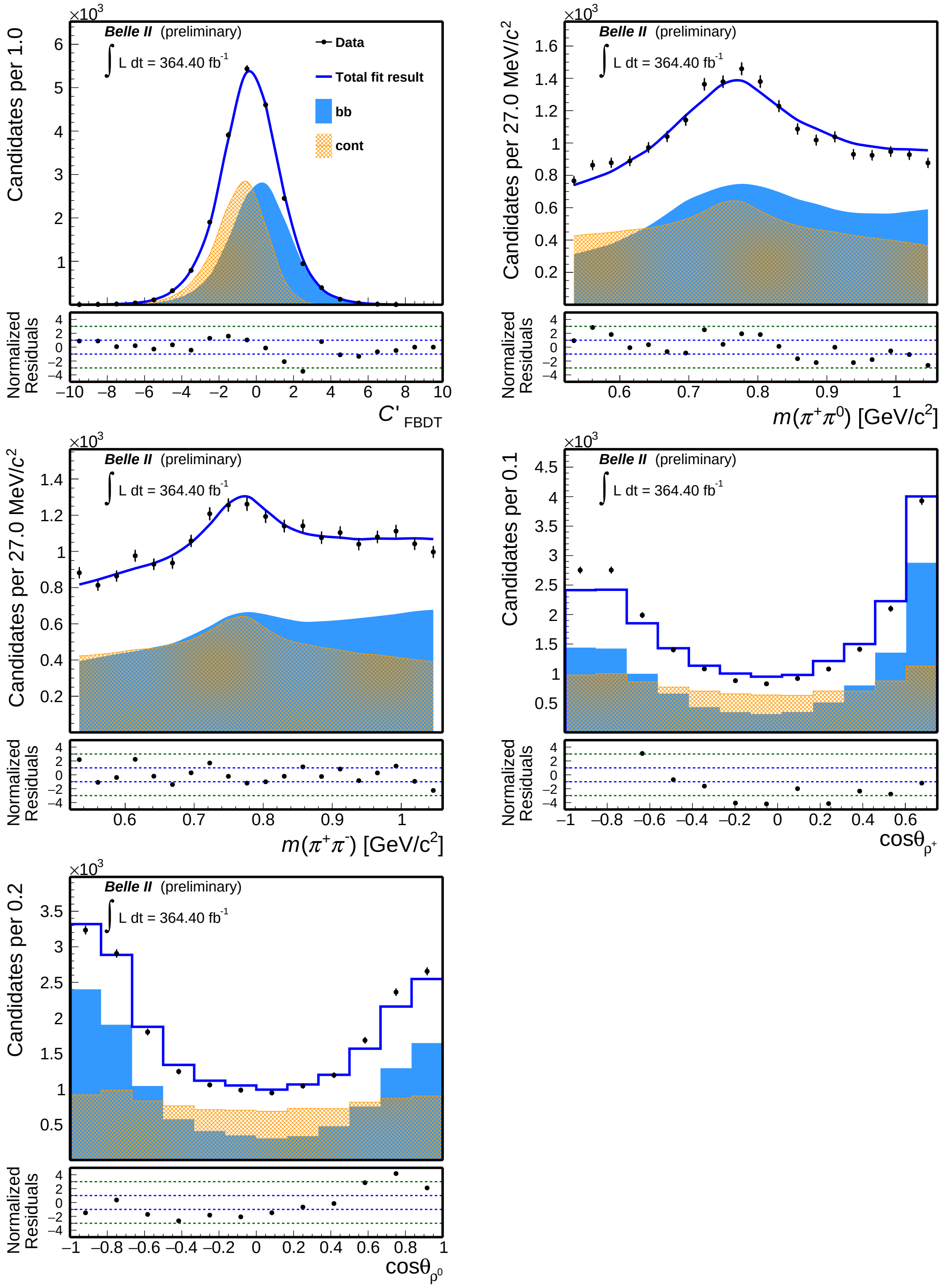


Figure 6.16: Distributions of (top left) transformed continuum suppression output, (top right) $m(\pi^+\pi^0)$, (middle left) $m(\pi^+\pi^-)$, (middle right) $\cos\theta_{\rho^+}$, and (bottom left) $\cos\theta_{\rho^0}$ candidates reconstructed in the Belle II data sidebands ($M_{bc} > 5.27$ GeV/c^2 and $-0.15 < \Delta E < 0.10$ GeV). Projections of a fit in which the continuum model is taken directly from simulation.

To check the reliability of the $B\overline{B}$ modeling, one would need a pure $B\overline{B}$ sample in data, with features similar to those of $B\overline{B}$ events in the $B^+ \to \rho^+\rho^0$ sample. Such a sample is not available. The closest approximation is the data sideband itself, statistically subtracted from the continuum component whose shape is taken from off-resonance data and proportions from the fit of sideband composition discussed above. This approach involves a circularity, as the fit of sideband composition assumes a possibly wrong $B\overline{B}$ model, which is the exactly the issue I am trying to address. However, I expect to converge to a stable and reliable picture with a small number of fit iterations refining both $B\overline{B}$ shape and sideband composition after each. Figures 6.17–6.18 show the results. Helicity angle and pion-momentum distributions are compared between the statistically pure $B\overline{B}$ sample resulting from the subtraction and simulation. The agreement is better than that observed originally in the continuum component, but discrepancies are present. If similar discrepancies are present in the $B^+ \to \rho^+\rho^0$ signal region and are not accounted for, they can leak into the signal composition biasing the fit results.

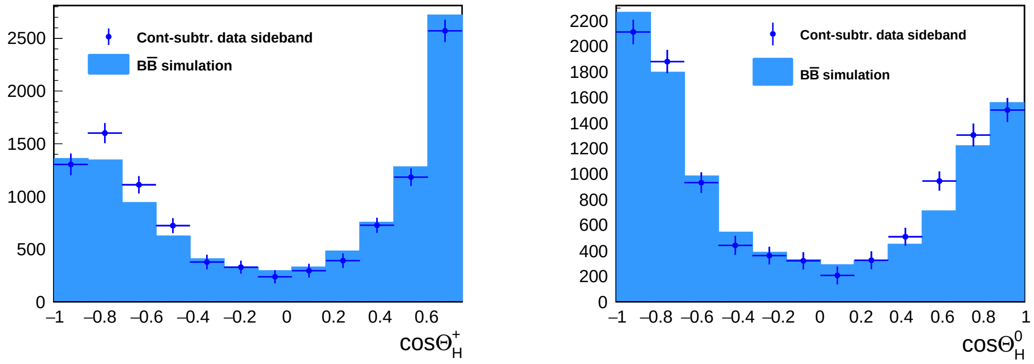


Figure 6.17: Distributions of (left) $\cos\theta_{\rho^+}$, and (right) $\cos\theta_{\rho^0}$ reconstructed in (solid) $B\overline{B}$ simulation and (points) sideband data after subtracting the continuum distributions according to the proportions found from the data sideband fit. Distributions are normalized to the same area.

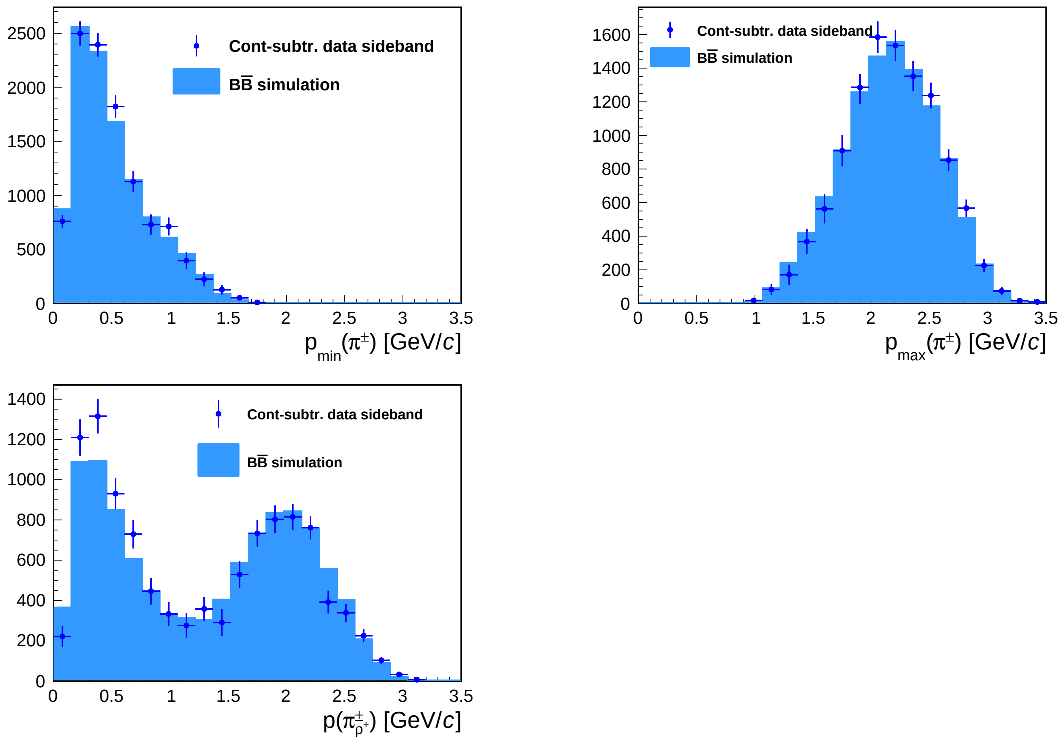


Figure 6.18: Momentum distributions of (top left) π from the ρ^0 with lower momentum, (top right) π from the ρ^0 with higher momentum, and (bottom left) π from the ρ^+ reconstructed in (solid) $B\overline{B}$ simulation and (points) sideband data after subtracting the continuum distributions according to the proportions found from the data sideband fit. Distributions are normalized to the same area.

The $B\overline{B}$ component contains B candidates misreconstructed from an admixture of many B decays: any significant proportion or shape mismodeling of such decays in the simulation could contribute the observed data-simulation mismodeling. An empiric option for partially mitigating the risk of large fit biases is to weigh the $B\overline{B}$ component of simulation to make it resembling the data. As weights, I use the ratios of the two-dimensional distributions of the momenta of the pions. The data-simulation weight for each B^+ candidate is the product of the weights coming from the ρ^+ and ρ^0 final-state weights (Fig. 6.19).

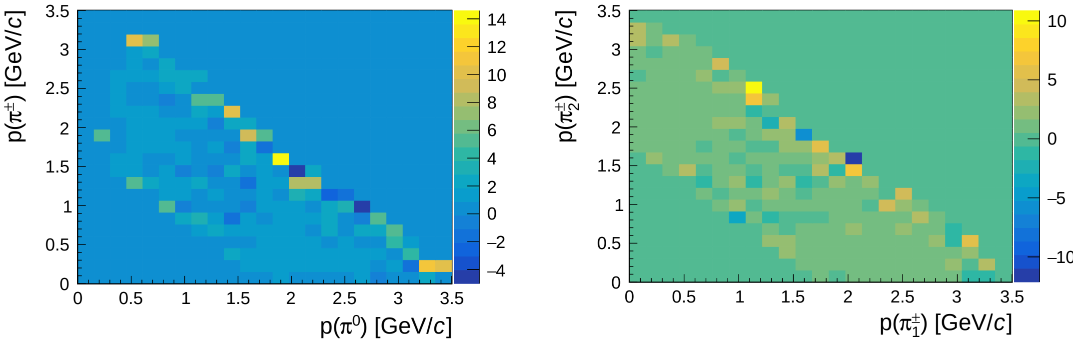


Figure 6.19: Two-dimensional data/simulation weight distributions for the pions coming from the (left) ρ^+ and (right) ρ^0 in the $B\bar{B}$ component. The presence of negative weights is due to the subtraction of the continuum component in data.

6.3.3 Blind closure test on partial sample

At this point, I have investigated an extensive set of plausible hypotheses that could explain the observed angular discrepancies. I also have addressed those that could be independently studied in control samples by modifying our continuum and $B\bar{B}$ models. While I am confident that this results in a globally improved model for our $B^+ \to \rho^+\rho^-$ fit, there is no guarantee that residual mismodelings are not present in that sample. After applying all the corrections discussed in the previous sections, a blind closure test on a subsample of signal $B^+ \to \rho^+\rho^-$ data is probably the only straightforward option to gather information on possible residual mismodelings without spoiling the blind status of this thesis' analysis. This test is achieved with a full but blind $B^+ \to \rho^+\rho^-$ fit of the sample collected until July 2021. In this fit I model the continuum component using directly the shapes from off-resonance data and adjust the simulated $B\bar{B}$ model using data weights as discussed. The fit is configured in such a way that numerical results remain undisclosed and only fit projections are inspected. The subset of data collected until July 2021 is a significant fraction of the Belle II sample, amounting to about 50%, which would seem questionable a choice. However, our choice does not pose a blinding concern as those data have already been used for a full unblinded preliminary result shown in the Winter 2022 conferences [79].

The projections of this fit are shown in Figure 6.20. The fit describes accurately the data in ΔE, C'_{FBDT}, and in the $m(\pi\pi)$ invariant masses. It also shows acceptable agreement for $\cos\theta_{\rho^+}$. These projections show significant improvement over those of the fit performed with all shapes taken from simulation and not corrected (Fig. 6.21). An improvement in the description of the $\cos\theta_{\rho^+}$ is especially visible. However, significant mismodeling still affects the $\cos\theta_{\rho^0}$ distribution. This indicates the presence of one or more causes of additional mismodelings, affecting particularly the ρ^0 angular distribution.

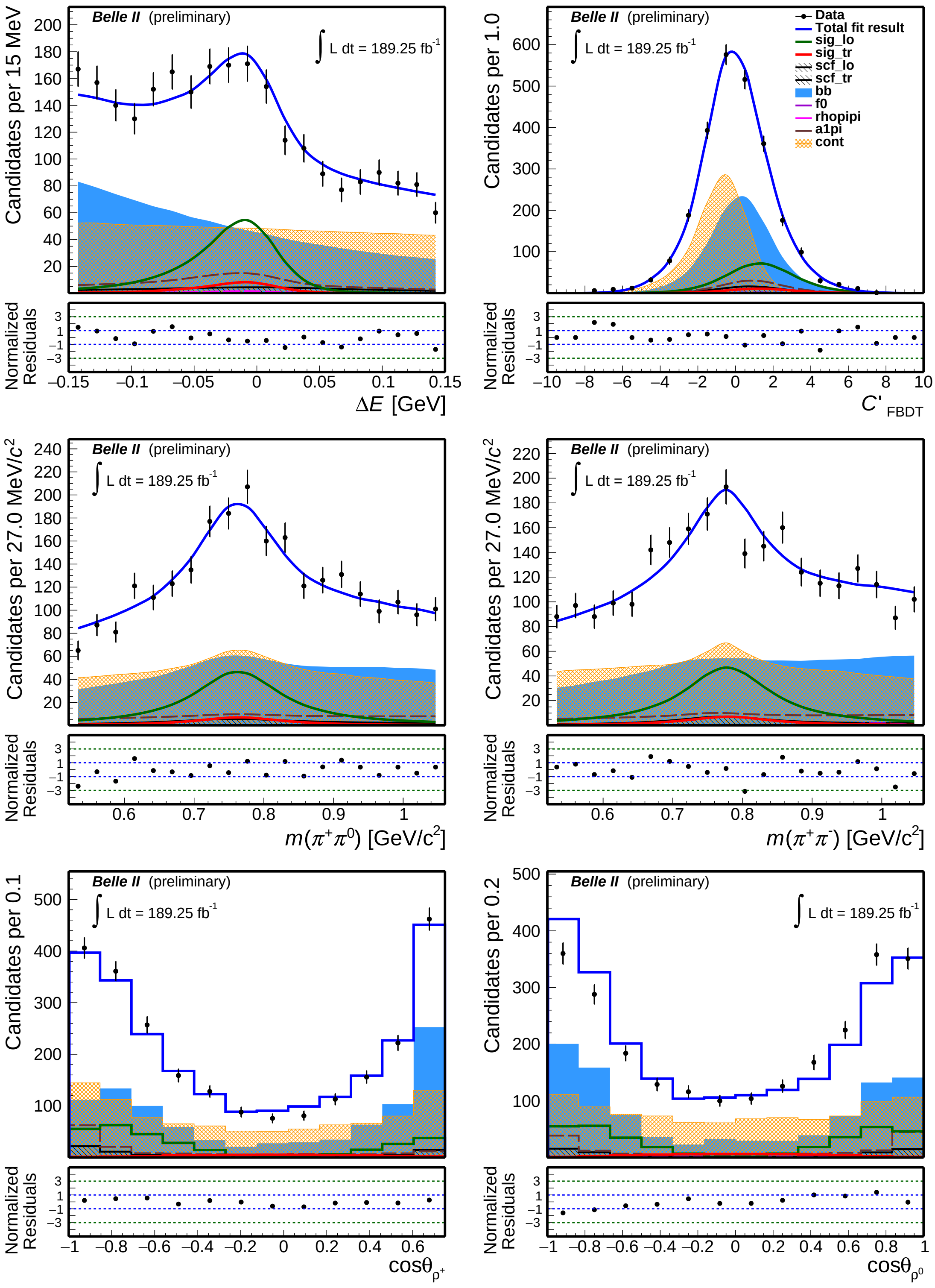


Figure 6.20: Distributions of (top left) ΔE, (top right) transformed continuum suppression output,, (middle left) $m(\pi^+\pi^0)$, (middle right) $m(\pi^+\pi^-)$ (bottom left) cosθ_{ρ^+}, and (bottom right) cosθ_{ρ^0} from $B^+ \to \rho^+\rho^0$ candidates reconstructed in the Belle II data set collected at the $\Upsilon(4S)$ resonance up to July 2021 and corresponding to 198 million $B\overline{B}$ pairs. Projections of a fit in which both continuum and $B\bar{B}$ models are adjusted following our studies on data are overlaid.

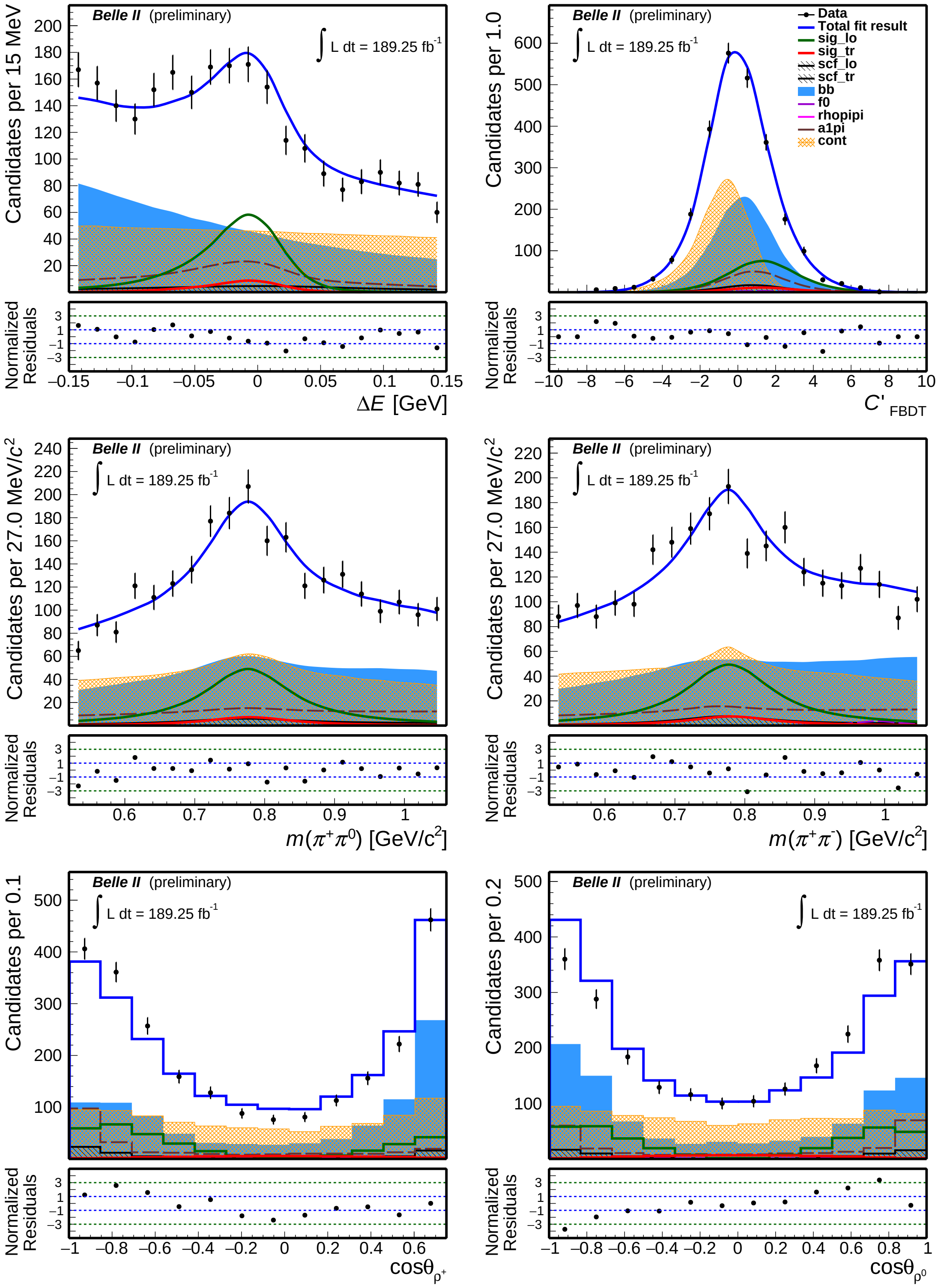


Figure 6.21: Distributions of (top left) ΔE, (top right) transformed continuum suppression output,, (middle left) $m(\pi^+\pi^0)$, (middle right) $m(\pi^+\pi^-)$ (bottom left) cosθ_{ρ^+}, and (bottom right) cosθ_{ρ^0} from $B^+ \to \rho^+\rho^0$ candidates reconstructed in the Belle II data set collected at the $\Upsilon(4S)$ resonance up to July 2021 and corresponding to 198 million $B\overline{B}$ pairs. Projections of a fit in which all shapes are taken from simulation are overlaid.

6.3.4 Additional checks on cos$\boldsymbol{\theta_{\rho^0}}$

The large mismodeling observed in quantities related to the ρ^0 suggests that a dedicated investigation of a ρ^0 signal from B decays might be informative. I reconstruct in data and in simulation the control channel $B^0 \to \bar{D}^0(\to K^+\pi^-)\rho^0(\to \pi^+\pi^-)$, the only known B decay channel offering a statistically significant sample of ρ^0 mesons in our analysis phase-space [10]. I apply the same baseline track selection, PID, and continuum suppression requirements as for the signal channel, and then the same additional ΔE-M_{bc}-$m(K\pi)$ restrictions as for the $B^+ \to \bar{D}^0\rho^+$ in Table 6.2. The resulting sample composition expected from simulation is different from the $B^0 \to \rho^+\rho^0$ composition, as shown in Table 6.4. Two important backgrounds contaminate the signal: correctly reconstructed non-resonant $B^0 \to \bar{D}^0\pi^+\pi^-$ decays, which have the same final state of the $B^0 \to \bar{D}^0\rho^0$, and combinatorial candidates reconstructed from other $B\overline{B}$ decays.

Figure 6.22 shows the $m(\pi^+\pi^-)$ and relevant helicity-angle distributions in simulation and data. Simulation shows general agreement with data for the $m(\pi^+\pi^-)$ observable, suggesting that major issues in sample composition modeling are unlikely. The distribution of cosθ_{ρ^0} show a large mismodelings: an excess of data affects the cos$\theta_{\rho^0} < 0$ region, compensated by a similarly sized deficit in the region $0 <$ cos$\theta_{\rho^0} < 0.7$. Both regions are dominated by signal with a sizable population of non-resonant $B^0 \to \bar{D}^0\pi^+\pi^-$ decays. Conversely, the region cos$\theta_{\rho^0} > 0.7$, which is expected to be dominated by other $B\bar{B}$ backgrounds, shows better agreement. While subtle instrumental effects cannot be ruled out, this phenomenology seems suggestive of the unaccounted for presence of interference between the $B^0 \to \bar{D}^0\pi^+\pi^-$ and signal $B^0 \to \bar{D}^0\rho^0(\to \pi^+\pi^-)$ components. Indeed, such interference has been measured by the LHCb collaboration [80] and is not yet included in the default Belle II simulation. It is not straightforward to estimate how including the LHCb findings in our model would improve our data-simulation comparison, thus confirming or excluding conclusively our ansatz on ρ^0 mismodeling being due to large interference. The LHCb analysis is based on an extended portion of the Dalitz plot with respect to ours, and modifying privately the Belle II simulation to properly include the amplitude structure observed by LHCb involves some nontrivial technical difficulties. In addition, even if I would include the proper interference structure in the $B^0 \to \bar{D}^0\rho^0(\to \pi^+\pi^-)$ fit and that would reconcile data and similation, the differences in composition between $B^0 \to \bar{D}^0\rho^0$ and $B^0 \to \rho^+\rho^0$ samples are such that it would be unrealistic to achieve a quantitative application of these findings to the target signal sample.

Nevertheless, the body of evidence accumulated with the many tests described in this chapter indicates that a plausible motivation for the persisting mismodeling observed in our data could be larger-than-expected interference effects between the resonant $B^0 \to \rho^+\rho^0$ signal and a variety of competing decay modes such as $B \to \rho\pi\pi$ decays, which contaminate our sample in proportions that are often unknown.

Component	Fraction
Signal	0.483
$B^0 \to \bar{D}^0\pi^+\pi^-$ non resonant	0.195
Other backgrounds	0.322

Table 6.4: Summary of the $B^0 \to \bar{D}^0(\to K^+\pi^-)\rho^0(\to \pi^+\pi^-)$ sample composition after applying all the selection in simulation.

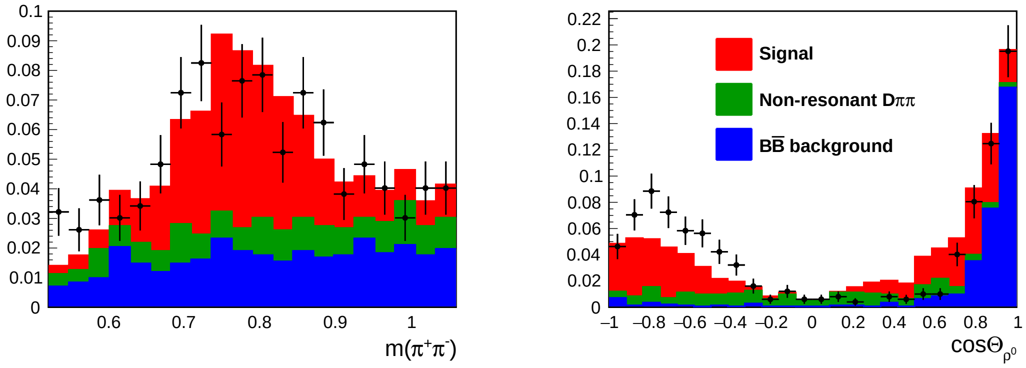


Figure 6.22: (Left) Invariant mass $m(\pi^+\pi^-)$ and (right) helicity angle $\cos\theta_{\rho^0}$ distributions of $B^+ \to \bar{D}^0\pi^+$ candidates reconstructed in (solid) simulation and (points) in the Belle II data set. Distributions are normalized to the same area. Different colors indicate the different components of the simulated sample.

6.4 Discussion

The outcome of the tests described in this chapter is summarized as follows. Using various control channels in data I addressed a number of minor-to-significant mismodeling of our fit model, which lead to a more accurate description of Belle II data. However, I could not find a sensible way of correcting one residual mismodeling that affects mostly the angular distribution of the ρ^0 decay products. There are strong indications that such angular mismodeling might be due to larger-than-expected interference effects between signal and subleading decays into the same final states.
In our plan for the first Belle II $B^+ \to \rho^+\rho^0$ analysis, we intended to treat the interference effects as part of the systematic uncertainties. This was motivated by previous $B^+ \to \rho^+\rho^0$ analyses [41] [42], where such interference had been neglected. In addition, qualitative considerations suggest that the impact of interference in $\rho^+\rho^0$ decays is likely to be closer to the relatively small impact observed in $\rho^+\rho^-$ decays rather than the more significant one observed in $\rho^0\rho^0$ decays. The rationale would be that that the $\rho^0\rho^0$ amplitude is color-suppressed and has to compete with other 4π $b \to u$ tree amplitudes that are color-favored, such as $a_1\pi$. Even in the reduced phase space of the analysis region, this can still produce competing amplitudes with comparable sizes, and thus conducive to large interference. Conversely, both $\rho^+\rho^0$ and $\rho^+\rho^-$ are color-favored, which offers a relative amplitude-size advantage in their analysis space over the other color-favored 4π states that have different lineshapes. To be competitive with the resonant signal, interfering decays that meet the final analysis selection most likely have to contain a ρ meson, like $\rho\pi\pi$ and $a_1\pi$. Since these have a direct counterpart in the $\rho^+\rho^-$ final state, where no large interference is observed, we took that to favor a similar situation of reduced interference in the case of our sample.

However, data show that this is unlikely to be the case prompting a reappraisal of our analysis plan. Any empirical solution aimed at reconciling data with simulation in a sufficiently conservative way would completely spoil the statistical power of the analysis. Blind tests performed in the necessary pessimistic case of attributing the full size of the discrepancy to a signal mismodeling show biases in the fit results amounting to multiple statistical standard deviations, which make them unaviable. A satisfactory and robust solution would be to expand the analysis to consider a more sophisticated modeling of the dipion masses and possibly additional decay angles, capable of capturing reliably the effects of different components and interference, which could then be incorporated as part of the fit results, if not a full Dalitz plot analysis. However, this is a major expansion of the analisys that would be incompatible with the scope and constraints of this thesis. Hence, we decided to leave the discussion open on this point and maintain the data blind so as to be able to bring this to final internal approval and submission to journal once a more reliable treatment of the interference will be finalized.

Chapter 7

Fit of $B^+ \to \rho^+\rho^0$ data

After validating the analysis on simulated samples and control data, I apply the fit to signal candidates in data. Since the analysis is still under collaboration review, I only show fit diagnostics, projections, and uncertainties while hiding the central values.

I perform a fit to the Belle II data set corresponding to 387 million of $B\overline{B}$ pairs and containing 5382 $B^+ \to \rho^+\rho^0$ candidates after the selection. Before the fit I apply the ΔE and $m(\pi^+\pi^0)$ shift factors from Sec. 6.2.3 to the relevant peaking probability-density functions. In addition, I model the helicity angles of the continuum component using off-resonance data, extended in the continuum-suppression output range $C_{\text{FBDT}} > 0.9$, and correct the helicity-angle model of the $B\overline{B}$ component using sideband data, following the findings in Sec. 6.3.2. The fit runs for 5 minutes with `HESSE` and for about one hour with `MINOS` on a Xeon E5-2640 v4 2.4 GHz CPU. Central values are undisclosed because the analysis is still under internal Belle II review. The results in terms of statistical uncertainties are

$$\begin{aligned} \mathcal{B}(B^+ \to \rho^+\rho^0) &= ([\text{xx}]^{+1.40}_{-1.39}) \times 10^{-6}, \\ f_L &= [\text{xx}]^{+0.017}_{-0.015}, \\ A_{CP} &= [\text{xx}] \pm 0.062. \end{aligned}$$

Table 7.1 summarizes the estimated correlations.

Full projections onto the fit observables are shown in Figs. 7.1–7.2. The sample is dominated by large continuum and $B\overline{B}$ background components. As expected, a signal peak is observed in the $\Delta E \approx 0$ region and in the dipion mass distributions. The C'_{FBDT} distribution shows separation between the continuum and the other components. All distributions are adequately modeled by the total fit function with the exception of the $\cos\theta_{\rho^0}$ distribution in both charges, as expected. For B^+ candidates the fit undershoots the data in the whole $\cos\theta_{\rho^0}>0.2$ range, and overshoots them elsewhere, while for B^- candidates the distribution of the discrepancy is similar, but less pronounced.

The observed mismodeling are sufficiently serious to prompt a thorough investigation of possible causes before proceeding with full unblinding of the data. The tests performed in Sec. 6.3 disfavor efficiency or sample-composition causes, suggesting that this mismodeling may be due to interference effects, not included in the simulation, between the signal and charmless B mesons decaying in the same final state through different intermediate states. Such effects were not reported in previous measurements of $B^+ \to \rho^+\rho^0$ decays even though sensitivities were similar to our analysis. The original strategy of this work was to improve over previous measurements by considering interference as a small perturbation through the systematic uncertainties, as supported by qualitative physics considerations discussed in the previous chapter. The size of the observed mismodelings indicates that this is unlikely to be a robust approach and a more detailed understanding of the causes is needed. Further studies on how to properly account for these effects are currently ongoing in the context of internal collaboration review and are not included in this thesis.

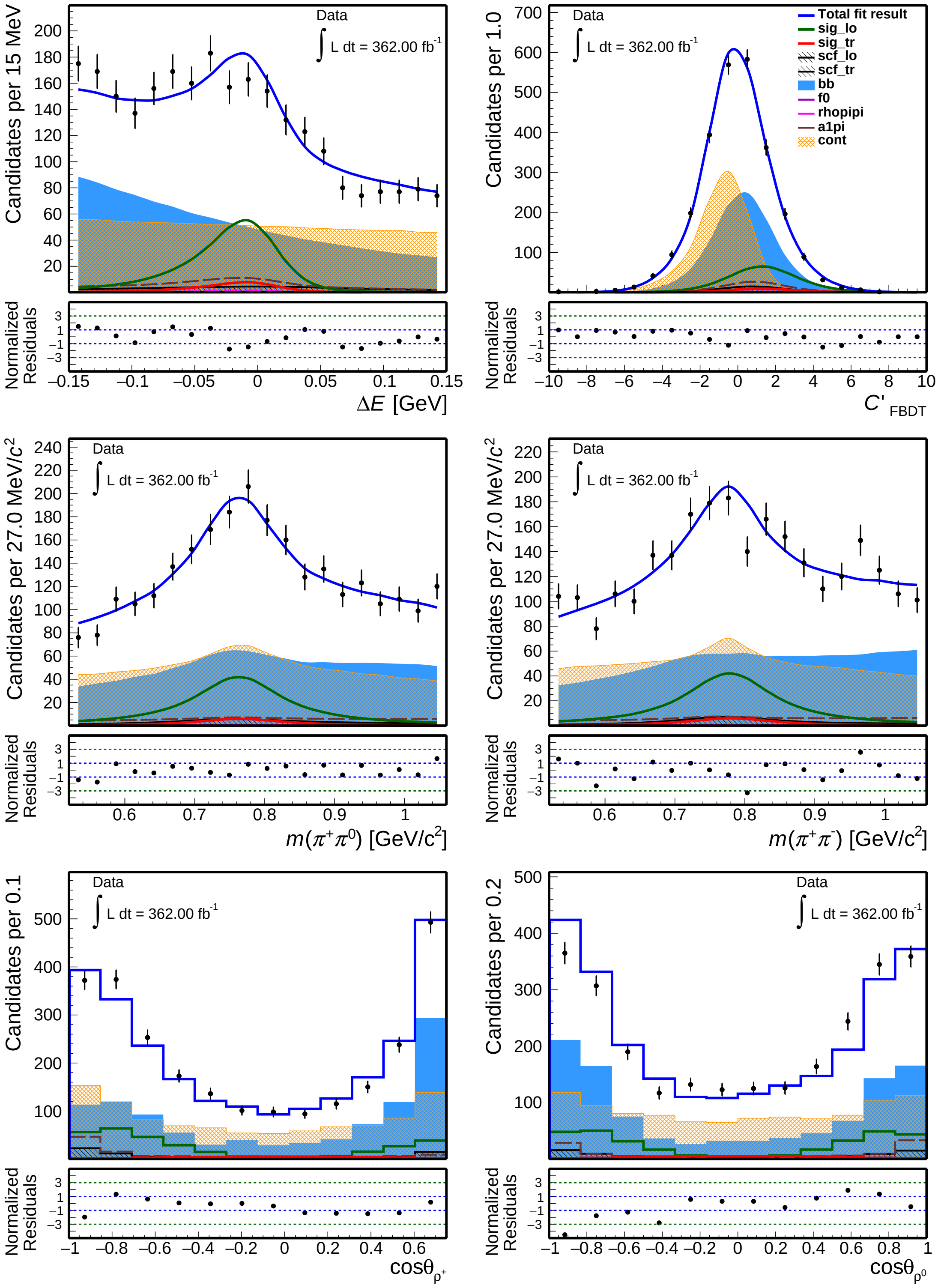


Figure 7.1: Distributions of (top left) ΔE, (top right) transformed continuum suppression output, (middle left) $m(\pi^-\pi^0)$, (middle right) $m(\pi^+\pi^-)$, (bottom left) $\cos\theta_{\rho^-}$, and (bottom right) $\cos\theta_{\rho^0}$ for $B^- \to \rho^-\rho^0$ candidates reconstructed in the Belle II data set collected at the $\Upsilon(4S)$ resonance and corresponding to 387 million $B\overline{B}$ pairs. Fit projections are overlaid.

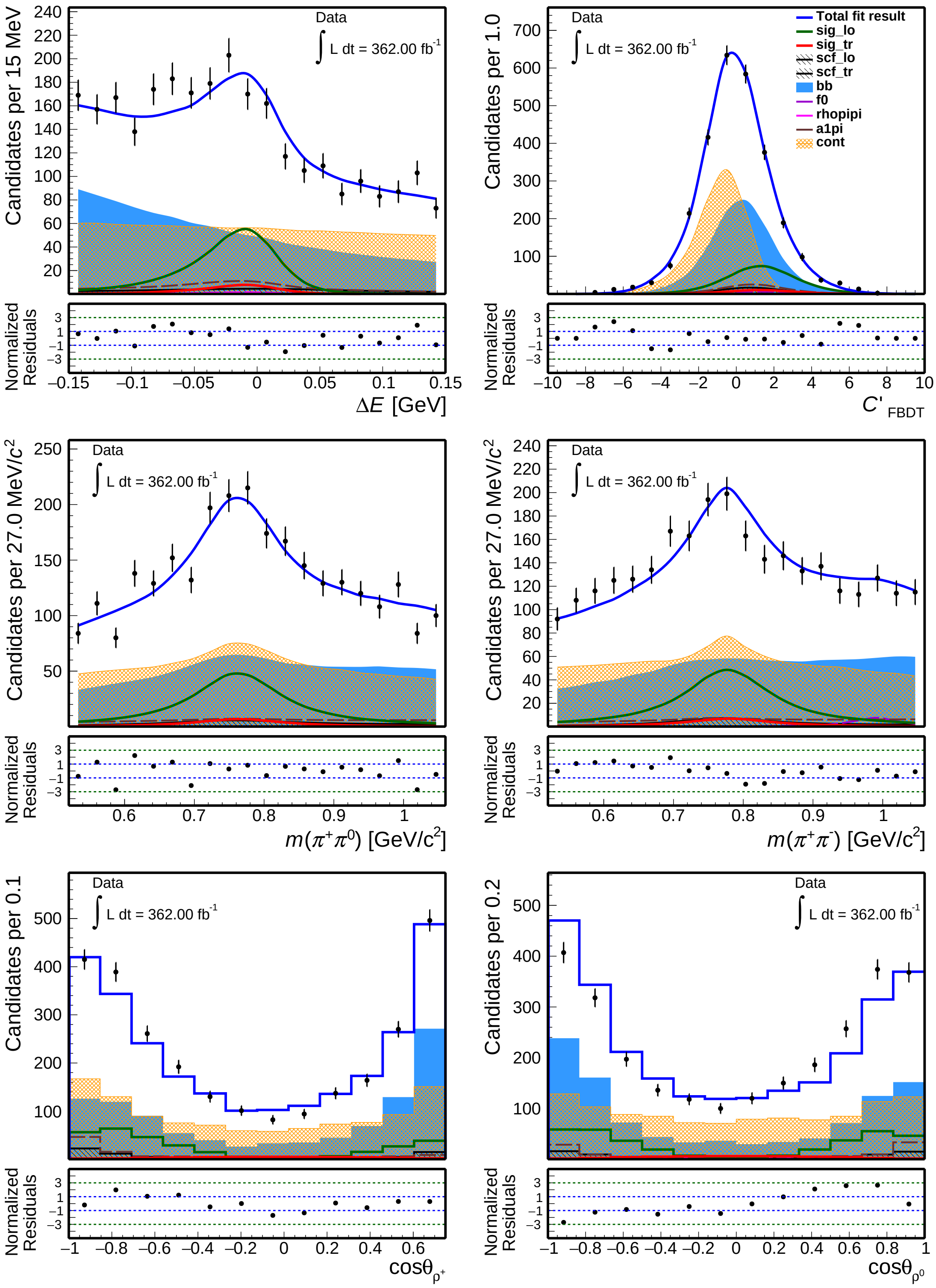


Figure 7.2: Distributions of (top left) ΔE, (top right) transformed continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $m(\pi^+\pi^-)$, (bottom left) $\cos\theta_{\rho^+}$, and (bottom right) $\cos\theta_{\rho^0}$ for $B^+ \to \rho^+\rho^0$ candidates reconstructed in the Belle II data set collected at the $\Upsilon(4S)$ resonance and corresponding to 387 million $B\overline{B}$ pairs. Fit projections are overlaid.

	$\mathcal{B}(B^+ \to \rho^+\rho^0)$	f_L	A_{CP}	$r_{\mathrm{sxf}\ (\rho^+\rho^0)_L}$	$r_{\mathrm{sxf}\ (\rho^+\rho^0)_T}$	$N_{B\overline{B}}$	$N_{f_0\pi\pi}$	$N_{\rho\pi\pi}$	$N_{a_1\pi}$	$A_{CP\ f_0\pi\pi}$	$A_{CP\ \rho\pi\pi}$	$A_{CP\ a_1\pi}$
$\mathcal{B}(B^+ \to \rho^+\rho^0)$	1	+7.2%	-1.4%	-0.9%	-0.1%	-35.2%	-7.8%	-20.7%	-15.4%	-7.1%	+10.5%	-0.7%
f_L		1	+3.9%	-0.4%	+0.1%	-11.0%	+1.4%	+32.0%	-3.9%	+5.9%	-16.9%	+4.3%
A_{CP}			1	<0.1%	<0.1%	-6.0%	-2.8%	+20.6%	+0.8%	+5.7%	+29.1%	-29.0%
$r_{\mathrm{sxf}\ (\rho^+\rho^0)_L}$				1	<0.1%	-0.6%	+0.1%	+0.1%	-0.3%	+0.1%	<0.1%	<0.1%
$r_{\mathrm{sxf}\ (\rho^+\rho^0)_T}$					1	-0.1%	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%
$N_{B\overline{B}}$						1	-13.1%	-9.8%	-21.7%	-9.6%	+3.3%	-0.6%
$N_{f_0\pi\pi}$							1	-14.0%	-7.3%	+57.4%	+7.7%	-2.2%
$N_{\rho\pi\pi}$								1	-5.4%	+5.6%	-47.6%	+10.6%
$N_{a_1\pi}$									1	-4.0%	+2.0%	-9.2%
$A_{CP\ f_0\pi\pi}$										1	+20.8%	-11.1%
$A_{CP\ \rho\pi\pi}$											1	+12.4%
$A_{CP\ a_1\pi}$												1

Table 7.1: Full correlation matrix obtained from the fit to data.

Chapter 8

Systematic uncertainties

This chapter describes the determination of the systematic uncertainties.

8.1 Introduction

I consider various sources of systematic uncertainties associated with assumptions made in the analysis or intrinsic uncertainties of external analysis inputs. Whenever a systematic uncertainty is associated to an arbitrary choice in the analysis, I evaluate its impact by using simplified simulated experiments that realize different incarnations for that choice and repeat the analysis on them using a model implementing our default choice. The distance between the averages of the results is used as systematic uncertainty as it avoids double counting of statistical fluctuations. Since the central values of our results are hidden, systematic uncertainties are reported as fractional contributions relative to the current world-average central values.

As discussed in Sec. 6.4, the set of systematic uncertainty presented in this chapter is incomplete, since an improved treatment of the interference effects is still in progress. The partial systematic uncertainty, including an evaluation based on the original strategy planned to account for interference effects, is indicated in the final results in brackets.

8.2 Charged-particle reconstruction efficiency

Since there are three tracks in the $B^+ \to \rho^+\rho^0$ final state, any uncertainty on track-finding efficiency, which is determined from simulation, is a source of systematic uncertainty in the measurement of branching ratio.

The track-finding efficiency is determined in Belle II by using $e^+e^- \to \tau^+\tau^-$ collision events, where one τ lepton decays as $\tau^+ \to \nu_\tau \ell^+ \bar{\nu}_\ell$, $(\ell = e, \mu)$ and the other one as $\tau^+ \to \nu_\tau \pi^+\pi^-\pi^+$ (charge-conjugated decays are reconstructed too). First, three-track candidates (lepton and two pions) are reconstructed, and if a fourth track is found in the event, that is combined with the others. The efficiency is measured using the numbers of signal events found in the two disjoint samples where either the fourth track is found or not. The procedure is performed in data and simulation, to determine a data-simulation efficiency ratio of $\varepsilon_{data}/\varepsilon_{MC} = 0.061 \pm 0.027(\text{stat}) \pm 0.231(\text{syst})$. Since this correction factor is consistent with unity, the track finding efficiency is not included as a correction to the calculation of the $B^+ \to \rho^+\rho^0$ branching fraction, but its uncertainty of 0.24% per track contributes to the systematic uncertainty. The uncertainties are added linearly resulting in a total systematic uncertainty of 0.72%.

8.2.1 Instrumental tracking asymmetry

To obtain the A_{CP} measurement, I fit separately positively and negatively charged $B^+ \to \rho^+\rho^0$ candidates. Any possible charge-dependent difference in selection or reconstruction efficiency could result in a spurious A_{CP} result. To assign a systematic uncertainty due to possible charge-asymmetry in the track-finding efficiency, I use the asymmetry determined in Belle II by using unbiased samples of tracks from $D^+ \to K^0_S\pi^+$ decays. The CP asymmetry of this channel is known [10], and any deviation from that value is attributed to instrumental effects.

The measured instrumental asymmetry is $A_{\text{det}} = -0.0078 \pm 0.0151$, compatible with zero. The 1.51% uncertainty of the asymmetry is included as systematic uncertainty on A_{CP}.

8.3 $B\overline{B}$ pair counting

For the calculation of the branching fraction, the number of $B\overline{B}$ pairs $(387{\pm}6){\times}10^6$ collected by Belle II is used, as this is more precise than the product of luminosity and production cross-section. The number of $B\overline{B}$ pairs is evaluated by counting the total number of events reconstructed at the $\Upsilon(4\text{S})$ collision energy (on-resonance) that pass the requirements described in Sec. 4.1.4. The same counting is performed also in off-resonance collisions, and then scaled to the on-resonance sample size. The number of $B\overline{B}$ pairs is obtained with the difference between these two counts. I assign an uncertainty of 3.14%, which results from the propagation of the uncertainties on the $\Upsilon(4\text{S})$ [77] branching fractions and on the $B\overline{B}$ pair counting.

8.4 Charged-particle identification and continuum-suppression selection efficiency

Selections based on charged-particle identification and continuum suppression are essential parts of the analysis. These introduce an efficiency that is estimated from simulation. Differences between the particle identification and continuum suppression distributions in data and simulation exist and need to be accounted for to derive corrections, if needed, and a systematic uncertainty.

I use $B^+ \to \bar{D}^0(\to K^+\pi^-)\pi^+$ decays to evaluate such efficiencies in data and simulation. I reconstruct the channel with the same requirements as for the signal, and then divide the resulting sample into two disjoint subsamples, made of the B^+ candidates that either pass or fail the continuum suppression and particle identification requirements. I simultaneously fit the ΔE distributions of the two disjoint samples, and define the efficiency as $\varepsilon = (N^{\text{sig}}_{\text{pass}})/(N^{\text{sig}}_{\text{pass}} + N^{\text{sig}}_{\text{not pass}})$, where $N^{\text{sig}}_{\text{(not) pass}}$ is the signal yield in the sample of candidates (not) passing the selection. This is done in simulation and data, and the resulting ratio of efficiencies is $R_{\text{data/MC}} = 1.0007 \pm 0.0007$. This value is compatible with unity, therefore the charged-particle identification and continuum-suppression selection efficiency is not included as a correction to the calculation of the $B^+ \to \rho^+\rho^0$ branching fraction, but its uncertainty of 0.07% is considered as a systematic uncertainty.

8.5 Neutral-pion reconstruction efficiency

The reconstruction and selection of the neutral pion is associated with an efficiency, which is determined from simulation. Belle II studies data-simulation discrepancies using $D^0 \to K^-\pi^+\pi^0$ decays, normalized to the $D^0 \to K^-\pi^+$ channel.

Data-simulation efficiency ratios are provided as functions of $p(\pi^0)$, and are all compatible with unity, therefore no correction to the calculation of the $B^+ \to \rho^+\rho^0$ branching fraction is needed. I evaluate the systematic uncertainty by averaging the ratio uncertainties weighted according to the $p(\pi^0)$ distribution of $B^+ \to \rho^+\rho^0$ signal decays. This results in a 4.51% systematic uncertainty contribution on the branching fraction.

8.6 Candidate selection

In the default analysis, I select one candidate per event by choosing the one with the highest value of vertex probability. The efficiency of this requirement is determined using simulation, and could deviate from the actual value in data. To assign a systematic uncertainty, I repeat the same analysis on the data sample with one candidate per event chosen randomly. I take as systematic uncertainties on the parameters the difference between the (blind) results obtained using the default candidate selection and the ones obtained choosing one random candidate per event, which are respectively 1.59% for $\mathcal{B}(B^+ \to \rho^+\rho^0)$, 0.79% for f_L, and 0.027% for A_{CP}

8.7 Fit model

Many ΔE, C'_{FBDT} and $m(\pi^+\pi^{0,-})$ shapes are described with various peaking analytical functions extracted from simulation. However, simulated samples have limited size, which reflects into model uncertainties. To evaluate the systematic uncertainty associated with this uncertainty, I generate batches of 1000 simplified simulated experiments each with a different model, based on variations of the shape parameters within their uncertainties, for the different components. I then fit the simulated samples with the default fit model. For each component, the variation of the mean of the distributions of the fit results with respect to the default results shown in Sec. 5.3 is taken as contribution to the systematic uncertainty.

An additional source of systematic uncertainty may come from the statistical uncertainty on the determination of the ΔE and $m(\pi^+\pi^0)$ shift parameters. To estimate that, I repeat the fit of the data sample shifting the value of the two parameters by $\pm 1\sigma$. The difference between the fit results with the various shift values is taken as systematic uncertainty and summed in quadrature with the other systematic uncertainty on the fit model.

The systematic uncertainty due to the description of the helicity angles is treated separately.

8.8 Fit bias

I study the estimator properties, as described in Sec. 5.3. The pull distributions show biases of $(4 \pm 3)\%$ for $\mathcal{B}(B^+ \to \rho^+\rho^0)$, $(-8 \pm 3)\%$ for f_L and $(1 \pm 3)\%$ for A_{CP}, which are consitent with zero. I include the uncertainties in the systematic uncertainty after rescaling them according to the statistical uncertainty.

8.9 Interference with four-pion final state B^+ decays

The original strategy for this measurement was to neglect interference effects between B^+ decays into four-pion final states and to asses a systematic uncertainty that accounts for to this assumption. I prepared this work following the measurements performed by the Belle experiment in $B^0 \to \rho^+\rho^-$ and $B^0 \to \rho^0\rho^0$ decays [40, 56].

Using the Belle experiment simulation, I generate sets of simplified simulated samples with various possible intensities and phases of interference between the signal channel and each of possible interfering channels. I assume any interference effect to be dominated by the interference between the longitudinally polarized signal $B^+ \to \rho^+\rho^0$ and one of the most abundant of the interfering channels, $B^+ \to a_1^+\pi^0$. The generator amplitude is modeled as

$$A = A_{\rho^+\rho^0} + Ce^{i\phi}A_{a_1^+\pi^0}, \tag{8.1}$$

where A_j is the amplitude of process j, C is the relative proportion between $B^+ \to \rho^+\rho^0$ and $B^+ \to a_1^+\pi^0$ yields, and ϕ is an interference phase. I generate simulated samples that implement interference by injecting various values of C and ϕ. Simulated samples include only $B^+ \to \rho^+\rho^0$ and $B^+ \to a_1^+\pi^0$ decays and are fitted with a simplified two-component four-dimensional model that only includes dipion masses and cosines of helicity angles and does not take interference into account, as our default fit.

Table 8.1 summarizes the results. From a comparison of the results obtained for a certain C value, the maximal variations of 1.2% for $\mathcal{B}(B^+ \to \rho^+\rho^0)$ and < 0.001 for f_L are included in the systematic uncertainty. However, given the findings discussed in the previous sections, the above treatment is probably a too simplistic approximation for the sample at hand, which will become outdated once a more detailed understanding of the causes of our mismodeling will be finalized.

C, ϕ	X_{sig}	f_L	$f_{a_1^+\pi^0}$
0.37, 0	386358 ± 2465	0.99989 ± 0.00029	0.0602 ± 0.0061
0.37, $\pi/2$	386123 ± 2471	0.99998 ± 0.00023	0.0645 ± 0.0060
0.37, π	390781 ± 2493	1.00060 ± 0.00036	0.0600 ± 0.0061
0.4, 0	375716 ± 2440	0.99953 ± 0.00036	0.06484 ± 0.0061
0.4, $\pi/2$	377454 ± 2458	0.99947 ± 0.00027	0.0730 ± 0.0061
0.4, π	372774 ± 2434	1.00007 ± 0.00030	0.0733 ± 0.0061
0.42, 0	361934 ± 2406	1.00015 ± 0.00029	0.0783 ± 0.0062
0.42, $\pi/2$	362665 ± 2411	0.99994 ± 0.00034	0.0807 ± 0.0062
0.42, π	366276 ± 2413	0.99942 ± 0.00033	0.0731 ± 0.0062

Table 8.1: Results of the interference study. C is the relative proportion between $B^+ \to \rho^+\rho^0$ and $B^+ \to a_1^+\pi^0$ yields used in the generator, ϕ is the interference phase, $X_{\text{sig}} = N_L/\varepsilon_L + N_T/\varepsilon_T$ is the numerator of the branching fraction expression (Eq. 5.2), $f_{a_1^+\pi^0}$ is the fraction of $B^+ \to a_1^+\pi^0$ decays extracted from the fit.

8.10 Summary of systematic uncertainties

Table 8.2 summarizes the partial systematic uncertainties for branching fraction, fraction of longitudinally polarized events, and *CP* asymmetry and compares them with the statistical uncertainty. As the central values of parameters of interest are kept hidden, systematic uncertainties are reported in fractional form relative to the world-average values [10]. As discussed in previous sections, these uncertainties have only an indicative value since they are subject to significant changes once a detailed understanding of the angular mismodeling will be available.

Source	$\mathcal{B}$	f_{L}	A_{CP}
Tracking	0.72%	-	-
Instrumental asymmetry correction	-	-	0.015
$N_{B\overline{B}}$	3.14%	-	-
PID and continuum-supp. eff.	0.07%	-	-
π^0 efficiency	4.51%	-	-
Best candidate selection	1.59%	0.79%	0.037
Fit model	0.16%	0.09%	0.005
Fit biases	0.23%	0.15%	0.0006
4π final-states inteference	1.2%	0.1%	-
Total	1.39×10^{-6} (5.8%)	0.008 (0.8%)	0.041
Statistical uncertainty	${}^{+1.40}_{-1.39}\times10^{-6}$ (5.8%)	0.017 (1.8%)	0.062

Table 8.2: Summary of systematic uncertainties. As the central values are hidden, systematic uncertainties are expressed as fractional percentages of the world-average values [10].

Chapter 9

Results

By combining the fit results with systematic uncertainties, I report the measurement of the $B^+ \to \rho^+\rho^0$ branching fraction,

$$\mathcal{B}(B^+ \to \rho^+\rho^0) \;=\; ([\mathrm{xx}]^{+1.40}_{-1.39}\ (\mathrm{stat}) \pm [1.39]\ (\mathrm{syst})) \times 10^{-6},$$

the fraction of longitudinally polarized $B^+ \to \rho^+\rho^0$ decays,

$$f_L \;=\; [\mathrm{xx}]^{+0.017}_{-0.015}\ (\mathrm{stat}) \pm [0.008]\ (\mathrm{syst}),$$

and the CP-violating charge-asymmetry,

$$A_{CP} \;=\; [\mathrm{xx}] \pm 0.062\ (\mathrm{stat}) \pm [0.041]\ (\mathrm{syst}).$$

Central values are hidden as the analysis is still under internal Belle II review. The systematic uncertainties are preliminary as they are subject to change once a more thorough understanding and correction of the angular mismodeling will be finalized. The current systematic uncertainties are (marginally) larger than the statistical uncertainty only for the $\mathcal{B}(B^+ \to \rho^+\rho^0)$ result, and dominated by the uncertainty associated with the neutral-pion reconstruction efficiency. The dominant contributions to the f_L and A_{CP} systematic uncertainties are associated with the candidate selection.

Table 9.1 shows comparison of these results with existing Belle [41] and BaBar [42] measurements. I also project the Belle results (from 2003) to the Belle II luminosity by applying a $\sqrt{85/387}$ factor to the statistical uncertainties to compare the analysis sensitivity independently of sample size.

The A_{CP} statistical sensitivity is improved by 10% (absolute) with respect to the Belle measurement but inferior to that of the current world-best result (by the BaBar collaboration [42]). The statistical sensitivity on f_L is a factor of two better than at Belle and comparable with the current world-best result. The achieved $\mathcal{B}(B^+ \to \rho^+\rho^0)$ sensitivity is improved by 40% with respect to Belle determination and comparable with the current best results. While uncertainties might change following the internal Belle II review, we expect this result to offer a substantive contribution to the global knowledge of α.

	Data fit	Belle 2003 [41]	Belle 2003 scaled	BaBar 2009 [42]
$\mathcal{B}$, 10^{-6}	$[\mathrm{xx}]^{+1.40}_{-1.39} \pm [1.39]$	$31.7 \pm 7.1^{+3.8}_{-6.7}$	$31.7 \pm 2.4^{+3.8}_{-6.7}$	$23.7 \pm 1.4 \pm 1.4$
f_L	$[\mathrm{xx}]^{+0.017}_{-0.015} \pm [0.008]$	$0.95 \pm 0.11 \pm 0.02$	$0.950 \pm 0.036 \pm 0.02$	$0.950 \pm 0.015 \pm 0.006$
A_{CP}	$[\mathrm{xx}] \pm 0.062 \pm [0.041]$	$0.00 \pm 0.22 \pm 0.03$	$0.000 \pm 0.073 \pm 0.03$	$-0.054 \pm 0.055 \pm 0.010$

Table 9.1: Comparison of results with previous measurements. The first contribution to the uncertainties is statistical, the second is systematic. The scaled 2003 Belle result only includes the statistical uncertainty.

Summary

This thesis reports on a Belle II measurement of the decay properties of $B^+ \to \rho^+\rho^0$ decays based on a sample of 387 million pairs of bottom-antibottom mesons produced in electron-positron collisions at the SuperKEKB accelerator. The result will contribute to the determination of α, one of the flavor-physics parameter that is currently limiting the sensitivity of some of the most incisive flavor-based tests of the Standard Model.

I measure the branching fraction, charge-parity-violating decay-rate asymmetry, and the fraction of longitudinally polarized $B^+ \to \rho^+(\to \pi^+\pi^0)\rho^0(\to \pi^+\pi^-)$ decays reconstructed in the full Belle II data set collected to date.

The analysis is challenging as the $B^+ \to \rho^+\rho^0$ decay has a relatively low branching fraction, $O(10^{-5})$, background-like final states, broad intermediate resonances, and requires to measure angular distributions of final-state particles. This results in three principal experimental demands.

- **Maximizing the signal-to-background ratio without introducing intractable biases** in the fit that determines the final results. To suppress the dominant background due to lighter quark-pairs, I develop a multivariate discriminator exploiting information about event topology, position, and quality of the reconstructed B-decay vertices, and flavor. In addition, I also address peaking backgrounds from bottom meson production, which are less pervasive in rate than continuum, but similarly harmful as their experimental signatures mirror signal.

- **Determining accurately the sample composition** to distinguish backgrounds from signals, and between the different signal angular-momentum properties. I devise a fit that properly accounts for the multidimensional dependences between the observables achieving a very detailed description of sample composition. For each component, I identify major dependences and ensure that the densities of dependent observables are modeled properly by studying the distributions of estimators on realistically simulated data.

- **Accounting for mismodelings** of the angular distributions. To understand the origin of this effect and limit its impact on the precision of my measurement, I pursue a thorough study based on various control samples reconstructed in collision data.

I achieve the reconstruction of $B^+ \to \rho^+\rho^0$ decays in the full Belle II data set and measure

$$\begin{aligned}
\mathcal{B}(B^+ \to \rho^+\rho^0) &= ([\mathrm{xx}]^{+1.40}_{-1.39}\ (\mathrm{stat}) \pm [1.39]\ (\mathrm{syst})) \times 10^{-6}, \\
f_L &= [\mathrm{xx}]^{+0.017}_{-0.015}\ (\mathrm{stat}) \pm [0.008]\ (\mathrm{syst}), \\
A_{CP} &= [\mathrm{xx}] \pm 0.062\ (\mathrm{stat}) \pm [0.041]\ (\mathrm{syst}).
\end{aligned}$$

The central values are hidden as the analysis is still under internal collaboration review. The systematic uncertainties are preliminary (hence the brackets) as they are subject to change once a more detailed understanding and correction of the causes of the observed angular mismodeling will be achieved. The statistical uncertainties on all of these result show significant improvement over the measurement of Belle, the predecessor experiment. The precisions in $\mathcal{B}(B^+ \to \rho^+\rho^0)$ and f_L are comparable with the current world-best result, even if based on a smaller sample, while the precision on A_{CP} is moderately inferior. A conclusive comparison of the systematic uncertainties is premature at this time, since a detailed understanding of the angular mismodeling is still in progress and has the potential to impact the final systematic precision, in one direction or the other. Based on the current, preliminary knowledge, all systematic uncertainties are significantly improved over those of the Belle measurement and comparable with those of the world-best result on all quantities, except for those on the *CP*-violating asymmetry, where Babar achieves better precision. Once systematic uncertainties associated with the observed angular mismodeling will be finalized, we expect this result to offer a substantive contribution to the global knowledge of α.

Appendices

Appendix A

$B^+ \to \rho^+\rho^0$ modeling

To determine the $B^+ \to \rho^+\rho^0$ sample composition in Belle II data, I perform a fit over the six-dimensional space of fit observables (see Sec. 5.1). To properly account for the multidimensional dependences between the observables, I study for each component the distributions of every observable conditional on each of the others to identify the major dependences. Examples of such distributions are shown in what follows (Figs. A.1–A.8).

Fit observable shapes are modeled from simulation using empirical analytical functions and histogram templates. Figures A.9–A.15 show modeling fit projections for the analytical shapes in all the components.

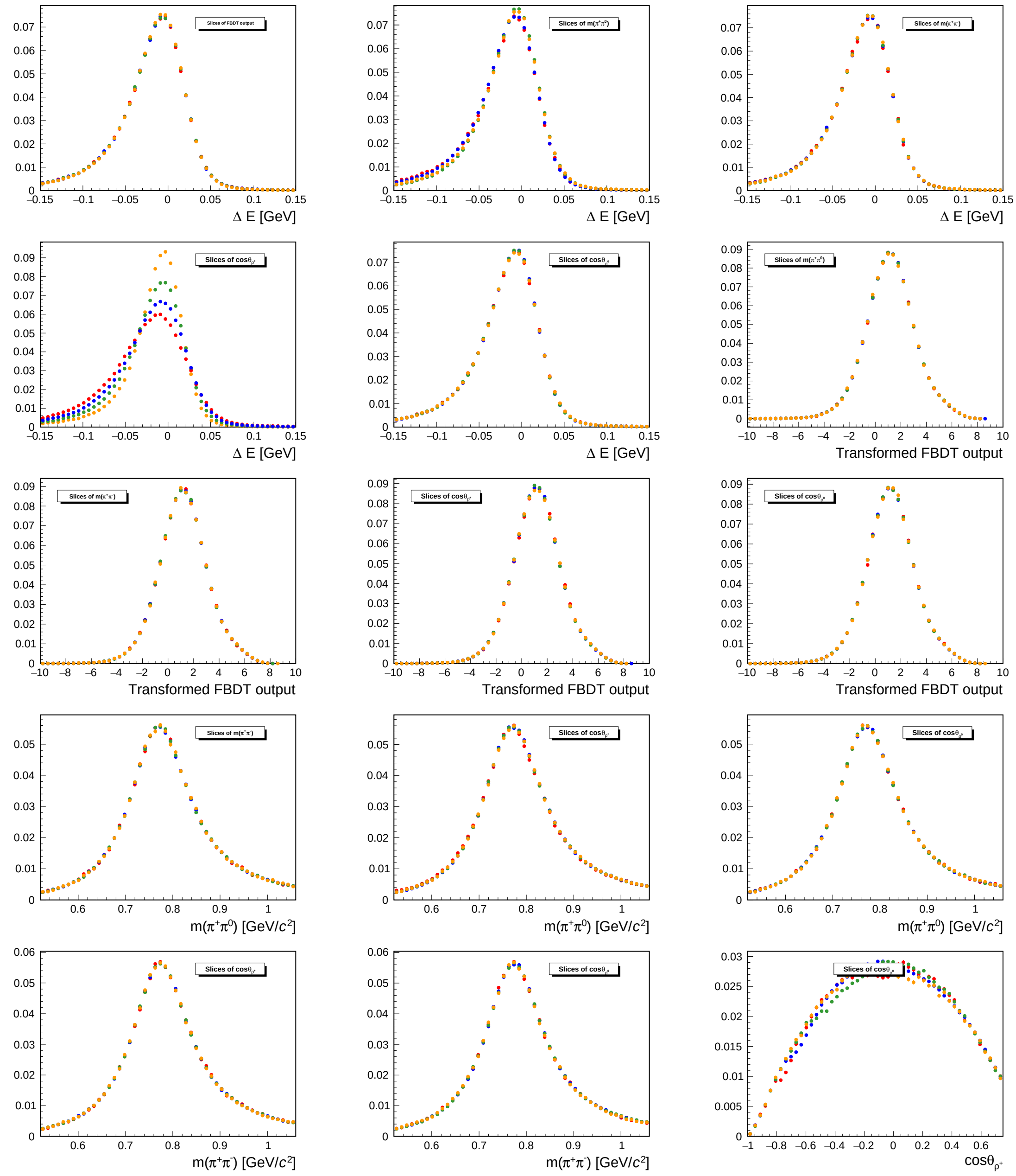


Figure A.1: Distributions of fit observables in slices of other fit observables for simulated transversely polarized signal decays. Distributions normalized to unity.

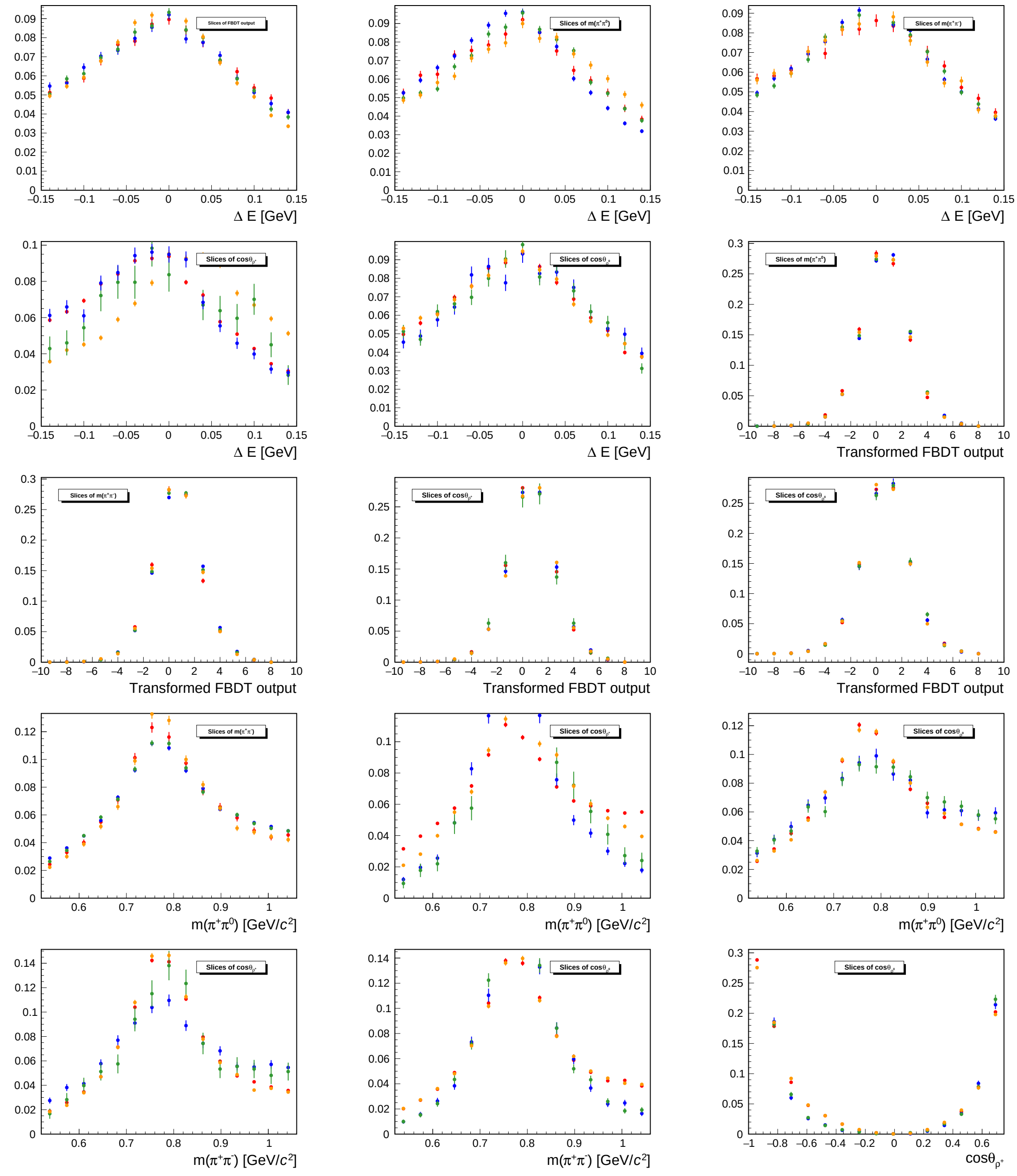


Figure A.2: Distributions of fit observables in slices of other fit observables for simulated longitudinally polarized self-cross-feed events. Distributions normalized to unity.

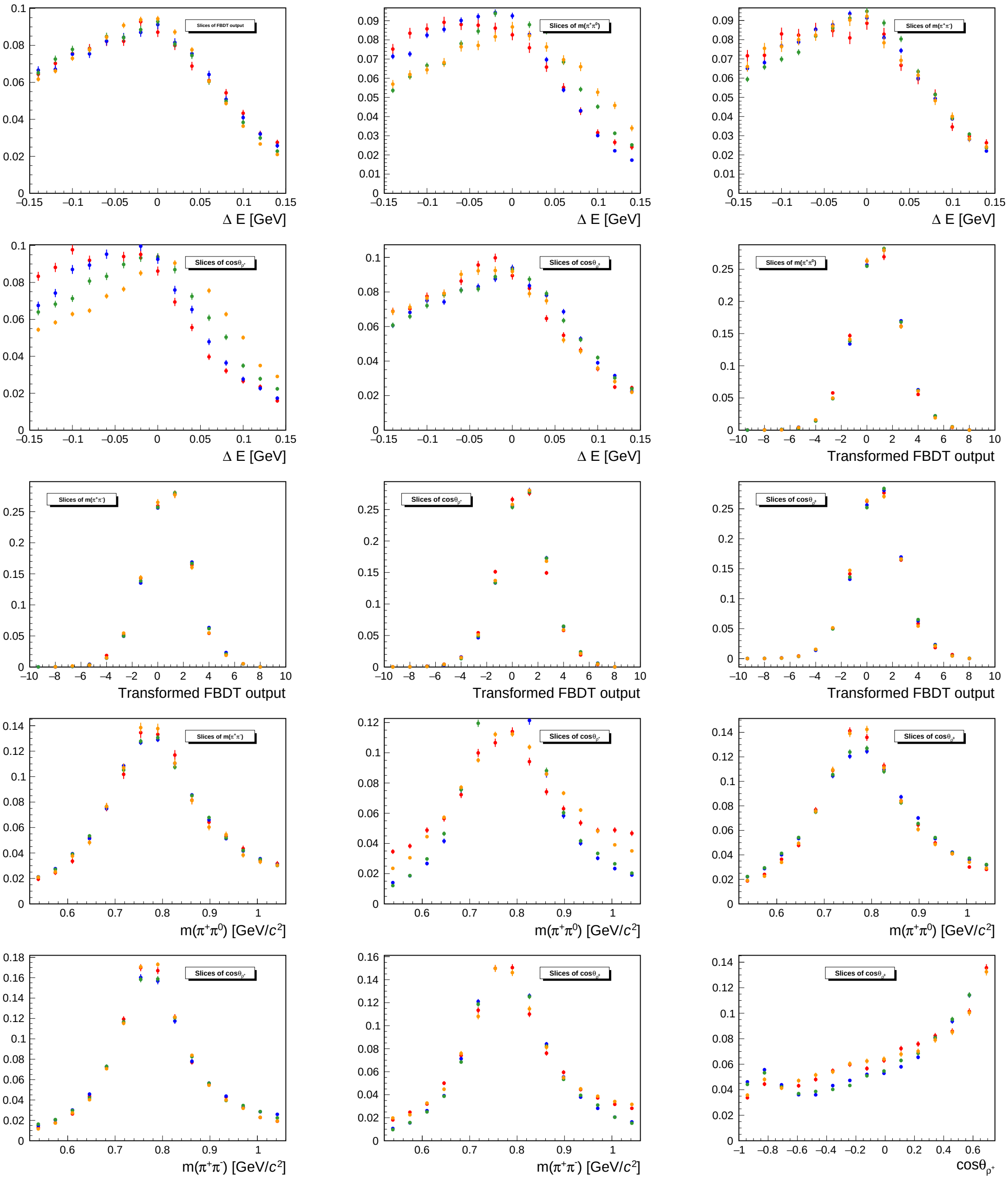


Figure A.3: Distributions of fit observables in slices of other fit observables for simulated transversely polarized self-cross-feed events. Distributions normalized to unity.

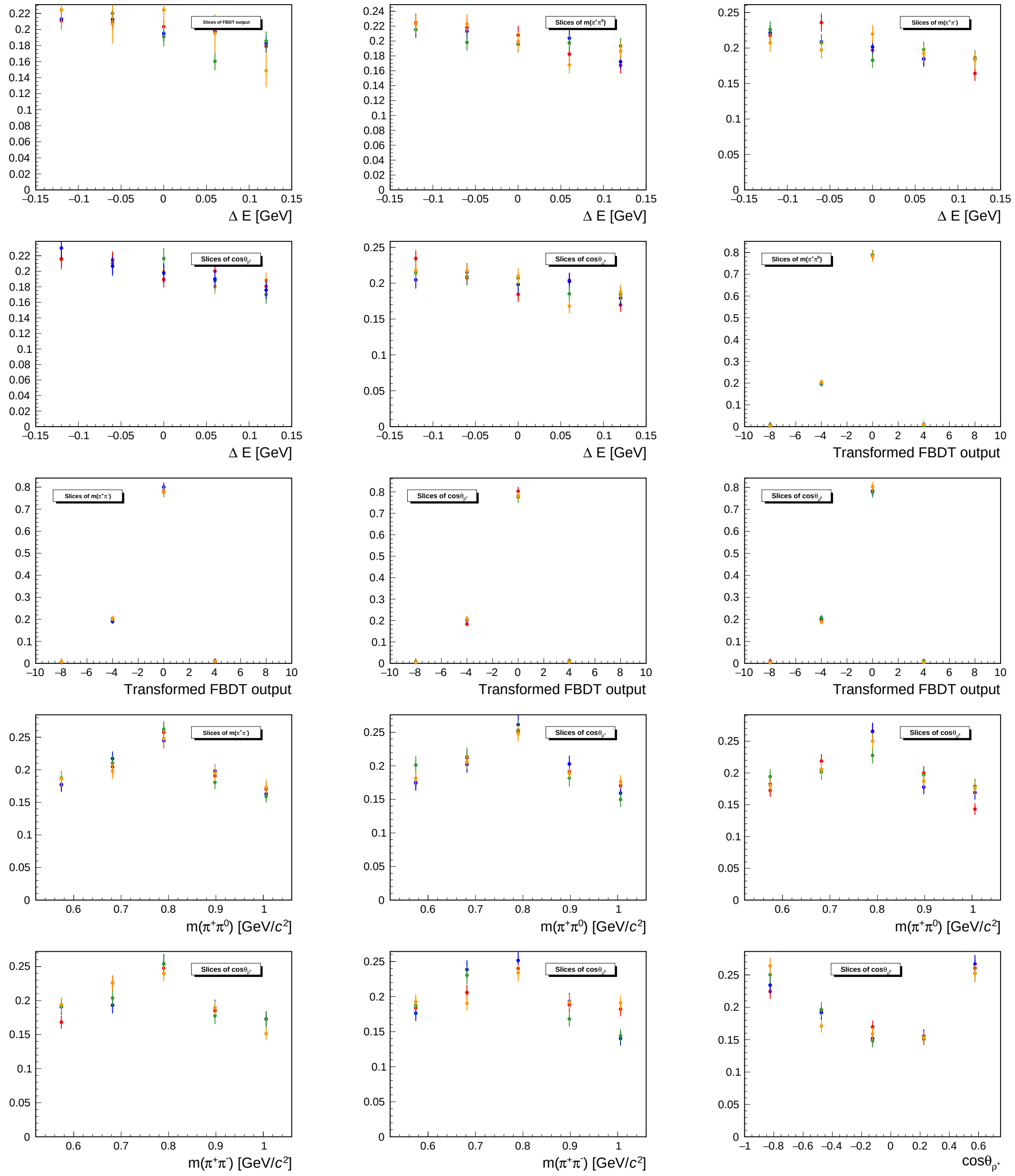


Figure A.4: Distributions of fit observables in slices of other fit observables for simulated continuum events. Distributions normalized to unity.

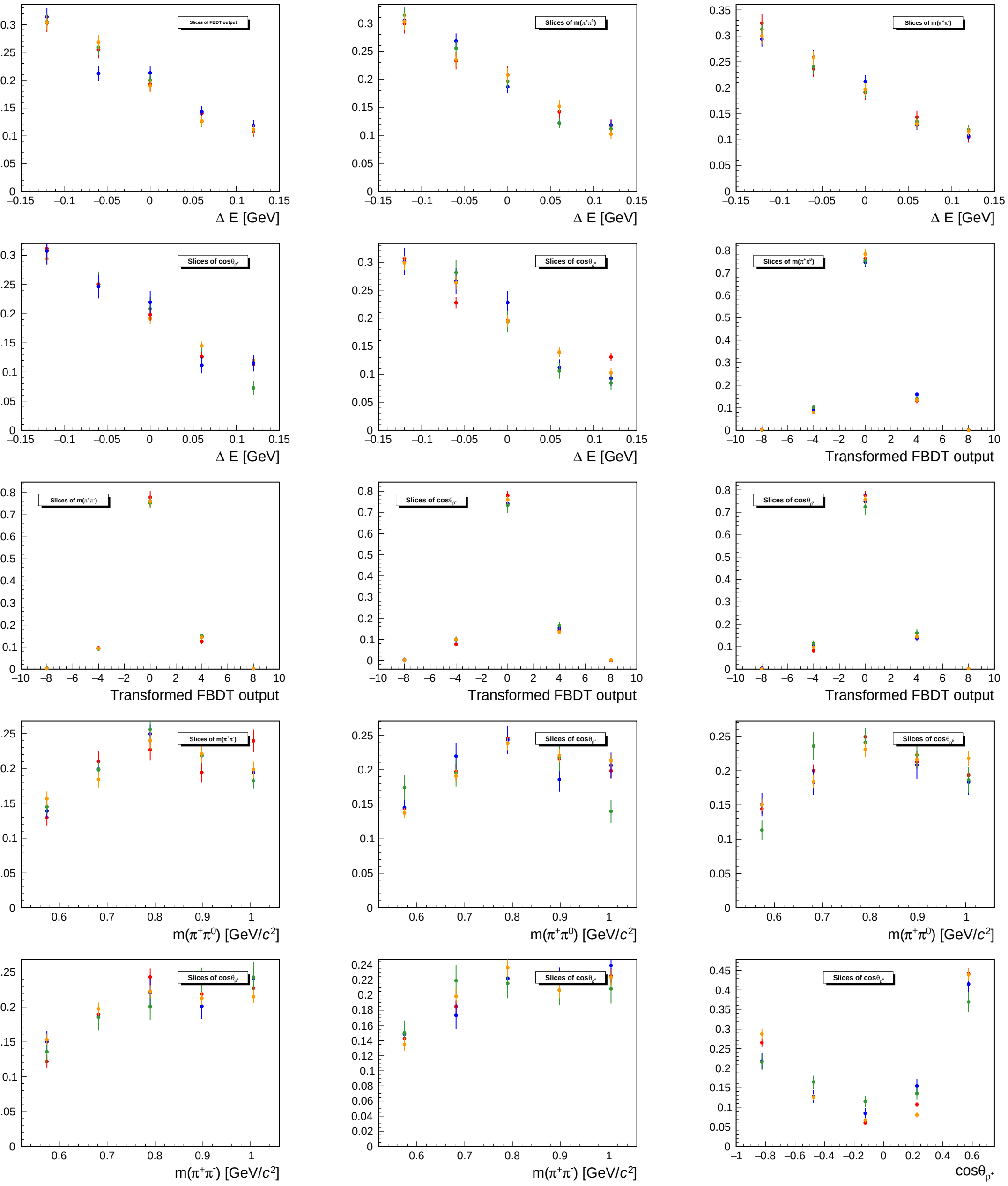


Figure A.5: Distributions of fit observables in slices of other fit observables for simulated $B\bar{B}$ events. Distributions normalized to unity.

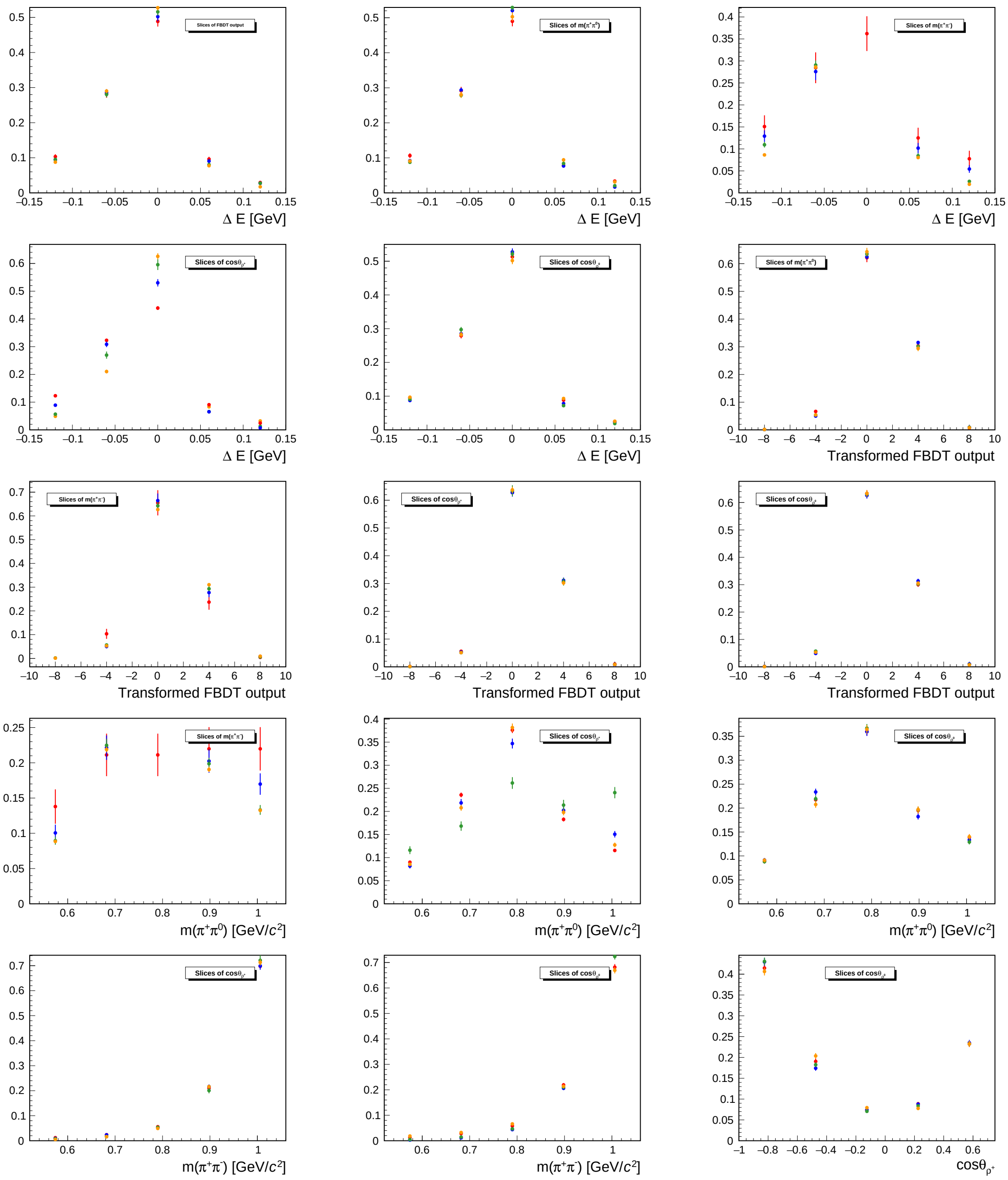


Figure A.6: Distributions of fit observables in slices of other fit observables for simulated $B \to f_0\pi\pi$ decays. Distributions normalized to unity.

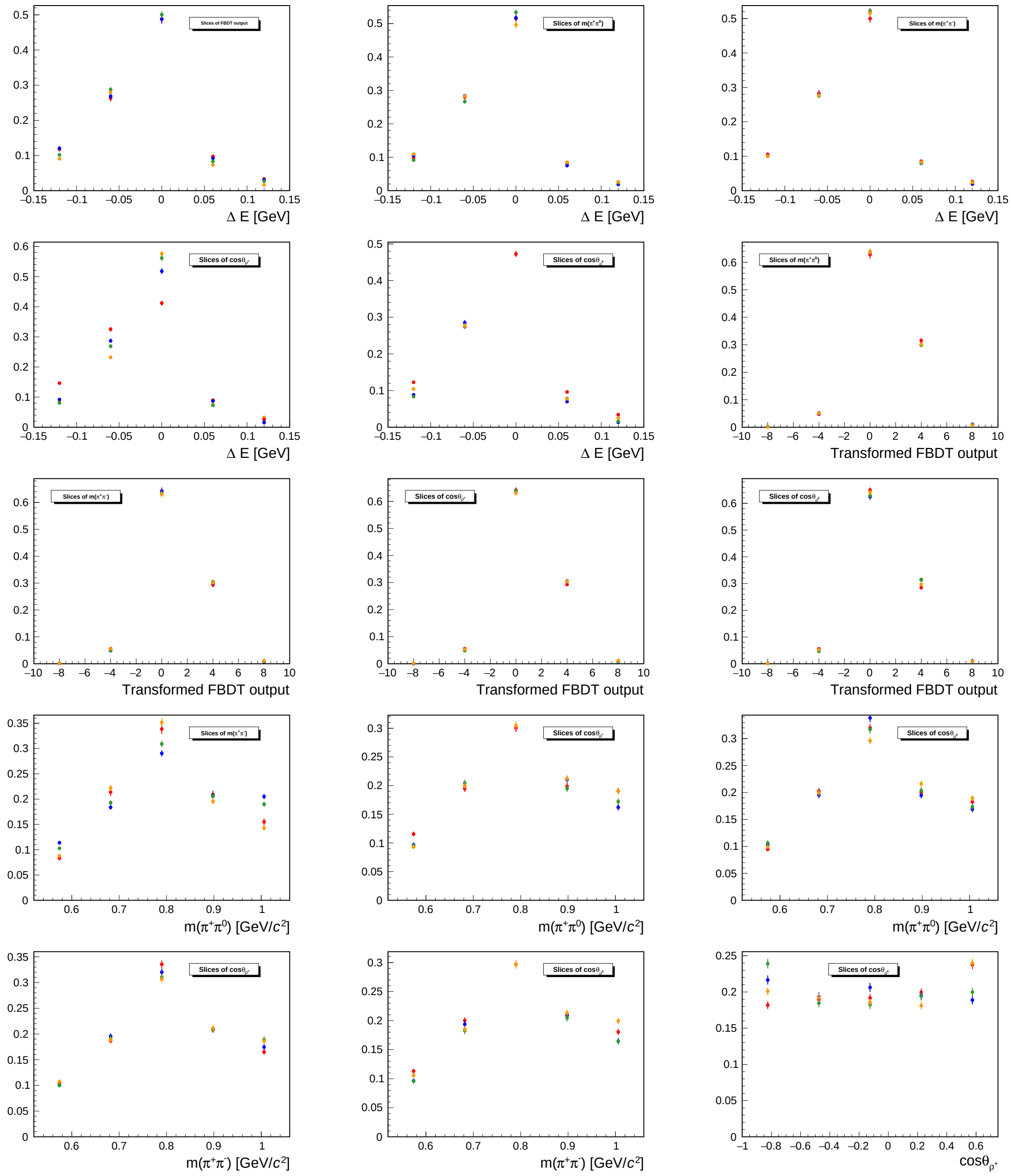


Figure A.7: Distributions of fit observables in slices of other fit observables for simulated $B \to \rho\pi\pi$ decays. Distributions normalized to unity.

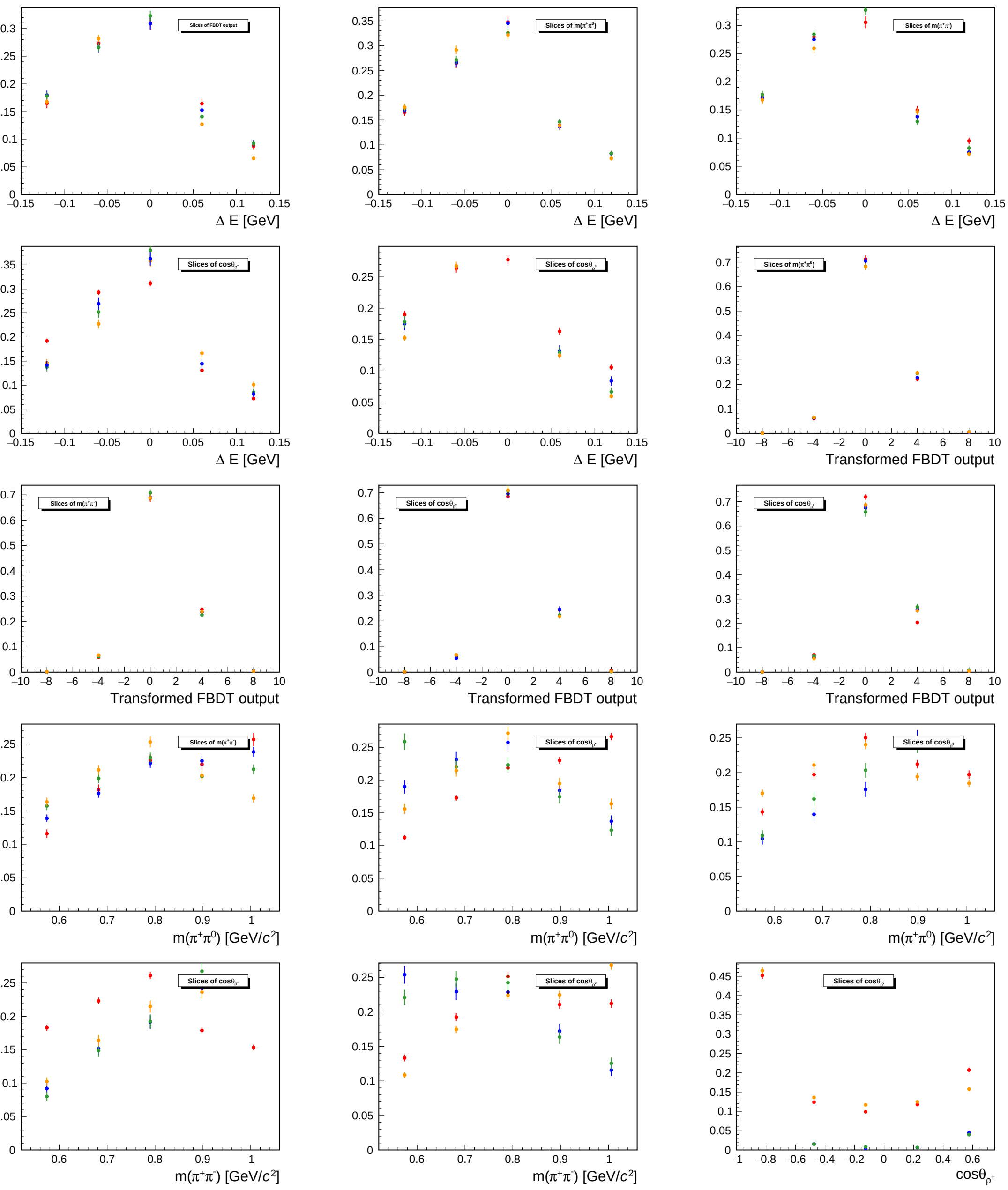


Figure A.8: Distributions of fit observables in slices of other fit observables for simulated $B \to a_1\pi$ decays. Distributions normalized to unity.

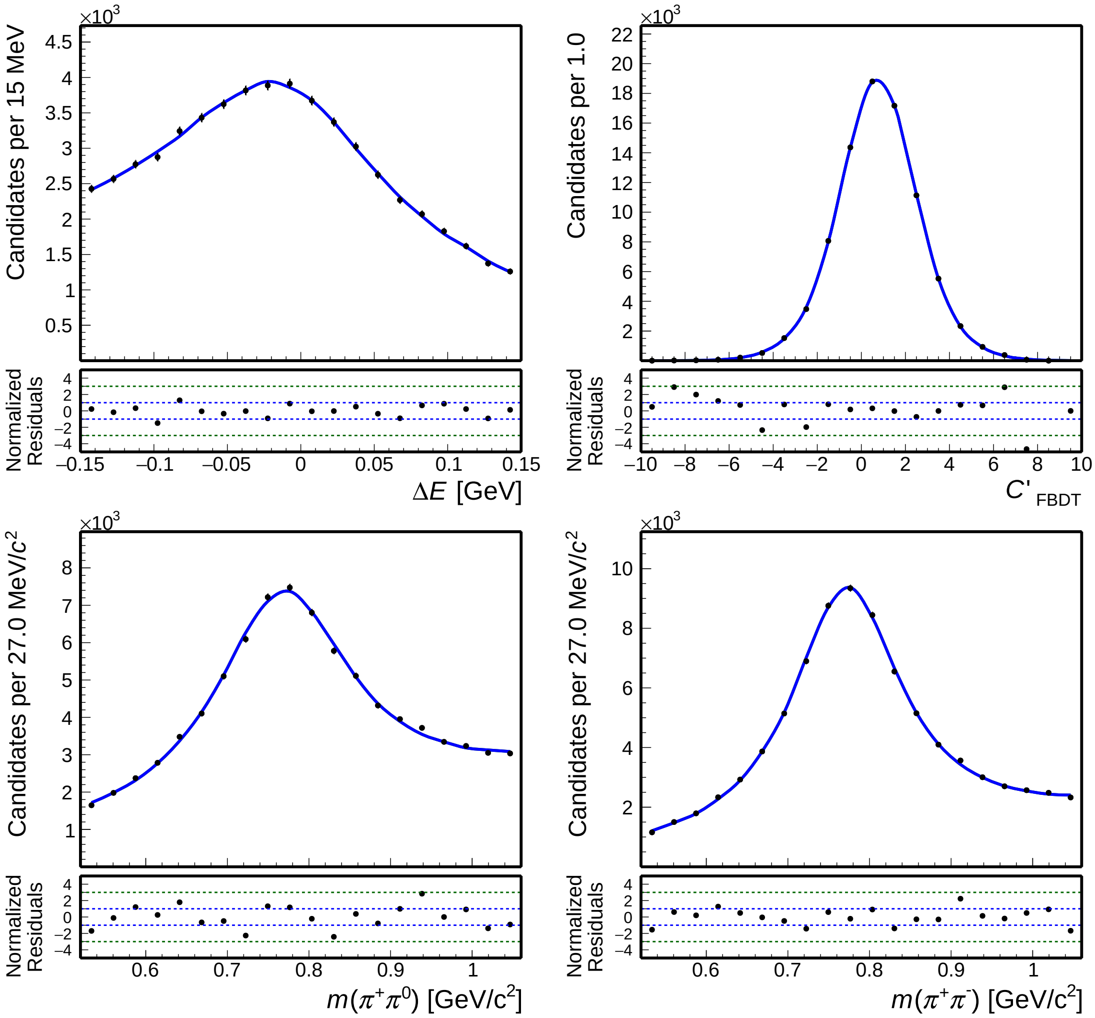


Figure A.9: Distributions of (top left) ΔE, (top right) continuum suppression output, (bottom left) $m(\pi^+\pi^0)$, and (bottom right) $m(\pi^+\pi^-)$ for simulated longitudinally polarized self cross feed. The corresponding model shapes are overlaid (solid).

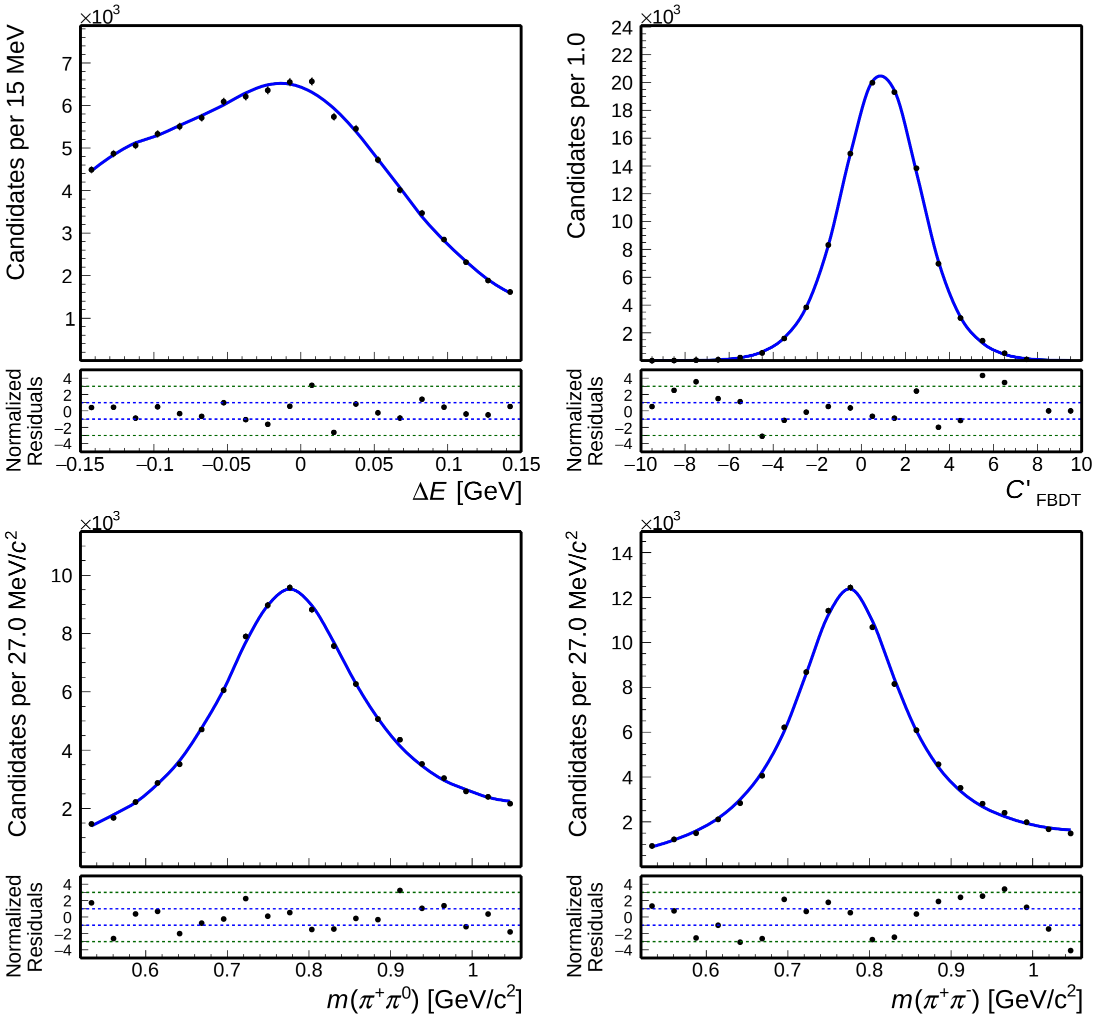


Figure A.10: Distributions of (top left) ΔE, (top right) continuum suppression output, (bottom left) $m(\pi^+\pi^0)$, and (bottom right) $m(\pi^+\pi^-)$ for simulated transversely polarized self cross feed. The corresponding model shapes are overlaid (solid).

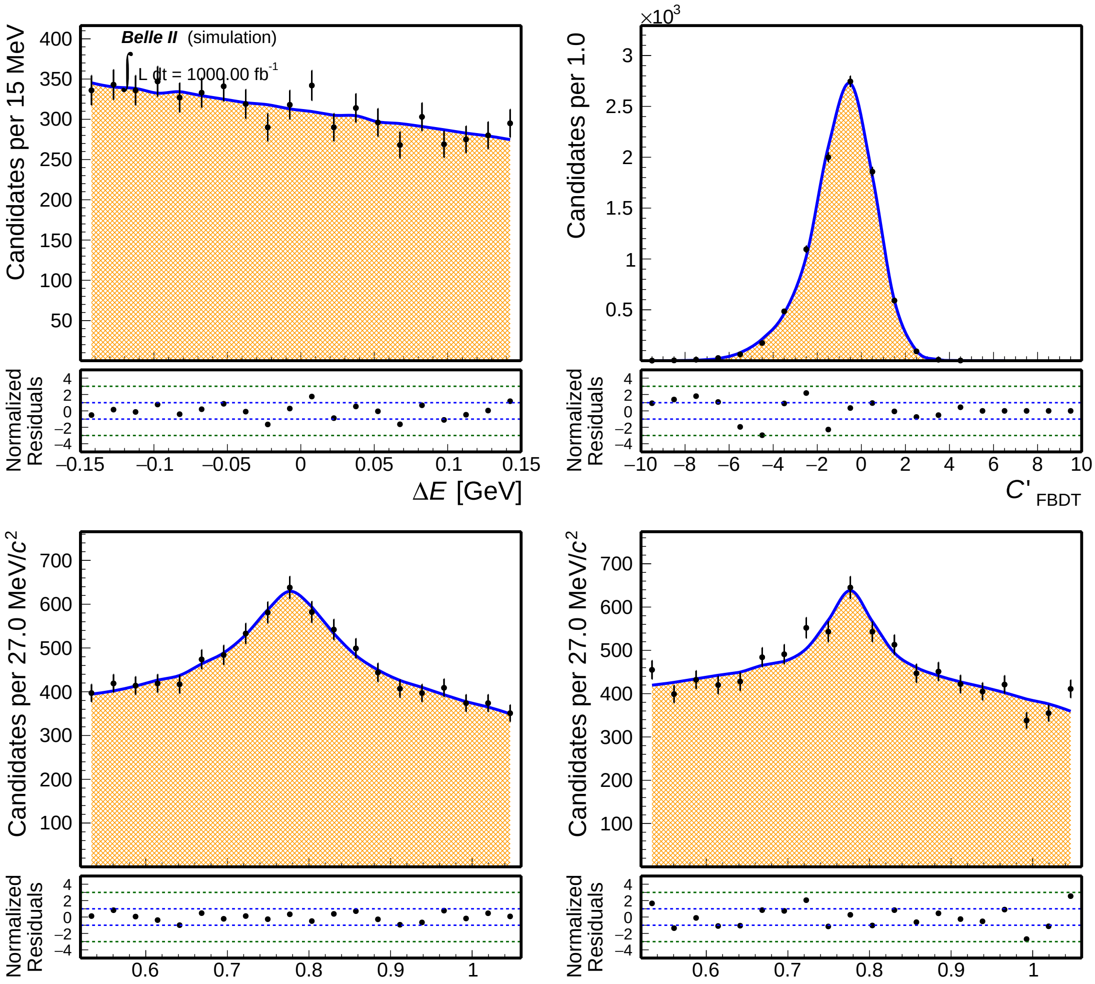


Figure A.11: Distributions of (top left) ΔE, (top right) continuum suppression output, (bottom left) $m(\pi^+\pi^0)$, and (bottom right) $m(\pi^+\pi^-)$ for simulated continuum. The corresponding model shapes are overlaid (solid).

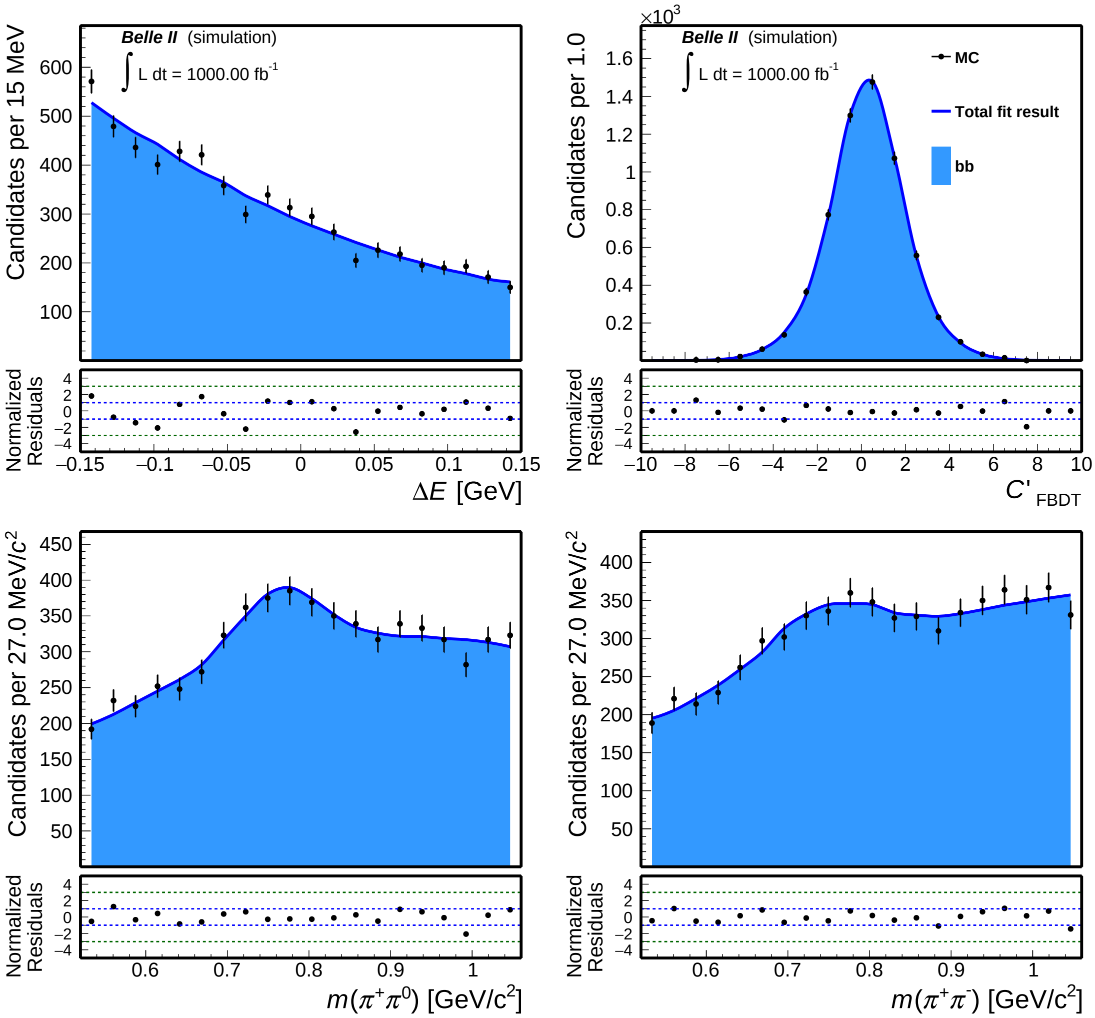


Figure A.12: Distributions of (top left) ΔE, (top right) continuum suppression output, (bottom left) $m(\pi^+\pi^0)$, and (bottom right) $m(\pi^+\pi^-)$ for simulated $B\overline{B}$ background. The corresponding model shapes are overlaid (solid).

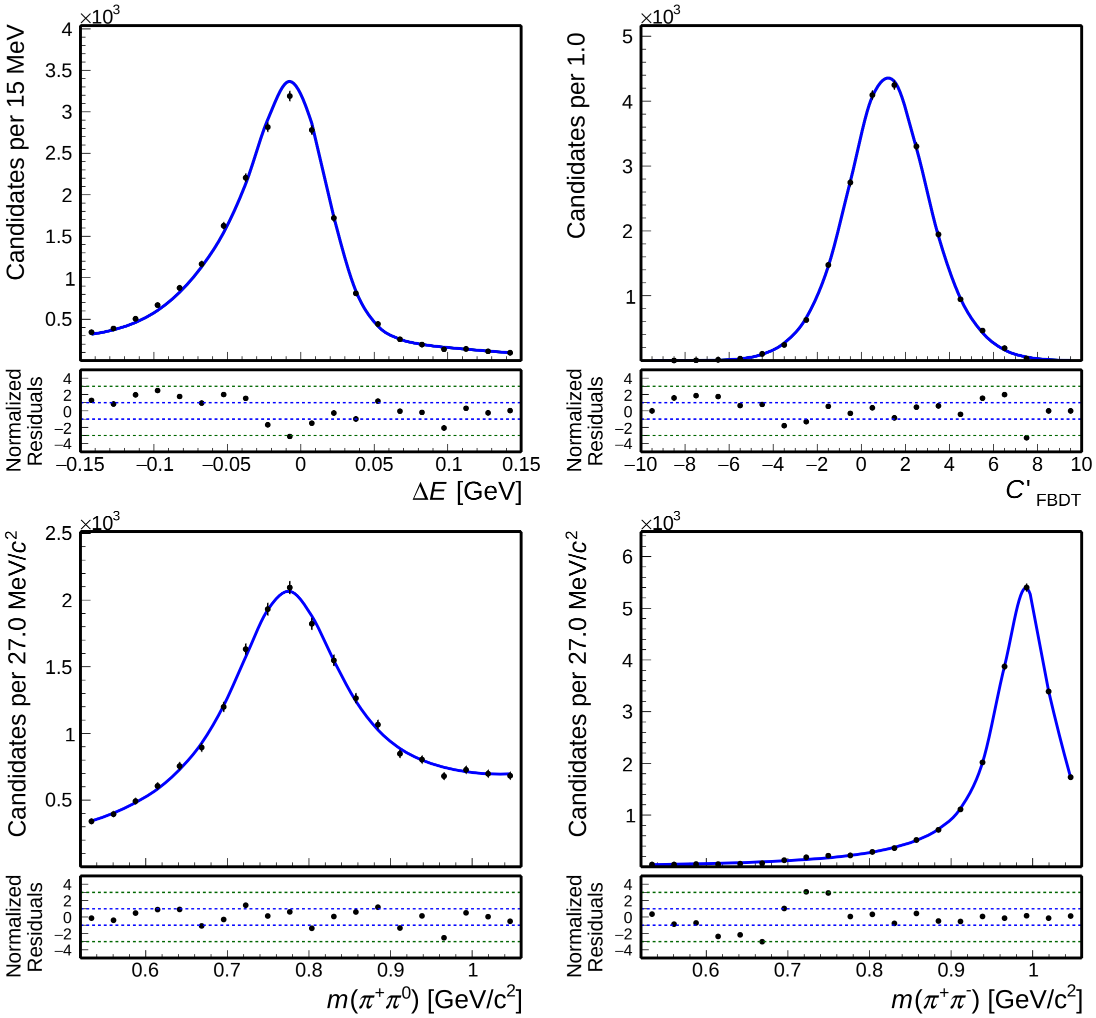


Figure A.13: Distributions of (top left) ΔE, (top right) continuum suppression output, (bottom left) $m(\pi^+\pi^0)$, and (bottom right) $m(\pi^+\pi^-)$ for simulated $B \to f_0\pi\pi$ decays. The corresponding model shapes are overlaid (solid).

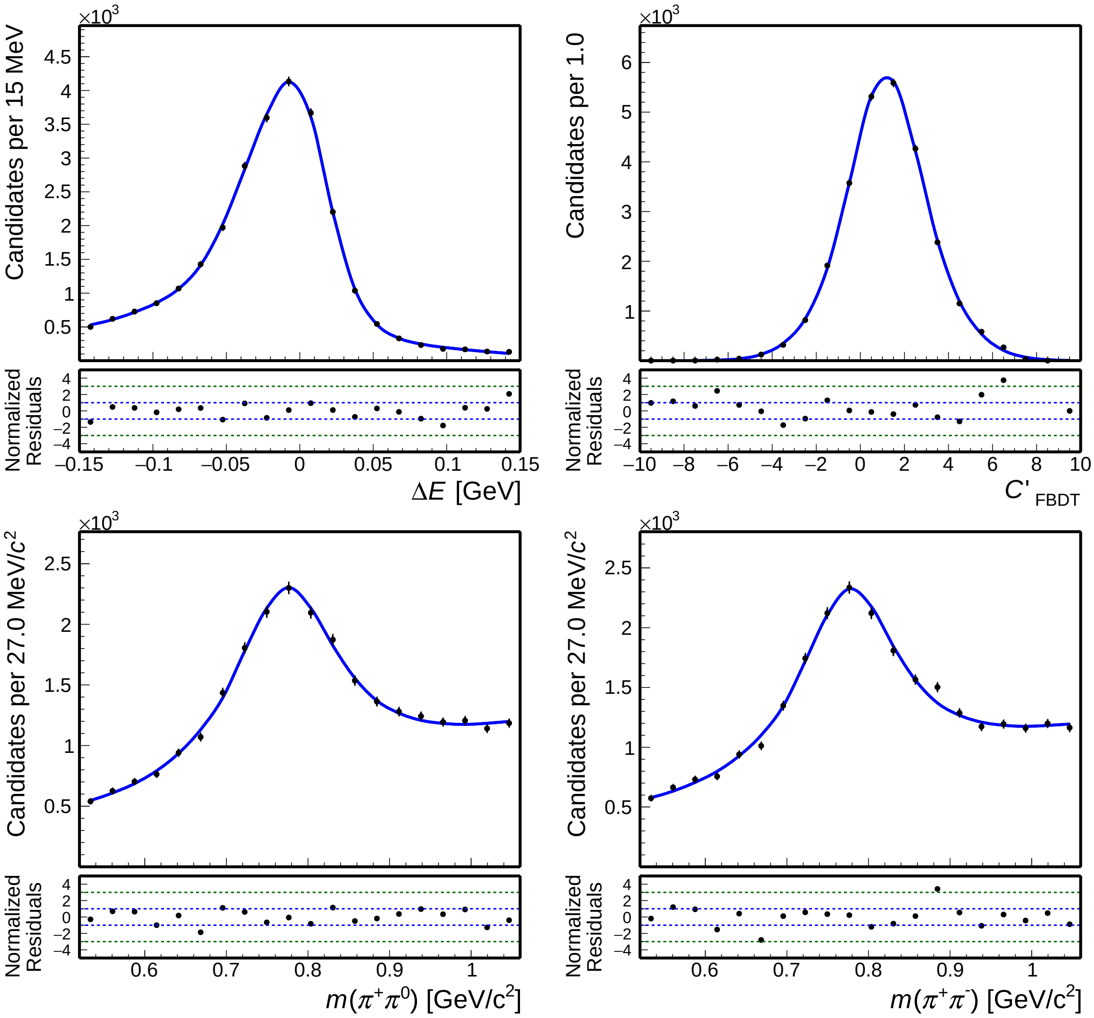


Figure A.14: Distributions of (top left) ΔE, (top right) continuum suppression output, (bottom left) $m(\pi^+\pi^0)$, and (bottom right) $m(\pi^+\pi^-)$ for simulated $B \to \rho\pi\pi$ decays. The corresponding model shapes are overlaid (solid).

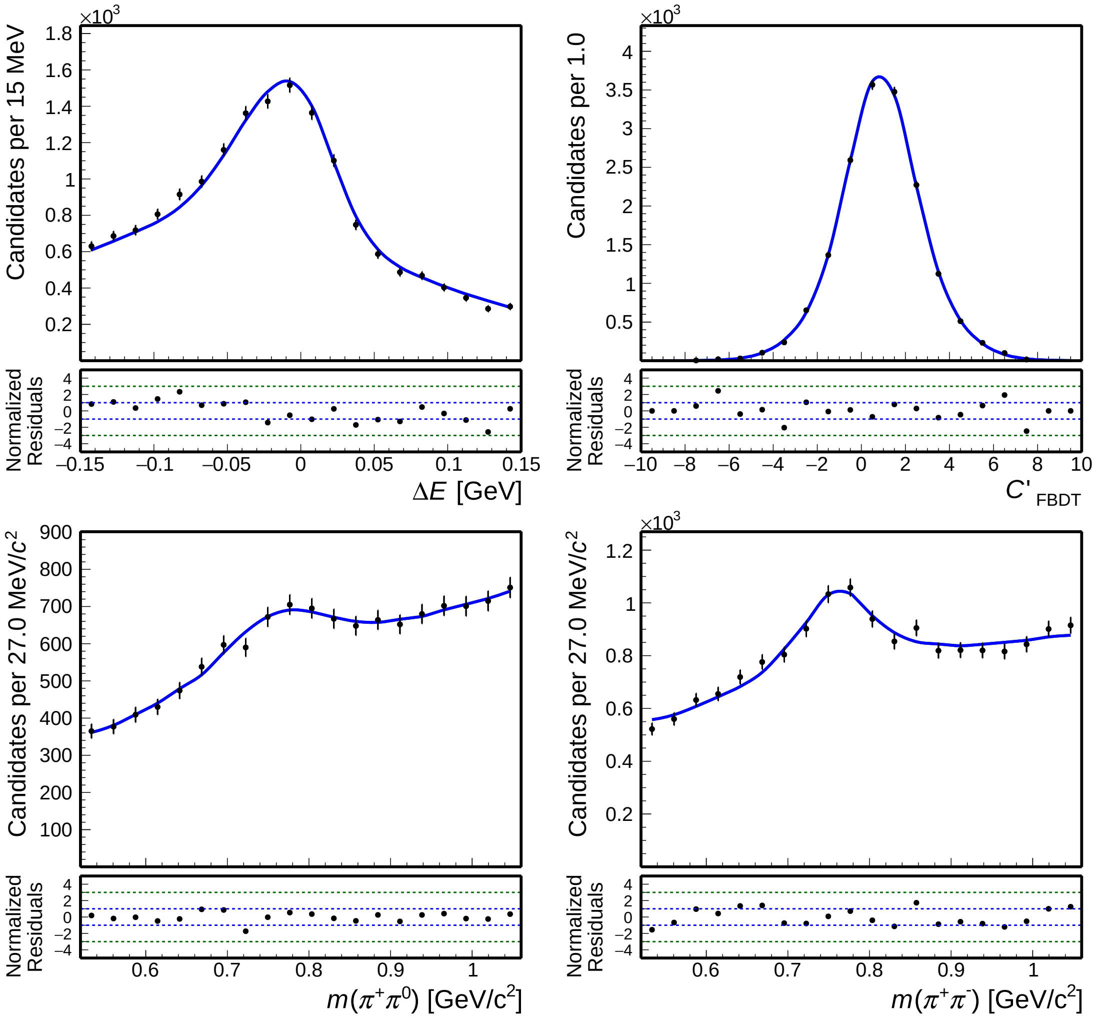


Figure A.15: Distributions of (top left) ΔE, (top right) continuum suppression output, (bottom left) $m(\pi^+\pi^0)$, and (bottom right) $m(\pi^+\pi^-)$ for simulated $B^+ \to a_1\pi$ decays. The corresponding model shapes are overlaid (solid).

Appendix B

Control channel modeling

To probe possible residual inconsistencies between data and simulation, I reproduce the key steps of the analysis on the abundant control sample $B^+ \to \overline{D}^0(\to K^+\pi^-)\rho^+(\to \pi^+\pi^0)$ decays, that shares relevant features with my signal sample. The fit shapes are extracted from simulation. Figures B.1-B.3 show modeling fit projections.

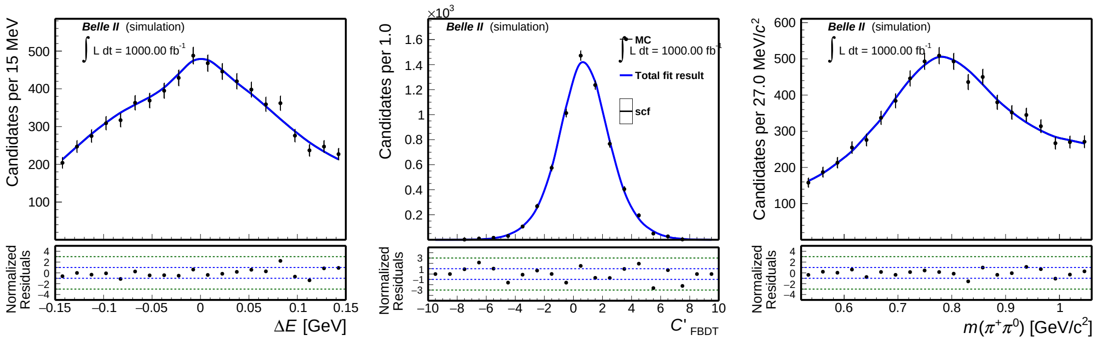


Figure B.1: Distributions of (left) ΔE, (middle) C', and (right) $m(\pi^+\pi^0)$ for simulated self cross-feed $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ candidates. Fit projections are overlaid (blue solid line).

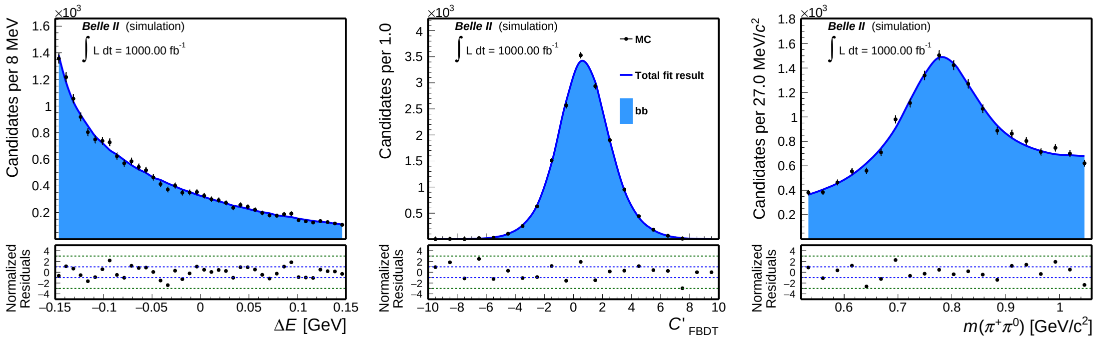


Figure B.2: Distributions of (left) ΔE, (middle) C', and (right) $m(\pi^+\pi^0)$ for $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ candidates reconstructed in simulated $B\overline{B}$ background sample. Fit projections are overlaid (blue solid line).

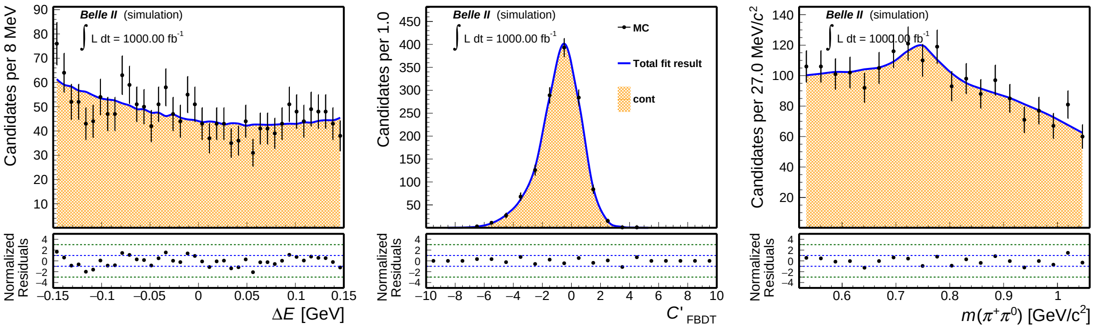


Figure B.3: Distributions of (left) ΔE, (middle) C', and (right) $m(\pi^+\pi^0)$ for $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ candidates reconstructed in simulated continuum sample. Fit projections are overlaid (blue solid line).

Bibliography

[1] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[2] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961) 579.

[3] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519** (1968) 367.

[4] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.

[5] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.

[6] G. Guralnik, C. Hagen, and T. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585.

[7] To a good approximation, the attractive force between quarks grows linearly with distance. As two quarks are separated, putting energy into the bounded system, it becomes energetically favorable for a new quark–antiquark pair to appear, rather than extending the distance further. As a consequence new bounded states are generated instead of separeted and free quarks.

[8] G. Lüders, *Proof of the TCP theorem*, Annals of Physics **2** (1957) 1.

[9] W. Pauli, *Exclusion Principle, Lorentz Group, and reversal of space-time and charge*, Niels Bohr and the Development of Physics (1955) 1.

[10] P. Zyla *et al.*, Particle Data Group, *Review of Particle Physics*, Prog. Theor. Exp. Phys. **2020** (2020) 083C01.

[11] M. Gell-Mann and M. Lévy, *The axial vector current in beta decay*, Il Nuovo Cimento **16** (1960) 705.

[12] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.

[13] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (1970) 1285.

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

[15] A general $N \times N$ complex matrix U is defined with $2N^2$ free parameters. The unitary condition $UU^\dagger = \mathbb{1}$ eliminates N^2 degrees of freedom. We can then redefine $(N-1)$ quark fields to absorb $(N-1)^2$ free parameters.
We end up with $2N^2 - N^2 - (2N-1) = (N-1)^2$ free parameters.

[16] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[17] C. Jarlskog, *Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Violation*, Phys. Rev. Lett. **55** (1985) 1039.

[18] C. Jarlskog, *A Basis Independent Formulation of the Connection Between Quark Mass Matrices, CP Violation and Experiment*, Z. Phys. C **29** (1985) 491.

[19] I. Dunietz, O. W. Greenberg, and D.-D. Wu, *A priori definition of maximal CP nonconservation*, Phys. Rev. Lett. **55** (1985) 2935.

[20] Y. Ahmis *et al.*, (Belle II Collaboration), *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, Eur. Phys. J. C77 (2017) 895.

[21] J. Charles *et al.*, (CKMfitter Group), *Unitarity Triangle plot*, Eur. Phys. J. C41, 1 (2005) [hep-ph/0406184], updated plots available at `http://ckmfitter.in2p3.fr` .

[22] H.-Y. Cheng and S. Oh, *Flavor SU(3) symmetry and QCD factorization in $B \to PP$ and PV decays*, JHEP **09** (2011) 024, `arXiv:1104.4144 [hep-ph]`.

[23] S.-H. Zhou, Q.-A. Zhang, W.-R. Lyu, and C.-D. Lü, *Analysis of Charmless Two-body B decays in Factorization Assisted Topological Amplitude Approach*, Eur. Phys. J. **C77** (2017) no. 2, 125, `arXiv:1608.02819 [hep-ph]`.

[24] N. G. Deshpande and X.-G. He, *Isospin structure of penguins and their consequences in B physics*, Phys. Rev. Lett. **74** (1995) 26–29, `arXiv:hep-ph/9408404 [hep-ph]`. [Erratum: Phys. Rev. Lett.74,4099(1995)].

[25] M. Gronau, O. F. Hernandez, D. London, and J. L. Rosner, *Electroweak penguins and two-body B decays*, Phys. Rev. **D52** (1995) 6374–6382, `arXiv:hep-ph/9504327 [hep-ph]`.

[26] J. Charles, O. Deschamps, S. Descotes-Genon, and V. Niess, *Isospin analysis of charmless B-meson decays*, Eur. Phys. J. **C77** (2017) no. 8, 574, `arXiv:1705.02981 [hep-ph]`.

[27] M. Gronau and D. London, *Isospin analysis of CP asymmetries in B decays*, Phys. Rev. Lett. **65** (1990) 3381–3384.

[28] H. J. Lipkin, Y. Nir, H. R. Quinn, and A. Snyder, *Penguin trapping with isospin analysis and CP asymmetries in B decays*, Phys. Rev. **D44** (1991) 1454–1460.

[29] M. Gronau, O. F. Hernandez, D. London, and J. L. Rosner, *Decays of B mesons to two light pseudoscalars*, Phys. Rev. **D50** (1994) 4529–4543, `arXiv:hep-ph/9404283 [hep-ph]`.

[30] M. Gronau, O. F. Hernandez, D. London, and J. L. Rosner, *Broken SU(3) symmetry in two-body B decays*, Phys. Rev. **D52** (1995) 6356–6373, `arXiv:hep-ph/9504326 [hep-ph]`.

[31] M. Beneke, M. Gronau, J. Rohrer, and M. Spranger, *A Precise determination of α using $B^0 \to \rho^+\rho^-$ and $B^+ \to K^*\rho^+$*, Phys. Lett. **B638** (2006) 68–73, `arXiv:hep-ph/0604005 [hep-ph]`.

[32] A. E. Snyder and H. R. Quinn, *Measuring CP asymmetry in $B \to \rho\pi$ decays without ambiguities*, Phys. Rev. **D48** (1993) 2139–2144.

[33] M. Gronau and J. Zupan, *On measuring alpha in $B(t) \to \rho^{\pm}\pi^{\mp}$*, Phys. Rev. **D70** (2004) 074031, `arXiv:hep-ph/0407002 [hep-ph]`.

[34] M. Gronau and J. Zupan, *Weak phase alpha from $B \to a_1(1260)\pi$*, Phys. Rev. **D73** (2006) 057502, `arXiv:hep-ph/0512148 [hep-ph]`.

[35] S. Laplace and V. Shelkov, *CP violation and the absence of second class currents in charmless B decays*, Eur. Phys. J. **C22** (2001) 431–438, `arXiv:hep-ph/0105252 [hep-ph]`.

[36] A. F. Falk, Z. Ligeti, Y. Nir, and H. Quinn, *Comment on extracting alpha from $B \to \rho\rho$*, Phys. Rev. **D69** (2004) 011502, `arXiv:hep-ph/0310242 [hep-ph]`.

[37] M. Gronau and J. L. Rosner, *Controlling ρ width effects for a precise value of α in $B \to \rho\rho$*, Phys. Lett. **B766** (2017) 345–350, `arXiv:1612.08524 [hep-ph]`.

[38] M. Pivk and F. R. Le Diberder, *Isospin constraints from / on $B \to \pi\pi$*, Eur. Phys. J. **C39** (2005) 397–409, `arXiv:hep-ph/0406263 [hep-ph]`.

[39] B. Aubert, M. Bona, Y. Karyotakis, J. P. Lees, V. Poireau, E. Prencipe, X. Prudent, V. Tisserand, J. G. Tico, E. Grauges, and et al., *Measurement of the branching fraction, polarization, and CP asymmetries in $B^0 \to \rho^0\rho^0$ decay, and implications for the CKM angle α*, Physical Review D **78** (Oct, 2008) . `http://dx.doi.org/10.1103/PhysRevD.78.071104`.

[40] P. Vanhoefer, J. Dalseno, C. Kiesling, A. Abdesselam, I. Adachi, H. Aihara, S. Al Said, K. Arinstein, D. Asner, H. Atmacan, and et al., *Study of $B^0 \to \rho^+\rho^-$ decays and implications for the CKM angle ϕ_2*, Physical Review D **93** (Feb, 2016) . `http://dx.doi.org/10.1103/PhysRevD.93.032010`.

[41] J. Zhang, M. Nakao, K. Abe, K. Abe, T. Abe, I. Adachi, H. Aihara, M. Akatsu, Y. Asano, T. Aso, and et al., *Observation of $B^+ \to \rho^+\rho^0$ Decays*, Physical Review Letters **91** (Nov, 2003) . `http://dx.doi.org/10.1103/PhysRevLett.91.221801`.

[42] B. Aubert, Y. Karyotakis, J. P. Lees, V. Poireau, E. Prencipe, X. Prudent, V. Tisserand, J. Garra Tico, E. Grauges, L. Lopez, and et al., *Improved Measurement of $B^+ \to \rho^+\rho^0$ and Determination of the Quark-Mixing Phase Angle α*, Physical Review Letters **102** (Apr, 2009) . `http://dx.doi.org/10.1103/PhysRevLett.102.141802`.

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

[44] P. Oddone, *Detector Considerations*, in D. H. Stork, proceedings of Workshop on Conceptual Design of a Test Linear Collider: Possibilities for a BB Factory (1987) 423.

[45] P. Raimondi *et al.*, *Beam-beam issues for colliding schemes with large Piwinski angle and crabbed waist*, `arXiv:physics/0702033`.

[46] T. Abe *et al.*, (Belle II Collaboration), *Belle II Technical Design Report*, `arXiv:1011.0352`.

[47] C. Marinas, DEPFET, *The Belle II pixel detector: High precision with low material*, Nucl. Instrum. Meth. **A731** (2013) 31.

[48] K. Adamczyk *et al.*, (Belle-II SVD Collaboration), *The Belle II silicon vertex detector assembly and mechanics*, Nucl. Instrum. Meth. **A845** (2017) 38.

[49] H. Hirano *et al.*, *A high-resolution cylindrical drift chamber for the KEK B-factory*, Nucl. Instrum. Methods Phys. Res. A **455** (12, 2000) 294–304.

[50] V. Aulchenko *et al.*, *Electromagnetic calorimeter for Belle II*, J. Phys. Conf. Ser. **587** (2015) 012045.

[51] K. Inami and I. Belle, *TOP counter for particle identification at the Belle II experiment*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **766** (2014) 5–8.

[52] S. Iwata *et al.*, *Particle identification performance of the prototype aerogel RICH counter for the Belle II experiment*, PTEP **2016** (2016) 033H01, `arXiv:1603.02503 [physics.ins-det]`.

[53] T. Kuhr *et al.*, (Belle-II Framework Software Group), *The Belle II Core Software*, Comput. Softw. Big Sci. **3** (2019) 1.

[54] BaBar, D. Boutigny *et al.*, *The BABAR physics book: Physics at an asymmetric B factory.* 10, 1998.

[55] G. Kramer and W. F. Palmer, *Branching ratios and* CP *asymmetries in the decay* $B \to \mathrm{VV}$, Phys. Rev. D **45** (Jan, 1992) 193–216. `https://link.aps.org/doi/10.1103/PhysRevD.45.193`.

[56] P. Vanhoefer, J. Dalseno, C. Kiesling, I. Adachi, H. Aihara, D. Asner, V. Aulchenko, T. Aushev, A. Bakich, A. Bala, and et al., *Study of* $B^0 \to \rho^0\rho^0$ *decays, implications for the CKM angle* ϕ_2 *and search for other* B^0 *decay modes with a four-pion final state*, Physical Review D **89** (Apr, 2014) . `http://dx.doi.org/10.1103/PhysRevD.89.072008`.

[57] B. Aubert, M. Bona, D. Boutigny, Y. Karyotakis, J. P. Lees, V. Poireau, X. Prudent, V. Tisserand, A. Zghiche, J. G. Tico, and et al., *Study of* $B^0 \to \rho^+\rho^-$ *decays and constraints on the CKM angle*, Physical Review D **76** (Sep, 2007) . `http://dx.doi.org/10.1103/PhysRevD.76.052007`.

[58] R. Aaij, B. Adeva, M. Adinolfi, A. Affolder, Z. Ajaltouni, S. Akar, J. Albrecht, F. Alessio, M. Alexander, S. Ali, and et al., *Observation of the* $B^0 \to \rho^0\rho^0$ *decay from an amplitude analysis of* $B^0 \to (\pi^+\pi^-)(\pi^+\pi^-)$ *decays*, Physics Letters B **747** (Jul, 2015) 468–478. `http://dx.doi.org/10.1016/j.physletb.2015.06.027`.

[59] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **552** (2005) no. 3, 566–575.

[60] J.-F. Krohn, F. Tenchini, P. Urquijo, F. Abudinén, S. Cunliffe, T. Ferber, M. Gelb, J. Gemmler, P. Goldenzweig, T. Keck, I. Komarov, T. Kuhr, L. Ligioi, M. Lubej, F. Meier, F. Metzner, C. Pulvermacher, M. Ritter, U. Tamponi, and A. Zupanc, *Global decay chain vertex fitting at Belle II*, Nuclear Instruments and Methods in

Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **976** (2020) 164269.

[61] Belle, J. Tanaka, *Kinematic fitting*, 2000. Belle note 194 (internal) (2000) .

[62] 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, Munich, Max Planck Inst., 2018.

[63] S. Brandt, C. Peyrou, R. Sosnowski, and A. Wroblewski, *The Principal axis of jets. An Attempt to analyze high-energy collisions as two-body processes*, Phys. Lett. **12** (1964) 57–61.

[64] D. M. Asner *et al.*, *Search for exclusive charmless hadronic B decays*, Phys. Rev. D **53** (Feb, 1996) 1039–1050. `https://link.aps.org/doi/10.1103/PhysRevD.53.1039`.

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

[66] S. H. Lee *et al.*, Belle, *Evidence for $B^0 \to \pi^0\pi^0$*, Phys. Rev. Lett. **91** (2003) 261801, `arXiv:hep-ex/0308040 [hep-ex]`.

[67] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, Journal of High Energy Physics **2006** (May, 2006) 026–026. `http://dx.doi.org/10.1088/1126-6708/2006/05/026`.

[68] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152–155.

[69] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, The European Physical Journal C **45** (Jan, 2006) 97–107. `http://dx.doi.org/10.1140/epjc/s2005-02396-4`.

[70] S. Agostinelli *et al.*, GEANT4, *GEANT4–a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250–303.

[71] B. Ward, S. Jadach, and Z. Was, *Precision calculation for ee- → 2f: the KK MC project*, Nuclear Physics B - Proceedings Supplements **116** (mar, 2003) 73–77. `https://doi.org/10.1016%2Fs0920-5632%2803%2980147-0`.

[72] Belle II, G. De Marino and K. Trabelsi, *HLT Skims*, Belle II note BELLE2-NOTE-TE-2020-018 (internal) .

[73] T. Keck, *FastBDT: A speed-optimized and cache-friendly implementation of stochastic gradient-boosted decision trees for multivariate classification*, 2016. `arXiv:1609.06119 [cs.LG]`.

[74] J. H. Friedman, *Greedy function approximation: A gradient boosting machine.*, Ann. Statist. **29** (10, 2001) 1189–1232. `https://doi.org/10.1214/aos/1013203451`.

[75] J. H. Friedman, *Stochastic gradient boosting*, Computational Statistics Data Analysis **38** (2002) no. 4, 367 – 378. `http://www.sciencedirect.com/science/article/pii/S0167947301000652`. Nonlinear Methods and Data Mining.

[76] F. James and M. Roos, *MINUIT: a system for function minimization and analysis of the parameter errors and corrections*, Comput. Phys. Commun. **10** (1975) no. CERN-DD-75-20, 343–367.

[77] S. Choudhury *et al.*, Belle Collaboration, *Measurement of the* B^+/B^0 *production ratio in* e^+e^- *collisions at the* $\Upsilon(4S)$ *resonance using* $B \to J/\psi(\ell\ell)K$ *decays at Belle*, Phys. Rev. D **107** (Feb, 2023) L031102. `https://link.aps.org/doi/10.1103/PhysRevD.107.L031102`.

[78] M. Alam *et al.*, CLEO, *Exclusive hadronic B decays to charm and charmonium final states*, Phys. Rev. D **50** (1994) 43–68, `arXiv:hep-ph/9403295`.

[79] F. Abudinén *et al.*, Belle-II, *Measurement of Branching Fraction and Longitudinal Polarization in* $B^0 \to \rho^+\rho^-$ *Decays at Belle II*, `arXiv:2208.03554 [hep-ex]`.

[80] R. Aaij *et al.*, LHCb, *Dalitz plot analysis of* $B^0 \to \overline{D}^0\pi^+\pi^-$ *decays*, Phys. Rev. D **92** (2015) no. 3, 032002, `arXiv:1505.01710 [hep-ex]`.