UNIVERSITÀ DEGLI STUDI DI TRIESTE

DIPARTIMENTO DI FISICA

XXXIII Ciclo del Dottorato di Ricerca in Fisica

MEASUREMENT OF THE BRANCHING FRACTION, LONGITUDINAL POLARIZATION FRACTION, AND CHARGE-PARITY VIOLATING ASYMMETRY IN $B^+ \to \rho^+\rho^0$ DECAYS AT BELLE

Settore scientifico-disciplinare: FIS/04

Candidato:
Eldar Ganiev

Coordinatore:
Prof. Francesco Longo
Supervisore di tesi:
Dott. Diego Tonelli

ANNO ACCADEMICO 2019–2020

Abstract

This is an experimental particle physics thesis aimed at 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 (e^+e^-) collisions produced by the energy-asymmetric KEKB collider at the $\Upsilon(4S)$ resonance collected by the Belle experiment and corresponding to 772 million pairs of bottom-antibottom mesons. Belle is an hermetic magnetic spectrometer surrounded by particle-identification detectors, calorimeters, and muon detectors designed to reconstruct the decay products of 10.58 GeV e^+e^- collisions.

The analysis is entirely developed using simulated and control-data samples. Only when all procedures are established the analysis is applied to the signal candidates in data. 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. Data are enriched in $B^+ \to \rho^+\rho^0$ events by a multivariate statistical-learning selection that suppresses the most prominent source of background. A multidimensional fit of sample composition then identifies statistically the signal. Special care is devoted to keep the impact of systematic uncertainties under control. A total of approximately 1000 $B^+ \to \rho^+\rho^0$ decays are reconstructed. The fit results, combined with acceptance and efficiency corrections determined from simulation, allow for determining the $B^+ \to \rho^+\rho^0$ branching fraction,

$$\mathcal{B}(B^+ \to \rho^+\rho^0) \;=\; ([\mathrm{xx}]^{+1.20}_{-1.19}\ (\mathrm{stat}) \pm 1.46\ (\mathrm{syst})) \times 10^{-6},$$

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

$$f_L \;=\; [\mathrm{xx}] \pm 0.025\ (\mathrm{stat}) \pm 0.021\ (\mathrm{syst}),$$

and the *CP*-violating charge-asymmetry,

$$A_{CP} \;=\; [\mathrm{xx}] \pm 0.052\ (\mathrm{stat}) \pm 0.015\ (\mathrm{syst}).$$

The central values are hidden as the analysis is still under internal collaboration review. The uncertainties are comparable with, or better than, the current world best results. Our results provide a 13% improvement on the global knowledge of α.

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, more general theory. Completing the Standard Model is the principal goal of today's particle physics.

A 'direct' approach, which broadly consists in searching for decay products of non-SM particles produced on mass-shell in high-energy collisions, has been traditionally fruitful. However, its current reach is limited by the collision energy of today's accelerators and by the large investments needed to further it in future. A complementary approach consists in performing precise measurements in lower-energy processes where virtual non-SM particles could contribute, and searching for deviations of experimental results from the SM predictions. The reach of such 'indirect' approach is not constrained by collision energy, but rather by the precision attainable, both in measurements and predictions.

The Belle experiment is an international collaboration of more than 400 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 the energy of 10.58 GeV from 1999 to 2010. Belle collected collision samples corresponding to $711\,\mathrm{fb}^{-1}$ of integrated luminosity, or, equivalently, 772 million pairs of quantum-correlated bottom-antibottom mesons (B, bound states of a b quark and a lighter quark). This is the world's largest sample of this kind and currently enables extremely precise tests of the SM.

This work focuses on improving the determination of α (the alternative notation ϕ_2 is also used), a fundamental parameter of quark dynamics that limits the 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 that offer enhanced sensitivity to constrain current global fits of SM quark dynamics. I report the most precise measurement of the decay properties of $B^+ \to \rho^+\rho^0$ decays, which are those that mostly limit the precision on the current knowledge of α. Decays $B^+ \to \rho^+(\to \pi^+\pi^0)\rho^0(\to \pi^+\pi^-)$ are reconstructed in the full Belle data set. 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. Difficulties in overcoming this challenge are those that prevented the Belle experiment from achieving this measurement since year 2003.

This thesis is structured as follows. Chapter 1 introduces the flavor sector of the Standard Model; Chapter 2 describes the Belle 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 extend 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

$$SU_C(3) \otimes SU_L(2) \otimes U_Y(1),$$

where $SU_C(3)$ is the standard unitary group that describes the strong interactions (quantum chromodynamics, QCD), and C stands for the color charge; $SU_L(2) \otimes U_Y(1)$ 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 $SU_C(3)$ 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 $SU_L(2) \otimes U_Y(1)$ 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.

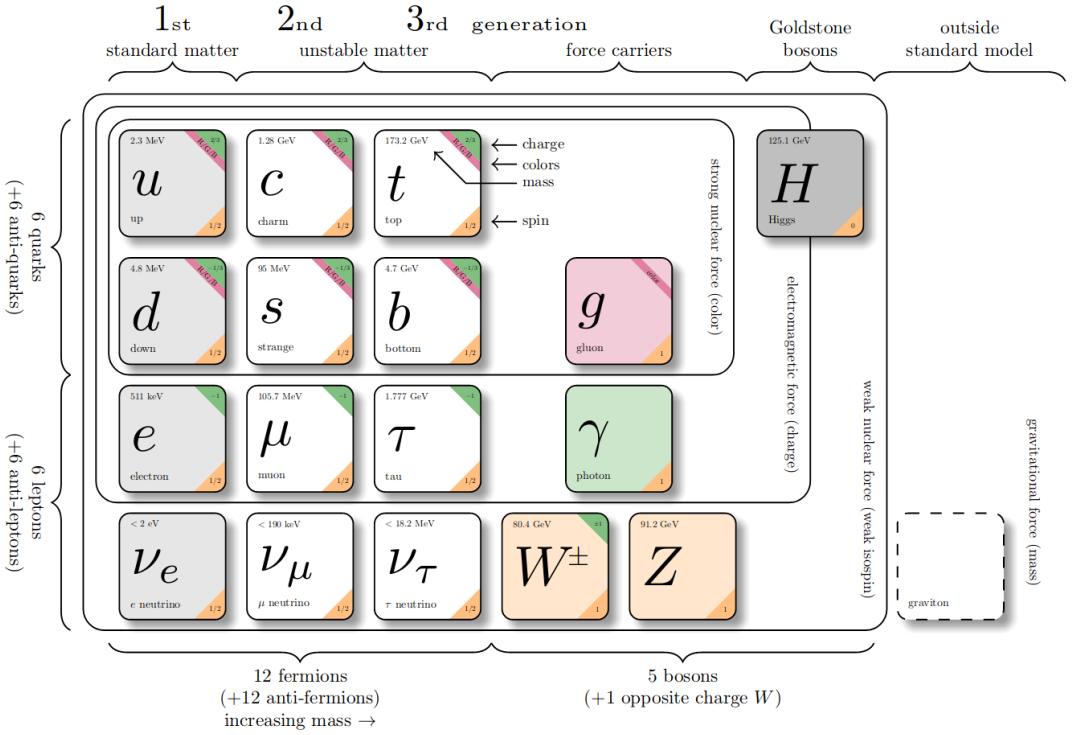


Figure 1.1: Scheme of particles and interactions in the Standard Model.

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 allows for *CP* violation. The unitarity condition $V_{\text{CKM}}V_{\text{CKM}}^\dagger = \mathbb{1}$ yields nine relations,

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

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 non-degenerate 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.

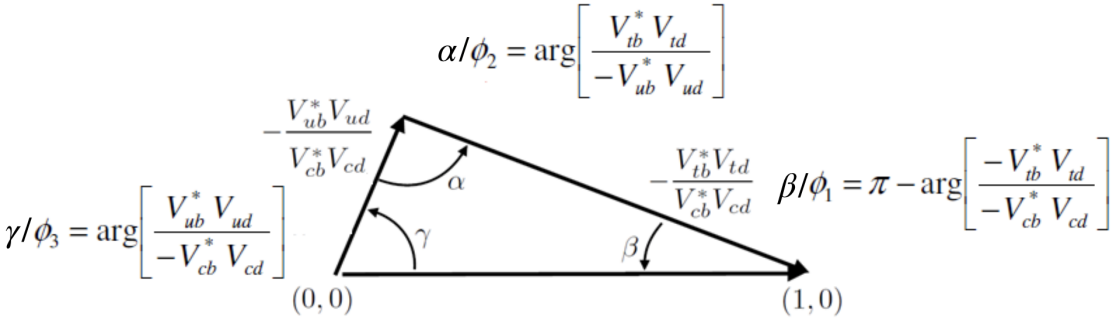


Figure 1.2: Graphical representation of the Unitarity Triangle.

Flavored neutral mesons $|M\rangle$ exhibit a phenomenon that is important in enriching 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 evolution paths to the simple particle decays, which interfere with the decay themselves thus enriching the dynamics and our instruments to study it.

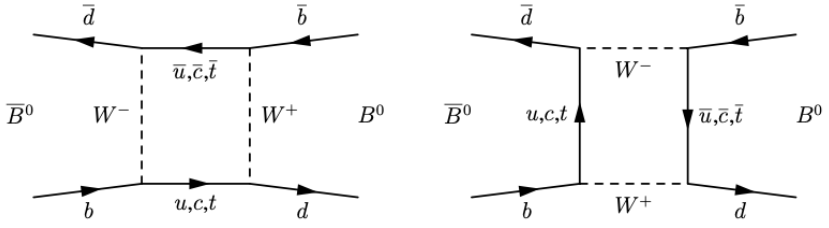


Figure 1.3: Leading order Feynman diagrams contributing to neutral $B^0 - \overline{B}^0$ mixing.

1.4 Flavor physics to extend 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. 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.

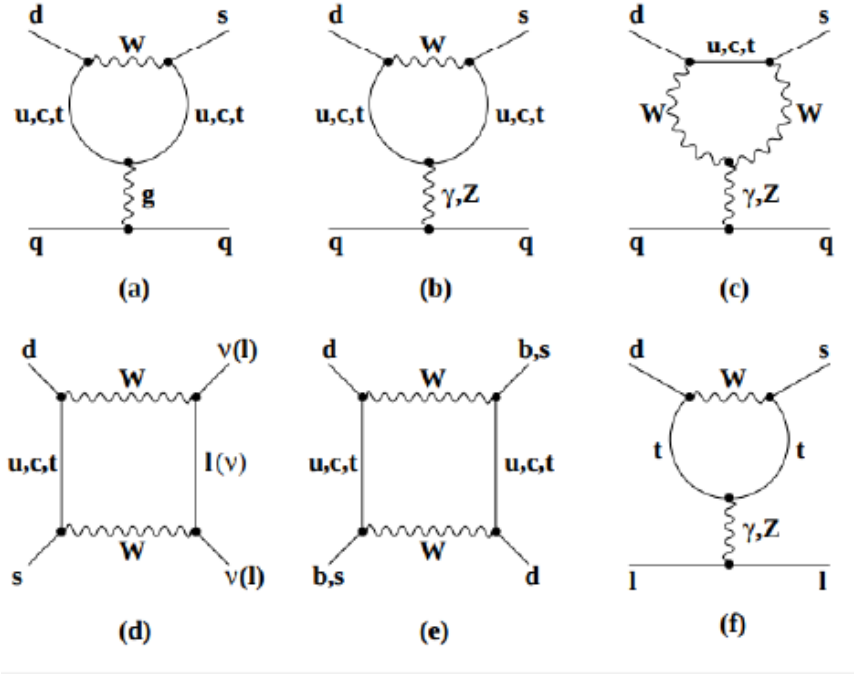


Figure 1.4: Examples of leading FCNC diagrams.

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 instruments 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. For this purpose, 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 $|\bar{f}\rangle$,

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

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

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

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

This is the only kind of *CP* violation that is possible for both neutral and charged mesons, and for baryons. The phase difference ϕ_D between the decay amplitudes A_f and $\bar{A}_{\bar{f}}$ is the sum of phases contributed by the couplings involved in the process. Since in general, different intermediate states contribute to a decay, the total decay amplitudes A_f and $\bar{A}_{\bar{f}}$ are written as

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

where the phases δ_i are *CP*-conserving phases, and the phases ϕ_i are *CP*-violating phases associated with the elements of the CKM matrix that appear in the amplitudes of a given decay. 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 intermediate states 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 $|A_i|$ and the *CP*-conserving phases δ_i are difficult to compute theoretically as they typically arise from strong-interaction amplitudes at low energy, which is 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 the SM predictions in a more reliable way.

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

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

In this case, *CP* violation arises as a consequence of a difference between the flavor oscillation rates $|M\rangle \to |\bar{M}\rangle$ and $|\bar{M}\rangle \to |M\rangle$.

c) *CP* violation by interference of mixing and decay is observed when the neutral mesons $|M\rangle$ and $|\bar{M}\rangle$ can evolve 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 sum of the decay and the mixing phases ϕ_D and ϕ_M 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, 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.

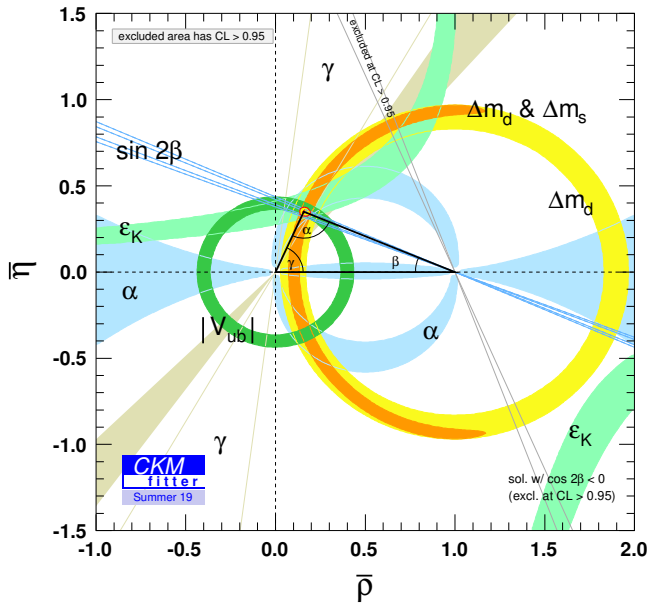


Figure 1.5: Current constrains on sides and angles of the Unitarity Triangle.

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, flavor physics emerges

as a very promising asset to search for non-SM in the next decade. In fact, 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, successor of Belle.

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].

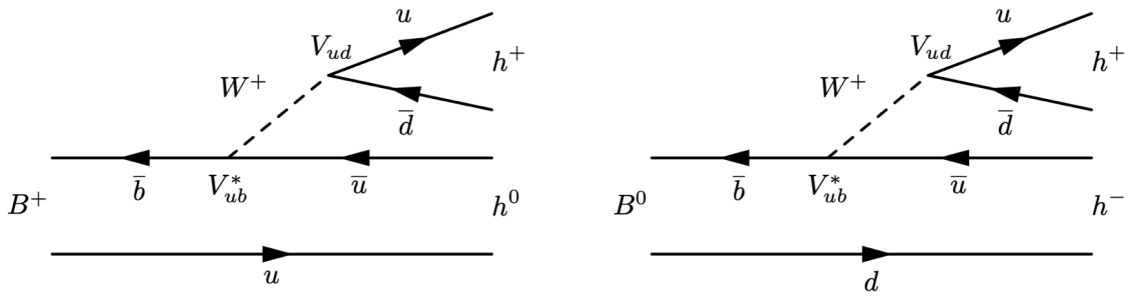
Considering, as an example, the decays of neutral B^0 mesons only, and assuming that only tree-level diagrams contribute (Fig. 1.6a, 1.6b), one obtains

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

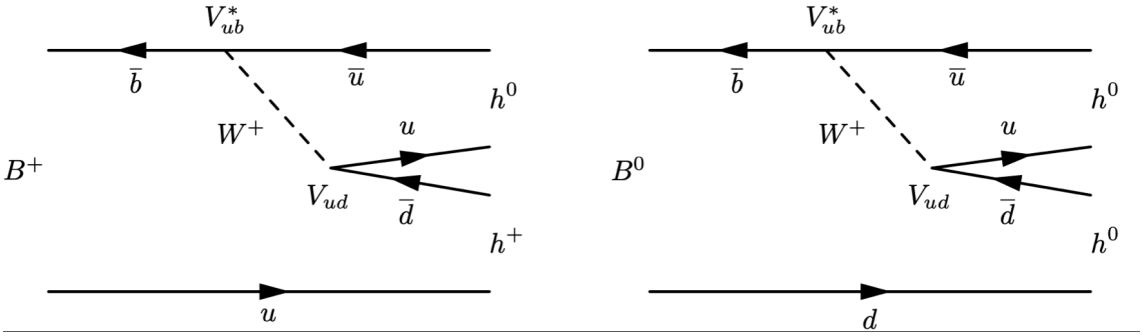
where the mixing and decay amplitudes are transparently shown. The $\frac{V_{ud}^* V_{ub}}{V_{ud} V_{ub}^*}$ term corresponds to the tree level decay amplitude, while $\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}$ term corresponds to the mixing amplitude. In the mixing amplitude, we consider only box diagrams with an internal top quark neglecting those with a charm and up quark, as the mixing amplitude goes with m_{up}^6 factor (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 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 even though within sizable uncertainties. If large penguin amplitudes 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 $B^0 \to h^0 h^0$, 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]

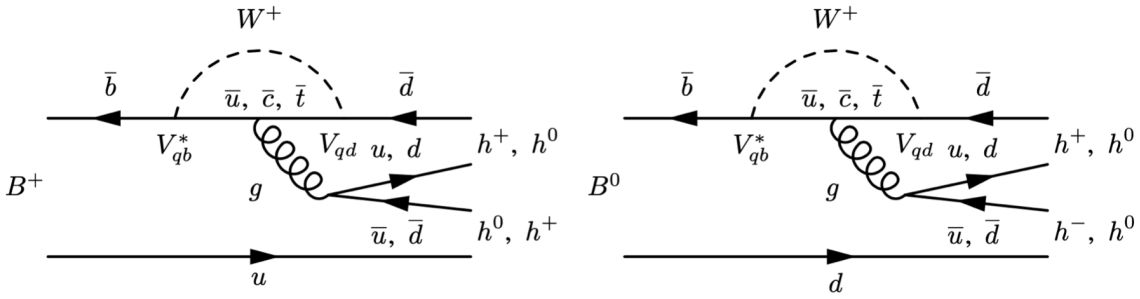
$$A^{ij} = \langle h_1^i h_2^j | \mathcal{H}_{\text{eff}} | B \rangle = V_{ud} V_{ub}^* (T_u^{ij} + P_u^{ij}) + V_{cd} V_{cb}^* P_c^{ij} + V_{td} V_{tb}^* P_t^{ij},$$



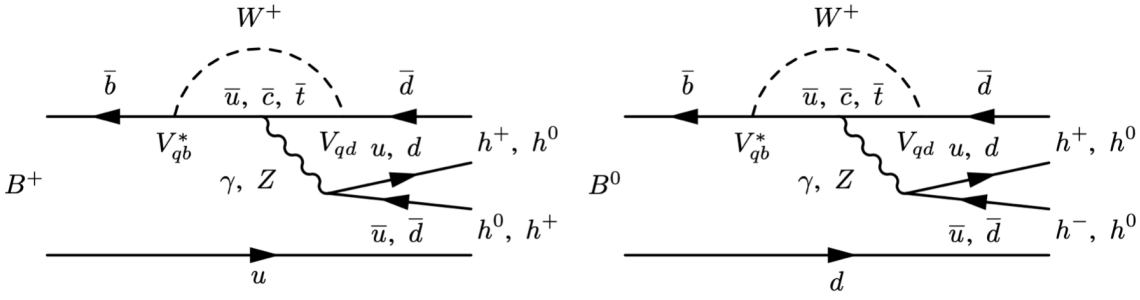
(a) Color-favored tree diagrams T.



(b) Color-suppressed tree diagrams C.



(c) QCD-penguin diagram P.



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

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 $\alpha = (86.2^{+4.4}_{-4.0})^\circ$ is dominated by the $B \to \rho\rho$ decays, and to a lesser extent, by the $B \to \pi\pi$ system [26]. The main goal of this thesis' work is to improve it.

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

- $\pi(u\bar{d})$, $\pi^0(\frac{u\bar{u}-d\bar{d}}{\sqrt{2}})$, $\pi^-(\bar{u}d)$ with $I_3 = -1, 0, +1$, and
- $\rho^+(u\bar{d})$, $\rho^0(\frac{u\bar{u}-d\bar{d}}{\sqrt{2}})$, $\rho^-(\bar{u}d)$ 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, the relations between decay amplitudes derived from the isospin symmetry can be 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 work. The three

relevant final states for $I_{\rm f} = 0$ 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

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

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

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

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

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 $A_{\Delta I,I_{\rm f}}$, where ΔI is the isospin shift and $I_{\rm f}$ is the final-state isospin. Hadronization and strong rescattering effects that are absent in the amplitudes $A_{\Delta I}$ are included into the amplitudes $A_{\Delta I,I_{\rm f}}$. The projection of the states $\mathcal{H}_{\rm eff}|B^{i+j}\rangle$ onto the final states yields

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

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

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

The *CP*-conjugated amplitudes $\bar{A}_{\Delta I,I_{\rm f}}$ carry strong phases identical to those of $A_{\Delta I,I_{\rm f}}$, 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. **??**). 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} it follows

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

where $c = \cos(\alpha - \delta)$ and $\bar{c} = \cos(\alpha + \delta - 2\alpha_{\text{eff}})$. The differing lifetimes for charged and neutral B mesons τ_{B^+} and τ_{B^0} 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.

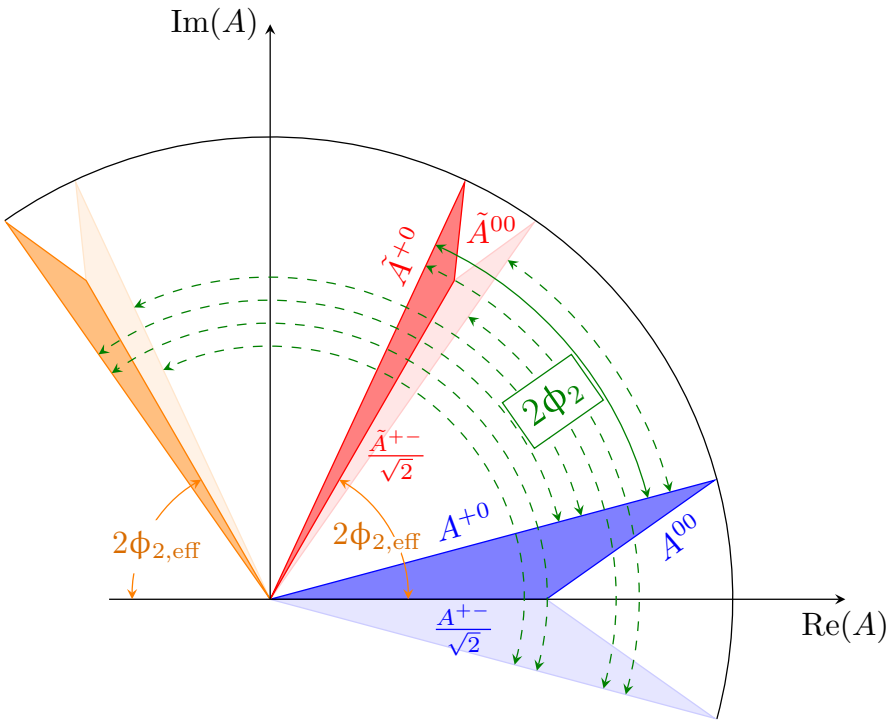


Figure 1.7: Geometrical representation of the isospin triangular relations (Eq. (1.17)) in the complex plane of $B^{i+j} \to h^i h^j$ amplitudes. The blue and the red shaded areas correspond to the isospin triangles. The angle between the *CP* conjugate charged amplitudes A^{+-} and $\tilde{A}^{+-}$ 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^{00}_{CP} 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 $B^0 \to h^0 h^0$,

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

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^{00}_{CP} 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 $S^{\rho^0\rho^0}_{CP}$ is available by BaBar with large uncertainties $S^{\rho^0\rho^0}_{CP} = 0.3 \pm 0.7 \pm 0.2$ [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 $f_{\rm L}$ of decays leading to longitudinally polarized mesons $\rho_{\rm L}\rho_{\rm L}$, 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 $f_{{\rm L},\rho^+\rho^0} = 0.950 \pm 0.016$ and $f_{{\rm L},\rho^+\rho^-} = 0.990^{+0.021}_{-0.019}$ [10]. Consequently, the isospin analysis of $B \to \rho\rho$ is performed only for decays leading to longitudinally polarized $\rho_{\rm L}$ mesons.

Chapter 2

The Belle experiment at the KEKB collider

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

2.1 The KEKB collider

KEKB is an electron-positron (e^+e^-) energy-asymmetric collider, designed to produce more than 20 $B\overline{B}$ pairs ($B^0\overline{B}^0$ and B^+B^- in approximately equal proportions) per second via decays of $\Upsilon(4S)$ mesons produced at threshold [43, 44]. 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 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.

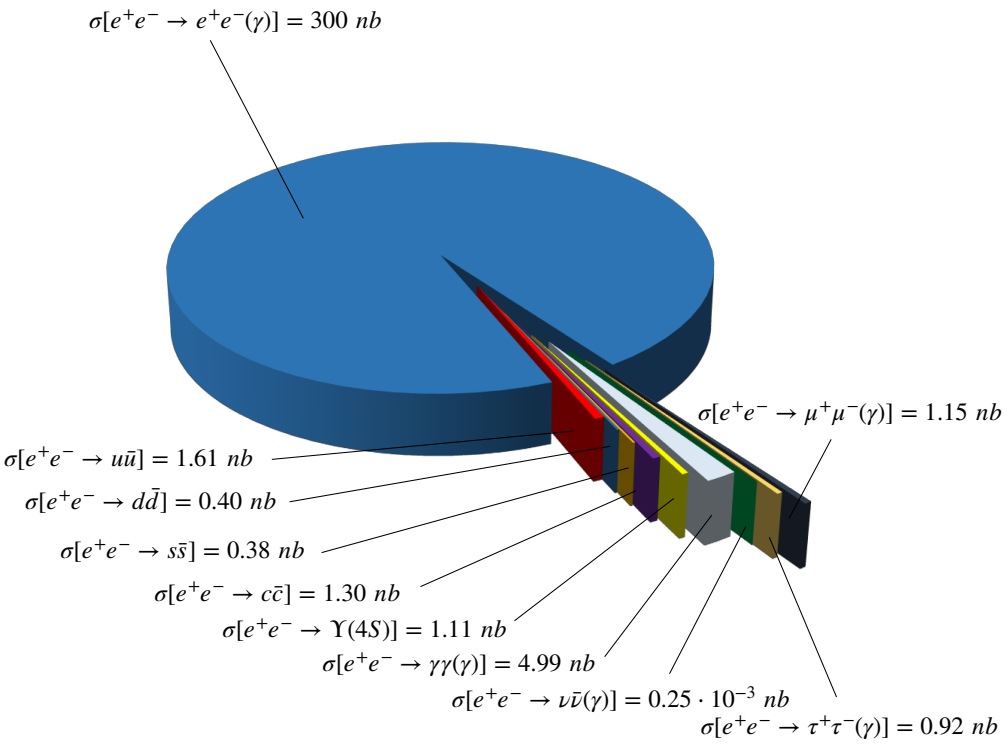


Figure 2.1: Cross sections of the main final states produced in e^+e^- collision at the $\Upsilon(4S)$ center-of-mass energy.

Figure 2.2 shows the hadron-production cross-section in e^+e^- collisions as a function 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 u, d, c, s).

Because the $\Upsilon(4S)$ 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 $\Upsilon(4S)$ 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 [45]. KEKB (Fig. 2.3) implements a 8–3.5 GeV energy-asymmetric double-ring design, which achieves a vertex displacement of about $200\,\mu$m.

Electrons are produced in a thermionic gun with a barium-impregnated tungsten cathode, then accelerated to 8 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 3.5 GeV with the linac and injected in the low-energy ring (LER). The HER and LER can store currents up to 1.1 A and 2.6

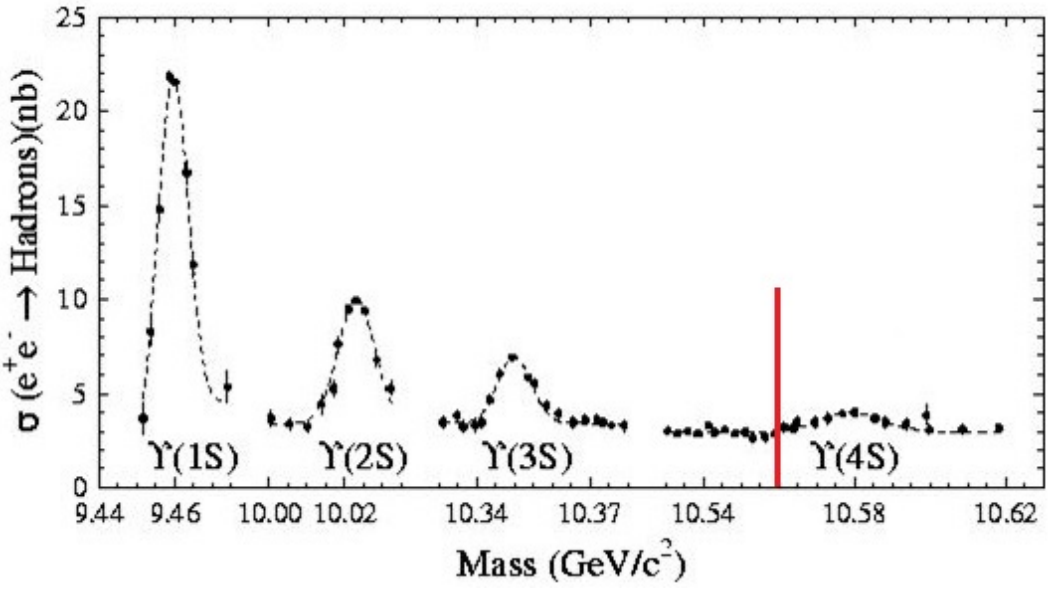


Figure 2.2: Hadron production cross section from e^+e^- collisions as a function of the final-state mass. The vertical red line indicates the $B\overline{B}$ production threshold.

A, respectively. The particle beam trajectories are controlled by dipole magnets for beam direction, quadrupole magnets to focus the core beam, and sextupole magnets to focus off-axis particles.

The electrons and positrons continuously collide at a single interaction point, around which the Belle detector was installed. The HER and LER intersect at a 22 mrad crossing angle, which provides an effective beam separation preventing parasitic collisions in the interaction region due to multiple bunches. In addition, the crossing angle prevents any need for separation magnets, which would increase beam-related background. In 2007, KEKB was upgraded with special RF cavities called 'crab-cavities' that tilt the particle bunches to compensate the crossing angle and thus providing head-on collisions that improved remarkably performance.

The performance of the KEKB 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 $\mathcal{L}$,

$$\mathcal{L} = \frac{N_{e^-} N_{e^+} f N_b}{4\pi\sigma_x^*\sigma_y^*} R_{\mathcal{L}}(\theta_x, \beta_x^*, \beta_y^*, \varepsilon_x, \varepsilon_y, \sigma_z).$$

The instantaneous luminosity depends on the following KEKB parameters (typical values for the relevant quantities are reported): the average number of electrons (positrons) per bunch $N_{e^-} = 3.3 \times 10^{10}$ ($N_{e^+} = 1.4 \times 10^{10}$); the revolution frequency $f = 99.4$ kHz; the number of bunches N_b; a measure of the transverse beam sizes $\sigma_x^* = 77$ μm and $\sigma_y^* = 2$ μm; and an empiric reduction factor $R_{\mathcal{L}} \approx 0.854$. This depends on the half crossing angle $\theta_x = 11$ mrad; the beta function, which is a measure for the transverse size of particle beam, $\beta_x^* = 0.033$ m and $\beta_y^* = 0.001$ m at the interaction point region; the beam emittance, which is a measure of the average spread of particle coordinates in position-and-momentum phase space, $\varepsilon_x = 1.8 \times 10^{-8}$ m and $\varepsilon_y = 3.6 \times 10^{-10}$ m; and longitudinal bunch size $\sigma_z = 0.4$ cm. The integral of instantaneous luminosity over time T, called integrated luminosity,

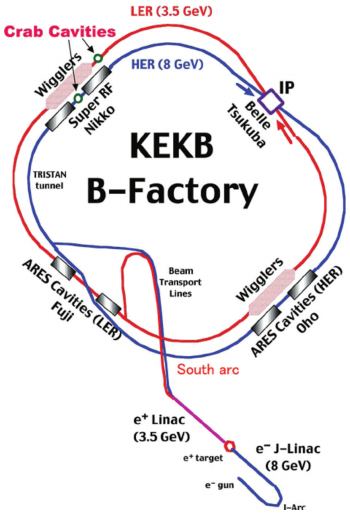


Figure 2.3: Illustration of the KEKB collider.

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

is a direct measure of the number of produced events of interest $N = \mathcal{L}_{int} \times \sigma$.

In 2009, KEKB broke the world luminosity record achieving an instantaneous luminosity of $2.11{\times}10^{34}$ cm^{-2}s^{-1}. During its operation from October 1999 to June 2010 KEKB delivered data corresponding to more than 1 ab^{-1} of integrated luminosity. The integrated luminosity collected by Belle for each collision energy is listed in Table 2.1 and is determined by measuring the cross sections for the quantum-electrodynamics processes Bhabha $e^+e^- \to e^+e^-$ $(n\gamma)$ scattering and $e^+e^- \to \gamma\gamma$. These two offer high production rate, accurate theoretical predictions for the cross sections, and simple event topology that can be simulated precisely and selected with essentially no background.

Resonance	On-peak luminosity, fb^{-1}	Off-peak luminosity, fb^{-1}	$B\overline{B}$ pairs $\times 10^6$
$\Upsilon(1S)$	5.7	1.8	102
$\Upsilon(2S)$	24.9	1.7	158
$\Upsilon(3S)$	2.9	0.25	11
$\Upsilon(4S)$	711	89.4	772
$\Upsilon(5S)$	121.4	1.7	7.1

Table 2.1: Summary of the luminosity integrated by Belle. Off-peak stands for the collision energy 60 MeV below the $\Upsilon(\mathrm{n}S)$ mass.

2.2 The Belle detector

Belle (Fig. 2.4) is a large-solid-angle, approximately cylindrical multi-purpose magnetic spectrometer surrounded by a calorimeter and particle-identification systems, installed around the KEKB interaction point. It is designed to determine energy, momentum, and identity a broad range of particles produced in 10.58 GeV e^+e^- collisions. Belle is approximately a cylinder ≈ 8 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 (Be) pipe, a four-layer silicon-strip vertex-detector (SVD), a central wire drift-chamber (CDC), aerogel threshold Cherenkov counters (ACC), time-of-flight scintillator counters (TOF), an array of CsI(Tl) crystals (ECL), a pair of BGO crystal arrays (EFC), a superconducting solenoidal magnet, and multiple layers of resistive plate counters (KLM). The performance of the Belle subsystems is summarized in the Table 2.2.

The main features of the Belle detector are robust tracking performance, which provides precisely reconstructed decay-vertices and high mass resolution; high-efficiency charged-particle identification and strong performance in neutral-particle reconstruction. A detailed description of Belle and its performance is in Ref. [46,47]. In the following, we 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 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.

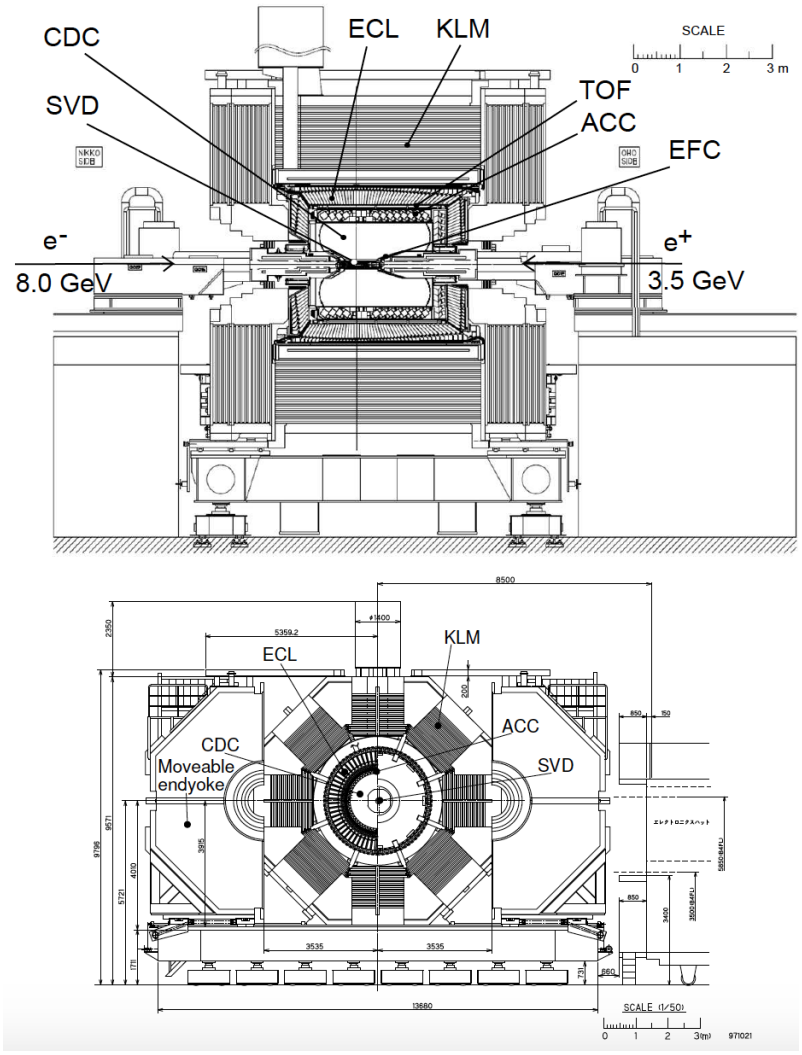


Figure 2.4: Longitudinal (top) and transverse (bottom) cross sections of the Belle detector.

Detector type	Name	Type	θ coverage	Performance
Tracking	SVD	Double-sided Si strip	[17°, 150°]	Single-hit resolution: 12 μm ($r\phi$)
		3/4 layers[a]		19 μm (z)
	CDC	Drift chamber	[17°, 150°]	Single-hit resolution: 130 μm ($r\phi$)
				200 - 1400 μm (z)
				$\sigma(\mathrm{d}E/\mathrm{d}x) \approx 7\%$
	(SVD, CDC)			$\sigma_{p_T}/p_T = 0.0019 p_T$ [GeV/c]
				$\oplus$ 0.0030/β
Particle identification (PID)	TOF	Time of flight scintillator	[34°, 130°]	$\sigma_t = 100$ ps
	ACC	Threshold Cherenkov aerogel counter	[17°, 127°]	Number of photo-electrons ≥ 6
Calorimetry	ECL	CsI(Tl)	[12.4°, 31.4°]	$\sigma_E/E \approx 1.7\%$[b]
			[32.2°, 128.7°]	
			[130.7°, 155.1°]	
Muon and K_L^0 detector	KLM	Resistive plate chamber	[20°, 155°]	$\sigma_\theta = \sigma_\phi = 30$ mrad for K_L^0
Trigger	L1	Hardware	Full	
	L3[c]	Software	Belle	
	L1+L3		acceptance	Max. rate ≈ 0.5 kHz[d]
				Physics mode eff. $> 99\%$[e]
PID	$\mu^\pm$ (KLM)			μ eff = 90%[f], π misID = 2%
	K/π (TOF, ACC, CDC)			K eff $\geq 85\%$, π misID $< 10\%$
	$e^\pm$ (ECL, CDC)			e eff = 90%, π misID $\approx 0.3\%$
Neutrals	(ECL)			$\sigma(m_{\pi^0}) = 4.8$ MeV/c^2 [g]

Table 2.2: Summary of the Belle detector main characteristics. [a]Until summer 2003 Belle used a 3-layer SVD. [b]For Bhabha events. [c]L3 trigger was operated partially from 2004 to 2007. [d]The maximum trigger rate is determined in the end of the DAQ chain. [e]For $B\overline{B}$ events. [f]For momenta above 0.8 GeV/c. [g]For π^0 reconstructed in hadronic events.

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 z-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 detector beam pipe was made of two beryllium cylinders, 0.5 mm thick, at radii of 20 mm and 23 mm, respectively. A 2.5 mm gap between the inner and outer walls of the pipe was filled with helium for cooling. The beam pipe was coated with a 20 μm gold sheet that absorbed low-energy photons, which could damage the silicon detector. The total thickness of the beam pipe was 1% of a radiation length.

In 2003, the inner detector was upgraded. The silicon vertex detector was moved closer to the interaction point to improve spatial resolution. The radius of the beam pipe was shrunk from 20 mm to 15 mm. Massive tantalum and tungsten masks were placed on each side of the beam-pipe central section to reduce beam background, which could appear due to the beam-pipe radius reduction and paraffin replaced helium for cooling.

2.2.3 Tracking system

Ability to reconstruct efficiently charged-particle trajectories (tracks) and measure precisely their momenta are essential for Belle 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, tracking is achieved through an integrated system consisting of up to four 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 was maintained in a cylindrical volume of 3.4 m diameter and 4.4 m length. The field was 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 surrounded all the sub-detectors up to the KLM. The iron yoke of the Belle detector served as the return path of magnetic flux. The magnetic field was mapped with the accelerator focusing quadrupole magnets located inside the Belle solenoid. The 4% non-uniformity of the magnetic field along the central axis was taken into account by using Kalman filtering [48].

2.2.3.2 Silicon vertex detector

Two different SVD detectors equipped Belle: the first version, SVD1, was used until 2003 and then replaced by the improved version, SVD2. The properties of SVD1 and SVD2 are summarized in Table 2.4. Both versions used double-sided silicon strip sensors. Each sensor was made of 300 μm thick n-type bulk with n-strips on one side and orthogonal p-strips on another. The p-strips were intended to measure $r-\phi$ coordinates, the n-strips the z-coordinate. In addition, p-stops were implanted between the n-strips to isolate each of them from another. The detector was operated under reverse bias of 75 V to reach the full depletion at the pn-junction. A charged particle passing through detector ionizes the

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.

silicon, freeing electron-hole pairs that drift under the electric field, inducing a signal at the strip electrodes implanted at both ends of the depletion region.

	SVD1	SVD2
Number of layers	3	4
Radii of layers (mm)	30.0/45.5/60.5	20.0/43.5/70.0/88.8
Ladders per layer	8/10/14	6/12/18/18
Pitch $r\phi$ (μm)	25	50 (65 for layer 4)
Pitch z (μm)	84	75 (73 for layer 4)

Table 2.4: Comparison of two SVD versions.

SVD1 consisted of three concentric cylindrical layers and covered the polar angle region, $23^\circ < \theta < 139^\circ$, which corresponded to 86% of the solid angle. Each layer was made of several ladders; each ladder contained two to four sensors. SVD1 consisted of 102 sensors and had 81920 readout channels. Figure 2.5 depicts the SVD1 configuration.

SVD2 consisted of four concentric cylindrical layers; the fourth layer allowed to reconstruct charged particles from SVD hits only. The angular coverage was extended to $17^\circ < \theta < 150^\circ$ (93% of the solid angle). SVD2 consisted of 246 sensors and had 110592 readout channels. The configuration of SVD2 is shown in Fig. 2.6. More details are in Ref. [49].

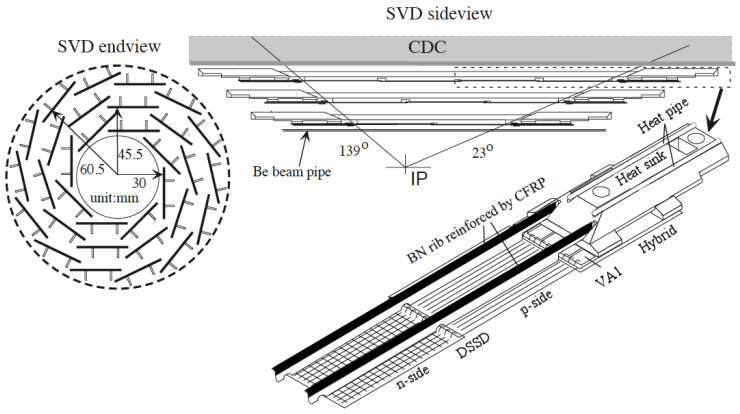


Figure 2.5: Schematic of the first version of the silicon vertex detector (SVD1) shown in a cross-section (left) and in a side view (right).

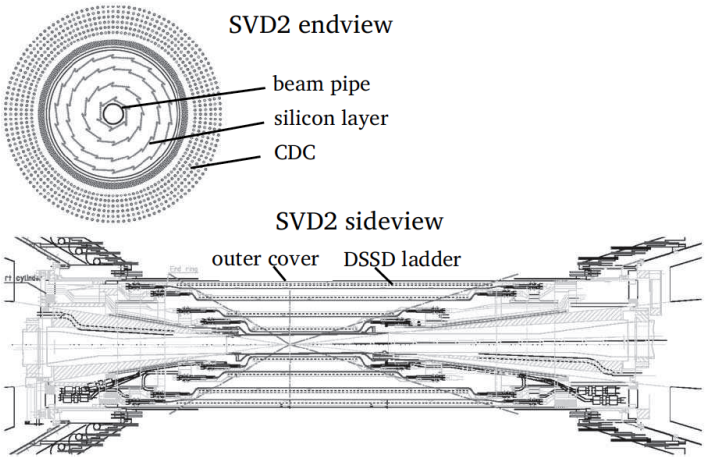


Figure 2.6: Schematic of the second version of the silicon vertex detector (SVD2) shown in a cross-section (top) and in a side view (bottom).

2.2.3.3 Central drift chamber

The CDC allowed to determine charged particles' momentum and charge by measuring their trajectories curved by the axial field. In addition, the CDC provided charged-particle identification by measurements of specific ionization energy-loss, dE/dx.

A charged particle passing through a drift chamber ionizes the gas, liberating electrons and positive ions. The electrons drift in an applied electric field towards the sense wires, where a high gradient causes their acceleration followed by secondary ionization that induces an electric signal. The difference between collision time and drift time is used to infer the particle trajectory.

The CDC structure is shown in Fig. 2.7. The CDC inner (outer) radius is 103.5 mm (874 mm). The CDC was asymmetric in the z-direction to provide the $17° < \theta < 150°$ angular coverage in asymmetric collisions. It contained 50 layers of sense wires, 32 axial layers and 18 small angle stereo layers, and 3 cathode strip layers forming 8400 drift cells. Each drift cell (Fig. 2.8) contained a positively biased sense wire surrounded by eight negatively biased wires strung along the z-axis. Cells were immersed in a low-Z gaseous mixture of 50% helium and 50% ethane. This choice of gas minimizes Coulomb scattering, which degrades

momentum resolution; the large ethane component provides good ionization energy-loss measurement resolution. Axial layers measured the $r - \phi$ coordinate, stereo layers along with cathode strip layers measured the z-coordinate. The spatial resolution in $r - \phi$ was 130 μm, and was better than 2 mm in z.

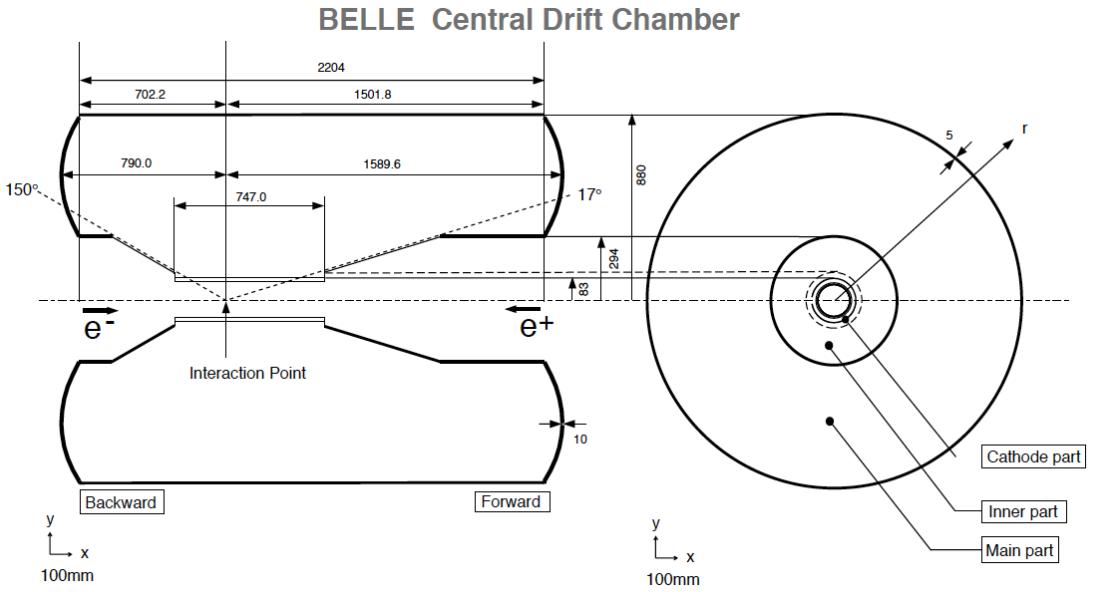


Figure 2.7: CDC view.

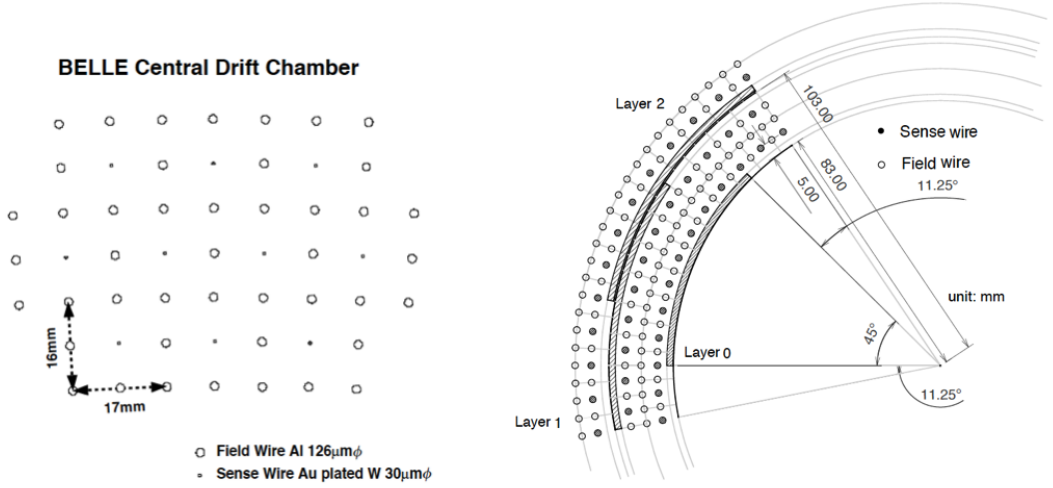


Figure 2.8: CDC cell structure.

The hit amplitude measured from the sense wire was used to determine the ionization energy loss per unit path length, dE/dx. The energy loss depends on the particle's velocity at a given momentum, therefore, dE/dx varies according to particle mass, as shown in Figure 2.9. The measurements of particle's ionization energy loss associated with each CDC hit were combined in a truncated mean. The resolution, determined in a sample of pions from K^0_S decays, was 7.8% which allowed separation of charged pions and kaons of momenta up to 0.8 GeV/c with a 3σ significance [50].

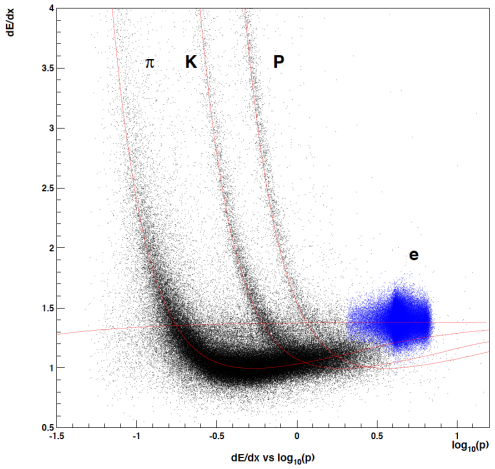


Figure 2.9: Truncated mean of dE/dx versus momentum. The dots are measurements taken during beam operation, the lines are expected distributions for each particle type. Momenta in GeV/c.

2.2.3.4 Tracking performance

Charged track reconstruction mainly relied on the CDC information. Momentum resolution was improved by combining CDC tracks with SVD hit information, especially for low-momentum particles. The transverse momentum resolution was

$$\sigma(p_T)/p_T = 0.0020 p_T[\mathrm{GeV}/c] \oplus 0.0029/\beta \tag{2.1}$$

as shown in Fig. 2.10. The momentum- and angle-dependent SVD1 (SVD2) impact parameter resolutions were

$$\begin{aligned} \sigma_{xy} &= 19 \oplus 54/(p\beta\sin^{3/2}\theta)\mu\mathrm{m}(\sigma_{xy} = 22 \oplus 36/(p\beta\sin^{3/2}\theta)\mu\mathrm{m}) \\ \sigma_z &= 42 \oplus 44/(p\beta\sin^{5/2}\theta)\mu\mathrm{m}(\sigma_z = 28 \oplus 32/(p\beta\sin^{5/2}\theta)\mu\mathrm{m}). \end{aligned} \tag{2.2}$$

A comparison of SVD1 and SVD2 impact parameter resolutions is shown in Fig. 2.11.

2.2.4 Electromagnetic calorimeter

The ECL 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 ECL layout is shown in Figure 2.12. 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 2.5 summarizes the geometrical parameters of each section.

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 consisted of a highly segmented array of 8736 cesium iodide crystals doped with thallium (CsI(Tl)). Thallium shifted the energy of the excitation light into the visible

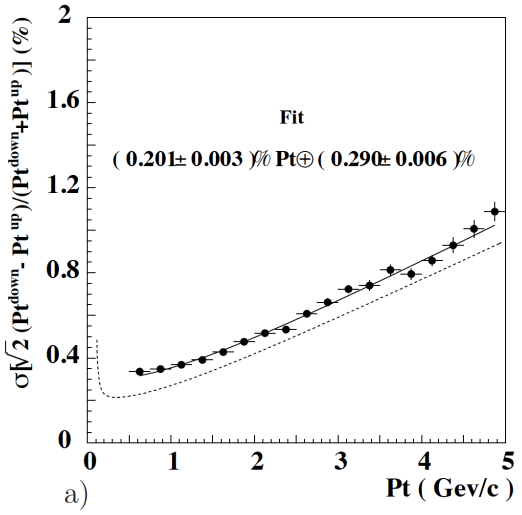


Figure 2.10: Transverse momentum resolution for cosmic rays. The blacks points are the data, black curve is the fit, and the dotted-curve is the idealized case $\beta = 1$.

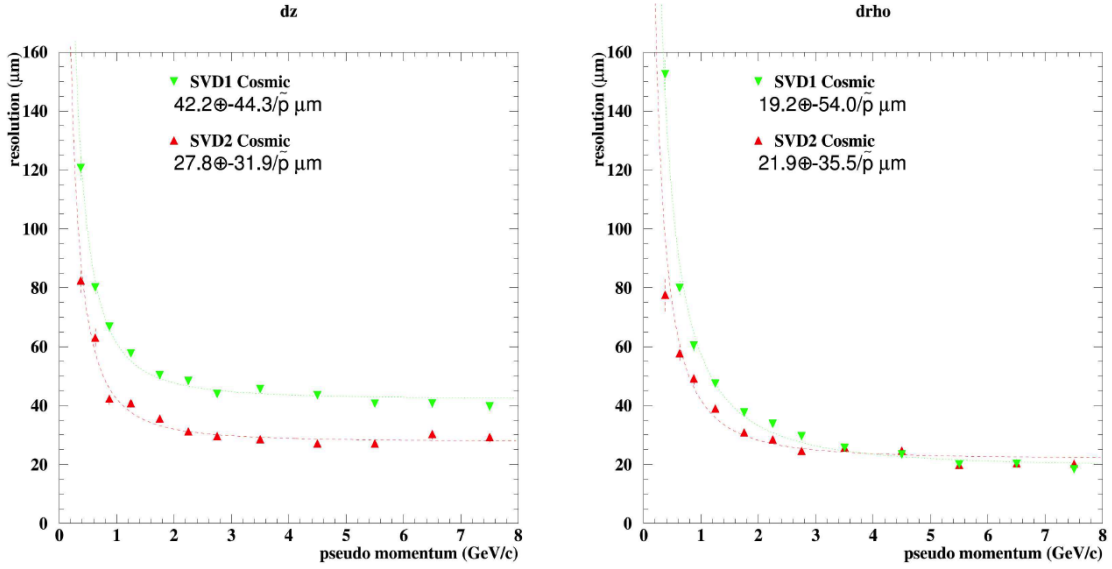


Figure 2.11: Angular and momentum dependent impact parameter resolutions for the two configurations of the vertex detector. The resolution is estimated with cosmic-ray muons recorded during one beam circulation.

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 2.5: ECL geometrical parameters.

spectrum. The light was detected by a independent pair of silicon PIN photodiodes and charge-sensitive preamplifiers installed at the outer end of each crystal.

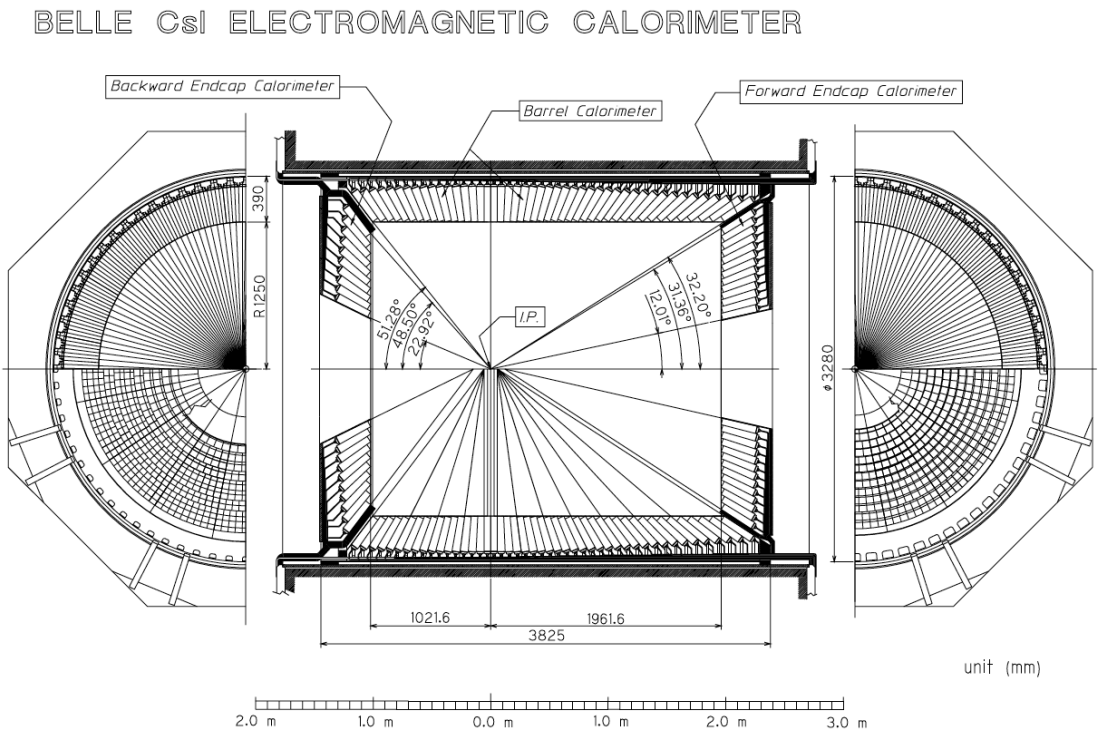


Figure 2.12: ECL layout.

Each crystal was arranged so as to point projectively to the nominal interaction point. The crystals were 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 $16.2X_0$) was chosen to minimize the fluctuations of shower leakages out of the rear of 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 had 55×55 mm^2 front face and 65×65 mm^2 rear face; the dimensions of the crystals in the endcap sections varied from 44.5 to 70.8 mm and from 54 to 82 mm for front and rear face, respectively. The energy and spatial resolutions were

$$\begin{aligned}\frac{\sigma(E)}{E} &= 1.34 \oplus \frac{0.066}{E[\mathrm{GeV}]} \oplus \frac{0.81}{\sqrt[4]{E[\mathrm{GeV}]}} \\ \sigma_{\mathrm{pos}} &= 0.27 \oplus \frac{3.4}{\sqrt{E[\mathrm{GeV}]}} \oplus \frac{1.8}{\sqrt[4]{E[\mathrm{GeV}]}},\end{aligned} \tag{2.3}$$

and are shown in Fig. 2.13. The π^0 mass resolution is 4.8 MeV/c^2 for low-momentum photons in hadronic events. More details are in Ref. [51].

2.2.5 Particle identification

Belle combined measurements of time-of-flight, Cherenkov radiation, and ionization energy loss in the drift chamber to identify charged particles.

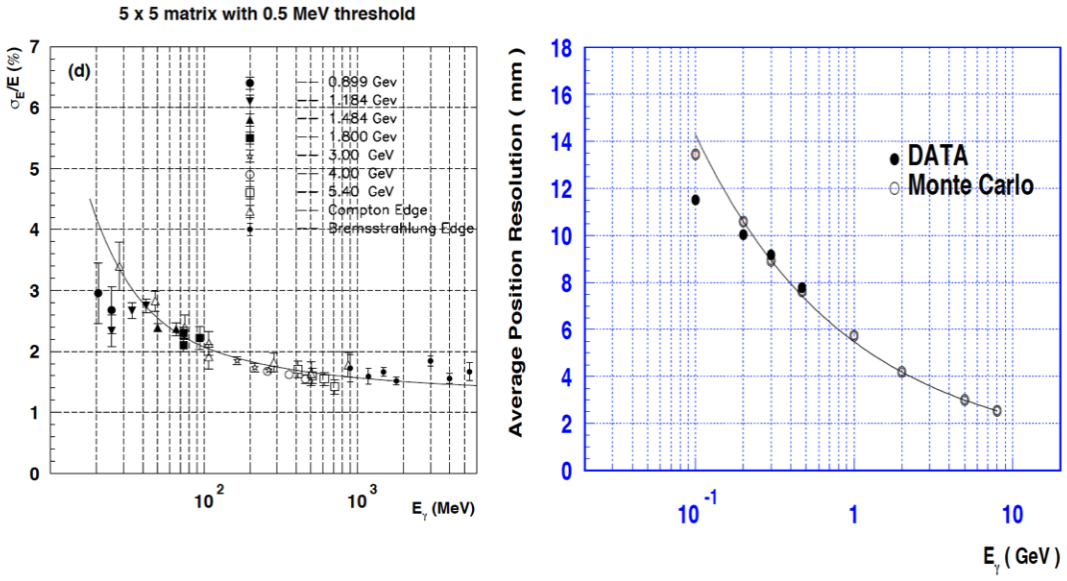


Figure 2.13: ECL energy resolution (left); the black dots correspond to photons of different energies reconstructed in data, the black line is the fitting curve. Position resolution (right); filled circles are the data, empty circles are the simulated data, and the black curve is the fitting curve.

2.2.5.1 Time-of-flight system

The TOF allowed to determine the particle's mass m by measuring the flight time T for a particle of known momentum p to travel the distance L from the interaction point to the TOF module,

$$m = \frac{p}{c}\sqrt{\left(\frac{cT}{L}\right)^2 - 1}, \tag{2.4}$$

where momentum information is provided by the CDC.

The TOF system consisted of 128 time-of-flight counters and 64 thin trigger scintillation counters (TSC), which provided fast timing signal for the trigger. The counters were composed of 64 modules concentrically arranged at 1.2 m radius, covering the polar region of $33° < \theta < 121°$. Each module contained two trapezoid-like shaped time-of-flight counters and one TSC counter separated by a radial gap of 1.5 cm. Scintillation light from a counter was collected by photomultipliers.

The achieved time resolution was 100 ps. The kaon-pion separation as a function of momentum is shown in Figure 2.14. More information is in Ref. [52].

2.2.5.2 Aerogel Cherenkov counter system

The ACC extended the Belle particle-identification capability to momenta beyond the reach of dE/dx and time-of-flight measurements. It distinguished charged pions from charged kaons of momenta 1.2 GeV/c to 3.5 GeV/c.

The ACC was based on detecting the Cherenkov light emitted by charged particles that travel through a medium faster than the speed of light in that medium. The emission angle with respect to the particle direction is inversely proportional to the particle's velocity. The ACC used aerogel with refractive indices such that only pions with momenta between 1 GeV/c and 4 GeV/c produce light.

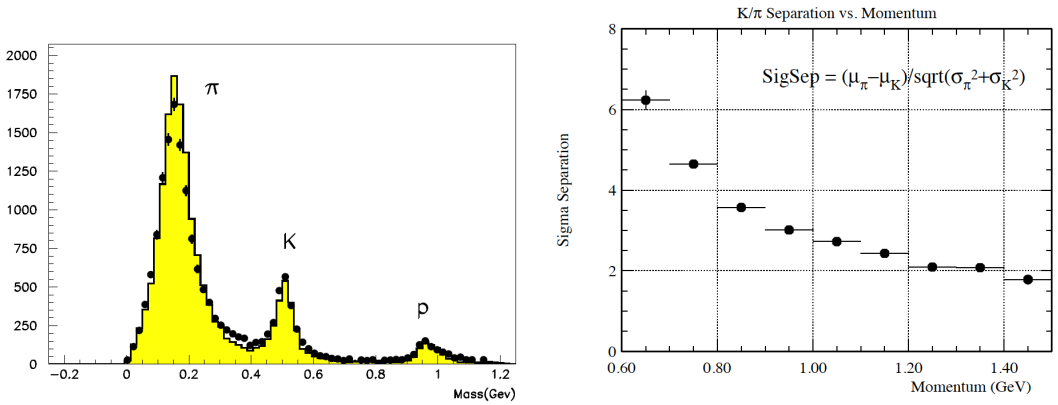


Figure 2.14: (Left) Mass distribution from TOF measurements of particle with momentum below 1.2 GeV/c. Black dots are data, the yellow filled histogram is the simulation. (Right) Kaon/pion separation in the TOF as a function of momentum.

The ACC was divided into a barrel and a forward endcap regions. The barrel compartment contained 960 counter modules segmented into 60 cells in the ϕ direction; the refraction index n varied from 1.010 to 1.028 depending on the polar angle. The endcap part contained 228 counter modules arranged into five concentric layers with index $n = 1.030$. Each module was made of five aerogel tiles stacked in a thin aluminium box of $12\times12\times12$ cm^3 volume. Cherenkov light was detected by one or two fine mesh-type photomultiplier tubes attached directly to the aerogel. Figure 2.15 shows the overall ACC configuration.

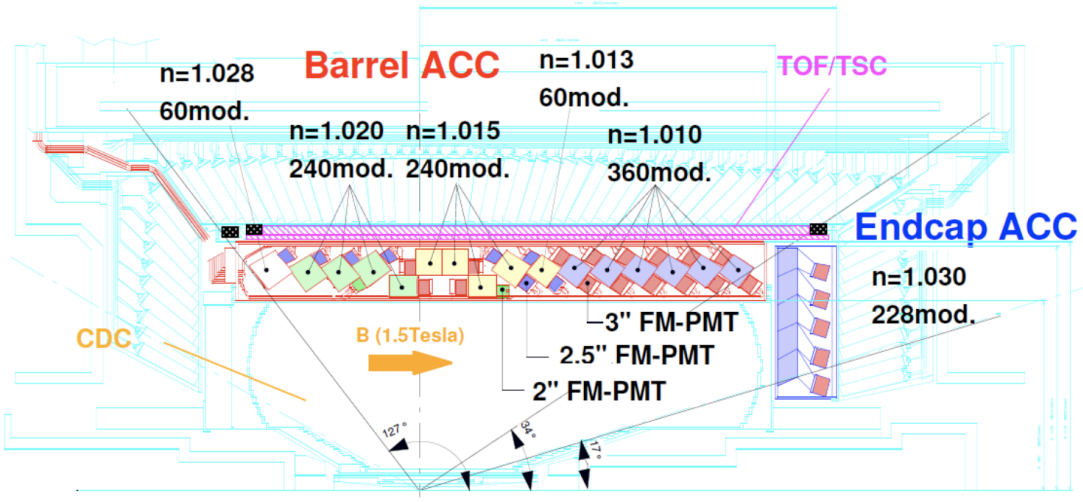


Figure 2.15: Side view of the ACC.

The kaon identification efficiency exceeded 80%, while the pion fake rate remained below 10%. The global Belle K/π separation is shown combining dE/dx measurement, time-of-flight, and ACC in Fig. 2.16. More details are in Ref. [53].

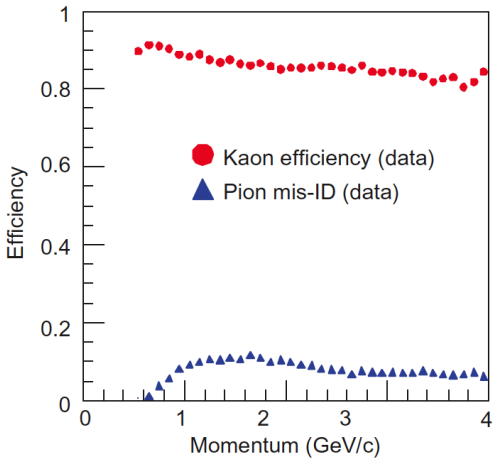


Figure 2.16: Kaon identification efficiency and pion fake rate as functions of momentum from the combination of all Belle PID detectors [54]. Red dots show kaon efficiency and blue triangles show pion fake rate.

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, the goal of the online event selection (trigger) was to distinguish them from background in real time, and to feed them to data acquisition system (DAQ). The physics of interest included hadronic, Bhabha, μ/τ-pair and two photon events. Preferably discarded events included 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 trigger was organized in a multi-stage logic that consisted of the level 1 (L1) hardware trigger, and the level 3 (L3) and level 4 (L4) software triggers. The SVD readout was outside of the standard trigger path and subjected to an early level of TOF-based trigger (Fig. 2.17).

The L1 trigger (Fig. 2.18) consisted of the sub-detector trigger systems and the central trigger system called global decision logic (GDL) [55]. The GDL used the information from all sub-detectors except for the SVD. The GDL received the signal from sub-systems 1.85 μs after bunch crossing and issued a decision 0.35 μs later. Sub-detector triggers were based on track and calorimeter energy information. The CDC provided z and $r-\phi$ track trigger signals. The TOF trigger system provided event timing signal and information on hit multiplicity and topology. The ECL provided two triggers based on cluster multiplicity and total energy deposition. These two categories of triggers were already sufficient for efficient signal-to-background discrimination. The KLM provided a high-efficiency trigger for muons. The GDL kept the hadronic events based on three independent triggers, which required either three or more charged track candidates, high levels of deposited energy in the ECL (with a veto on the ECL Bhabha trigger) or four isolated neutral clusters in the ECL. The overall efficiency for hadronic events was estimated to exceed 99%.

The L3 trigger was based on software algorithms and was used to further reduce the number of events to be stored. It required at least one track to originate in the interaction point region and an impact parameter in z smaller than 5.0 cm, using an ultra-fast track finder. All events that passed the L3 trigger were stored. The trigger could retain 99% of the physics events of interest, while reducing overall event rates by 50–60%.

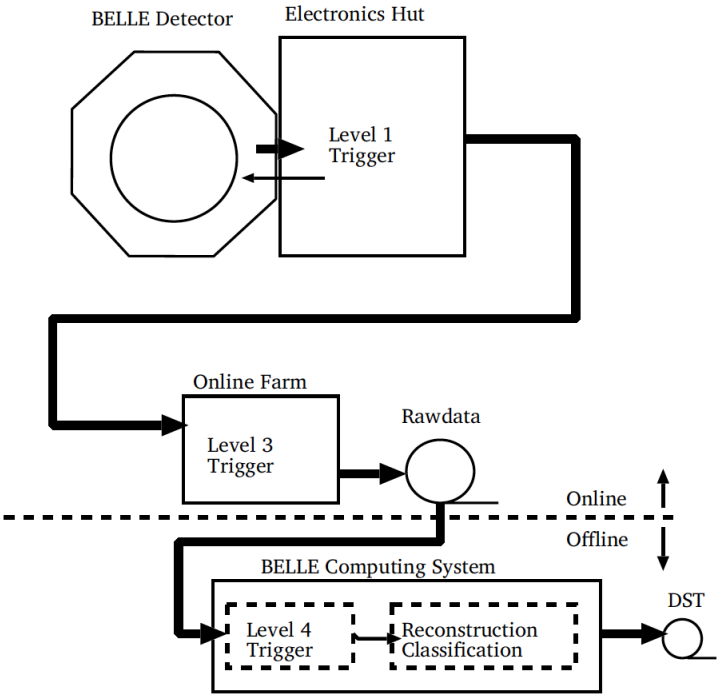


Figure 2.17: Overview of the Belle trigger data flow.

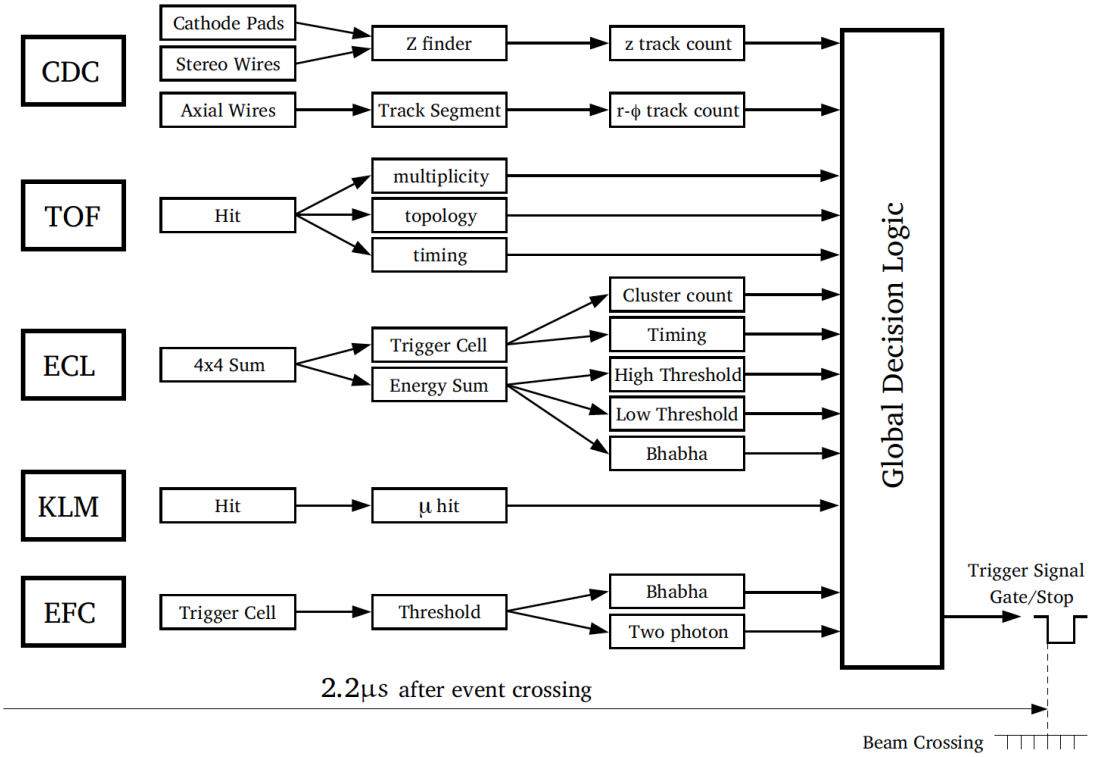


Figure 2.18: Schematic view of the L1 trigger logic.

The L4 trigger was similar to the L3 but imposed stricter requirements. An event was required to have at least one track with $p_T > 300$ MeV/c and impact parameter in r and

z smaller than 1 cm and 4 cm, respectively. The L4 trigger rejected around 78% of the events, while keeping almost 100% of the $B\overline{B}$ events.

The trigger accept rates for physics processes of interest at the design instantaneous luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$ are given in Table 2.6.

Process	Rate, Hz
$e^+e^- \to \Upsilon(4S)$	12
$e^+e^- \to q\bar{q}(q = u, d, s, c)$	28
$e^+e^- \to \mu^+\mu^-/\tau^+\tau^-$	16
$e^+e^- \to e^+e^-$ (Bhabha scattering) $\theta_{lab} > 17°$	4.4
$e^+e^- \to \gamma\gamma$ $\theta_{lab} > 17°$	0.24
Two photon events ($\theta_{lab} > 17°$ & $p_T \geq 0.1$ GeV/c)	≈ 35

Table 2.6: Trigger accept rates of various physics processes at $10^{34}\text{cm}^{-2}\text{s}^{-1}$ luminosity. Bhabha and $\gamma\gamma$ rates are pre-scaled by a factor of 100 because of their large cross section.

The total trigger rate of the events of interest was around 100 Hz. Figure 2.19 shows the Belle trigger rate change over time.

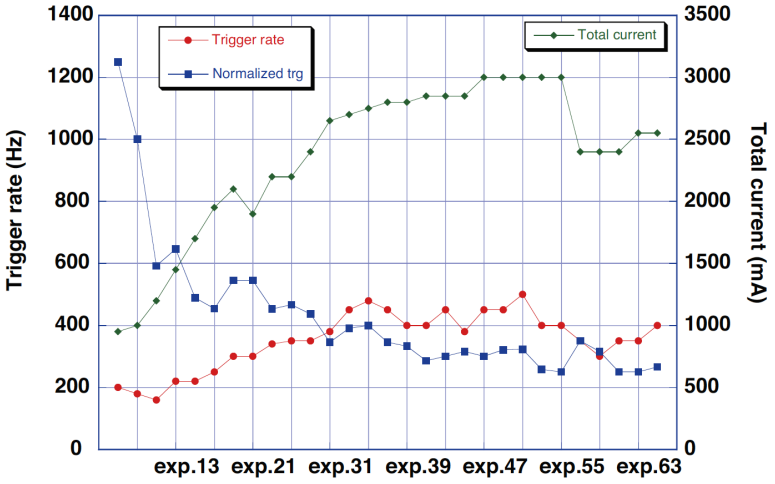


Figure 2.19: Average trigger rate as a function of the experiment number [47].

The DAQ collected the information read out from the sub-detectors for the events that passed the L1 trigger with a rate up to 500 Hz and a dead time of less than 10%. A schematic of the DAQ system is shown in Fig. 2.20. The DAQ was segmented into 7 sub-systems running in parallel, each handling the data from a sub-detector. The analog signals read out from the sub-detectors were digitized by time-to-digital converters. Most of the detectors except for the SVD and KLM were read out using charge-to-time conversion. The SVD signal was converted using flash analog-to-digital instead of time-to-digital converter. When the sequence control received a GDL signal, the data from every sub-detector were sent to the event builder, which combined detector-by-detector data into event-by-event data. The event data was then transferred to the L3 trigger, which wrote it to the raw data storage. The L4 trigger was then applied to the raw data. Events satisfying its criteria were converted to fully reconstructed data, stored in DST files. The size of a DST of a typical hadronic event was 100 kB.

The large amount of information stored in DST files were reduced into mini-DST [56] 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.

We analyze the data passed through the "HadronBJ" skim [57], which imposes a set of loose requirements to isolate hadronic meson events. The mini-DST files are produced through the Belle analysis software framework (BASF) [58], and are converted and analyzed using Belle II analysis software framework (BASF2) [59, 60].

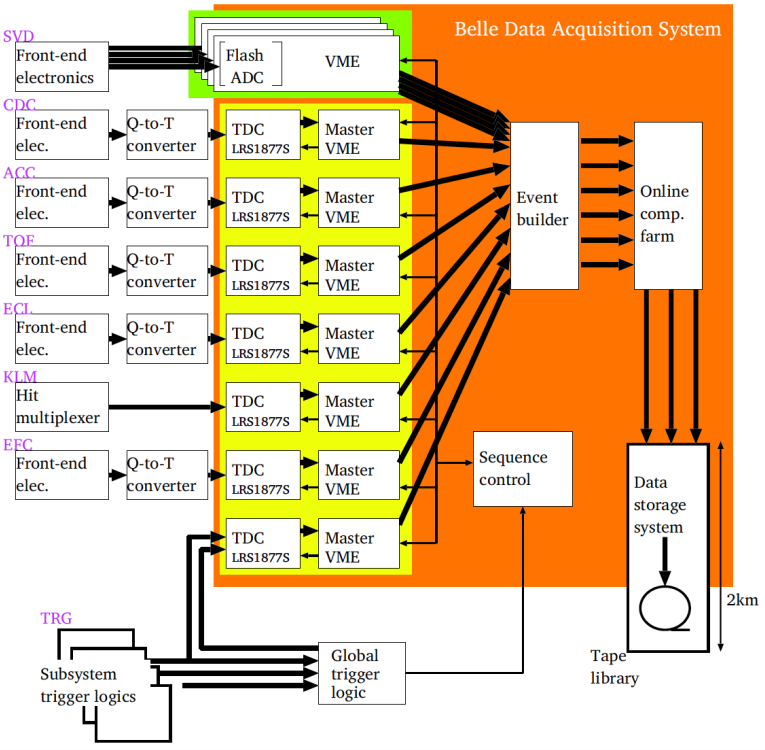


Figure 2.20: Schematic view of the Belle DAQ system.

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 both ρ meson spin projections are along their momenta and helicity λ equals zero while transversely polarized decays proceed through the $L = 1, 2$ states, implying that ρ meson helicities equal ± 1.

The different polarizations are distinguished experimentally through the decay angles of the pions with respect to each other and the ρ mesons. We use the helicity basis to define the angles. 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 the spherical harmonics [61, 62] as

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

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

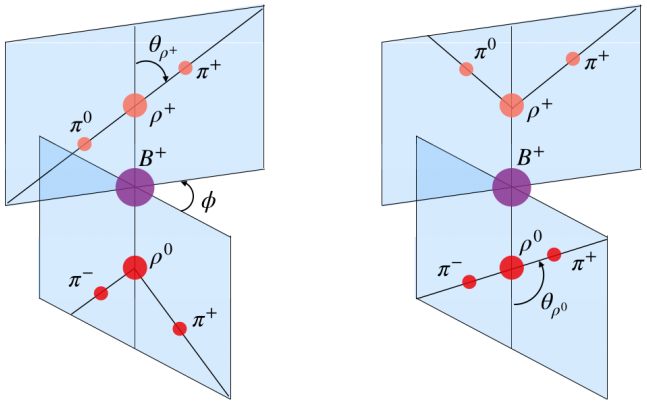


Figure 3.1: Geometric definition of the helicity basis angles.

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

The integration over 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

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

where $|H|^2 \equiv |H_0|^2 + |H_+|^2 + |H_-|^2$. Simulation, and various previous Belle analyses, show that the azimuthal acceptance for signal is uniform, supporting the derivation of Eq. (3.4).

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 performed 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 $B^0 \to \rho^+\rho^-$ and $B^0 \to \rho^0\rho^0$ decays are based on the full data set, corresponding to 772×10^6 $B\overline{B}$ pairs, while the $B^+ \to \rho^+\rho^0$

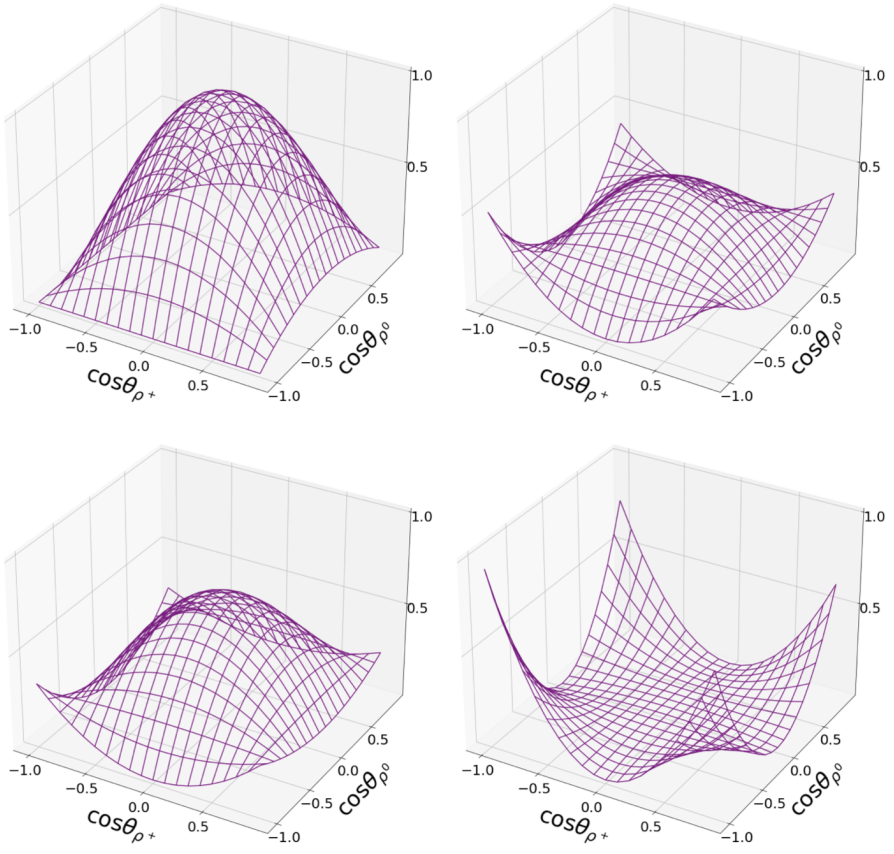


Figure 3.2: Theoretical two-dimensional angular distributions for $f_L = 0$, $1/3$, $1/2$, and 0.95, respectively.

measurement is only based on $85{\times}10^6$ $B\overline{B}$ pairs, just about 10% of the sample. The BaBar measurements on the $B^0 \to \rho^0\rho^0$ and $B^+ \to \rho^+\rho^0$ 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 is based on $383.6{\times}10^6$ $B\overline{B}$ pairs. The LHCb experiment studied only the $B^0 \to \rho^0\rho^0$ decay with a sample corresponding to 3 fb^{-1} of integrated luminosity.

The $B^+ \to \rho^+\rho^0$ decay is the only among $B \to \rho\rho$ decays that was not studied using the full Belle data set. This has become the major limiting factor for the precision of the CKM angle α. I therefore decide to target the measurement of the properties of the $B^+ \to \rho^+\rho^0$ decay in the full Belle data set.

3.3 *B* decay reconstruction at Belle

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

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,

	N_{sig}	$\mathcal{B} \times 10^{-6}$	f_L	$\mathcal{S}_{CP}$	$\mathcal{A}_{CP}$
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 [39–42, 63–65].

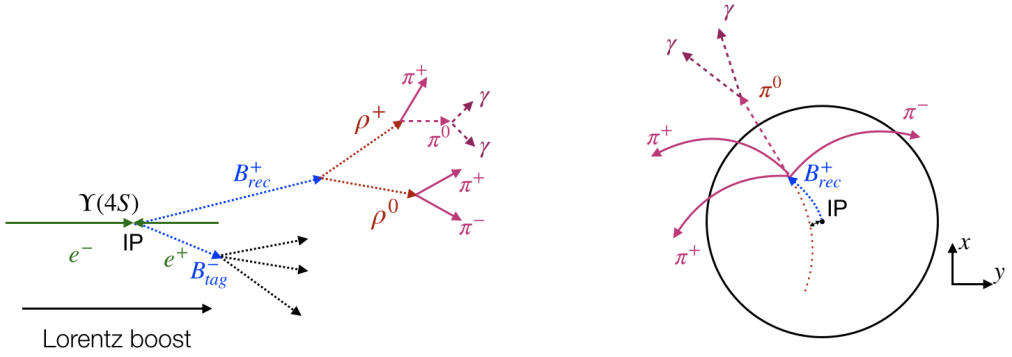


Figure 3.3: Sketch of a $B^+ \to \rho^+(\to \pi^+\pi^0(\to \gamma\gamma))\rho^0(\to \pi^+\pi^-)$ decay in the (left) longitudinal and (right) transverse views.

as opposed to QED events, and save them onto permanent memory.

The 1.5 ps B meson lifetime, combined with a typical momentum of 2 GeV/c in the laboratory frame, results in a flight length of about 150–200 μm. This allows reconstructing the decay position for background discrimination and for measurements of quantities depending 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 1–2 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 neutral pion kinematic properties 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 `RAVE` [66] 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` [67]). 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, where light mesons, such as pions, originated from light quark-antiquark pairs produced in the e^+e^- collision, form random combinations that accidentally meet the reconstruction and selection requirements;

B **decays** (other than signal), 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 [68]. 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 exploits the available event information to construct discriminating variables at particle-, candidate-, and event-level to separate signal from background.

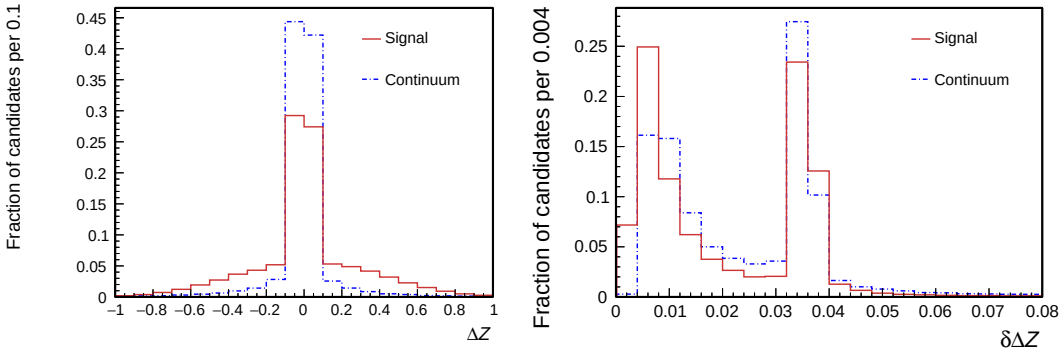


Figure 3.4: Distribution of (left) ΔZ and (right) $\delta\Delta Z$ in simulated signal $B^+ \to \rho^+\rho^0$ and continuum events. Distributions normalized to unity.

3.3.2 Particle-level variables

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

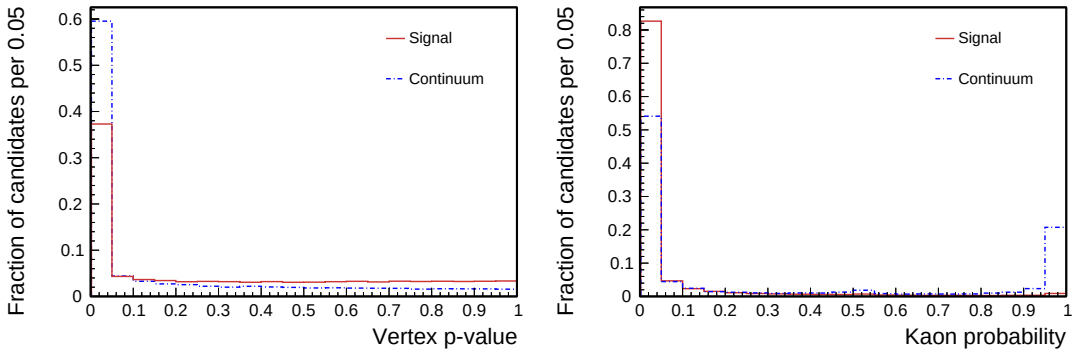


Figure 3.5: Distribution of (left) vertex quality p-value and (right) probability of pion to be reconstructed as kaon in simulated signal $B^+ \to \rho^+\rho^0$ and continuum events. Distributions normalized to unity.

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 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 KEKB or Belle 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 at Belle 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 $M_{\rm bc}$, defined as

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

  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, $M_{\rm bc}$ 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 $M_{\rm bc}$ 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 that drops to zero at the kinematic limit of half collision energy. This makes $M_{\rm bc}$ 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 an example of a two-dimensional ΔE–$M_{\rm bc}$ distribution for simulated $B^0 \to K^+\pi^-$ decays reconstructed from a realistic simulated e^+e^- sample and their one-dimensional projections, which are chosen for their exemplifying value. Continuum background is represented in blue, $B^0 \to K^+\pi^-$ events are represented in red, and B-background $B^0 \to \pi^+\pi^-$ events, where a pion is misidentified as a kaon, are represented in green. In the $M_{\rm bc}$ distribution, both $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$ decays peak at the B mass, while in the ΔE distribution the $B^0 \to \pi^+\pi^-$ component is shifted from the signal peak by about 0.05 GeV.

$M_{\rm bc}$ and ΔE are correlated by common inputs, the beam energy, the measured momentum of charged particles, and the reconstructed energy of neutrals. The correlation becomes severe if the final-state particles contain high-energy photons, because of photon-energy leakage in the calorimeter. Correlation is reduced by a modified $M_{\rm bc}$ quantity that includes a correction of the magnitude of the ρ^+ momentum but retains its direction to constrain the reconstructed B energy to be the beam energy, $M'_{\rm bc} = \sqrt{s/4 - |\vec{p'}_B^{\,*}|^2}$, where $\vec{p'}_B^{\,*} = \vec{p}_{\rho^0}^{\,*} + \frac{\vec{p}_{\rho^+}^{\,*}}{|\vec{p}_{\rho^+}^{\,*}|}\sqrt{(\sqrt{s}/2 - E_{\rho^0}^{*})^2 - m_{\rho^+}^2}$, with $\vec{p}_{\rho^0}^{\,*}$ being ρ^0 momentum, $\vec{p}_{\rho^+}^{\,*}$ the ρ^+ momentum, $E_{\rho^0}^{*}$ the ρ^0 energy, and m_{ρ^+} the known ρ^+ mass [10].

Figure 3.7 shows the two-dimensional distributions of ΔE as a function of $M_{\rm bc}$ and of $M'_{\rm bc}$ for simulated $B^+ \to \rho^+\rho^0$ decays and background. In the simulated background

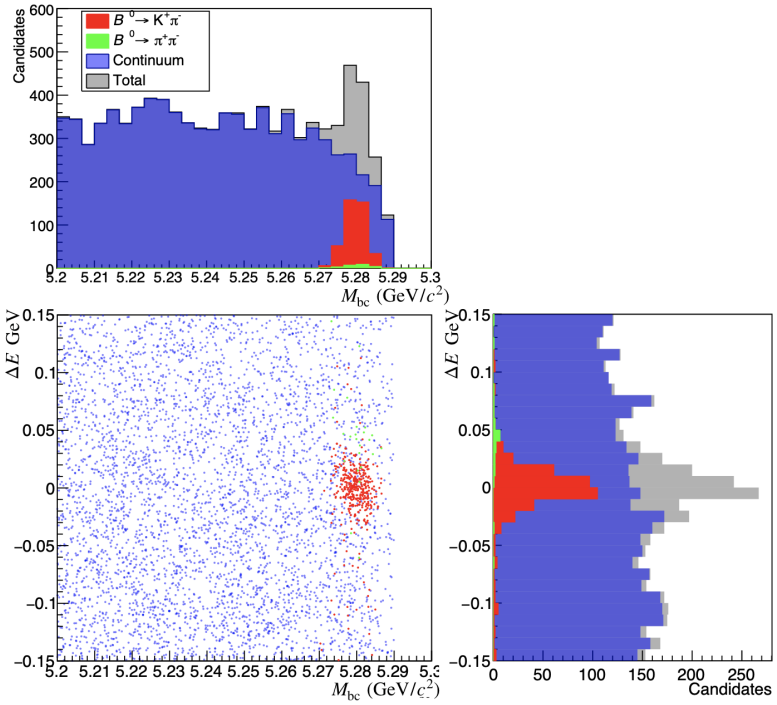


Figure 3.6: Distribution of ΔE as a function of $M_{\rm bc}$ for simulated $B^0 \to K^+\pi^-$ decays. Continuum background is represented in blue, $B^0 \to K^+\pi^-$ decays are represented in red, and $B^0 \to \pi^+\pi^-$ decays where a pion is misidentified as a kaon are represented in green.

sample, a -3.2% Pearson correlation exists between ΔE and $M_{\rm bc}$ (bottom left plot), which reduces to 0.3% for $M'_{\rm bc}$ (bottom right plot). In the simulated signal sample, a 7% Pearson correlation exists for both pairs of observables (top plots). As the dependence between ΔE and $M_{\rm bc}$ is reduced, I use the modified beam constrained mass $M'_{\rm bc}$ in the analysis.

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 3.8 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 $\Upsilon(4S)$ 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

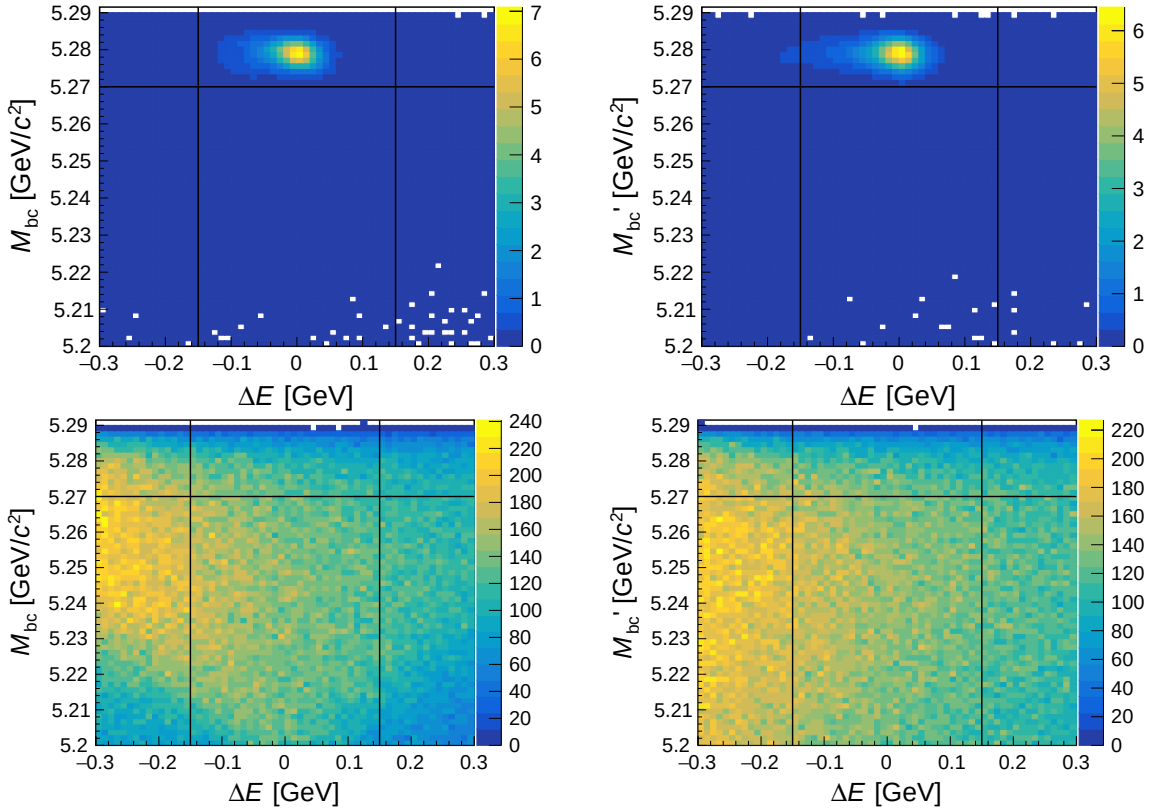


Figure 3.7: Distributions of (left) $M_{\rm bc}$ and (right) $M'_{\rm bc}$ as functions of ΔE for $B^+ \to \rho^+\rho^0$ candidates reconstructed in simulated (top) signal and (bottom) background events.

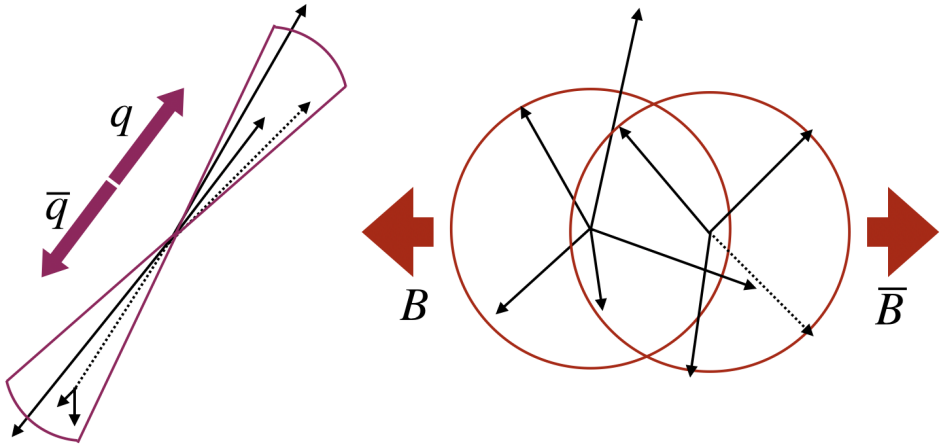


Figure 3.8: Event-shape sketch for a (left) pencil-like continuum and (right) sphere-like $B\overline{B}$ event.

from continuum and it is exploited through several variables. Relevant examples include

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. 3.9).

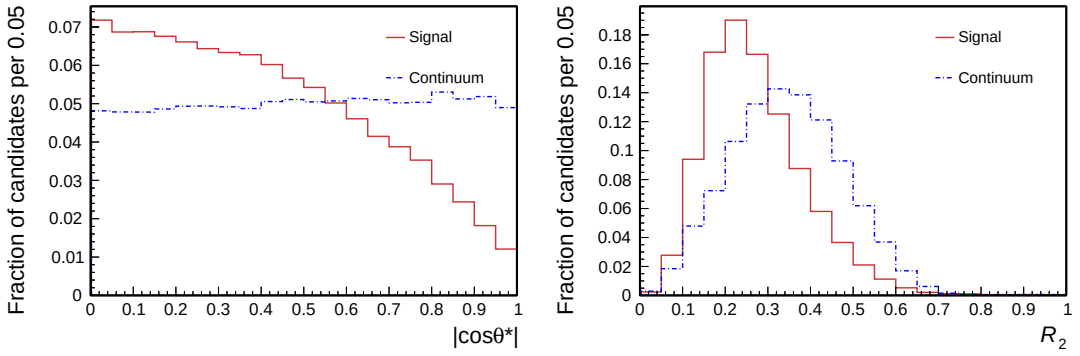


Figure 3.9: Distribution of (left) $|\cos\theta^*|$ and (right) R_2 for simulated signal and continuum data. Distributions normalized to unity.

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 [69]. 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, $T_{\rm sig}$ (Fig. 3.10, left panel), and for the momenta of all the other particles in the event, $T_{\rm tag}$. Continuum events tend to cluster at higher values of the $T_{\rm sig}$ distribution, unlike signal.

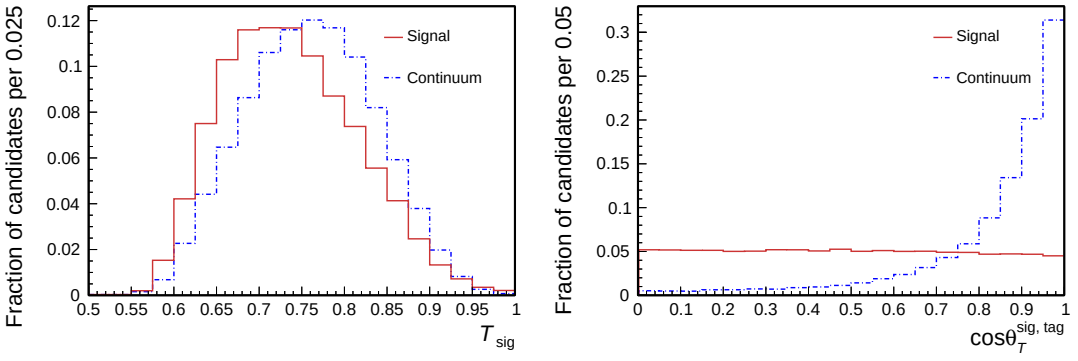


Figure 3.10: Distribution of (left) $T_{\rm sig}$ and (right) $|\cos\theta_{\rm T}|$ in simulated signal and continuum data. Distributions normalized to unity.

Thrust angles. The cosine of the angle between the $\mathbf{T}_{\rm sig}$ and $\mathbf{T}_{\rm tag}$ axes, $|\cos\theta_{\rm T}|$, is also used to discriminate between $B\overline{B}$ and continuum events offering one of the most powerful discriminators. Figure 3.10 (right) shows its distribution for simulated signal and continuum events. Since the momenta of the decay products of B^+ and B^- are isotropically distributed, $\mathbf{T}_{\rm sig}$ and $\mathbf{T}_{\rm tag}$ are randomly distributed, and $|\cos\theta_{\rm T}|$ follows a uniform distribution. For $q\overline{q}$ events, particle momenta are collimated, resulting in strongly directional $\mathbf{T}_{\rm sig}$ and $\mathbf{T}_{\rm tag}$ vectors and yielding a $|\cos\theta_{\rm T}|$ distribution peaked at 1.

CLEO cones. The CLEO collaboration introduced nine variables capable of offering additional discrimination against continuum [70]. "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. 3.11, 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. 3.11, 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 [71]. The most discriminating example 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. 3.9).

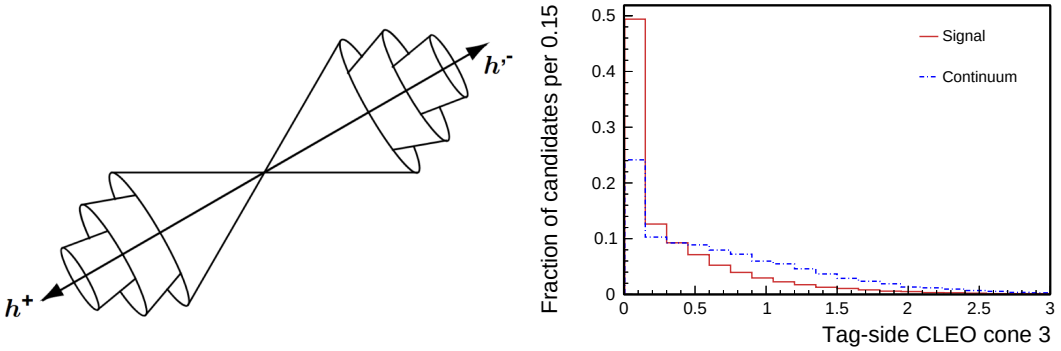


Figure 3.11: (Left) sketch of the first three CLEO cones. (Right) distribution of the third CLEO cone for simulated signal and continuum data. Distributions normalized to unity.

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) [72]. 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 daughters, 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 3.12 shows the distribution of the KSFW moments for $\ell = 2$, which offer the higher discrimination power.

Missing mass and transverse energy. The missing mass squared is defined as

$$M^2_{\mathrm{miss}} = \left(\sqrt{s} - \sum_{i=1}^{N} E^*_i\right)^2 - \sum_{i=1}^{N} |\mathbf{p}^*_i|^2, \tag{3.11}$$

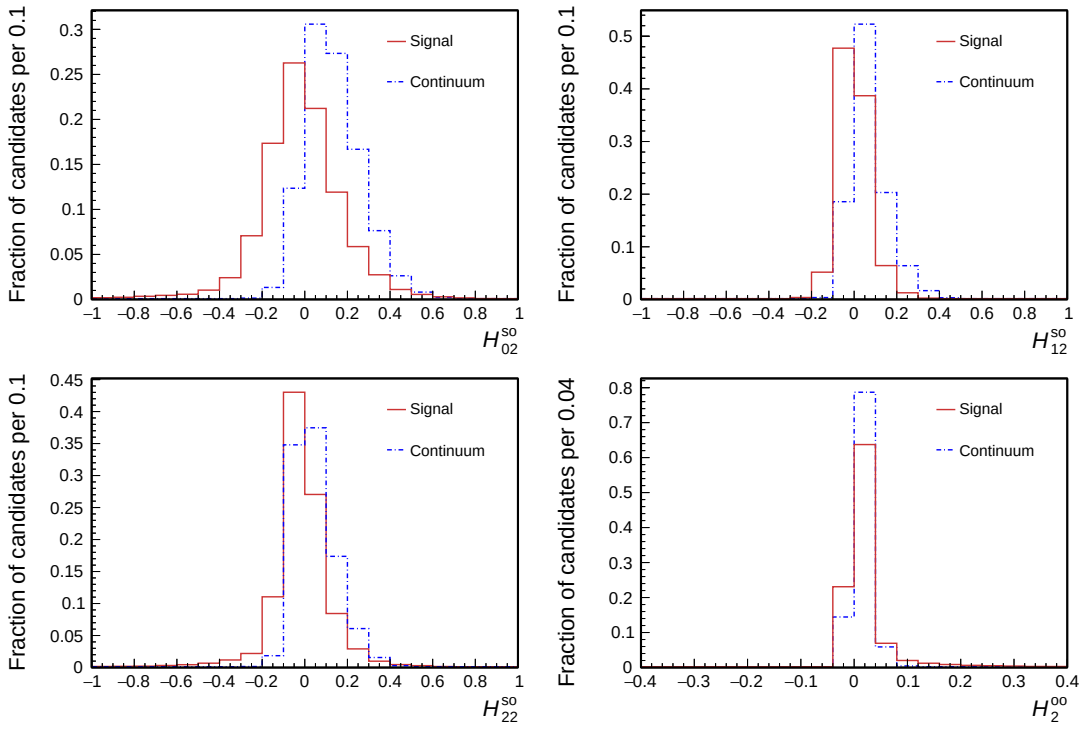


Figure 3.12: Distributions of the $\ell = 2$ KSFW moments for simulated signal and continuum events. Distributions normalized to unity.

where E_i^* and $\mathbf{p}_i^*$ are the energy and the momentum of the ith particle in the $\Upsilon(4S)$ frame, and N is the total number of particles in the event.

The transverse energy is the scalar sum of the transverse momenta of all particles in the event

$$E_{\rm T} = \sum_{i=1}^{N} |(\mathbf{p}_{\rm T}^*)_i|. \tag{3.12}$$

Transverse energies in continuum events are typically lower than in $B\overline{B}$ events. The distributions of $M_{\rm miss}^2$ and $E_{\rm T}$ are shown in Fig. 3.13.

These variables provide additional background discrimination for events that contain decays involving neutrinos, the majority of which are semileptonic B decays. Hence the squared missing mass and transverse energy variables are also useful to separate $B\overline{B}$ from continuum events.

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 ($r = 0$ or 1), proportional to the probability that the identification is correct, where zero means no flavor discrimination, and one means unambiguous flavor assignment. Hence, continuum events are expected to

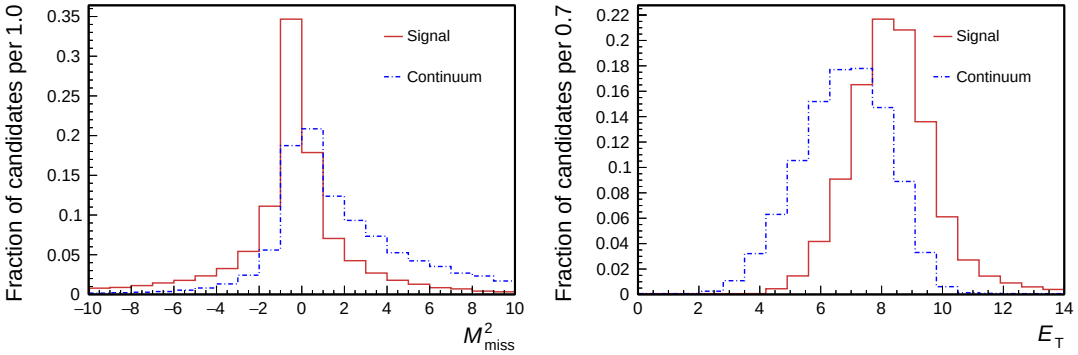


Figure 3.13: Distribution of (left) the M^2_{miss} and (right) E_{T} for simulated signal and continuum events. Distributions normalized to unity.

cluster close to zero of $q \cdot r$ distribution, while B events peak at $q \cdot r = \pm 1$ (Fig. 3.14, top left panel).

Additional discriminating flavor-related variables include (Fig. 3.14),

|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^*.

3.4 Analysis overview

This work aims at measuring the branching fraction, charge-parity 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 tuning the measurement procedure towards the expected result, reducing the chances for bias.

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 40 kinematic, decay-time, and final-state topology 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 average statistical precision on the final 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

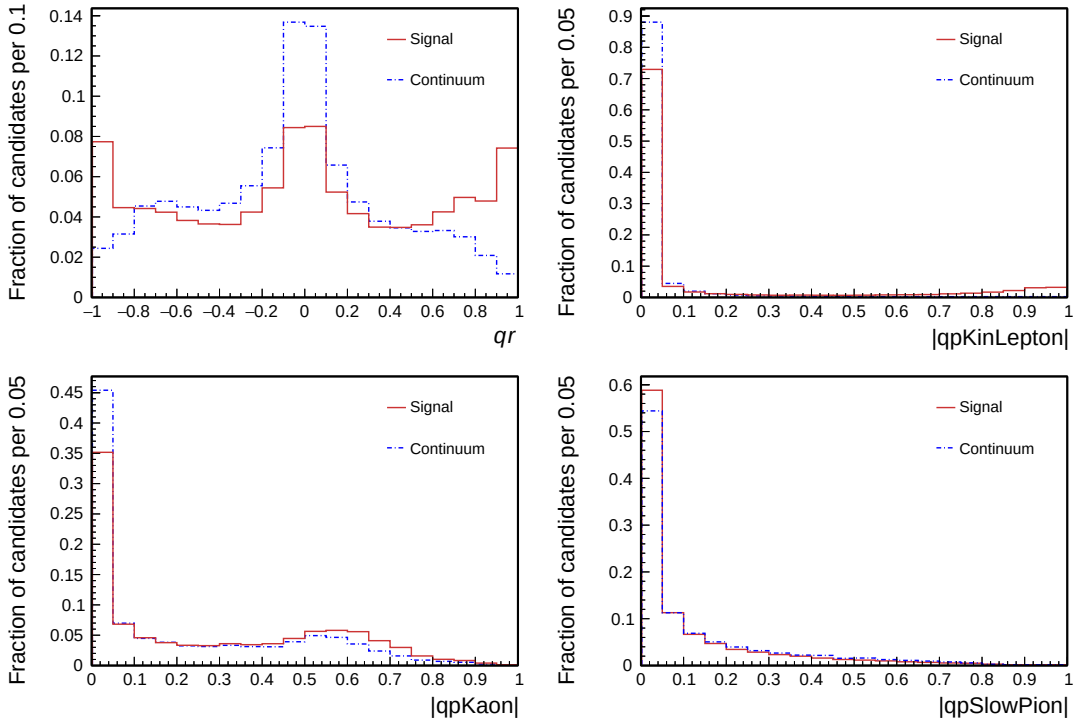


Figure 3.14: Distributions of (top left) qr, (top right) |qpKinLepton|, (bottom left) |qpKaon|, and (bottom right) |qpSlowPion| in simulated signal and continuum events. Distributions normalized to unity.

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 and lack of a reliable model is the main obstacle that prevented the Belle experiment to update its 2003 early result to date. The outcome is a robust model as shown by tests of the estimator properties using simulated data under various experimental conditions.

To probe possible residual undetected issues and possible effects that are not reproduced accurately by the simulation, I implement the key steps of the analysis on well-known abundant control samples that share relevant features with the signal decay.

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 data sample. The flow of the analysis is shown in Fig. 3.15.

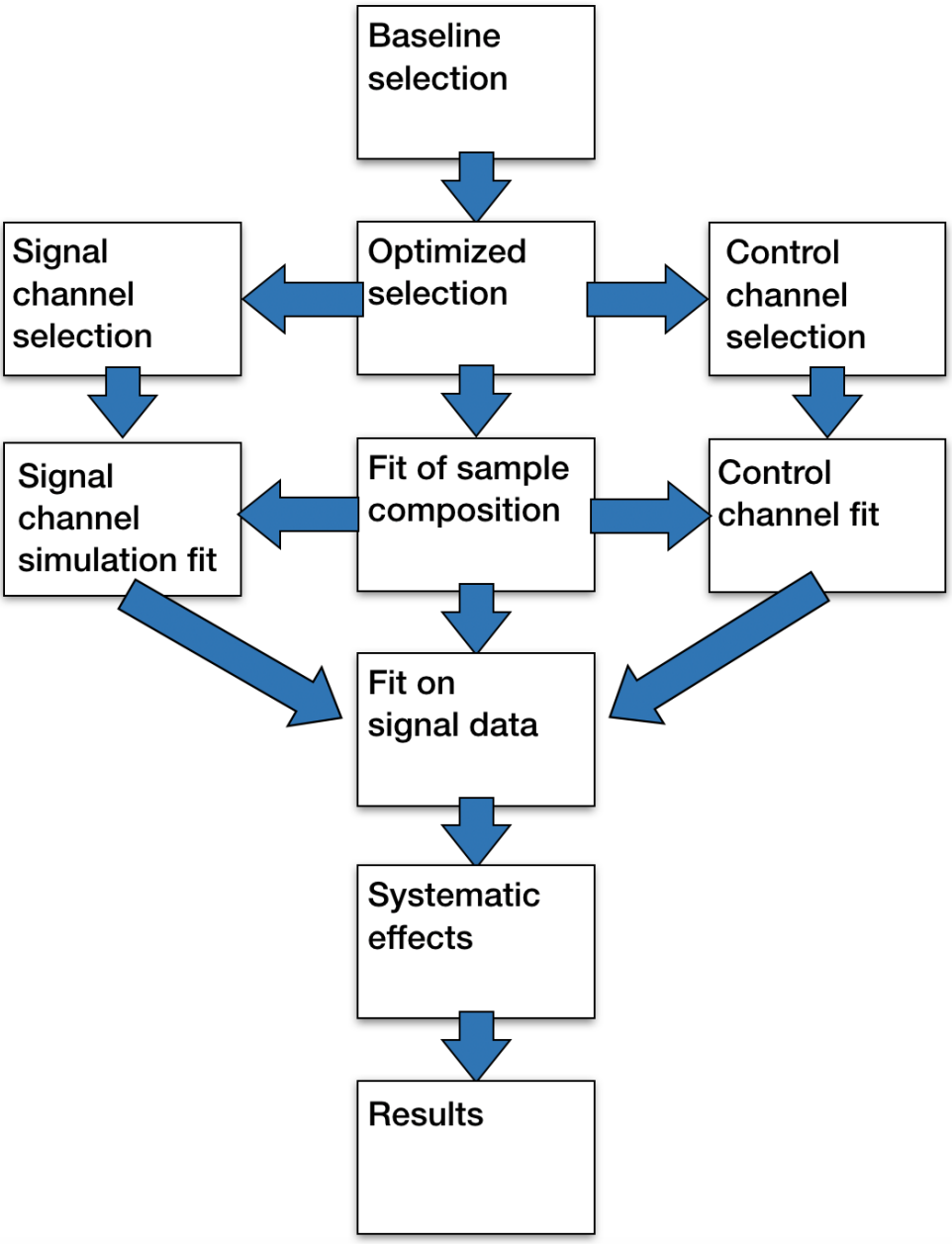


Figure 3.15: Block-diagram illustrating main steps of the analysis.

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 experiment operated from June 1999 to July 2010, accumulating electron-positron collision data corresponding to an integrated luminosity of 1.04 ab^{-1}. I use the full data set collected at the energy of the $\Upsilon(4S)$ resonance mass, corresponding to 711 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"), and corresponding to 89.4 fb^{-1} of integrated luminosity, to characterize and study continuum background. All data are processed with the final set of Belle calibrations and alignment constants, reconstructed using the final version of reconstruction algorithm, and converted into Belle II data-format (`ROOT`)[1].

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 software packages that 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 observed quantities. 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. Truth matching procedure is useful

[1] Version details for collaborators: `b2bii` version `release-03-01-00`

to optimize selection requirements, calculate signal efficiency, classify sample components, and for many other consistency checks.

Figure 4.1 shows a sketch of the generation sequence for an hadronic event in Belle. 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` [73]. Finally, the decay of the heavy hadron (top right corner in Fig. 4.1) is generated according to a certain `EvtGen` model [74]. The photon emission by final-state charged particles is generated by `PHOTOS` [75].

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 decay 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. It is therefore essential to properly model them in simulation.

The sample is then passed through the standard Belle detector simulation model `GSIM`, based on the CERN `GEANT3` package [76], 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 50×10^6 $B^+ \to \rho^+\rho^0$ decays for both polarizations, which are about a factor 2400 more than expected in data. Such a large amount of simulated data is needed to sample independent ensembles of hundreds of simulated experiments for studies of estimator properties.

For continuum background studies, I use centrally produced simulated samples corresponding to six times of the full Belle data set size. These samples include $e^+e^- \to u\overline{u}$, $e^+e^- \to d\overline{d}$, $e^+e^- \to s\overline{s}$, and $e^+e^- \to c\overline{c}$ events generated using `Pythia` and `PHOTOS`, and passed through `GSIM`. `Pythia` simulates parton interactions and parton showers according to phenomenological models.

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, which are split in

- **a favored B decay sample**, which contains B decays with branching fractions $\mathrm{O}(10^{-3})$ due to $b \to c$ transitions; the size of the sample corresponds to $10\times$ the full Belle data set;
- **a rare charmless B decay sample**, which contains B decays with branching fractions $\mathrm{O}(10^{-5})$ due to $b \to q$ ($q = u, d, s$) transitions; the size of the sample corresponds to $50\times$ the full Belle data set;
- **a rare charmonium B decay sample**, which contains B decays with branching fractions $\mathrm{O}(10^{-4})$ with $c\overline{c}$ (charmonium) in the final state; the sample size corresponds to $100\times$ the full Belle data set;
- **a rare semileptonic B decay sample**, which contains B decays with branching fractions $\mathrm{O}(10^{-4})$ following $b \to u\ell\nu$ transitions; the sample size corresponds to $20\times$

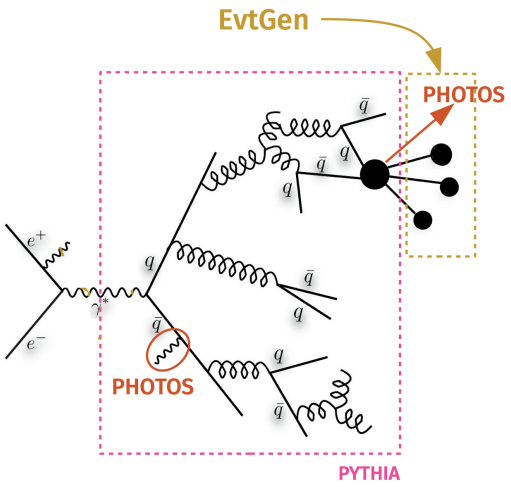


Figure 4.1: Sketch of the various portions of each simulation package models of an hadronic event in Belle.

the full Belle data set;

These samples are centrally generated and processed as the experimental data.

Exclusive simulated samples each dedicated to a specific decay mode are also produced for peaking background and control 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. These are used mostly for prototyping the analysis and assessing systematic uncertainties.

4.1.3 Basic data structures

The full Belle raw data is about 2 PB in size. As many physics topics are studied, these 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 further reduced in size so that each collaborator can access them. The most relevant of such subsets for our analysis is the hadronic stream.

4.1.4 Hadronic stream 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 2.6). 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.

Decay channel	f_L	N ($\times 10^6$)	EvtGen model
$B^+ \to \rho^+(\to \pi^+\pi^0)\rho^0(\to \pi^+\pi^-)$	1	50	PHOTOS SVV_HELAMP
$B^+ \to \rho^+(\to \pi^+\pi^0)\rho^0(\to \pi^+\pi^-)$	0	50	PHOTOS SVV_HELAMP
$B^+ \to \overline{D}^0(\to K^+\pi^-\pi^0)\pi^+$		20	PHOTOS PHSP
$B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$		20	PHOTOS PHSP
$B^+ \to \overline{D}^0(\to K^+\pi^-)\rho^+(\to \pi^+\pi^0)$		20	PHOTOS SVS
$B^0 \to \rho^+\rho^-$	1	12	PHOTOS SVV_HELAMP
$B^+ \to K^{*0}\rho^+$	0.48	12	PHOTOS SVV_HELAMP
$B^+ \to \rho^+ f_0$		12	PHOTOS SVS
$B^+ \to a_1^+\pi^0$		12	PHOTOS SVS
$B^0 \to a_1^\pm\pi^+$		12	PHOTOS SVS
$B^+ \to a_1^0\pi^+$		12	PHOTOS SVS
$B^+ \to \rho^0\pi^+\pi^0$		12	PHOTOS PHSP
$B^+ \to \rho^+\pi^+\pi^-$		12	PHOTOS PHSP
$B^+ \to \rho^-\pi^+\pi^+$		12	PHOTOS PHSP
$B^+ \to a_1^+\rho^0$	1	12	PHOTOS SVV_HELAMP
$B^0 \to a_1^\pm\rho^\mp$	1	12	PHOTOS SVV_HELAMP
$B^+ \to a_1^0\rho^+$	1	12	PHOTOS SVV_HELAMP
$B^+ \to f_0\pi^+\pi^0$		12	PHOTOS PHSP
$B^+ \to \pi^+\pi^-\pi^+\pi^0$		12	PHOTOS PHSP
$B^+ \to \eta'\rho^+$		2	PHOTOS SVS
$B^+ \to \omega\rho^+$	0.90	2	PHOTOS SVV_HELAMP

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.

This is achieved with a centralized loose selection based on a minimum number of charged particles and calorimeter clusters meeting basic quality criteria [57]. Charged particles are required to have $p_T > 100$ MeV/c and impact parameters $|dz| < 5$ cm and $|dr| < 2$ cm to suppress tracks coming from beam-gas events. Tracks satisfying these criteria are called good tracks. ECL clusters are required to have energy $E_\gamma > 50$ MeV deposited in the barrel region to suppress beam-induced background events and ECL noise; clusters satisfying these criteria are called good clusters. To make it into the hadronic stream an event should meet the following conditions (all quantities refer to the center-of-mass frame.)

- Three or more good tracks originated from a primary vertex with $|dr| < 1.5$ cm and $|dz| < 3.5$ cm, 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.
- One or more good ECL cluster within the fiducial volume of $-0.7 < \cos\theta < 0.9$ and total energy $0.18 < E_{\rm tot}/\sqrt{s} < 0.8$, to suppress QED events (as most of the final state

particles scatter at small angles), two-photon, τ-pair, and beam-gas events. The lower bound of the ECL cluster energy requirement is chosen to reduce two-photon and beam-gas events due to the typically low energy released in these processes, and τ-pairs due to the presence of at least two neutrinos, which carry a relatively large fraction of collision energy that is not reconstructed. The upper bound is chosen to reduce Bhabha events, whose electrons and positrons deposite large amounts of energy in the ECL.

- Collected visible energy $E_{\rm vis} \geq 0.2\sqrt{s}$, where $E_{\rm vis} = (\Sigma|\vec{p}_{\rm track}| + \Sigma E_\gamma)$. This requirement suppresses beam-gas and two-photon events because they occur at much smaller energies than $\sqrt{s}$.
- $HJM/E_{\rm vis} > 0.25$ or $HJM > 1.8$ GeV/c^2, where HJM stands for heavy jet mass, that is the invariant mass of particles found in hemispheres perpendicular to the thrust axis (Fig. 4.2 and Sec. 3.3.4). The requirement on HJM is the most effective to remove τ-pair events since HJM for such events peaks around the τ lepton mass. The logical "or" with the $HJM/E_{\rm vis}$ requirement is needed to recover $q\overline{q}$ events, as a significant fraction of them is reduced by the HJM requirement alone.

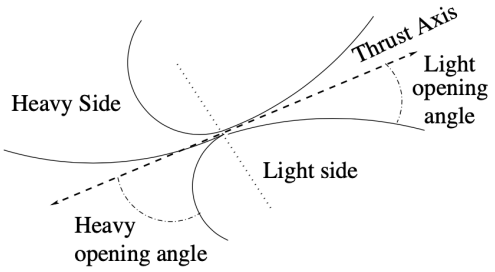


Figure 4.2: Sketch illustrating the heavy-jet-mass variable.

After applying these criteria, about 99% (79%) of all $B\overline{B}$ ($q\overline{q}$) events are retained, with a rejection of 95% of non-hadronic events as shown in Table 4.2.

	$B\overline{B}$	$q\overline{q}$	$\tau^+\tau^-$	QED	$\gamma\gamma$	Beam gas
σ, nb	1.09	2.62	0.05	0.001	0.04	0.11
ϵ	0.991	0.795	0.049	0.00002	0.004	> 0.0054

Table 4.2: Cross sections and efficiencies for various processes selected through the hadronic stream selection. QED includes Bhabha, radiative Bhabha and μ-pair production.

4.2 Signal 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 good tracks with impact parameters $|dz| < 4$ cm and $|dr| < 0.2$ cm to suppress tracks originated from beam-gas events.

This selection retains 80 % of correctly truth-matched pions in simulated signal decays. The π^+ (π^-) candidate multiplicity per event decreases from 5.13 (5.03) to 3.30 (3.30) as shown in Fig. 4.3.

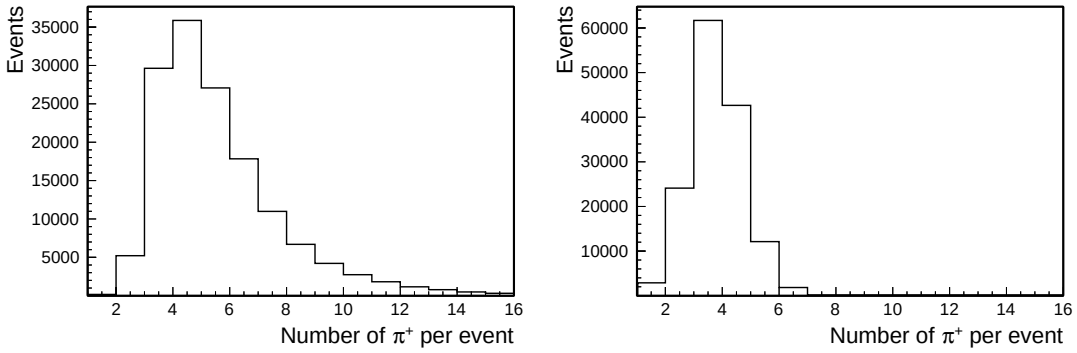


Figure 4.3: Distribution of π^+ multiplicities (left) before and (right) after the selection, respectively, in simulated $B^+ \to \rho^+\rho^0$ decays. The distributions for π^- particles are consistent and not shown.

4.2.2 Neutral pions

To reconstruct π^0 candidates, I require photons with energies $E_\gamma > 50$ (100) MeV in the barrel (endcap) section of the ECL to reduce beam-induced background and ECL noise. I then require the diphoton mass to meet $118 < m(\gamma\gamma) < 150$ 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` [67]). Only candidates with fit χ^2 smaller 50 are retained. The requirements on the diphoton mass and kinematic fit quality suppress misreconstructed neutral-pion candidates as shown in Fig. 4.4. True neutral pions tend to cluster at lower χ^2 values and in the 118–150 MeV/c^2 mass range, unlike misreconstructed pions.

This selection retains 75% of correctly truth-matched neutral pions from simulated signal decays. The π^0 candidate multiplicity per event decreases from 22.2 to 4.0 candidates as shown in Fig. 4.5.

4.2.3 ρ mesons

Candidate π^+ and π^- (π^0) are paired in a kinematic fit to form ρ^0 (ρ^+) 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. 4.6), which retains 90% of all signal ρ mesons according to simulation.

In addition, I require the ρ^+ helicity angle to meet $\cos\theta_{\rho^+} < 0.8$ 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 $\cos\theta_{\rho^+} > 0.8$ events I remove a large fraction of continuum background events as shown in Fig. 4.8.

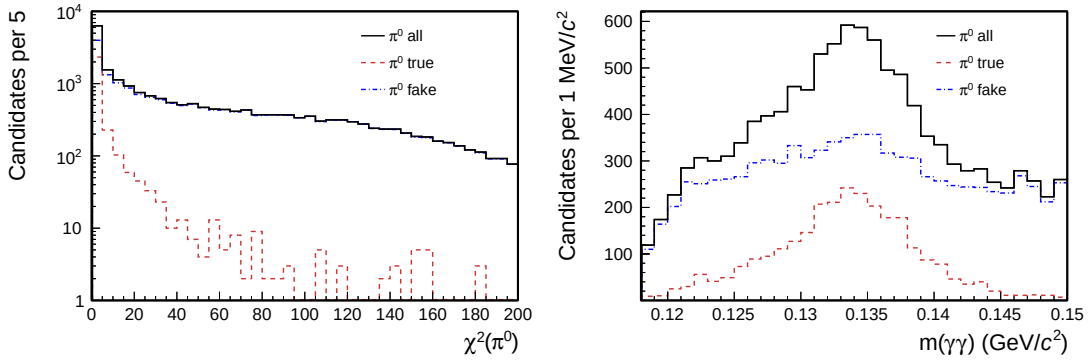


Figure 4.4: Distribution of (left) χ^2 of the diphoton mass-constrained vertex fit and (right) diphoton mass in simulated signal data. The black solid line corresponds to all π^0 candidates reconstructed in the signal sample, the red solid line corresponds to truth-matched π^0 candidates, and the blue solid line corresponds to non-truth-matched π^0 candidates.

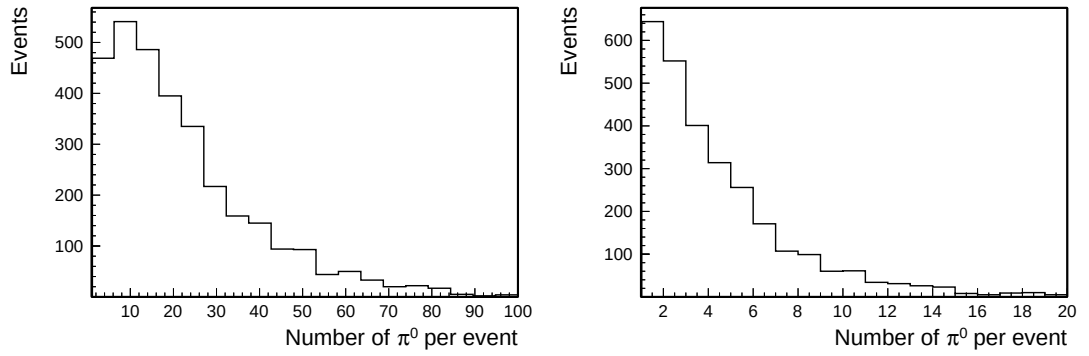


Figure 4.5: Distribution of π^0 multiplicity (left) before and (right) after the selection in simulated signal decays.

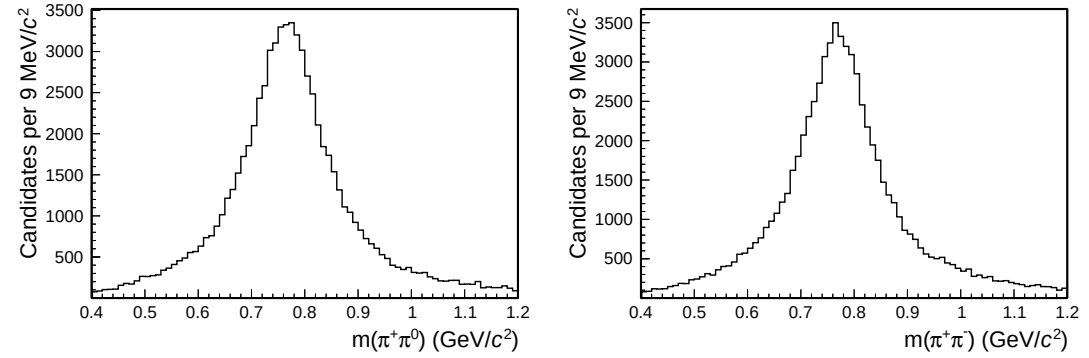


Figure 4.6: Distribution of (left) $\pi^+\pi^0$ and (right) $\pi^+\pi^-$ mass reconstructed in simulated signal decays.

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. We 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. 4.7). The $M'_{\rm bc}$ range narrowly encloses the sig-

nal region to reduce background, as we choose not to include $M'_{\rm bc}$ among the fit observables because it depends on ΔE. 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.

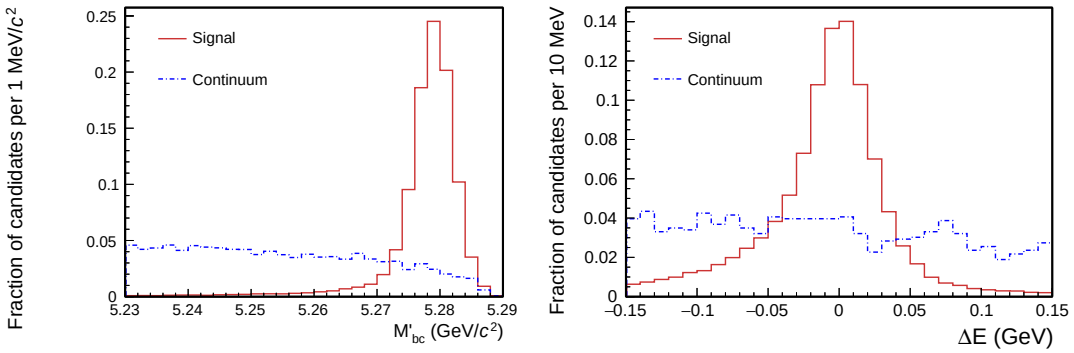


Figure 4.7: Distribution of (left) $M_{\rm bc}$ and (right) ΔE for simulated signal and continuum events. Distributions normalized to unity.

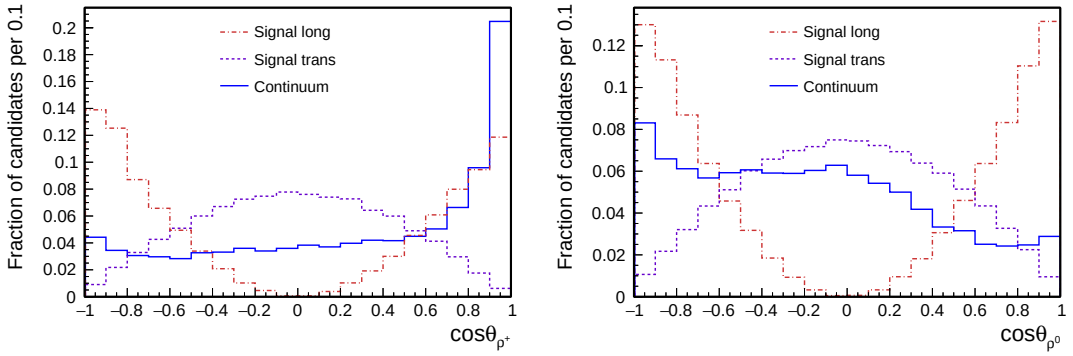


Figure 4.8: Distribution of the cosines of the helicity angles for (left) ρ^+ meson and (right) ρ^0 meson from simulated signal decays. Distributions normalized to unity.

After the selection, multiple B meson candidates per event are reconstructed in a fraction of events as shown in Fig. 4.9. In average, 1.12 (1.03) longitudinally (transversely) polarized 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 97.8%, to be compared with the 95.4% figure associated with a random choice.

4.3 Continuum background

The above selection suppresses events from various background processes. But the resulting sample is still dominated by events from direct production of light-pair of quarks and other B meson decays, as shown by Fig. 4.10 of the ΔE distribution in a realistically simulated sample resulting from the baseline selection. No visible signal peak is observed at $\Delta E \approx 0$ region. Further background suppression is needed.

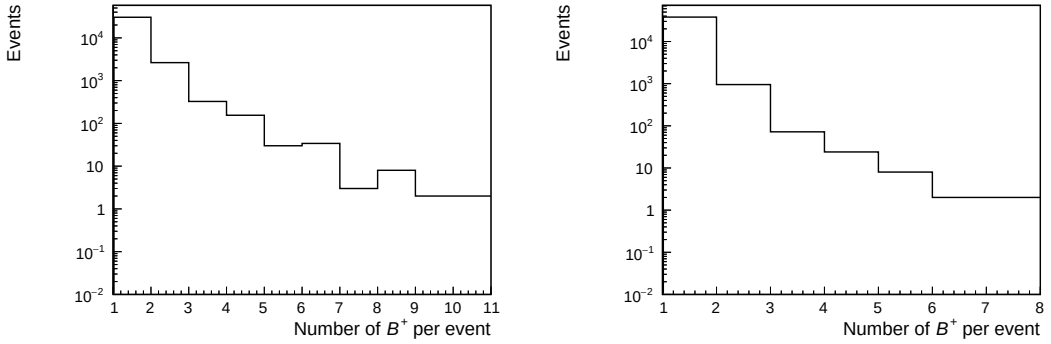


Figure 4.9: Distributions of B candidate multiplicity in simulated (left) longitudinally and (right) transversely polarized signal decays. The vertical axis is log-scaled.

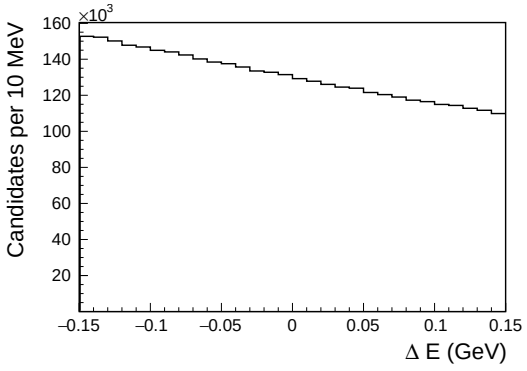


Figure 4.10: Distribution of ΔE in a realistically simulated sample selected with the baseline suppression.

At the $\Upsilon(4S)$ resonance, more than 3/4 of all hadronic events come from $e^+e^- \to q\overline{q}$ ($q = u, d, s, c$) 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 continuum, I combine non-linearly various discriminating variables in a boosted decision tree (FBDT) [77], 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.

The decision tree is a specific type of supervised classifier that approximates the optimal

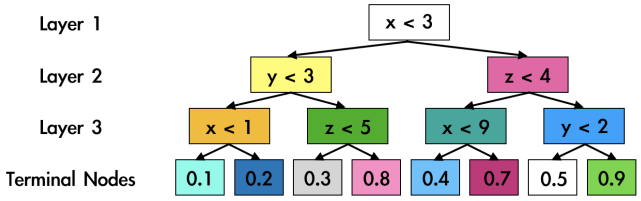


Figure 4.11: Schematic example of a three-layer decision tree. At each node of the tree a binary decision is made until a terminal node is reached. The numbers in the terminal node correspond to a probability of the test data-point to be signal.

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. 4.11 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 [78].

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 [79] 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 50 observables, I first extensively 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 exclude the magnitude of the thrust axis in the signal side, $T_{\rm sig}$ (see Sec. 3.3.4), as it shows a large correlation of 60% with dipion masses. I then remove variables that are not sufficiently powerful to discriminate signal B events. An example are CLEO cone variables (see Sec. 3.3.4) obtained using all particles in the event, as they are over 90% correlated with those obtained using only tag-side particles and provide less separation power. Hence, I converge to a subset of 38 discriminating variables whose correlation with relevant fit observables is below $\approx 15\%$ as shown in Fig. 4.29.

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 (more details in Chapter 6.) I also inspect off-resonance simulated and collision data to further check the consistency of discriminating-observable distributions for continuum background, which is expected to be the dominant contributor to the $B^+ \to \rho^+\rho^0$ sample composition. I compare simulation and data separately for background events in the sideband region $5.211 < M_{\rm bc} < 5.224$ GeV/c^2 and signal events in the region $M_{\rm bc} > 5.27$ 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. Figures 4.12–4.21 show data-simulation distributions of the chosen 38 discriminating observables. Good agreement is observed in most distributions, with the exceptions of on-resonance R_2 and $\cos\theta_{\rm T}^{\rm sig,tag}$. We fail to identify any obvious simulation sample-composition issue that could explain those discrepancies, but we choose to keep those inputs in the FBDT. If anything, the observed discrepancies could induce a minor suboptimality in the choice of the final selection but are unlikely to introduce biases in the final results.

4.3.2 Classifier training and test

I train the FBDT using 3×10^6 truth-matched signal events (with same proportions of longitudinally and transversely polarized and of positively and negatively charged B candidates) and 3×10^6 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 classifier in three configurations: *a*) excluding all flavor-tagging variables and topological variables that have more than 3% correlations with the fit observables; *b*) excluding the topological variables with correlations but including flavor-tagging variables; and *c*) using all 38 variables. Figure 4.22 shows ROC curves for the three configurations of input variables: the blue curve corresponds to configuration (*a*), the orange curve to (*b*), and the green curve to (*c*). The ROC curve corresponding to the configuration with all 38 variables has the largest AUC value. I therefore use this configuration throughout the analysis.

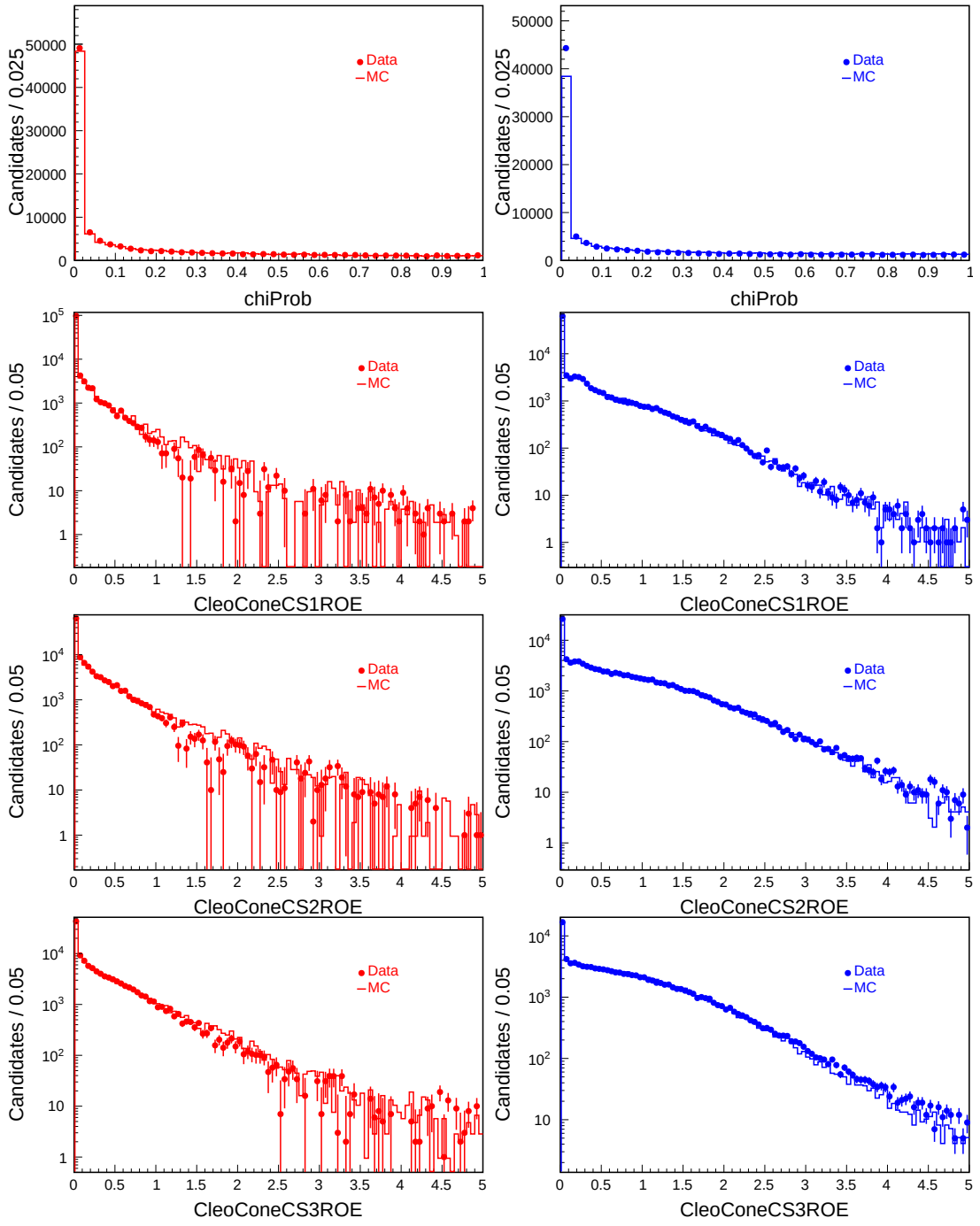


Figure 4.12: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (1/10).

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 4.23 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.

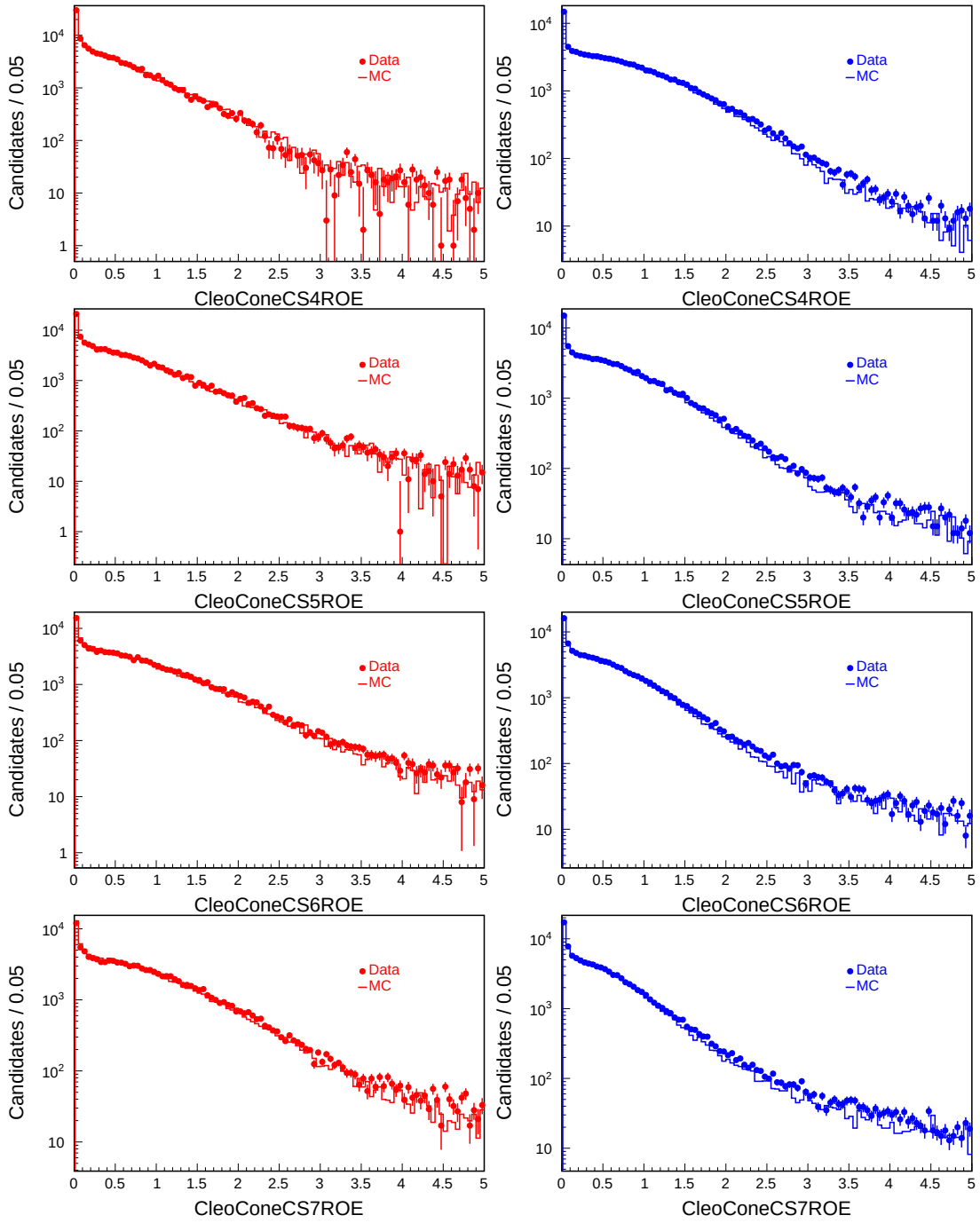


Figure 4.13: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (2/10).

4.3.3 Selection optimization

The ROC curve (Fig. 4.22) 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 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

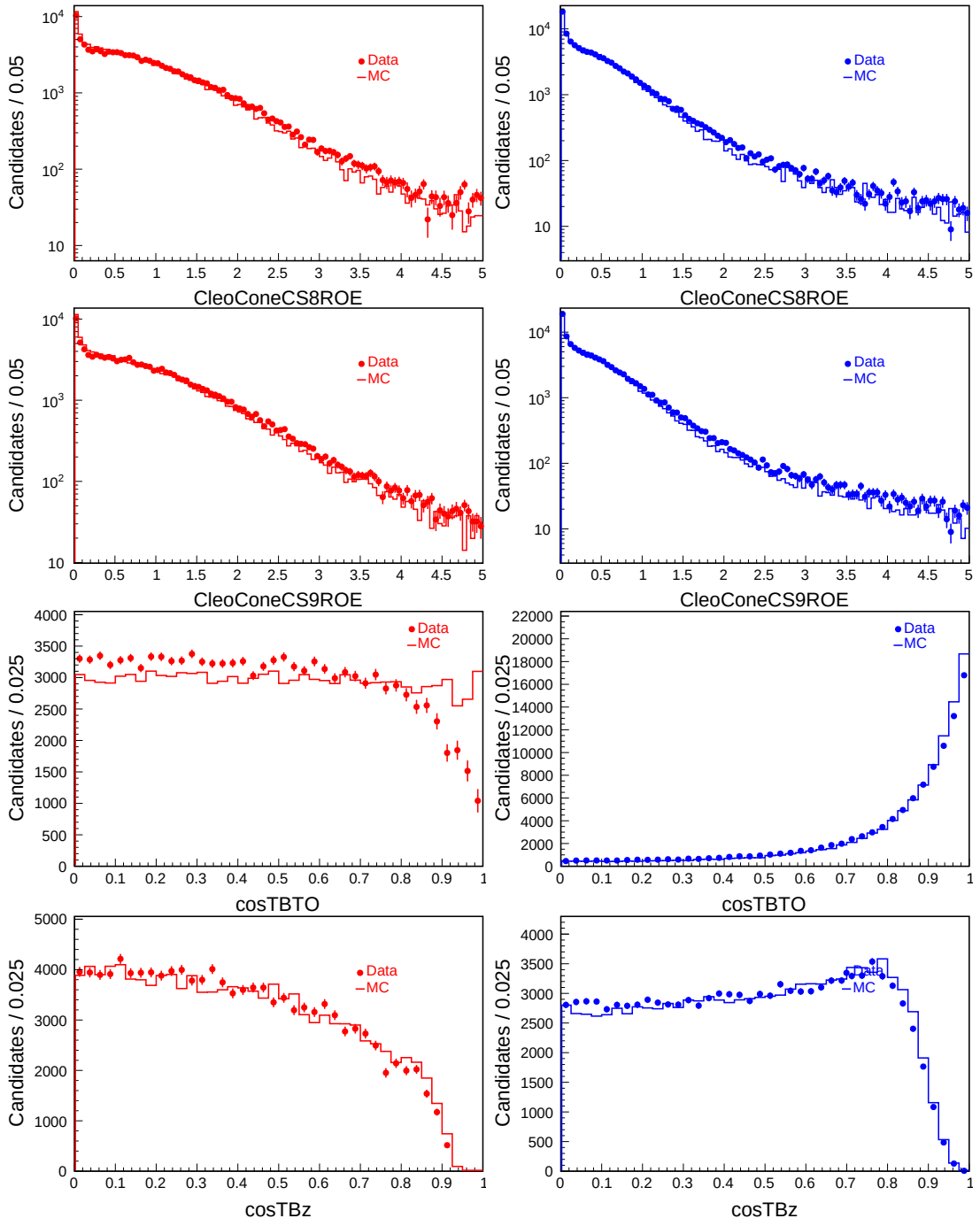


Figure 4.14: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (3/10).

statistical power of the measurement as suppressing excessively background may not leave enough information to the fit to separate signal efficiently. I therefore optimize the selection by minimizing the average expected statistical variance of the parameter of interest, $f_L \times \mathcal{B}(B^+ \to \rho^+\rho^0)$. This is 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

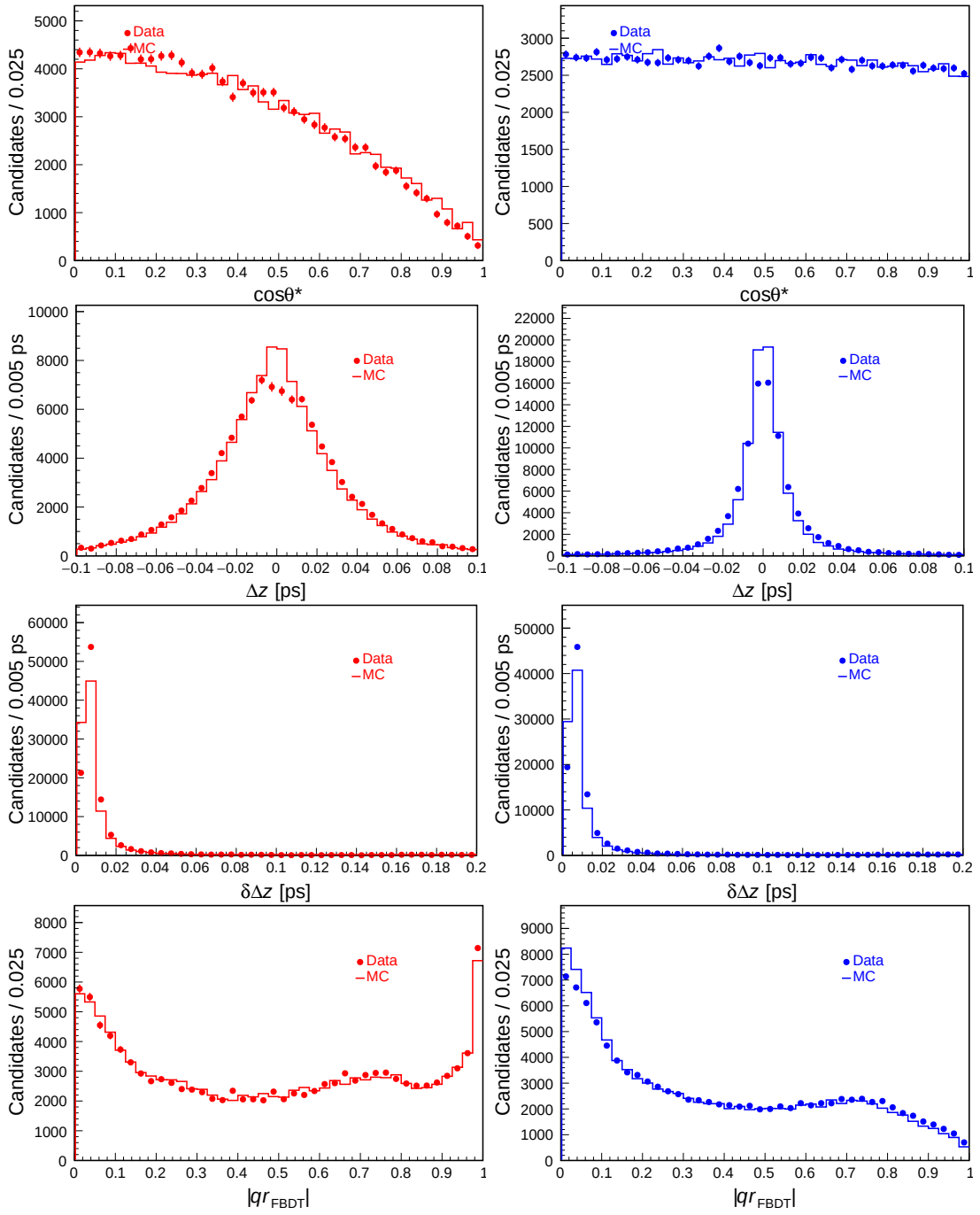


Figure 4.15: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (4/10).

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.

I also 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,

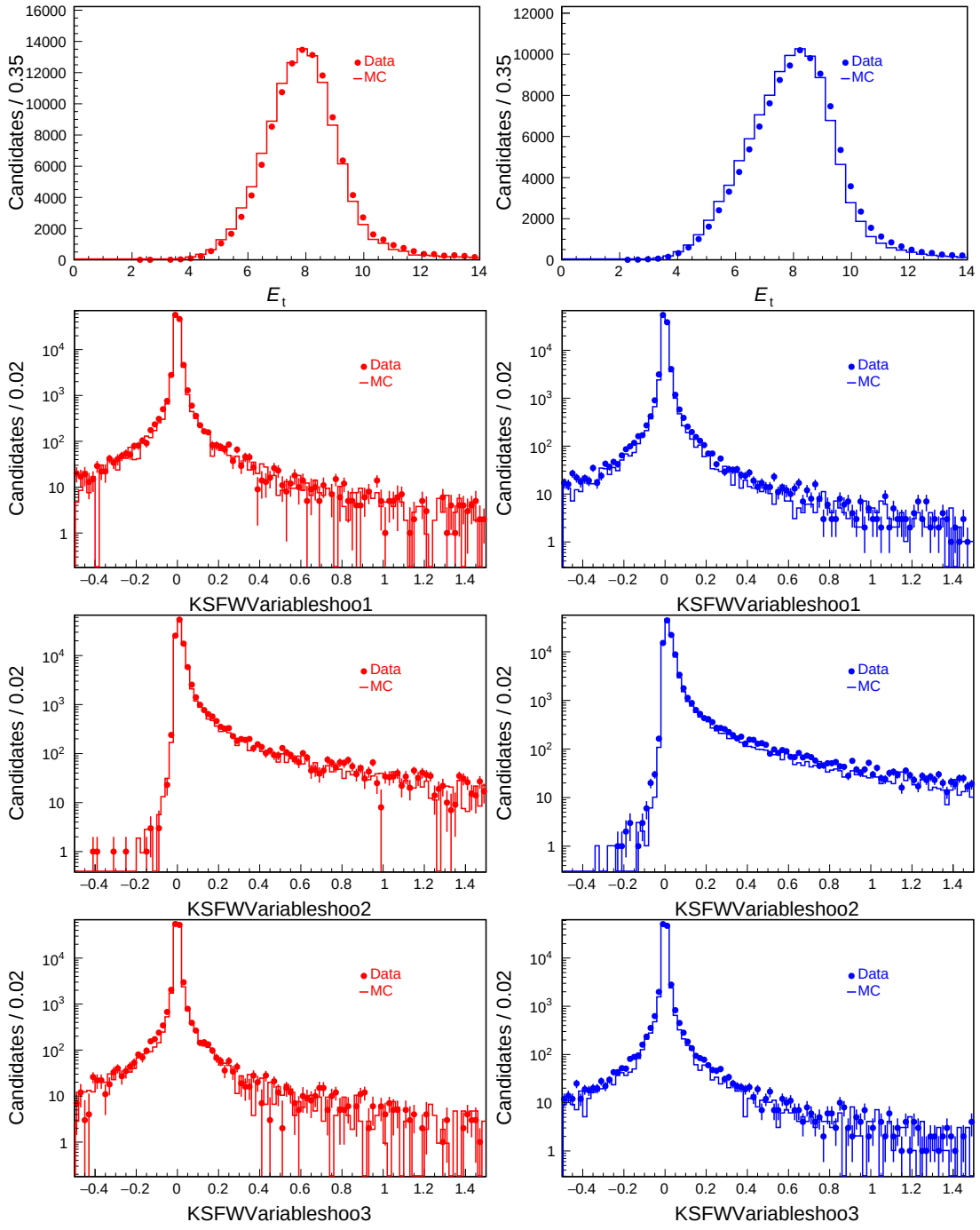


Figure 4.16: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (5/10).

which is an input to the continuum suppression FBDT, continuum suppression and particle identification can 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

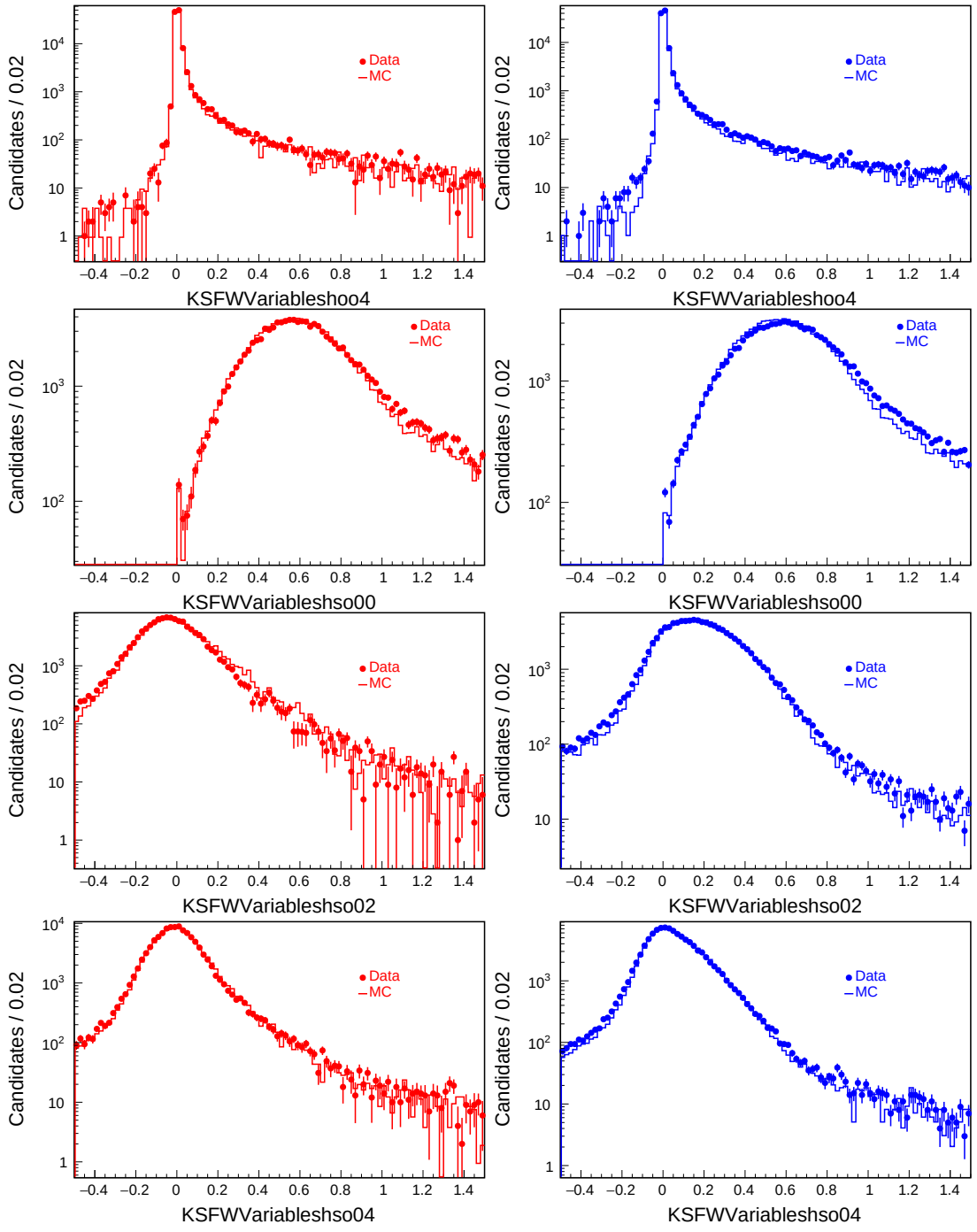


Figure 4.17: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (6/10).

likelihood is the same as used in the full sample-composition fit but the number of components is reduced. I include only signal and continuum background neglecting (both in generation and fitting) the 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

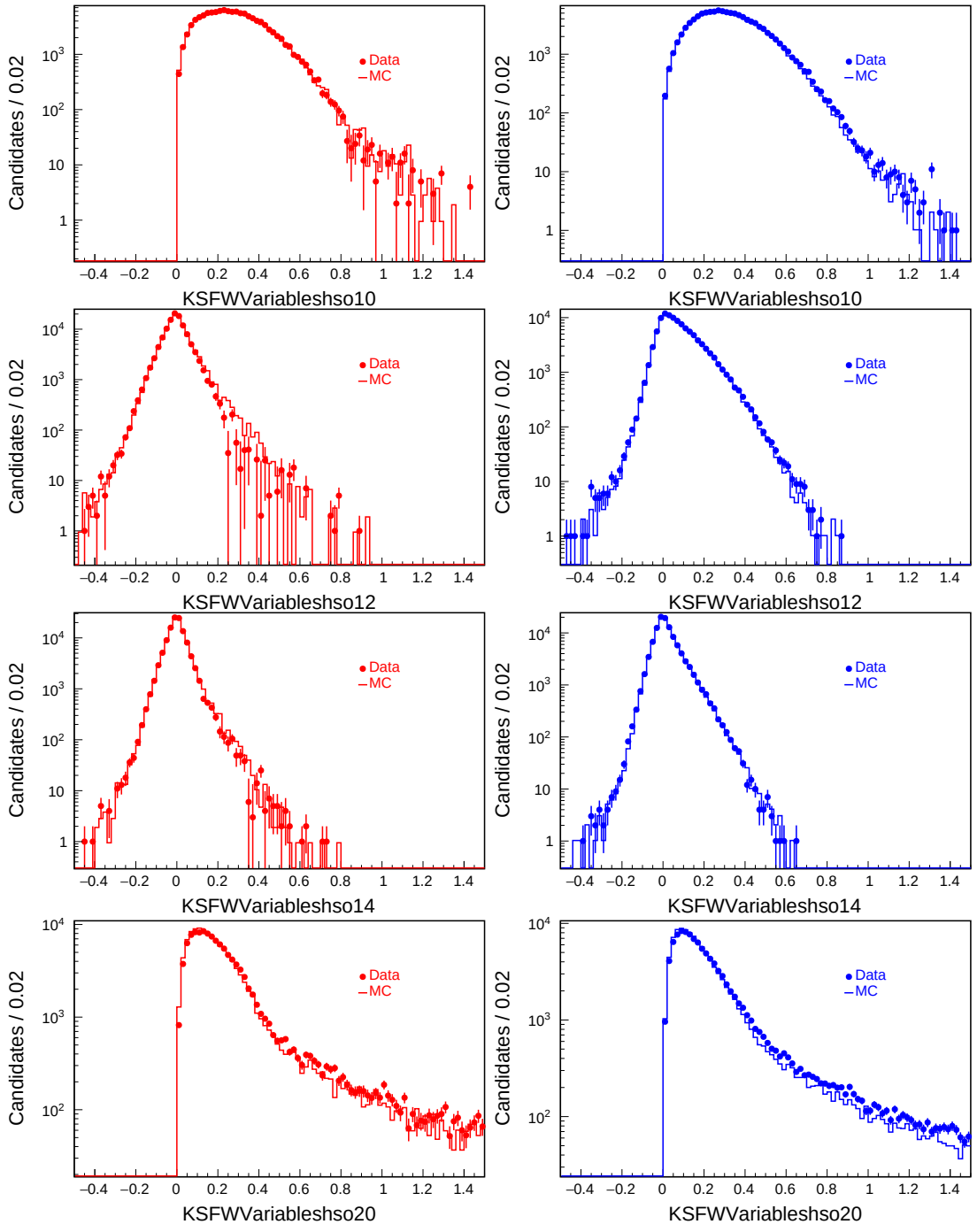


Figure 4.18: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (7/10).

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, 500 toy samples are generated. In generation, the yields of each component are drawn according to Poisson distributions mean valued at the yields expected in data. Signal distributions are drawn from simulated samples. Continuum background distributions are generated according to

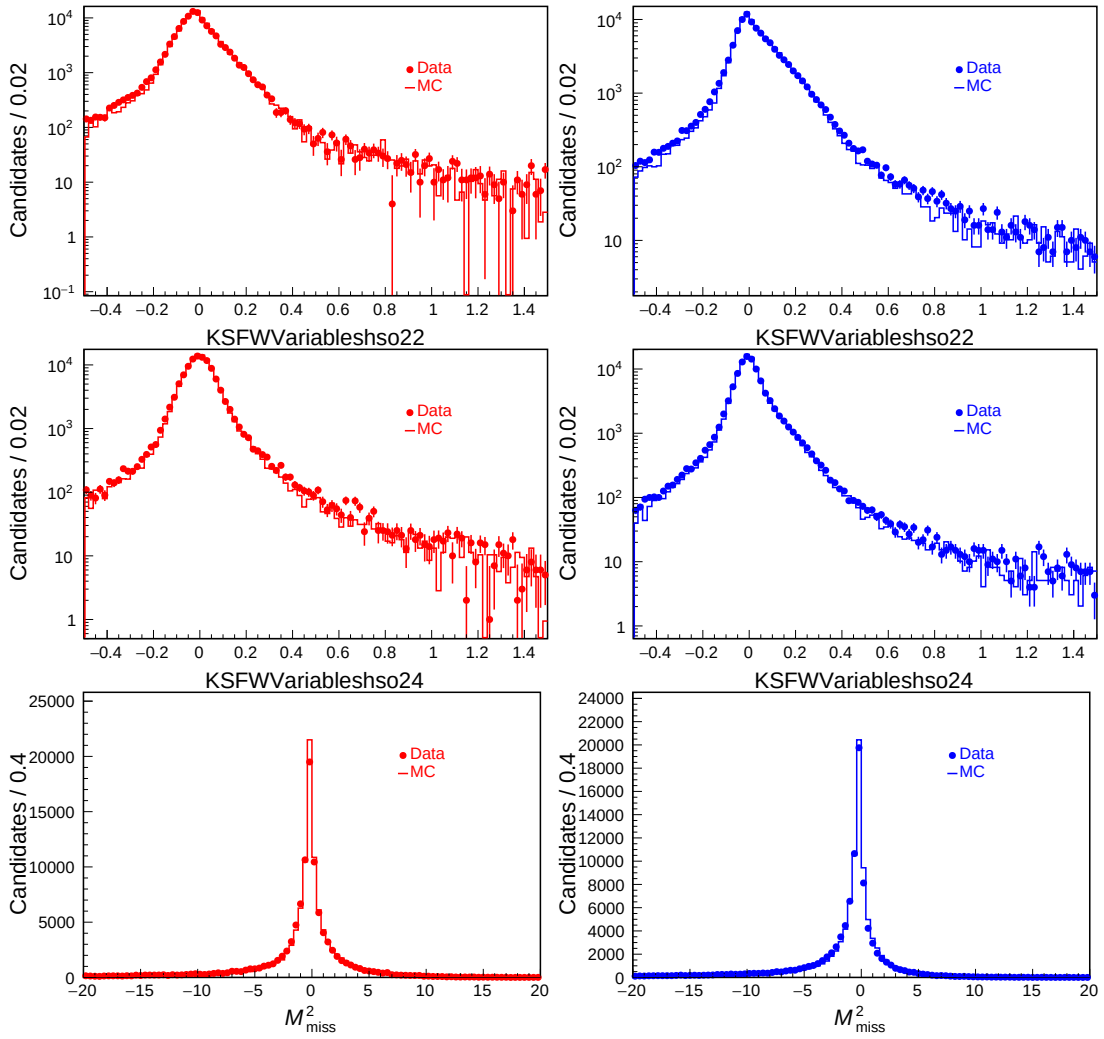


Figure 4.19: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (8/10).

fits of realistic simulated continuum distributions.

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 $f_L \times \mathcal{B}(B^+ \to \rho^+\rho^0)$, where θ_i is the true value and $\hat{\theta}_i$ is the fit estimate. Figure 4.24 shows an example of the residual distribution. The width of the residual is the desired average expected statistical uncertainty.

The results are shown in Fig. 4.25. The resulting optimal values are an upper bound at 0.8 for the probability of a pion to be identified as a kaon and a range of thresholds between 0.55 and 0.92 for the continuum suppression output; I choose the requirement $C_{\text{FBDT}} > 0.85$.

In addition, I check whether backgrounds where electrons and muons are misidentified as pions degrade the expected resolution by including a generic $B\overline{B}$ component, which is dominated by semileptonic B decays. I fit the composition of toy samples selected with and without a loose lepton identification requirement. This study shows that backgrounds where electrons and muons are misidentified as pions impact negligibly the final precision.

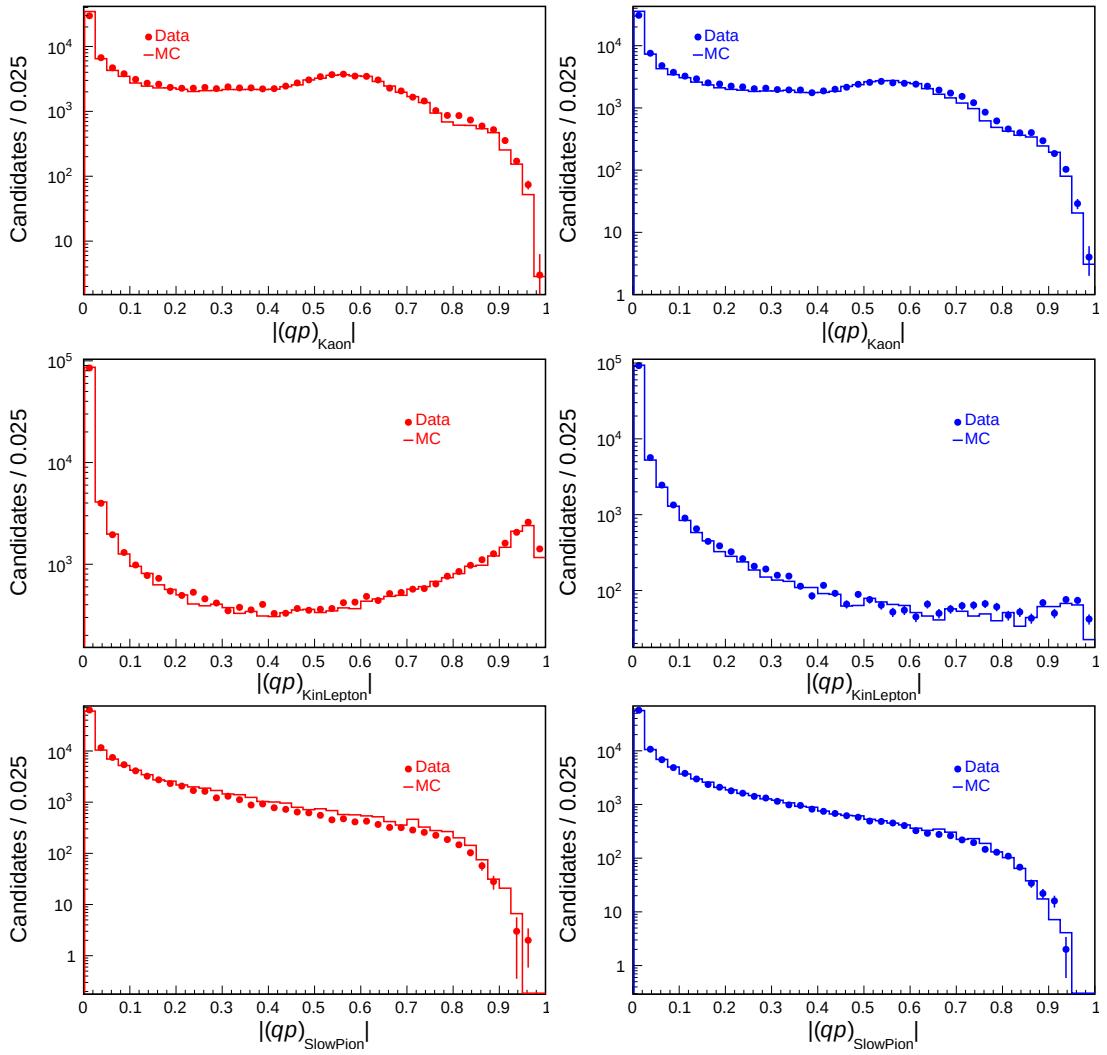


Figure 4.20: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (9/10).

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 fraction of self-cross feed candidates is 37.5% for longitudinally polarized and 21.0% for transversely polarized $B^+ \to \rho^+\rho^0$ decays. This is a potential limitation for the analysis reach. Large cross-feed both decreases signal efficiency and adds complexity to the fit of sample composition. To further investigate the self-cross feed source, I classify self-cross feed candidates as shown in Fig. 4.26. Nearly 70% of the self-cross feed candidates are associated with a misreconstructed π^0 where one of the photon candidates is a random cluster in the ECL, or with technical inefficiencies in associating to genuine signal simulated events that undergo gamma conversions before the calorimeter.

I therefore explore variations of the requirements on the diphoton invariant mass, energy of the photons, absolute value of the cosine of the π^0-helicity angle, χ^2 of the π^0 mass-

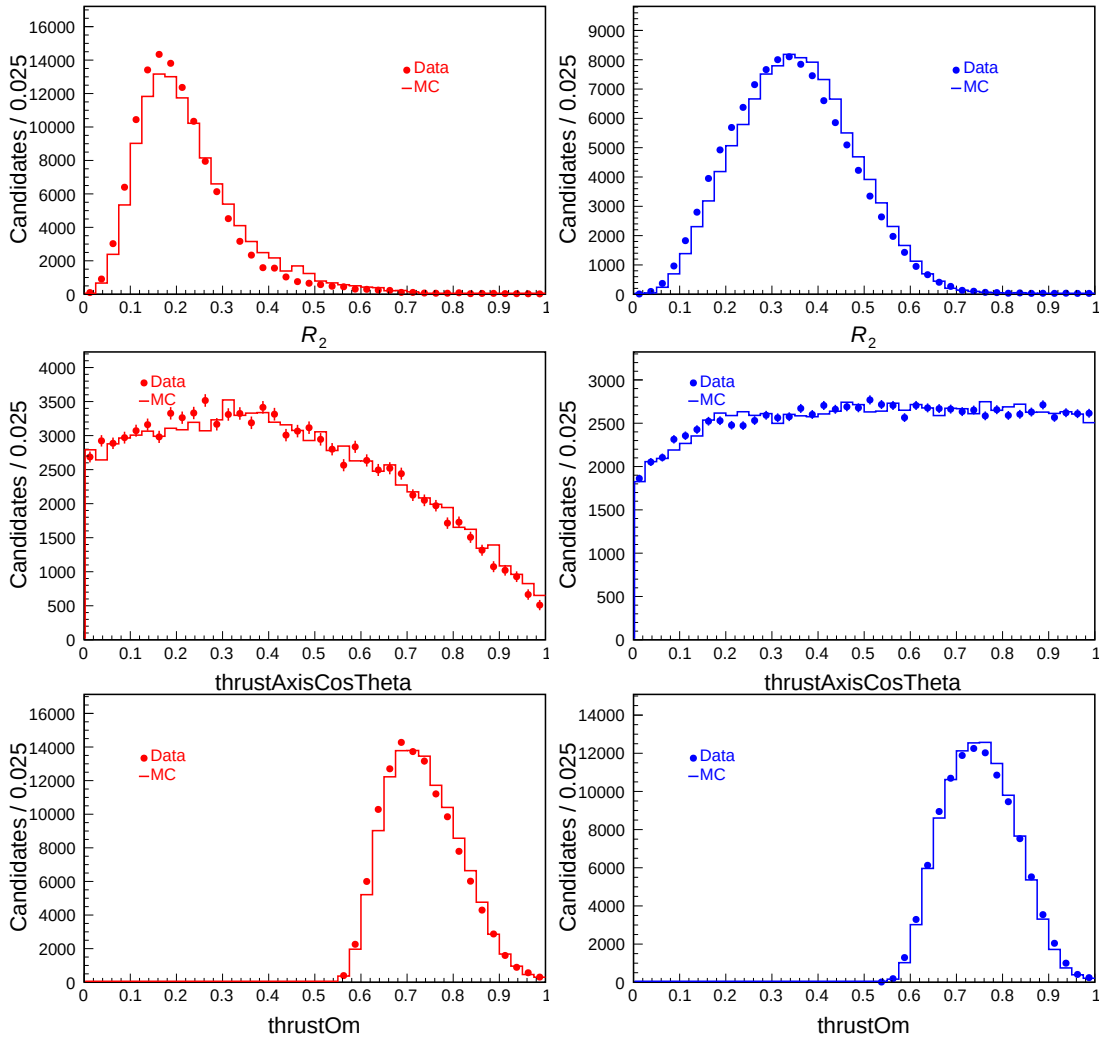


Figure 4.21: Comparison of continuum suppression inputs in (circles) data and (solid) simulation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in (left) on-resonance and (right) off-resonance samples (10/10).

constrained fit, and π^0 momentum in the $\Upsilon(4S)$ frame to reduce the π^0 misreconstruction rate. No significant improvement over the standard π^0 selection is found, which implies that no obvious way of reducing self-cross feed exists based on such higher-level π^0 variables and a careful modeling of self-cross feed is needed in the sample composition.

4.5 *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 distinguish $B\overline{B}$ background from B decays that tend to peak where signal peaks (peaking background) and B decays that do not peak (non-peaking background).

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

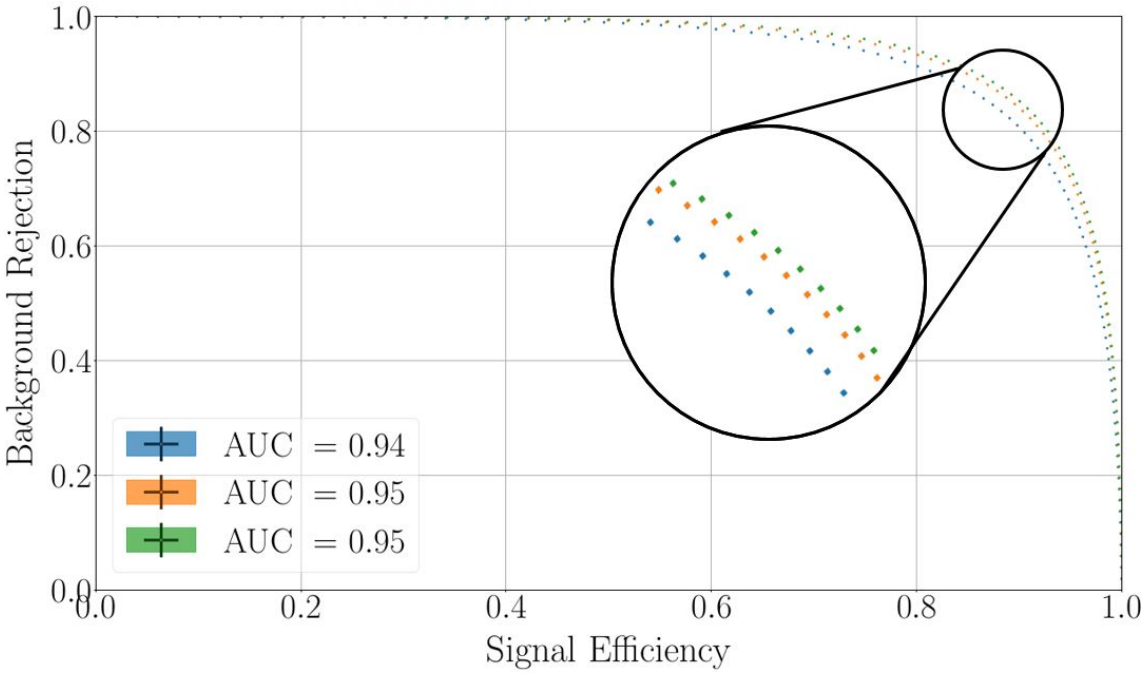


Figure 4.22: Receiver operating characteristic for three configurations of the continuum-suppression classifier: (blue) excluding flavor-tagging variables and topological variables with correlations of 3% or greater with the fit observables; (orange) excluding topological variables with correlations of 3% or greater but including flavor-tagging variables; and (green) using all 38 variables.

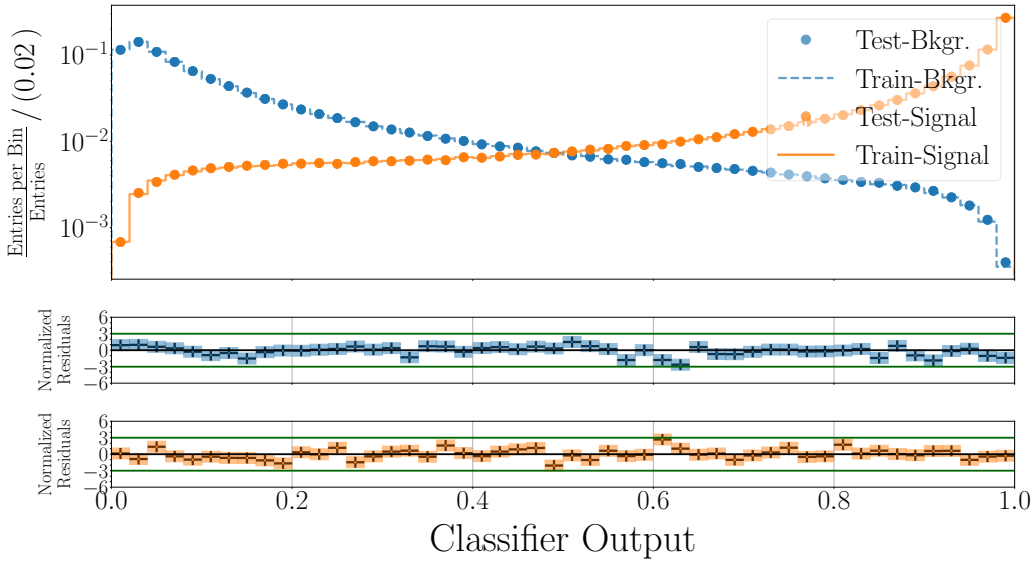


Figure 4.23: Output of the FBDT classifier on the training and testing samples.

$B\overline{B}$ backgrounds. Favored B decays contribute significantly to the sample composition due to their large branching fractions. A major contribution comes from $B \to DX \to \pi^+\pi^-\pi^+\pi^0$ decays where D identifies a generic charmed meson and X a generic decay partner. These backgrounds can generate ΔE distributions peaked under, or in proximity, of the $B^+ \to \rho^+\rho^0$ peak, and $M_{\rm bc}$ distributions peaked under the $B^+ \to \rho^+\rho^0$ decays. As the fit that deter-

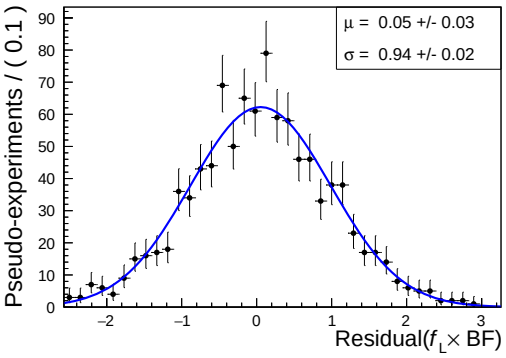


Figure 4.24: Residual distribution for $f_L \times \mathcal{B}(B^+ \to \rho^+\rho^0)$ obtained from the fit of simplified simulated $B^+ \to \rho^+\rho^0$ samples.

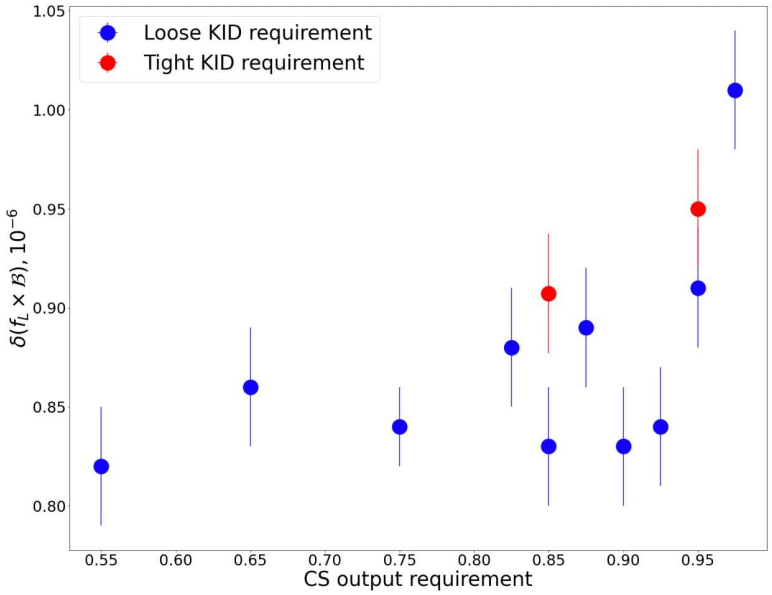


Figure 4.25: Average expected statistical uncertainty on $f_L \times \mathcal{B}(B^+ \to \rho^+\rho^0)$ as a function of continuum suppression and pion-identification requirements as resulting from analyses of simplified simulated experiments.

mines the signal yields generally assumes the background to be smoothly distributed, such backgrounds may be mistakenly absorbed into signal and bias the results if not taken into account.

4.5.1 Charm vetoes

To identify and remove $B \to DX \to \pi^+\pi^-\pi^+\pi^0$ backgrounds, I inspect the distributions of the two- and three-body masses of intermediate states, where the narrow peaks associated with properly reconstructed D decays allow for straightforward vetoes, as shown in Fig. 4.27. Four peaking structures are visible. A peak is evident at $1.86\,\mathrm{GeV}/c^2$, the known D^0 mass in both 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, as is a peak at $1.78\,\mathrm{GeV}/c^2$, which corresponds to $B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+\pi^0$ and $B^+ \to \overline{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays where the pion mass is mistakenly assigned to the kaon tracks resulting in an invariant mass shift.

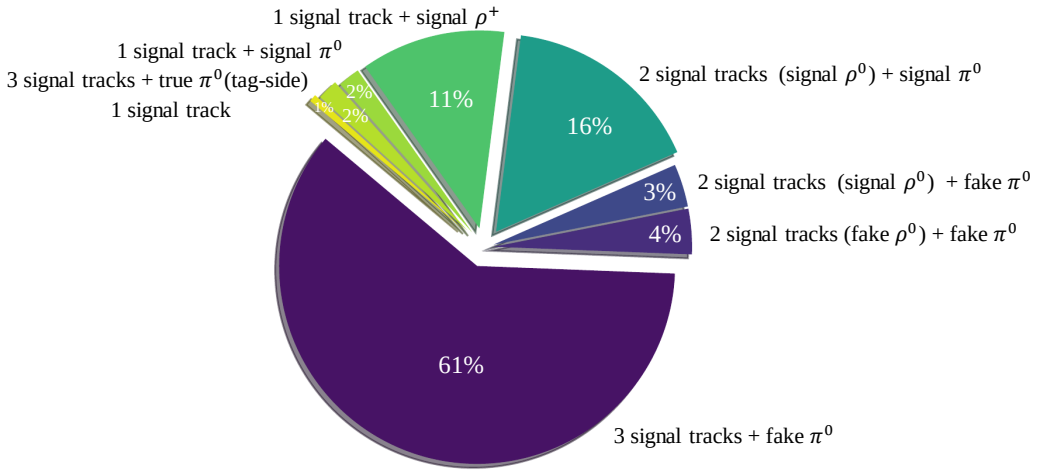


Figure 4.26: Composition of self-cross-feed candidates assuming an equal amount of longitudinally and transversely polarized events.

I study the ΔE distributions to check if these backgrounds mimic signal (Fig. 4.28). While backgrounds from two-body D meson decays do not have a peaking structure, backgrounds from three-body D meson decays tend to peak in the ΔE signal region. I veto them by imposing $|m(h^{\pm}\pi^{\mp}) - m_{D^0}| > 0.18\,\mathrm{GeV}/c^2$ and $|m(h^{\pm}\pi^{\mp}\pi^0) - m_{D^0}| > 0.32\,\mathrm{GeV}/c^2$, where h is a pion or kaon and m_{D^0} corresponds to the known D^0 mass value $1.86\,\mathrm{GeV}/c^2$.

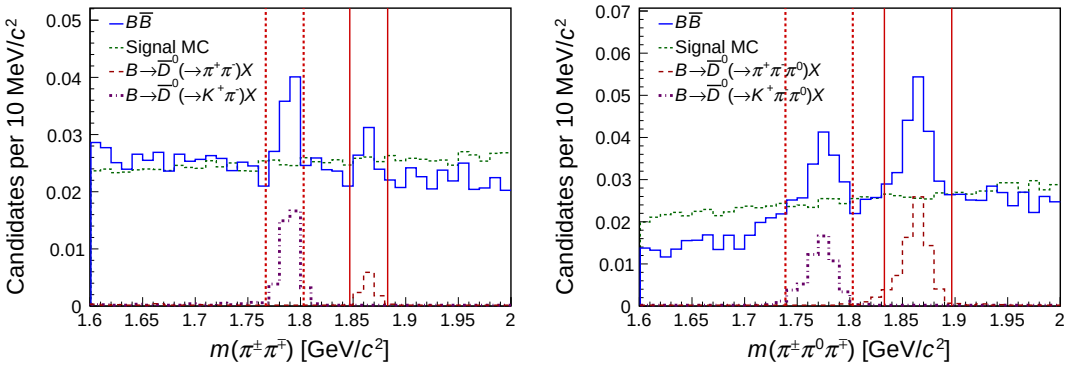


Figure 4.27: Mass distributions for (left) two- and (right) three-body combinations of final-state pions in $B^+ \to \rho^+\rho^0$ decays reconstructed from simulated signal and $B\overline{B}$ samples. The distributions for candidates reconstructed in favored B decay events are compared with the distributions for truth-matched signal candidates (in same proportions of longitudinal and transverse polarization). The red vertical lines show the veto regions.

4.5.2 Rare B decays

Rare 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 rare B peaking backgrounds should be included and properly modeled in the fit of sample composition.

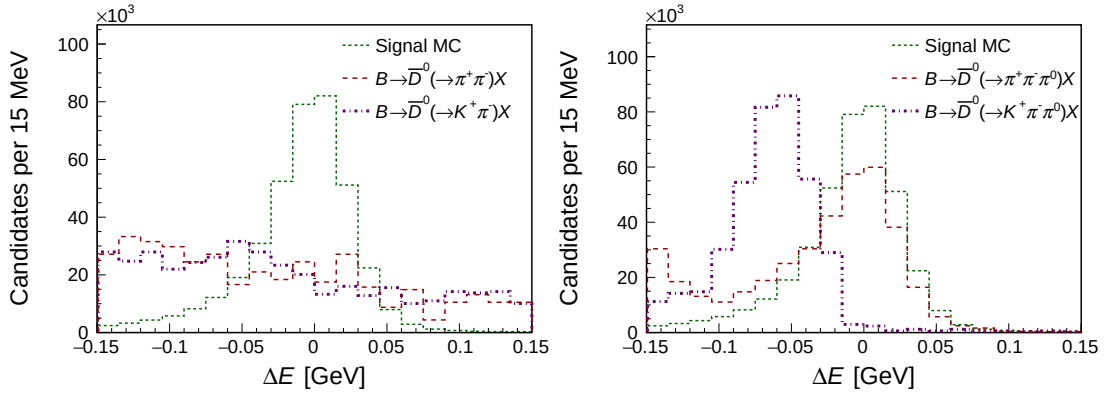


Figure 4.28: ΔE distributions for simulated charmed peaking background decays and truth-matched signal candidates.

I first study $B^+ \to \rho^+\rho^0$ candidates reconstructed in the simulated generic samples of most insidious rare charmless B decays 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 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}, except for the $B^+ \to a_1^+\rho^0$ and $B^+ \to a_1^0\rho^+$ modes. For these, I use phenomenological predictions of relations between the $B^+ \to a_1^+\rho^0$, $B^0 \to a_1^+\rho^-$, and $B^+ \to a_1^0\rho^+$ branching fractions [80–83]. Through these, the observed upper limit for $\mathcal{B}(B^0 \to a_1^+\rho^-)$, $< 61\times10^{-6}$ provides conservative guesses for $\mathcal{B}(B^+ \to a_1^+\rho^0) = \mathcal{B}(B^+ \to a_1^0\rho^+) = 30\times10^{-6}$.

In the fit, I also include combinatorial background from rare B decays. This component corresponds to random candidates from events that contain other rare B decay modes not listed in Table 4.3. From simulation, I expect 1070 such candidates.

Another possible source of peaking backgrounds are rare semileptonic and charmonium B decays. I count $B^+ \to \rho^+\rho^0$ candidates reconstructed in simulated rare semileptonic and charmonium B generic samples and passed through the full selection, as summarized in Table 4.4. The expected yield of individual components are below 5% of the expected $B^+ \to \rho^+\rho^0$ yield and distributions smooth enough to be absorbed in the inclusive $B\overline{B}$ component. Hence I exclude rare semileptonic and charmonium B decay modes from the fit.

	Channel	$\mathcal{B}\,[10^{-6}]$	Exp. yield
Known	$B^0 \to \rho^+\rho^-$	27.7 ± 1.9	212
	$B^+ \to a_1(1260)^+\pi^0$	26 ± 7	153
	$B^+ \to K^{*0}\rho^+$	9.2 ± 1.5	115
	$B^0 \to a_1(1260)^\mp\pi^\pm$	26 ± 5	165
	$B^+ \to a_1(1260)^0\pi^+$	20.4 ± 5.8	95
	$B^+ \to \eta'\rho^+$	9.7 ± 2.2	43
	$B^+ \to \omega\rho^+$	15.9 ± 2.1	25
With limit	$B^+ \to \rho^+ f_0(980)$	< 2.0	< 90
	$B^0 \to a_1(1260)^\mp\rho^\pm$	< 61	< 234
Unknown	$B^+ \to \rho^0\pi^+\pi^0$	10	105
	$B^+ \to \rho^+\pi^+\pi^-$	10	97
	$B^+ \to \rho^0\pi^+\pi^+$	10	20
	$B^+ \to a_1(1260)^+\rho^0$	30	78
	$B^+ \to a_1(1260)^0\rho^+$	30	57
	$B^+ \to f_0(980)\pi^+\pi^0$	10	57
	$B^+ \to \pi^+\pi^-\pi^+\pi^0$	10	33

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 $\mathcal{B} = 10^{-5}$.

Decay	Evts. in 1×lumi.
$B^+ \to X_u^+\, l^+\nu$	43
$B^0 \to X_u^0\, l^+\nu$	59
$B^+ \to X^+\, J/\psi$	27
$B^0 \to X^0\, J/\psi$	16
$B^+ \to X^+\, \psi(2S)$	3
$B^0 \to X^0\, \psi(2S)$	3

Table 4.4: Expected yields for rare $J/\psi X$, $\psi(2S)X$, and semileptonic B decays passing the full $B^+ \to \rho^+\rho^0$ selection as determined from simulation.

Object	Requirement
Charged pions	$\|dz\| < 4$ cm
	$\|dr\| < 0.2$ cm
	$\mathrm{p_T} > 100$ MeV/c
	$\mathcal{R}(K,\pi) < 0.8$
Neutral pion	$E_\gamma > 50$ MeV for barrel ECL
	$E_\gamma > 100$ MeV for endcap ECL
	$118 < m(\gamma\gamma) < 150$ MeV/c^2
	vertex fit $\chi^2 < 50$
Charged ρ meson	$0.52 < m(\pi^+\pi^0) < 1.06$ GeV/c^2
	$\cos\theta^{hel}_{\rho^+} < 0.8$
Neutral ρ meson	$0.52 < m(\pi^+\pi^-) < 1.06$ GeV/c^2
B meson	$M_{\mathrm{bc}} > 5.27$ GeV/c^2
Best candidate selection	$\chi^2(\pi^0, \mathrm{vertex})$
Continuum suppression	$C_{\mathrm{FBDT}} > 0.85$
Charm veto	$1.84 < m(\pi\pi) < 1.88$ GeV/c^2
	$1.83 < m(\pi\pi\pi^0) < 1.90$ GeV/c^2
	$1.84 < m(K\pi) < 1.88$ GeV/c^2
	$1.83 < m(K\pi\pi^0) < 1.90$ GeV/c^2

Table 4.5: Summary of selection.

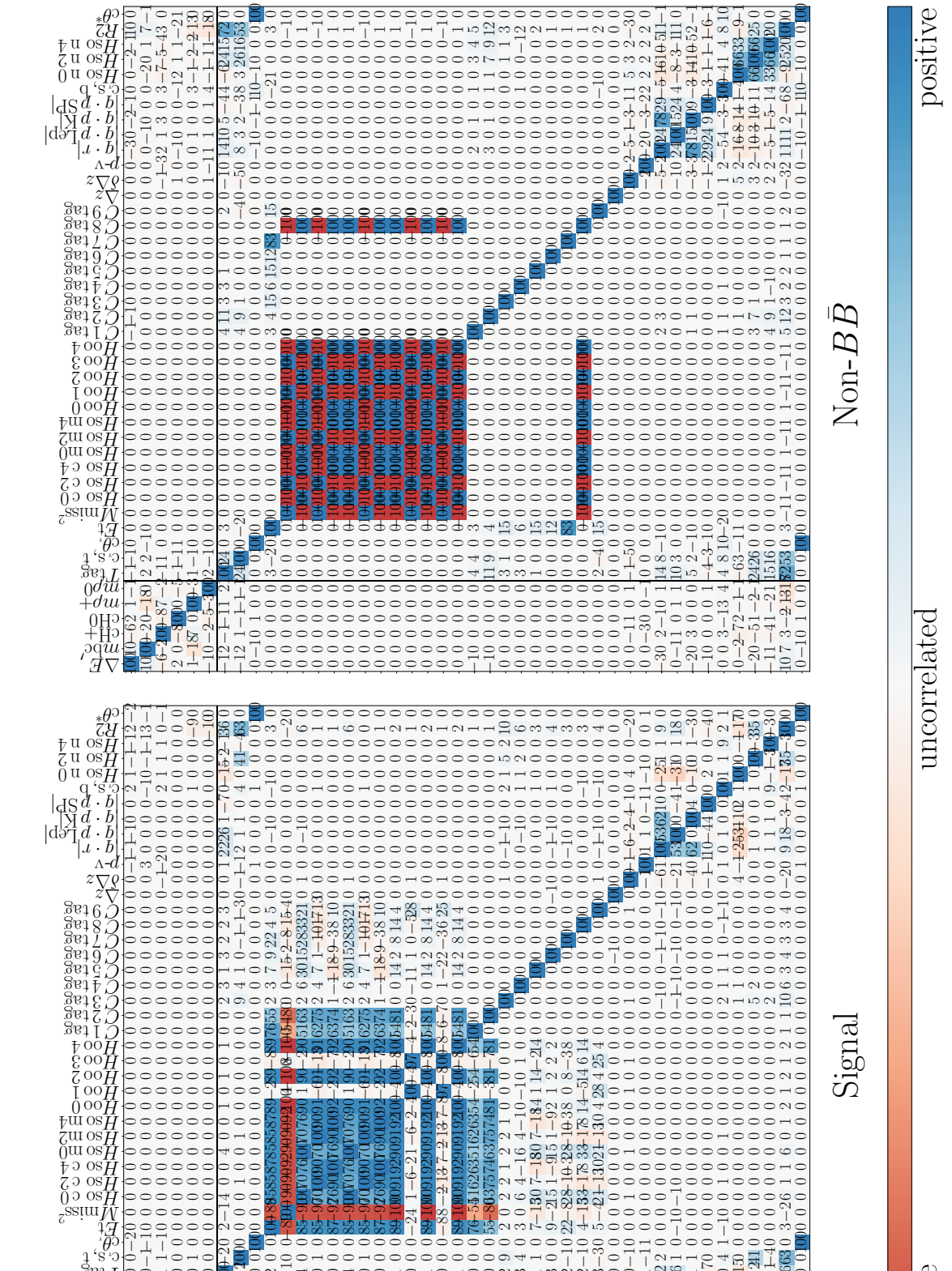


Figure 4.29: Linear correlation coefficients between training variables and fit observables.

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` [84], 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 because of its non-negligible contribution, as shown in Section 4.4. Continuum background dominates the sample composition. $B\overline{B}$ backgrounds are divided into 18 fit components, one inclusive component of non-peaking $B\overline{B}$ background, one inclusive component of combinatorial rare B background, and 16 exclusive peaking background components.

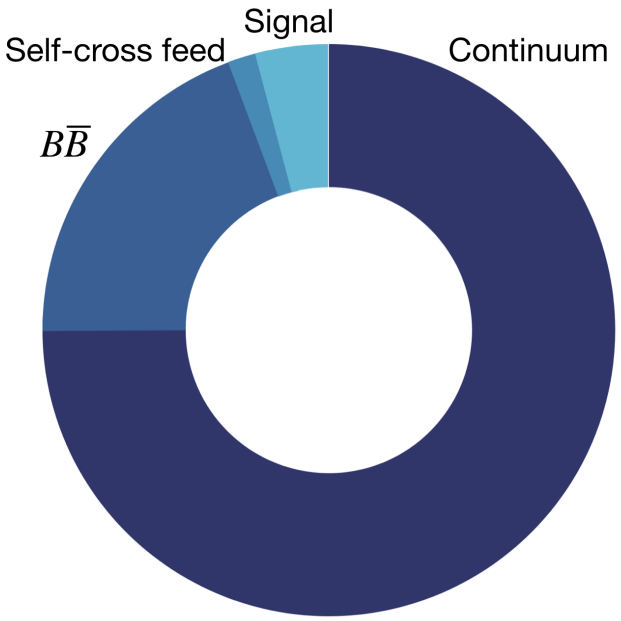


Figure 5.1: Pie-chart illustrating the $B^+ \to \rho^+\rho^0$ sample composition (assuming $f_L = 0.95$) resulting from the final selection as expected by simulation.

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

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 dependences.

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 $C'_{\mathrm{FBDT}} = \log\frac{C_{\mathrm{FBDT}} - C_{\mathrm{FBDT}_{\mathrm{min}}}}{C_{\mathrm{FBDT}_{\mathrm{max}}} - C}$, which is more easily described by an analytical model without degrading discriminating power.

Component	Expected yield
Longitudinally polarized $B^+ \to \rho^+\rho^0$	1400
Transversely polarized $B^+ \to \rho^+\rho^0$	140
Longitudinally polarized self-cross feed	580
Transversely polarized self-cross feed	30
Continuum background	29100
Non-peaking $B\overline{B}$ background	4950
Combinatorial rare B background	1070
$B^+ \to f_0(980)\rho^+$	90
$B^+ \to a_1(1260)^+\pi^0$	150
$B^0 \to a_1(1260)^\mp\pi^\pm$	170
$B^+ \to a_1(1260)^0\pi^+$	100
$B^0 \to \rho^+\rho^-$	210
$B^+ \to K^*(892)^0\rho^+$	120
$B^+ \to a_1(1260)^+\rho^0$	80
$B^0 \to a_1(1260)^\mp\rho^\pm$	230
$B^+ \to a_1(1260)^0\rho^+$	60
$B^+ \to \rho^0\pi^+\pi^0$	110
$B^+ \to \rho^+\pi^+\pi^-$	100
$B^+ \to \rho^-\pi^+\pi^+$	20
$B^+ \to f_0(980)\pi^+\pi^0$	60
$B^+ \to \pi^+\pi^-\pi^+\pi^0$	30
$B^+ \to \omega\rho^+$	30
$B^+ \to \eta'\rho^+$	40

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

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). The dip at $|\cos\theta_{\rho^0}| \approx 0.85$ is due to the charm vetoes.

In summary, the fit uses six discriminating observables:

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

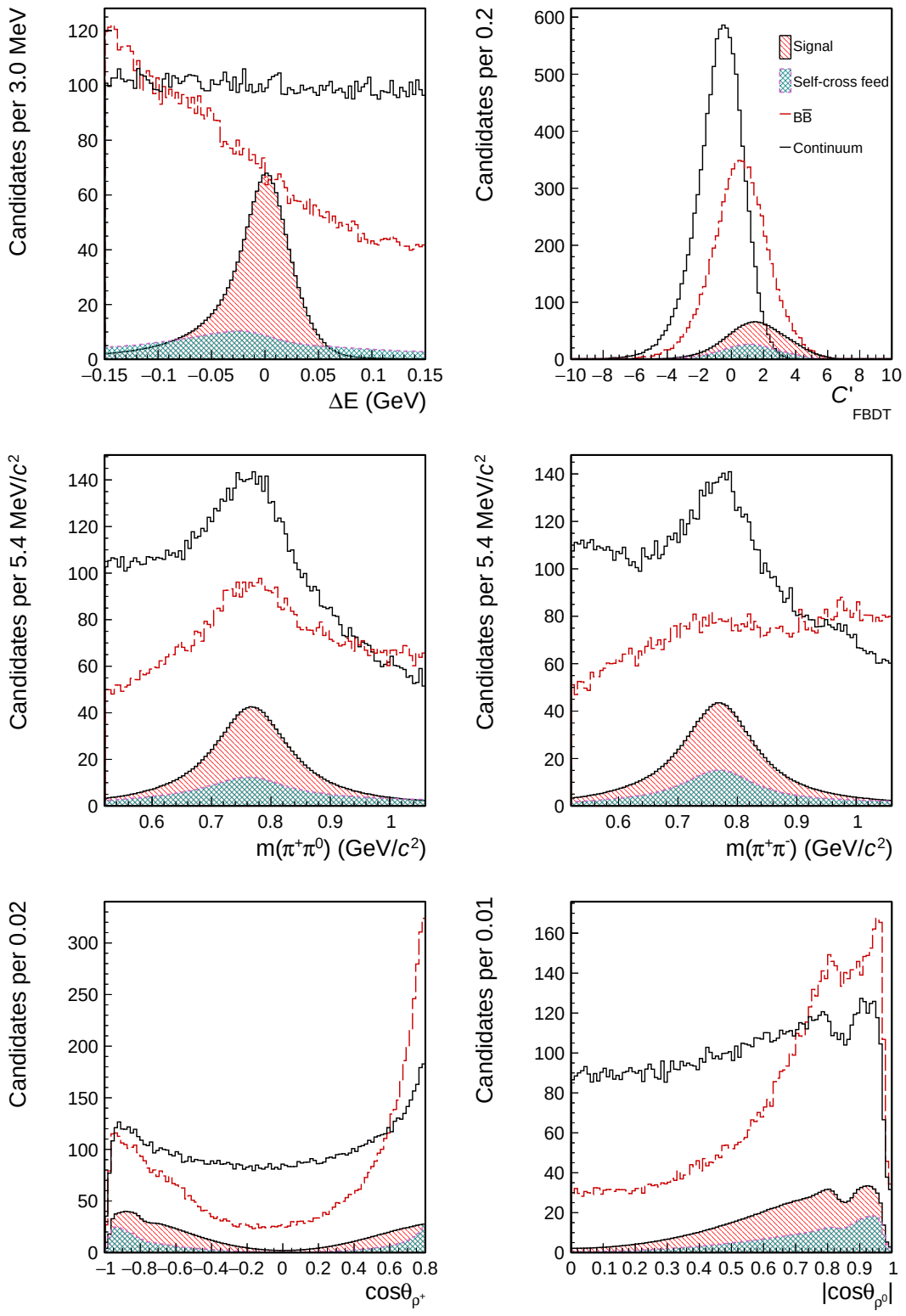


Figure 5.2: Distributions of (top left) ΔE, (top right) C', (middle left) $m(\pi^+\pi^0)$, (middle right) $m(\pi^+\pi^-)$, (bottom left) $\cos\theta_{\rho^+}$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated $B^+ \to \rho^+\rho^0$ components. Proportions do not mirror realistic expected yields and are chosen to maximize graphical visibility.

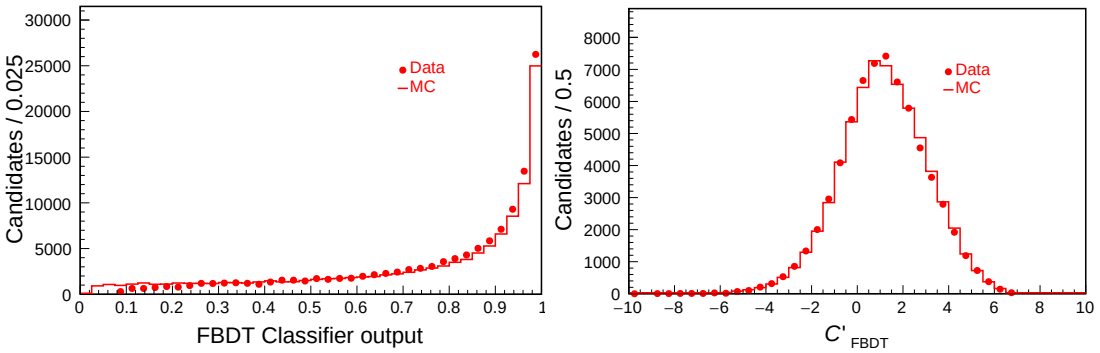


Figure 5.3: Distributions of (left) FBDT classifier output and (right) its log-transformation for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays reconstructed in collision and simulated data.

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 $N_{B\overline{B}}$,

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

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

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

The reconstruction and selection efficiency of reconstructed $B \to f$ decays ε is determined by using simulated samples as

$$\varepsilon = \frac{N}{N_{\text{gen}}}, \tag{5.5}$$

where N is the number of $B \to f$ decays meeting the selection criteria (see Tab. 4.5) and N_{gen} is the total number of generated $B \to f$ decays. Table 5.2 summarizes the resulting selection efficiencies.

In total Belle recorded $(771.6\pm10.6)\times10^6$ $B\overline{B}$ pairs [85]. As the ratio of charged to neutral $B\bar{B}$ pair production is consistent with one, $f_{B^+B^-}/f_{B^0\bar{B}^0} = 1.06 \pm 0.03$, the total number of B^+B^- pairs used in this analysis is around 397×10^6.

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 ratio, fractions of non-peaking $B\overline{B}$,

combinatorial rare B background, and branching fractions of $B^+ \to f_0\rho^+$, $B^+ \to \rho^0\pi^+\pi^0$, $B^+ \to \rho^+\pi^+\pi^-$, $B^+ \to \rho^-\pi^+\pi^+$, $B^+ \to f_0\pi^+\pi^0$, $B^+ \to \pi^+\pi^-\pi^+\pi^0$, $B^+ \to a_1^+\pi^0$, $B^0 \to a_1^+\pi^-$, $B^+ \to a_1^0\pi^+$, $B^+ \to K^{*0}\rho^+$, $B^0 \to \rho^+\rho^-$, $B^+ \to \omega\rho^+$, $B^0 \to a_1^+\rho^-$, and $B^+ \to \eta'\rho^+$ decays.

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 accuracy of this assumption is assessed during analysis validation using control decay modes.

The branching fractions 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. I use a phenomenological prediction for the proportions between the $B^+ \to a_1^+\rho^0$, $B^0 \to a_1^+\rho^-$, and $B^+ \to a_1^0\rho^+$ branching fractions [80–83], $\mathcal{B}(B^+ \to a_1^+\rho^0) : \mathcal{B}(B^0 \to a_1^+\rho^-) : \mathcal{B}(B^+ \to a_1^0\rho^+) = 1{:}2{:}1$, to constrain their relative rates and use only one free fit parameter $\mathcal{B}(B^0 \to a_1^+\rho^-)$. All fit parameters are summarized in Table 5.2.

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, as is the signal $M_{\rm bc}$ distribution, 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 all joint p.d.f. and 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 $\mathcal{L}_i$:

$$\mathcal{L}(\vec{\theta}|\vec{x}) \equiv \mathcal{L}(\vec{\theta}) = \prod_{i=1}^{N} \mathcal{L}_i(\vec{\theta}). \tag{5.6}$$

Each single event likelihood has the following expression

$$
\begin{aligned}
\mathcal{L}_i = f_{(\rho^+\rho^0)_L} p_i^{(\rho^+\rho^0)_L} + f_{(\rho^+\rho^0)_T} p_i^{(\rho^+\rho^0)_T} \\
+ r_{\text{sxf }(\rho^+\rho^0)_L} f_{(\rho^+\rho^0)_L} p_i^{\text{sxf }(\rho^+\rho^0)_L} \\
+ r_{\text{sxf }(\rho^+\rho^0)_T} f_{(\rho^+\rho^0)_T} p_i^{\text{sxf }(\rho^+\rho^0)_T} \\
+ f_{B\overline{B}} p_i^{B\bar{B}} + f_{f_0\rho^+} p_i^{f_0\rho^+} + f_{a_1^+\pi^0} p_i^{a_1^+\pi^0} + \\
f_{a_1^+\pi^-} p_i^{a_1^+\pi^-} + f_{a_1^0\pi^+} p_i^{a_1^0\pi^+} + \\
f_{K^{*0}\rho^+} p_i^{K^{*0}\rho^+} + f_{\rho^+\rho^-} p_i^{\rho^+\rho^-} + \\
f_{a_1^+\rho^0} p_i^{a_1^+\rho^0} + f_{a_1^+\rho^-} p_i^{a_1^+\rho^-} + \\
f_{a_1^0\rho^+} p_i^{a_1^0\rho^+} + f_{\rho^0\pi^+\pi^0} p_i^{\rho^0\pi^+\pi^0} + \\
f_{\rho^+\pi^+\pi^-} p_i^{\rho^+\pi^+\pi^-} + f_{\rho^-\pi^+\pi^+} p_i^{\rho^-\pi^+\pi^+} + \\
f_{f_0\pi^+\pi^0} p_i^{f_0\pi^+\pi^0} + f_{\pi^+\pi^-\pi^+\pi^0} p_i^{\pi^+\pi^-\pi^+\pi^0} + \\
f_{\omega\rho^+} p_i^{\omega\rho^+} + f_{\eta'\rho^+} p_i^{\eta'\rho^+} + \\
f_{\text{comb. rare }B} p_i^{\text{comb. rare }B} + (1 - f_{\text{sum}}) p_i^{\text{cont}},
\end{aligned}
\tag{5.7}
$$

where p_i^y indicates the six-dimensional probability density function of the generic component y (cont means continuum), and f_{sum} is the sum of all independent fractions. The detailed parametrization of the fractions is summarized in Table 5.2.

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 for estimating properly the uncertainty from the fluctuation of the total number of events.

5.2.1 Identification of dependences 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, dependences between the fit observables need to be identified and modeled.

For each component I identify the major dependences 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 $\cos\theta_{\rho^+}$ in slices of ΔE is obtained. The shape of the $\cos\theta_{\rho^+}$ distribution changes as a function of ΔE, 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.

Component	Fit parameter	Status	Selection efficiency, ε_i	Parametrization in the likelihood
$B^+ \to \rho^+\rho^0$ ($f_L = 1$)	$f_L\mathcal{B}(B^+ \to \rho^+\rho^0)$	Float	0.08	$f_{(\rho^+\rho^0)_L} = f_L\mathcal{B}(B^+ \to \rho^+\rho^0)N_{B\overline{B}}\varepsilon_L/N$
$B^+ \to \rho^+\rho^0$ ($f_L = 0$)	$(1-f_L)\mathcal{B}(B^+ \to \rho^+\rho^0)$	Float	0.1492	$f_{(\rho^+\rho^0)_T} = (1-f_L)\mathcal{B}(B^+ \to \rho^+\rho^0)N_{B\overline{B}}\varepsilon_T/N$
$B^+ \to \rho^+\rho^0$ ($f_L = 1$) sxf	$r_{\text{sxf }(\rho^+\rho^0)_L}$	Gaussian constrained	-	$r_{\text{sxf }(\rho^+\rho^0)_L} f_L\mathcal{B}(B^+ \to \rho^+\rho^0)N_{B\overline{B}}\varepsilon_L/N$
$B^+ \to \rho^+\rho^0$ ($f_L = 0$) sxf	$r_{\text{sxf }(\rho^+\rho^0)_T}$	Gaussian constrained	-	$r_{\text{sxf }(\rho^+\rho^0)_T}(1-f_L)\mathcal{B}(B^+ \to \rho^+\rho^0)N_{B\overline{B}}\varepsilon_T/N$
$B\bar{B}$	$f_{B\overline{B}}$	Float	-	$f_{B\overline{B}}$
$B^+ \to f_0\rho^+$	$f_{f_0\rho^+}$	Float	0.0614	$f_{f_0\rho^+}$
$B^+ \to a_1^+\pi^0$	$\mathcal{B}(B^+ \to a_1^+\pi^0)$	Gaussian constrained	0.0075	$f_{a_1^+\pi^0} = \mathcal{B}(B^+ \to a_1^+\pi^0)N_{B\overline{B}}\varepsilon_{a_1^+\pi^0}/N$
$B^0 \to a_1^+\pi^-$	$\mathcal{B}(B^0 \to a_1^+\pi^-)$	Gaussian constrained	0.0034	$f_{a_1^+\pi^-} = \mathcal{B}(B^0 \to a_1^+\pi^-)N_{B\overline{B}}\varepsilon_{a_1^+\pi^-}/N$
$B^+ \to a_0^+\pi^+$	$\mathcal{B}(B^+ \to a_1^0\pi^+)$	Gaussian constrained	0.006	$f_{a_1^0\pi^+} = \mathcal{B}(B^+ \to a_1^0\pi^+)N_{B\overline{B}}\varepsilon_{a_1^0\pi^0}/N$
$B^+ \to K^{*0}\rho^+$	$\mathcal{B}(B^+ \to K^{*0}\rho^+)$	Gaussian constrained	0.0162	$f_{K^{*0}\rho^+} = \mathcal{B}(B^+ \to K^{*0}\rho^+)N_{B\overline{B}}\varepsilon_{K^{*0}\rho^+}/N$
$B^0 \to \rho^+\rho^-$	$\mathcal{B}(B^0 \to \rho^+\rho^-)$	Gaussian constrained	0.0114	$f_{\rho^+\rho^-} = \mathcal{B}(B^0 \to \rho^+\rho^-)N_{B\overline{B}}\varepsilon_{\rho^+\rho^-}/N$
$B^+ \to a_1^+\rho^0$	$\mathcal{B}(B^0 \to a_1^+\rho^-)$	Float	0.0034	$f_{a_1^+\rho^0} = \mathcal{B}(B^0 \to a_1^\pm\rho^\mp)/2N_{B\overline{B}}\varepsilon_{a_1^+\rho^0}/N$
$B^0 \to a_1^+\rho^-$	$\mathcal{B}(B^0 \to a_1^+\rho^-)$	Float	0.0026	$f_{a_1^+\rho^-} = \mathcal{B}(B^0 \to a_1^\pm\rho^\mp)N_{B\overline{B}}\varepsilon_{a_1^+\rho^-}/N$
$B^+ \to a_1^0\rho^+$	$\mathcal{B}(B^0 \to a_1^+\rho^-)$	Float	0.0025	$f_{a_1^0\rho^+} = \mathcal{B}(B^0 \to a_1^\pm\rho^\mp)/2N_{B\overline{B}}\varepsilon_{a_1^0\rho^+}/N$
$B^+ \to \rho^0\pi^+\pi^0$	$f_{B^+\to\rho^0\pi^+\pi^0}$	Float	0.0136	$f_{B^+\to\rho^0\pi^+\pi^0}$
$B^+ \to \rho^+\pi^+\pi^-$	$f_{B^+\to\rho^+\pi^+\pi^-}$	Float	0.0126	$f_{B^+\to\rho^+\pi^+\pi^-}$
$B^+ \to \rho^-\pi^+\pi^+$	$f_{B^+\to\rho^-\pi^+\pi^+}$	Float	0.0026	$f_{B^+\to\rho^-\pi^+\pi^+}$
$B^+ \to f_0\pi^+\pi^0$	$f_{B^+\to f_0\pi^+\pi^0}$	Float	0.0073	$f_{B^+\to f_0\pi^+\pi^0}$
$B^+ \to \pi^+\pi^-\pi^+\pi^0$	$f_{B^+\to\pi^+\pi^-\pi^+\pi^0}$	Float	0.0043	$f_{B^+\to\pi^+\pi^-\pi^+\pi^0}$
$B^+ \to \omega\rho^+$	$\mathcal{B}(B^+ \to \omega\rho^+)$	Gaussian constrained	0.0021	$f_{\omega\rho^+} = \mathcal{B}(B^+ \to \omega\rho^+)N_{B\overline{B}}\varepsilon_{\omega\rho^+}/N$
$B^+ \to \eta'\rho^+$	$\mathcal{B}(B^+ \to \eta'\rho^+)$	Gaussian constrained	0.0057	$f_{\eta'\rho^+} = \mathcal{B}(B^+ \to \eta'\rho^+)N_{B\overline{B}}\varepsilon_{\eta'\rho^+}/N$
Combinatorial rare B	$N_{\text{comb. rare B}}$	Float	-	$f_{\text{comb. rare }B} = N_{\text{comb. rare }B}/N$
Continuum*	-	-	-	-

Table 5.2: Fit parameter summary. r_{sxf} is the self-cross-feed-to-signal-yield ratio, f_i indicates the fraction of component i, $\mathcal{B}(X)$ is the branching fraction of decay X, ε_i is the selection efficiency component i, $N_{B\overline{B}}$ is the number of $B\overline{B}$ pairs, N_i is the yield of component i, and N is the total number of events in the fitting sample.

* In a non-extended fit, there are $n-1$ independent fractions, therefore the continuum yield is determined from the normalization condition $f_n = 1 - \sum_{j=1}^{n-1} f_j$.

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 dependences are found between the ΔE, $m(\pi^+\pi^0)$, and $\cos\theta_{\rho^+}$ 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 dependences is in Appendix A. Tables 5.3–5.4 show matrices that summarize observed dependences between fit observables in the simulated signal sample. For self-cross-feed components, a small dependence between ΔE and C'_{FBDT} is also observed. Moreover, for the longitudinally polarized self-cross-feed component, the $m(\pi^+\pi^-)$ and $|\cos\theta_{\rho^0}|$ observables depend on each other, as shown in Fig. A.2. All observed dependences between fit observables in the simulated self-cross feed sample are summarized in Tables 5.5–5.6.

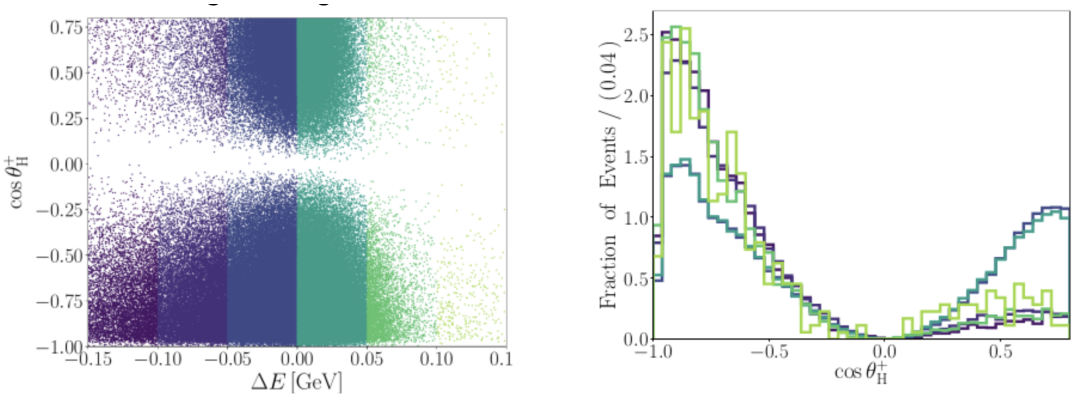


Figure 5.4: Example of dependences between fit observables: distributions of (left) $\cos\theta_{\rho^+}$ as a function of ΔE and (right) $\cos\theta_{\rho^+}$ in "slices" of ΔE for simulated longitudinally polarized signal decays.

	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 dependences between fit observables in the simulated sample of longitudinally polarized signal $B^+ \to \rho^+\rho^0$ decays. Squares indicate strong dependences, circles indicate mild dependences.

	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 dependences between fit observables in the simulated sample of transversely polarized signal $B^+ \to \rho^+\rho^0$ decays. Squares indicate strong dependences, circle indicate mild dependences.

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

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

For continuum, dependences between dipion mass and corresponding helicity observables are shown in Fig. A.4 and Table 5.7. Non-peaking $B\overline{B}$ background shows multiple dependences (Table 5.8). The most prominent dependences are between ΔE, $\cos\theta_{\rho^+}$, and $|\cos\theta_{\rho^0}|$, as shown in Fig. A.5. I also inspect "sliced" distributions for most of the 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 dependences as for continuum are observed for the $B^+ \to a_1^+\pi^0$ decay mode. Many peculiar dependences

are observed for the $B^+ \to K^{*0}\rho^+$ decay mode, such as dependences between ΔE, $|\cos\theta_{\rho^0}|$, $\cos\theta_{\rho^+}$,and $m(\pi^+\pi^-)$ (Fig. A.8).

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

Table 5.7: Matrix representing dependences between fit observables in the simulated sample of continuum events. Squares indicate strong dependences, circles indicate mild dependences.

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

Table 5.8: Matrix representing dependences between fit observables in the simulated sample of non-peaking $B\overline{B}$ background events. Squares indicate strong dependences, circles indicate mild dependences.

This study allows identifying the factorization of multidimensional p.d.f.'s that approximates better the real dependences present in the sample. Not all dependences 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 dependences 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 dependences 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. Difficulties in identifying an appropriate fit model is the main reason why the Belle collaboration failed to update its early 2003 measurement. The model is described in the following sections.

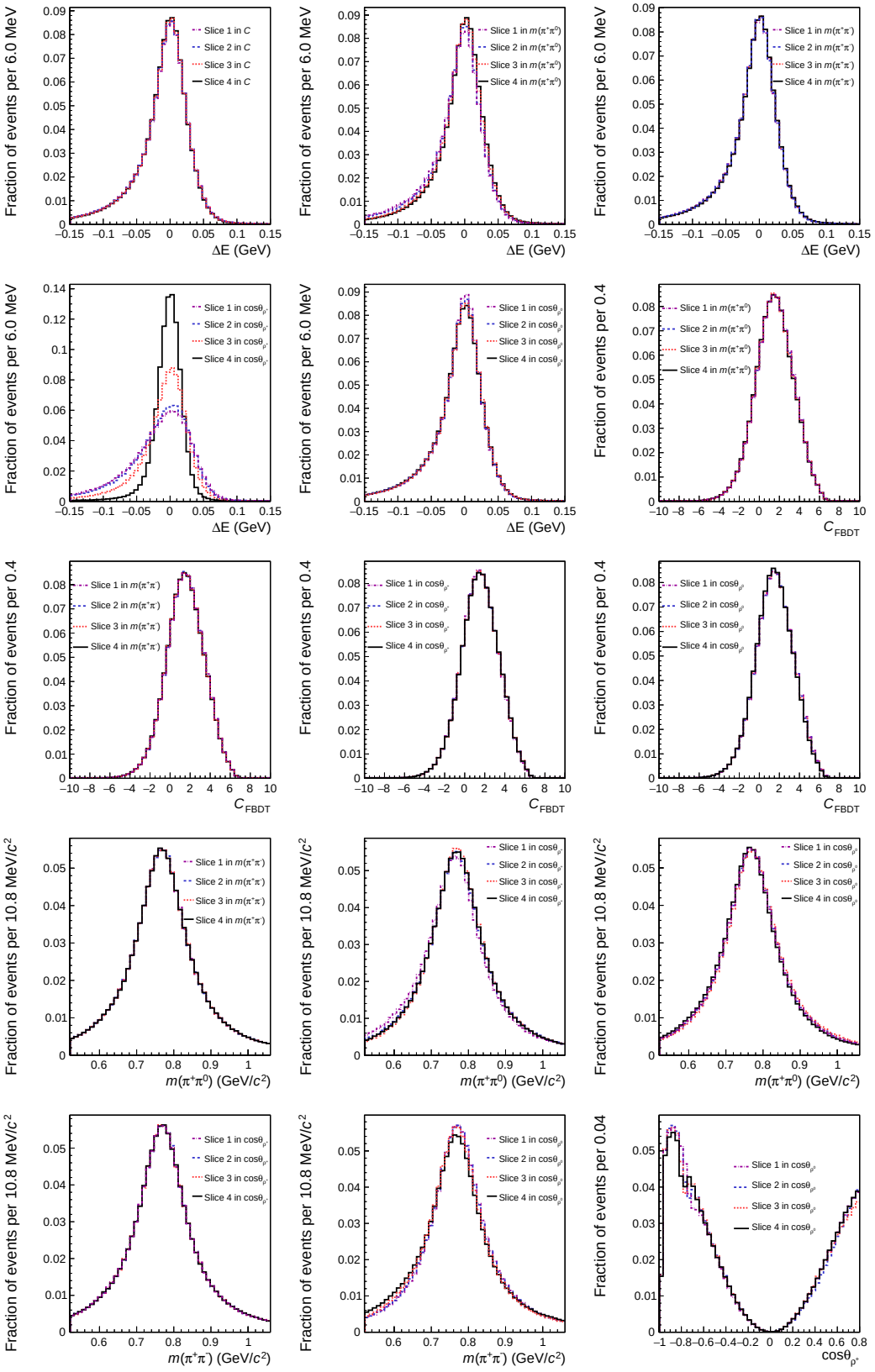


Figure 5.5: Example of dependence studies: distributions of fit observables in slices of other fit observables for simulated longitudinally polarized signal decays. Distributions normalized to unity. The full set of distributions in Appendix A.

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,

$$p_j = p_j(\Delta E, m(\pi^+\pi^0), \cos\theta_{\rho^+}) p_j(C') p_j(m(\pi^+\pi^-)) p_j(|\cos\theta_{\rho^0}|), \tag{5.8}$$

where index j corresponds to the longitudinally or transversely polarized signal; $p(\Delta E, m(\pi^+\pi^0), \cos\theta_{\rho^+})$ is a joint p.d.f. that describes the three-dimensional distribution of the ΔE, $m(\pi^+\pi^0)$, and $\cos\theta_{\rho^+}$ observables ("correlated term"), as shown in Figs. 5.5 and A.1; $p(C')$ is the p.d.f. that describes the continuum suppression output ("continuum suppression term"); $p(m(\pi^+\pi^-))$ describes the ρ^0 invariant mass ("ρ^0 mass term"); and $p(|\cos\theta_{\rho^0}|)$ describes the absolute value of the ρ^0 helicity angle ("ρ^0 helicity term").

Figure 5.7–5.8 shows one-dimensional projections of the empirical binned templates used to describe the correlated and ρ^0 helicity terms.

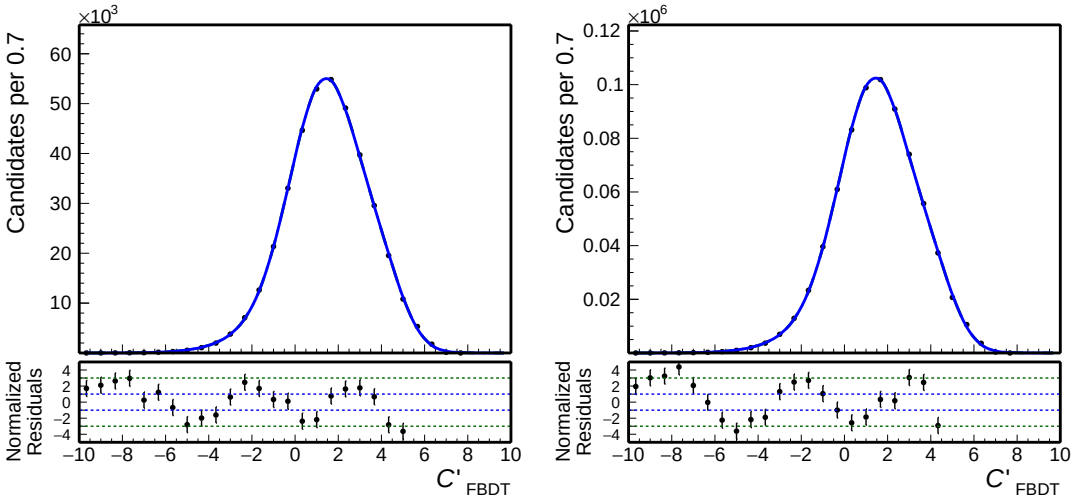


Figure 5.6: Example of fit modeling: distributions of continuum suppression output for (left) simulated longitudinally polarized and (right) transversely polarized signal. The corresponding model shapes are overlaid (solid). Normalized residuals for C'_{FBDT} values larger than 5.0 take values between -5 and -7 and are hence not displayed. This mismodeling of the upper C'_{FBDT} tail does not spoil the general qualitative agreement and is accounted for in C'_{FBDT}-shape systematic uncertainties. Full set of distributions in Appendix A.

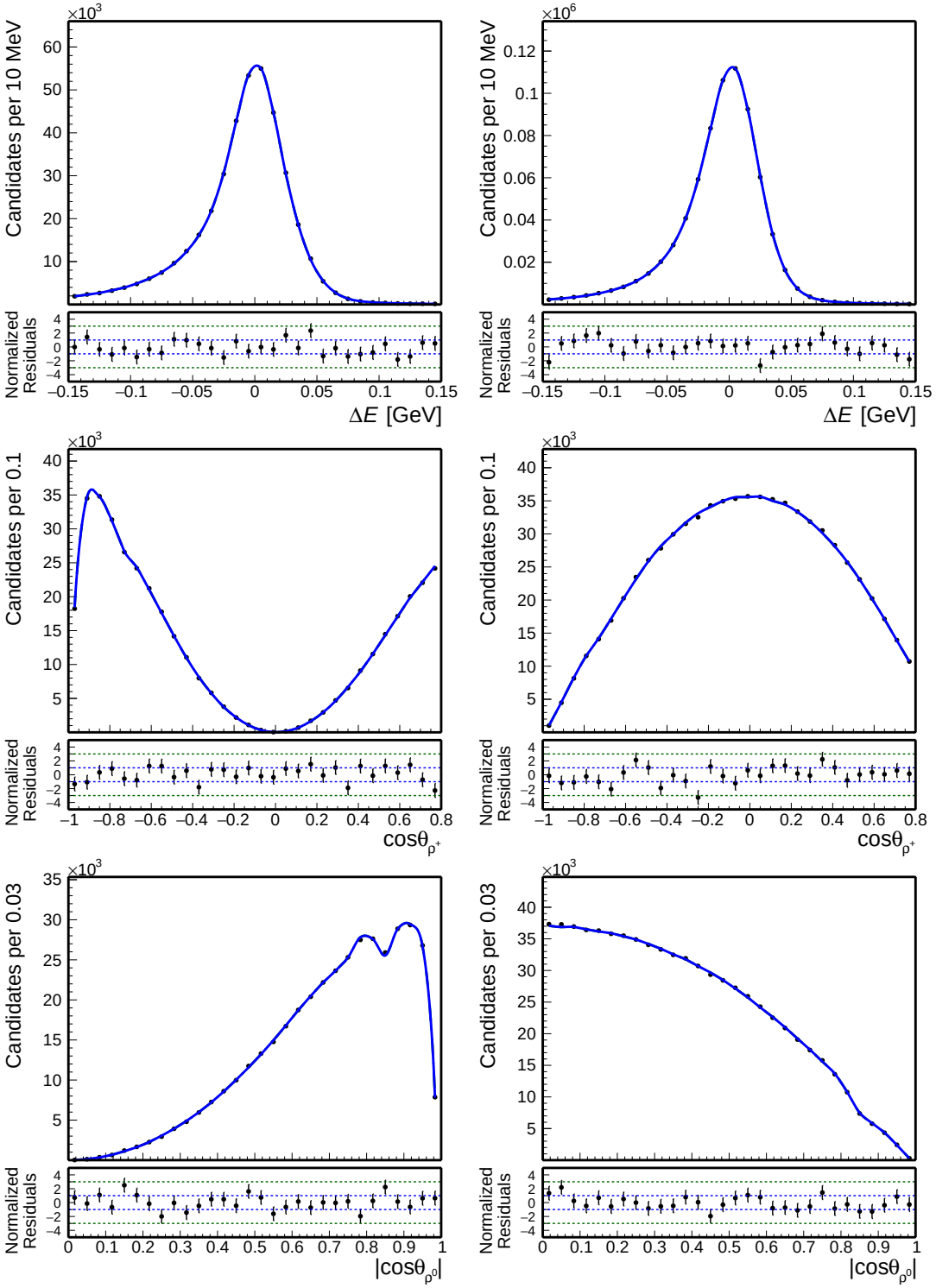


Figure 5.7: Example of fit modeling: distributions of (top) ΔE, (middle) $\cos\theta_{\rho^+}$, and (bottom) $|\cos\theta_{\rho^0}|$ for simulated longitudinally polarized (left) and (right) transversely polarized signal. The corresponding model shapes are overlaid (solid). Full set of distributions in Appendix A.

The p.d.f. of the continuum suppression term is modeled empirically as the sum of three

Gaussian functions,

$$p(C') \propto f_1 \frac{e^{-\frac{1}{2}(\frac{C'-\mu_1}{\sigma_1})}}{\sigma_1\sqrt{2\pi}} + f_2 \frac{e^{-\frac{1}{2}(\frac{C'-\mu_2}{\sigma_2})}}{\sigma_2\sqrt{2\pi}} + (1-f_1-f_2)\frac{e^{-\frac{1}{2}(\frac{C'-\mu_3}{\sigma_3})}}{\sigma_3\sqrt{2\pi}}, \tag{5.9}$$

where the weight f_i, mean μ_i, and width σ_i of each Gaussian are determined from a fit in simulation. The shapes are derived separately for each of the components. Figure 5.6 shows that the functional form chosen to describe C' distributions is adequate.

The p.d.f. of the ρ^0 mass term for the signal components is modeled using the physics-motivated Breit-Wigner function,

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

where m_{ρ^0} and Γ_{ρ^0} are determined from a fit in simulation. Figure 5.8 shows that the model is adequate.

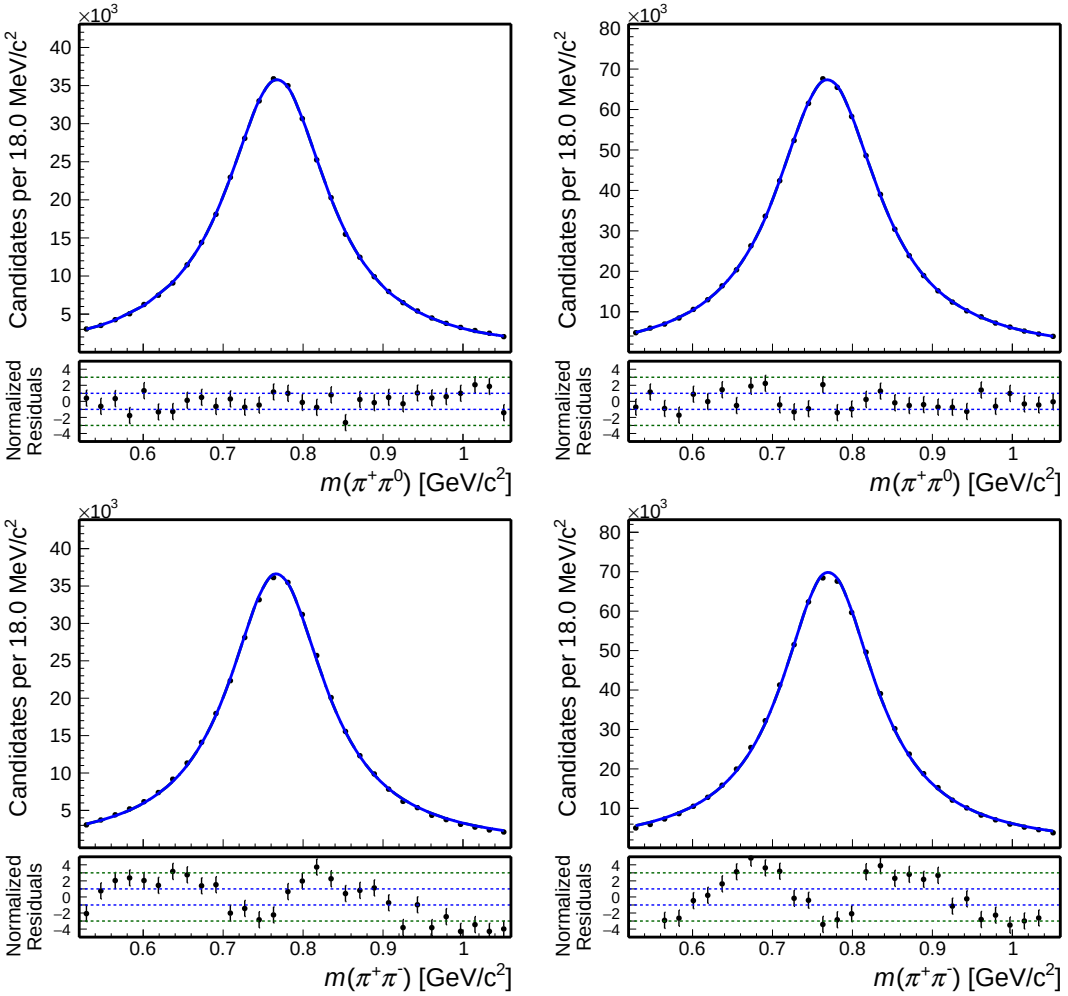


Figure 5.8: Distributions of (top) $m(\pi^+\pi^0)$ and (bottom) $m(\pi^+\pi^-)$ for simulated (left) longitudinally polarized and (right) transversely polarized signal. The corresponding model shapes are overlaid (solid). Full set of distributions in Appendix A.

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,

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

where $p(\Delta E, m(\pi^+\pi^0), \cos\theta_{\rho^+})$ is a joint p.d.f. that describes the three-dimensional distribution of the ΔE, $m(\pi^+\pi^0)$, and $\cos\theta_{\rho^+}$ observables ("ΔE correlated term"), as shown in Fig. A.2; $p(C')$ is the p.d.f. that describes the continuum suppression output ("continuum suppression term"); $p(m(\pi^+\pi^-), |\cos\theta_{\rho^0}|)$ is a joint p.d.f. that describes the two-dimensional distribution of the $m(\pi^+\pi^-)$ and $|\cos\theta_{\rho^0}|$ observables ("ρ^0 correlated term").

Figure A.9 shows the corresponding one-dimensional projections of the model; it also shows that a sum of three Gaussian functions (Eq. (5.9)) is adequate to describe the continuum suppression output distribution.

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

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

where $p(\Delta E, m(\pi^+\pi^0), \cos\theta_{\rho^+})$ is a joint p.d.f. that describes the three-dimensional distribution of the ΔE, $m(\pi^+\pi^0)$ and $\cos\theta_{\rho^+}$ distributions ("correlated term"), as shown in Figures A.3; $p(C')$ is a one-dimensional p.d.f. that describes the continuum suppression output distribution ("continuum suppression term"); $p(m(\pi^+\pi^-))$ describes the ρ^0 invariant mass distribution ("ρ^0 mass term"); and $p(|\cos\theta_{\rho^0}|)$ describes the ρ^0 helicity angle ("ρ^0 helicity term").

The correlated and ρ^0 helicity terms are modeled with an empirical binned templates (Fig. A.10). Figure A.10 (top right panel) shows that a sum of three Gaussian functions (Eq. (5.9)) is adequate to describe the continuum suppression term. The p.d.f. of the ρ^0 mass term is described with a Breit-Wigner functions (Eq. (5.10)), shown to be adequate in. Figure A.10 (bottom left panel).

5.2.4 Continuum p.d.f.

The p.d.f. for continuum is factorized into the product of four terms,

$$p_{\rm cont} = p(\Delta E) p(C') p(m(\pi^+\pi^0), \cos\theta_{\rho^+}) p(m(\pi^+\pi^-), |\cos\theta_{\rho^0}|), \tag{5.13}$$

where $p(\Delta E)$ is the p.d.f. that describes the ΔE distribution ("ΔE term"); $p(C')$ is a p.d.f. that describes the C' distribution ("continuum suppression term"); $p(m(\pi^+\pi^0), \cos\theta_{\rho^+})$ is a joint p.d.f. that describes the two-dimensional $m(\pi^+\pi^0)$ and $\cos\theta_{\rho^+}$ distribution ("ρ^+ correlated term"); and $p(m(\pi^+\pi^-), |\cos\theta_{\rho^0}|)$ is a joint p.d.f. that describes the two-dimensional $m(\pi^+\pi^-)$ and $|\cos\theta_{\rho^0}|$ distributions ("ρ^0 correlated term"). I choose such factorization based on the sliced distributions of Fig. A.4, where a dependence between ρ mass and its helicity angle is observed.

The ρ^+ and ρ^0 correlated terms are modeled with empirical binned templates (Fig. A.11). The ΔE term is modeled with a second-order polynomial function,

$$p(\Delta E) \propto a\Delta E^2 + b\Delta E + 1, \tag{5.14}$$

where a and b are determined from a modeling fit in simulation. Figure A.11 (top left panel) illustrates that the chosen function is adequate. Figure A.11 (top right panel) shows that a sum of three Gaussian functions (Eq. (5.9)) is adequate to describe the continuum suppression term.

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

The non-peaking portion of $B\bar{B}$ background p.d.f. is factorized into four terms,

$$p(B\bar{B}) = p(\Delta E, \cos\theta_{\rho^+}, |\cos\theta_{\rho^0}|)p(C')p(m(\pi^+\pi^0))p(m(\pi^+\pi^-)), \tag{5.15}$$

where $p(\Delta E, \cos\theta_{\rho^+}, |\cos\theta_{\rho^0}|)$ is a joint p.d.f. that describes the three-dimensional distribution of the ΔE, $\cos\theta_{\rho^+}$, and $|\cos\theta_{\rho^0}|$ observables ("correlated term"), as shown in Fig. A.5; $p(C')$ describes the C' distribution ("continuum suppression term"); $p(m(\pi^+\pi^0))$ and $p(m(\pi^+\pi^-))$ describe the $m(\pi^+\pi^0)$ and $m(\pi^+\pi^-)$ distributions ("mass term"), respectively. The model is obtained from simulated samples of favored B decays.

The correlated term is modeled with empirical binned template (Fig. A.12). The p.d.f. of the continuum suppression term is modeled with a sum of three Gaussian functions (Eq. (5.9)). Figure A.12 (top right panel) shows that the model is adequate. The p.d.f. of the ρ^+ mass term is modeled with a fifth-order polynomial function,

$$\begin{aligned} p(m(\pi^+\pi^0)) \propto am(\pi^+\pi^0)^5 + bm(\pi^+\pi^0)^4 + cm(\pi^+\pi^0)^3 + \\ dm(\pi^+\pi^0)^2 + em(\pi^+\pi^0) + 1, \end{aligned} \tag{5.16}$$

where a, b, c, d, and e are the free parameters in the modeling fit in simulation. The ρ^0 mass term is described with a second-order polynomial function,

$$p(m(\pi^+\pi^-)) \propto am(\pi^+\pi^-)^2 + bm(\pi^+\pi^-) + 1, \tag{5.17}$$

where a and b are the free parameters in the modeling fit in simulation. Figure A.12 shows that the chosen models are adequate.

5.2.6 P.d.f. summary

I determine the p.d.f.'s of every rare B background component separately. I summarize the model choices in Table 5.9 not to overload the text with model details. One-dimensional projections of rare B background 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 branching fractions of decay modes for which an external measurement exists, $B^+ \to a_1^+\pi^0$, $B^0 \to \rho^+\rho^-$, $B^0 \to a_1^+\pi^-$, $B^+ \to K^{*0}\rho^+$, $B^+ \to a_1^0\pi^+$, $B^+ \to \omega\rho^+$, and $B^+ \to \eta'\rho^+$, are also Gaussian constrained to the observed values. The fractions 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. 8.12).

Figure 5.9 shows example projections obtained from a fit to a realistic simulated sample (Fig. 5.10 shows the same projections with vertical axis in log scale). 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 $C' \approx 0$, while the peak corresponding to $B\overline{B}$ events (including signal and

Component	ΔE	C'	$m(\pi^+\pi^0)$	$m(\pi^+\pi^-)$	$\cos\theta_{\rho^+}$	$\lvert\cos\theta_{\rho^0}\rvert$
$B^+ \to \rho^+\rho^0$ ($f_L = 1$)	3D hist (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	3G	3D hist (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	BW	3D hist (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	1D hist ($\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to \rho^+\rho^0$ ($f_L = 0$)	3D hist (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	3G	3D hist (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	BW	3D hist (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	1D hist ($\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to \rho^+\rho^0$ ($f_L = 1$) sxf	3D hist (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	3G	3D hist (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	2D hist ($m(\pi^+\pi^-)$,$\lvert\cos\theta_{\rho^0}\rvert$)	3D hist (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	2D hist ($m(\pi^+\pi^-)$,$\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to \rho^+\rho^0$ ($f_L = 0$) sxf	3D hist (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	3G	3D hist (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	BW	3D hist (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	1D hist ($\lvert\cos\theta_{\rho^0}\rvert$)
Continuum	Pol2	3G	2D hist ($m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	2D hist ($m(\pi^+\pi^-)$,$\lvert\cos\theta_{\rho^0}\rvert$)	2D hist ($m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	2D hist ($m(\pi^+\pi^-)$,$\lvert\cos\theta_{\rho^0}\rvert$)
$B\bar{B}$	3D hist (ΔE, $\cos\theta_{\rho^+}$,$\lvert\cos\theta_{\rho^0}\rvert$)	3G	Pol5	Pol3	3D hist (ΔE, $\cos\theta_{\rho^+}$,$\lvert\cos\theta_{\rho^0}\rvert$)	3D hist (ΔE, $\cos\theta_{\rho^+}$,$\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to f_0\rho^+$	2D hist (ΔE, $\cos\theta_{\rho^+}$)	3G	BW	3G	2D hist (ΔE, $\cos\theta_{\rho^+}$)	1D hist ($\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to a_1^+\pi^0$	2D hist (ΔE, $\cos\theta_{\rho^+}$)	3G	Pol5	2D hist ($m(\pi^+\pi^-)$,$\lvert\cos\theta_{\rho^0}\rvert$)	2D hist (ΔE, $\cos\theta_{\rho^+}$)	2D hist ($m(\pi^+\pi^-)$,$\lvert\cos\theta_{\rho^0}\rvert$)
$B^0 \to a_1^+\pi^-$	3D hist (ΔE, $\cos\theta_{\rho^+}$,$\lvert\cos\theta_{\rho^0}\rvert$)	3G	2G	Pol5	3D hist (ΔE, $\cos\theta_{\rho^+}$, $\lvert\cos\theta_{\rho^0}\rvert$)	3D hist (ΔE, $\cos\theta_{\rho^+}$,$\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to a_1^0\pi^+$	3D hist (ΔE, $\cos\theta_{\rho^+}$,$\lvert\cos\theta_{\rho^0}\rvert$)	3G	Pol2 + BW	Pol5	3D hist (ΔE, $\cos\theta_{\rho^+}$,$\lvert\cos\theta_{\rho^0}\rvert$)	3D hist (ΔE, $\cos\theta_{\rho^+}$,$\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to K^{*0}\rho^+$	3G	3G	BW	3D hist ($m(\pi^+\pi^-)$, $\cos\theta_{\rho^+}$,$\lvert\cos\theta_{\rho^0}\rvert$)	3D hist ($m(\pi^+\pi^-)$, $\cos\theta_{\rho^+}$,$\lvert\cos\theta_{\rho^0}\rvert$)	3D hist ($m(\pi^+\pi^-)$, $\cos\theta_{\rho^+}$,$\lvert\cos\theta_{\rho^0}\rvert$)
$B^0 \to \rho^+\rho^-$	3D hist (ΔE, $\cos\theta_{\rho^+}$, $\lvert\cos\theta_{\rho^0}\rvert$)	3G	2G	Pol5	3D hist (ΔE, $\cos\theta_{\rho^+}$, $\lvert\cos\theta_{\rho^0}\rvert$)	3D hist (ΔE, $\cos\theta_{\rho^+}$, $\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to a_1^+\rho^0$	3D hist (ΔE, $\cos\theta_{\rho^+}$, $\lvert\cos\theta_{\rho^0}\rvert$)	3G	2G	Pol2 + BW	3D hist (ΔE, $\cos\theta_{\rho^+}$, $\lvert\cos\theta_{\rho^0}\rvert$)	3D hist (ΔE, $\cos\theta_{\rho^+}$, $\lvert\cos\theta_{\rho^0}\rvert$)
$B^0 \to a_1^+\rho^-$	3D hist (ΔE, $\cos\theta_{\rho^+}$, $\lvert\cos\theta_{\rho^0}\rvert$)	3G	2G	2G	3D hist (ΔE, $\cos\theta_{\rho^+}$, $\lvert\cos\theta_{\rho^0}\rvert$)	3D hist (ΔE, $\cos\theta_{\rho^+}$, $\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to a_1^0\rho^+$	3D hist (ΔE, $\cos\theta_{\rho^+}$, $\lvert\cos\theta_{\rho^0}\rvert$)	3G	2G	2G	3D hist (ΔE, $\cos\theta_{\rho^+}$, $\lvert\cos\theta_{\rho^0}\rvert$)	3D hist (ΔE, $\cos\theta_{\rho^+}$, $\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to \rho^0\pi^+\pi^0$	3G	3G	Pol3	Pol2 + BW	1D hist ($\cos\theta_{\rho^+}$)	1D hist ($\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to \rho^+\pi^+\pi^-$	3G	3G	2G	Pol3	1D hist ($\cos\theta_{\rho^+}$)	1D hist ($\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to \rho^-\pi^+\pi^+$	3G	3G	2G	Pol3	1D hist ($\cos\theta_{\rho^+}$)	1D hist ($\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to f_0\pi^+\pi^0$	3G	3G	Pol3	3G	1D hist ($\cos\theta_{\rho^+}$)	1D hist ($\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to \pi^+\pi^-\pi^+\pi^0$	3G	3G	Pol3	Pol3	1D hist ($\cos\theta_{\rho^+}$)	1D hist ($\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to \omega\rho^+$	3D hist (ΔE, $\cos\theta_{\rho^+}$,$\lvert\cos\theta_{\rho^0}\rvert$)	3G	2G	3G	3D hist (ΔE, $\cos\theta_{\rho^+}$,$\lvert\cos\theta_{\rho^0}\rvert$)	3D hist (ΔE, $\cos\theta_{\rho^+}$,$\lvert\cos\theta_{\rho^0}\rvert$)
$B^+ \to \eta'\rho^+$	3D hist (ΔE, $m(\pi^+\pi^-)$,$\lvert\cos\theta_{\rho^0}\rvert$)	3G	2G	3D hist (ΔE, $m(\pi^+\pi^-)$,$\lvert\cos\theta_{\rho^0}\rvert$)	1D hist ($\cos\theta_{\rho^+}$)	3D hist (ΔE, $m(\pi^+\pi^-)$,$\lvert\cos\theta_{\rho^0}\rvert$)
Combinatorial rare B	1D hist (ΔE)	1D hist (C_{FBDT})	1D hist ($m(\pi^+\pi^0)$)	1D hist ($m(\pi^+\pi^-)$)	1D hist ($\cos\theta_{\rho^+}$)	1D hist ($\lvert\cos\theta_{\rho^0}\rvert$)

Table 5.9: Model summary. The shorthand "nD hist" means n-dimensional histogram, "3G (2G)" means sum of three (two) Gaussian functions, "BW" indicates a Breit-Wigner, "Poln" indicates a polynomial of n-th degree.

background) is shifted toward $C' \approx 1$. 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.

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,

$$\text{Pull}(\theta_i) = \frac{\hat{\theta}_i - \theta_i}{\hat{\sigma}_{\hat{\theta}_i}}, \tag{5.18}$$

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 300 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 5.11 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 differ from zero. Biases of 20% for $\mathcal{B}(B^+ \to \rho^+\rho^0)$ and -30% for f_L are observed. I consider these remarkable improvements over the much larger biases up to 200% observed with oversimplified assumptions on the model as in Belle's 2003 measurement. The biases observed here are sufficiently small to be included in the systematic uncertainties without spoiling excessively the precision of the measurement.

Figure 5.12 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, we expect to significantly improve the precision of $\mathcal{B}(B^+ \to \rho^+\rho^0)$.

5.3.1 Dependence of bias on the $B^+ \to \rho\pi\pi$ branching fraction

Biases depend on the number of $B^+ \to \rho\pi\pi$ decays in the sample which is unfortunately unknown as this decay is unobserved. To render the measurement robust against this uncertainty I repeat the pull studies assuming various values of $B^+ \to \rho\pi\pi$ branching

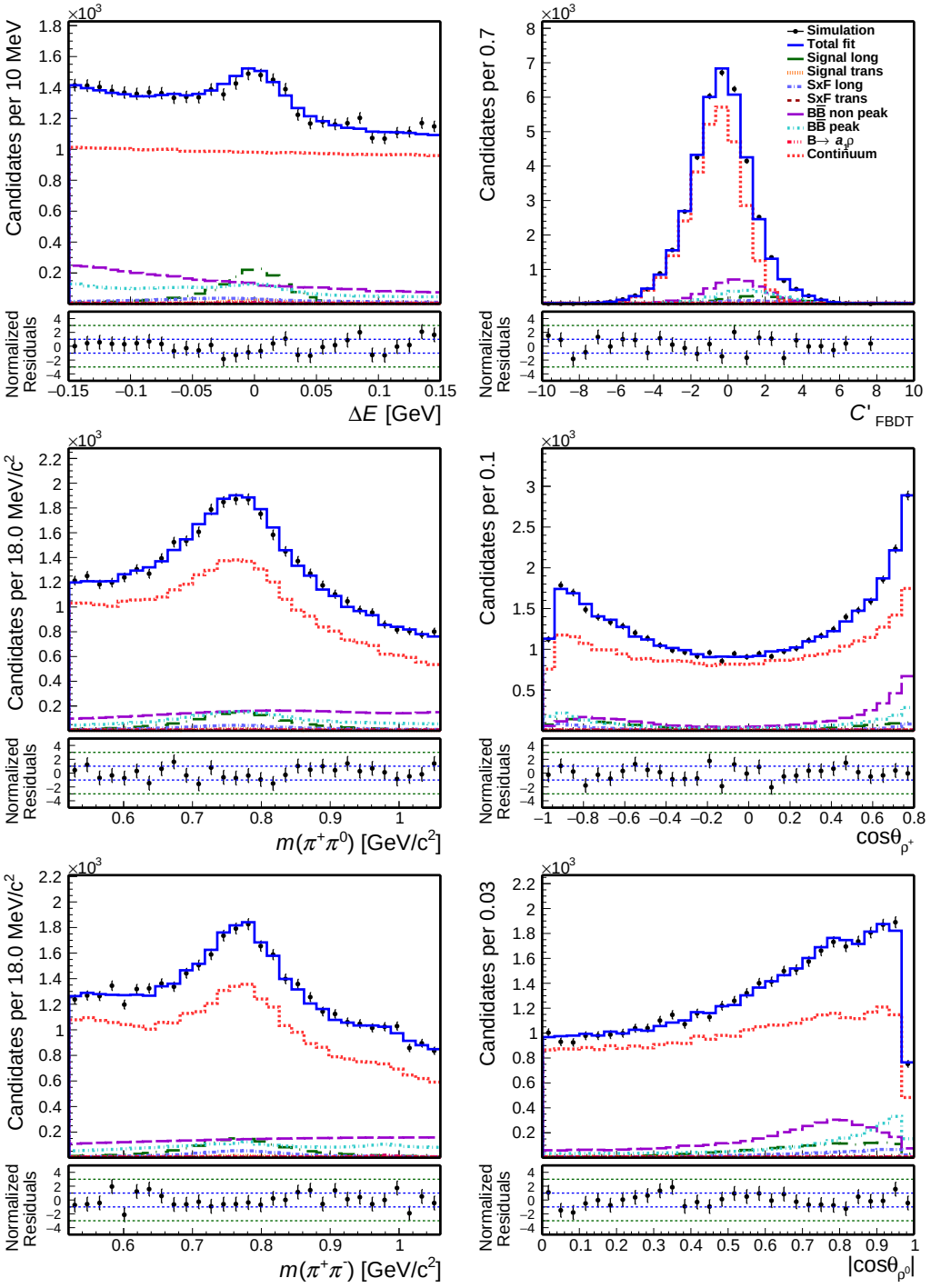


Figure 5.9: Example of fit projections: distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for a realistic simulated Belle sample with fit projections overlaid.

fractions, taken to be equal across isospin-partner channels, in a range from 0 to 10^{-5}, as shown in Figure 5.13. The bias depends linearly on the assumed true value of the $B^+ \to \rho\pi\pi$ branching fraction, but its value stays within an acceptable level.

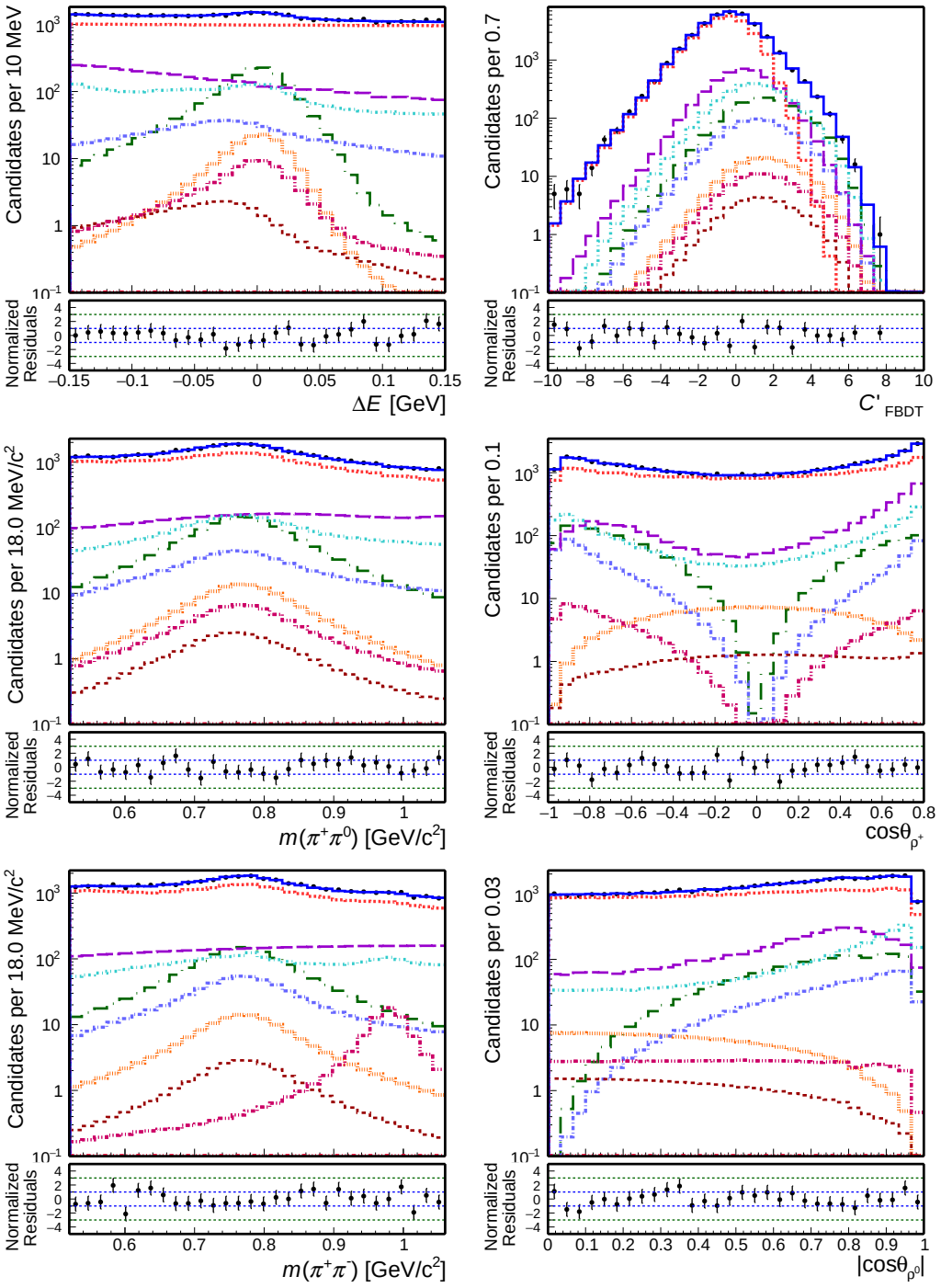


Figure 5.10: Example of fit projections: distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for a realistic simulated Belle sample with fit projections overlaid. The vertical axis is in log-scale.

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

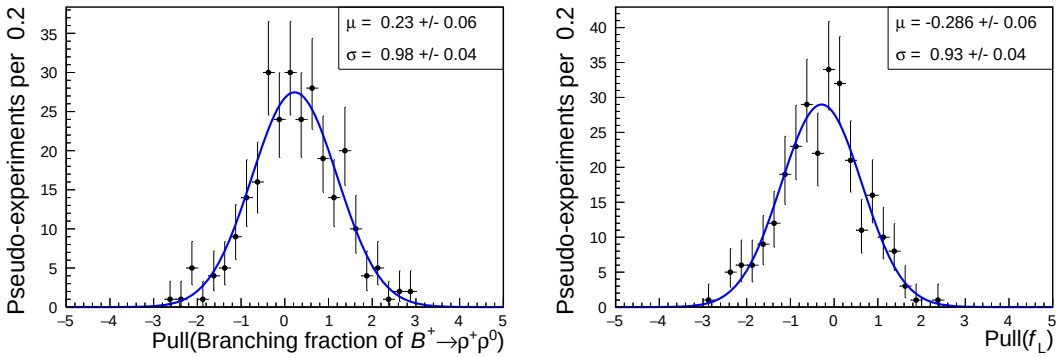


Figure 5.11: Pull distributions of (left) branching fraction and (right) fraction of longitudinally polarized $B^+ \to \rho^+\rho^0$ decays. Results of fits to Gaussian functions are overlaid (solid).

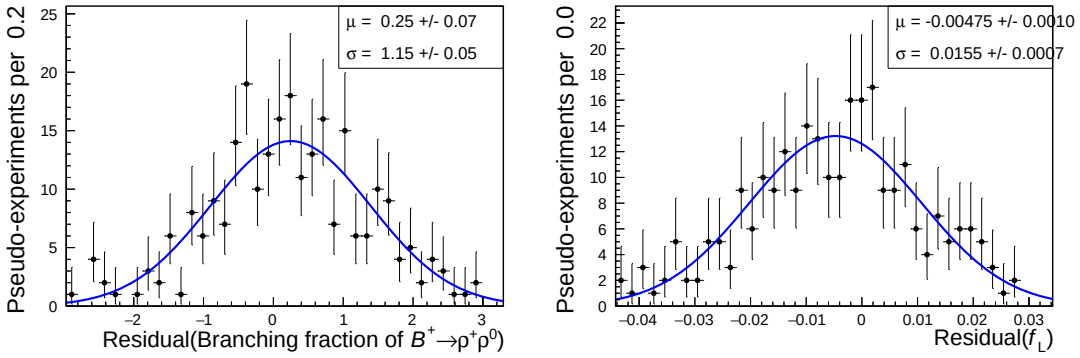


Figure 5.12: Residual distributions of (left) branching fraction and (right) fraction of longitudinally polarized $B^+ \to \rho^+\rho^0$ decays. Results of fits to Gaussian functions are overlaid (solid).

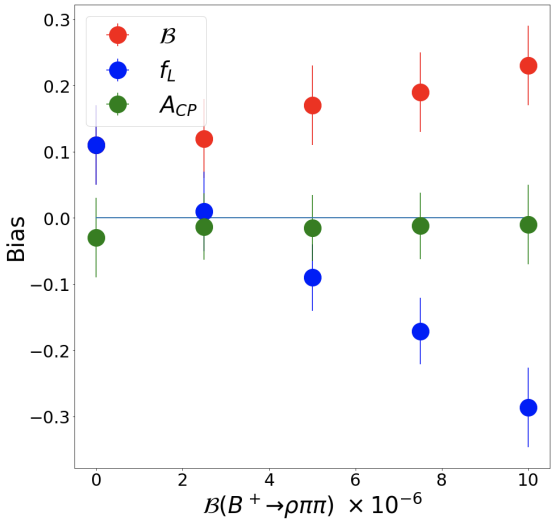


Figure 5.13: Observed bias in the estimate of branching fraction and fraction of longitudinally polarized $B^+ \to \rho^+\rho^0$ decays as a function of $B^+ \to \rho\pi\pi$ fraction in simulated samples.

candidates,

$$\mathcal{L}_i = \mathcal{L}_i^+ + \mathcal{L}_i^-, \tag{5.19}$$

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

$$\begin{aligned}
\mathcal{L}_i^+ = N^+[f_{\rm sig}^+(f_{(\rho^+\rho^0)_L} p_i^{(\rho^+\rho^0)_L} + f_{(\rho^+\rho^0)_T} p_i^{(\rho^+\rho^0)_T}) \\
+f_{\rm sig}^+(r_{\text{sxf }(\rho^+\rho^0)_L} f_{(\rho^+\rho^0)_L} p_i^{\text{sxf }(\rho^+\rho^0)_L} \\
+r_{\text{sxf }(\rho^+\rho^0)_T} f_{(\rho^+\rho^0)_T} p_i^{\text{sxf }(\rho^+\rho^0)_T}) \\
+f_{B\overline{B}}^+(f_{B\overline{B}} p_i{}^{B\bar{B}} + f_{f_0\rho^+} p_i^{f_0\rho^+} + f_{a_1^+\pi^0} p_i^{a_1^+\pi^0} \\
+f_{a_1^+\pi^-} p_i^{a_1^+\pi^-} + f_{a_1^0\pi^+} p_i^{a_1^0\pi^+} \\
+f_{K^{*0}\rho^+} p_i^{K^{*0}\rho^+} + f_{\rho^+\rho^-} p_i^{\rho^+\rho^-} \\
+f_{a_1^+\rho^0} p_i^{a_1^+\rho^0} + f_{a_1^+\rho^-} p_i^{a_1^+\rho^-} \\
+f_{a_1^0\rho^+} p_i^{a_1^0\rho^+} + f_{\rho^0\pi^+\pi^0} p_i^{\rho^0\pi^+\pi^0} \\
+f_{\rho^+\pi^+\pi^-} p_i^{\rho^+\pi^+\pi^-} + f_{\rho^-\pi^+\pi^+} p_i^{\rho^-\pi^+\pi^+} \\
+f_{f_0\pi^+\pi^0} p_i^{f_0\pi^+\pi^0} + f_{\pi^+\pi^-\pi^+\pi^0} p_i^{\pi^+\pi^-\pi^+\pi^0} \\
+f_{\omega\rho^+} p_i^{\omega\rho^+} + f_{\eta'\rho^+} p_i^{\eta'\rho^+} \\
+f_{\text{comb. rare}B} p_i^{\text{comb. rare}B}) + (1 - f_{\rm sum}^+) p_i^{\rm cont}],
\end{aligned} \tag{5.20}$$

and

$$\begin{aligned}
\mathcal{L}_i^- = N^-[(1 - f_{\rm sig}^+)(f_{(\rho^+\rho^0)_L} p_i^{(\rho^+\rho^0)_L} + f_{(\rho^+\rho^0)_T} p_i^{(\rho^+\rho^0)_T}) \\
+(1 - f_{\rm sig}^+)(r_{\text{sxf }(\rho^+\rho^0)_L} f_{(\rho^+\rho^0)_L} p_i^{\text{sxf }(\rho^+\rho^0)_L} \\
+r_{\text{sxf }(\rho^+\rho^0)_T} f_{(\rho^+\rho^0)_T} p_i^{\text{sxf }(\rho^+\rho^0)_T}) \\
+(1 - f_{B\overline{B}}^+)(f_{B\overline{B}} p_i{}^{B\bar{B}} + f_{f_0\rho^+} p_i^{f_0\rho^+} + f_{a_1^+\pi^0} p_i^{a_1^+\pi^0} \\
+f_{a_1^+\pi^-} p_i^{a_1^+\pi^-} + f_{a_1^0\pi^+} p_i^{a_1^0\pi^+} \\
+f_{K^{*0}\rho^+} p_i^{K^{*0}\rho^+} + f_{\rho^+\rho^-} p_i^{\rho^+\rho^-} \\
+f_{a_1^+\rho^0} p_i^{a_1^+\rho^0} + f_{a_1^+\rho^-} p_i^{a_1^+\rho^-} \\
+f_{a_1^0\rho^+} p_i^{a_1^0\rho^+} + f_{\rho^0\pi^+\pi^0} p_i^{\rho^0\pi^+\pi^0} \\
+f_{\rho^+\pi^+\pi^-} p_i^{\rho^+\pi^+\pi^-} + f_{\rho^-\pi^+\pi^+} p_i^{\rho^-\pi^+\pi^+} \\
+f_{f_0\pi^+\pi^0} p_i^{f_0\pi^+\pi^0} + f_{\pi^+\pi^-\pi^+\pi^0} p_i^{\pi^+\pi^-\pi^+\pi^0} \\
+f_{\omega\rho^+} p_i^{\omega\rho^+} + f_{\eta'\rho^+} p_i^{\eta'\rho^+} \\
+f_{\text{comb. rare}B} p_i^{\text{comb. rare}B}) + (1 - f_{\rm sum}^-) p_i^{\rm cont}],
\end{aligned} \tag{5.21}$$

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.

I verify that probability density functions used in Section 5.1 are adequate to describe distributions of both positively and negatively charged candidates.

The fit model describes well the distributions obtained for both charges, as shown in Figs. 5.14–5.15 obtained from a simultaneous fit to the realistic simulated sample of positively and negatively charged $B^+ \to \rho^+\rho^0$ candidates.

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 estimated (Fig. 5.16, left plot). By inspecting the residual distribution (Fig. 5.16, right plot), I expect to obtain the world-leading result for A_{CP} of $B^+ \to \rho^+\rho^0$ decay.

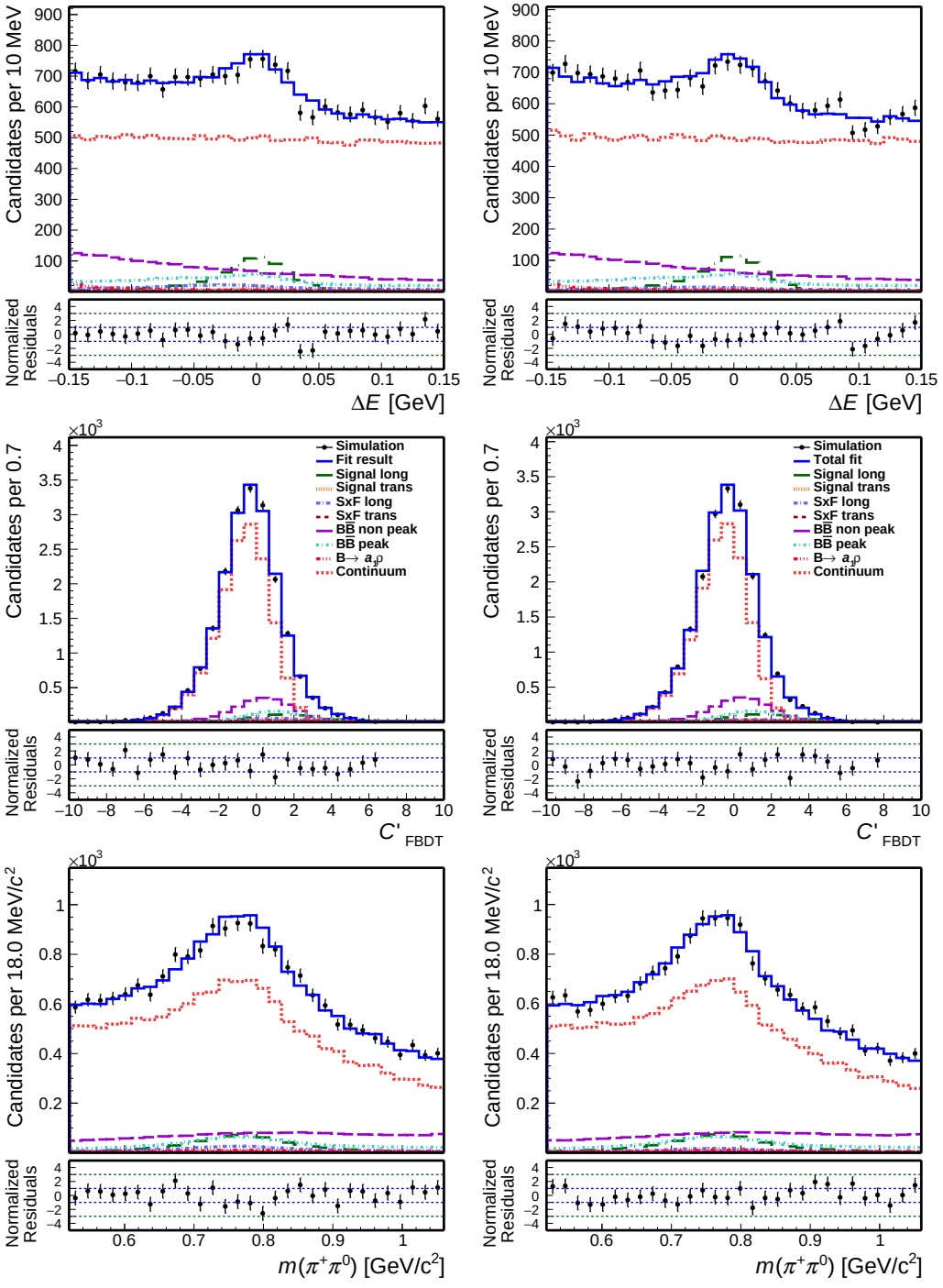


Figure 5.14: Example of fit projections: distributions of (top) ΔE, (middle) continuum suppression output, and (bottom) $m(\pi^+\pi^0)$ for a realistic simulated Belle sample of (left) negatively and (right) positively charged $B^+ \to \rho^+\rho^0$ candidates with fit projections overlaid.

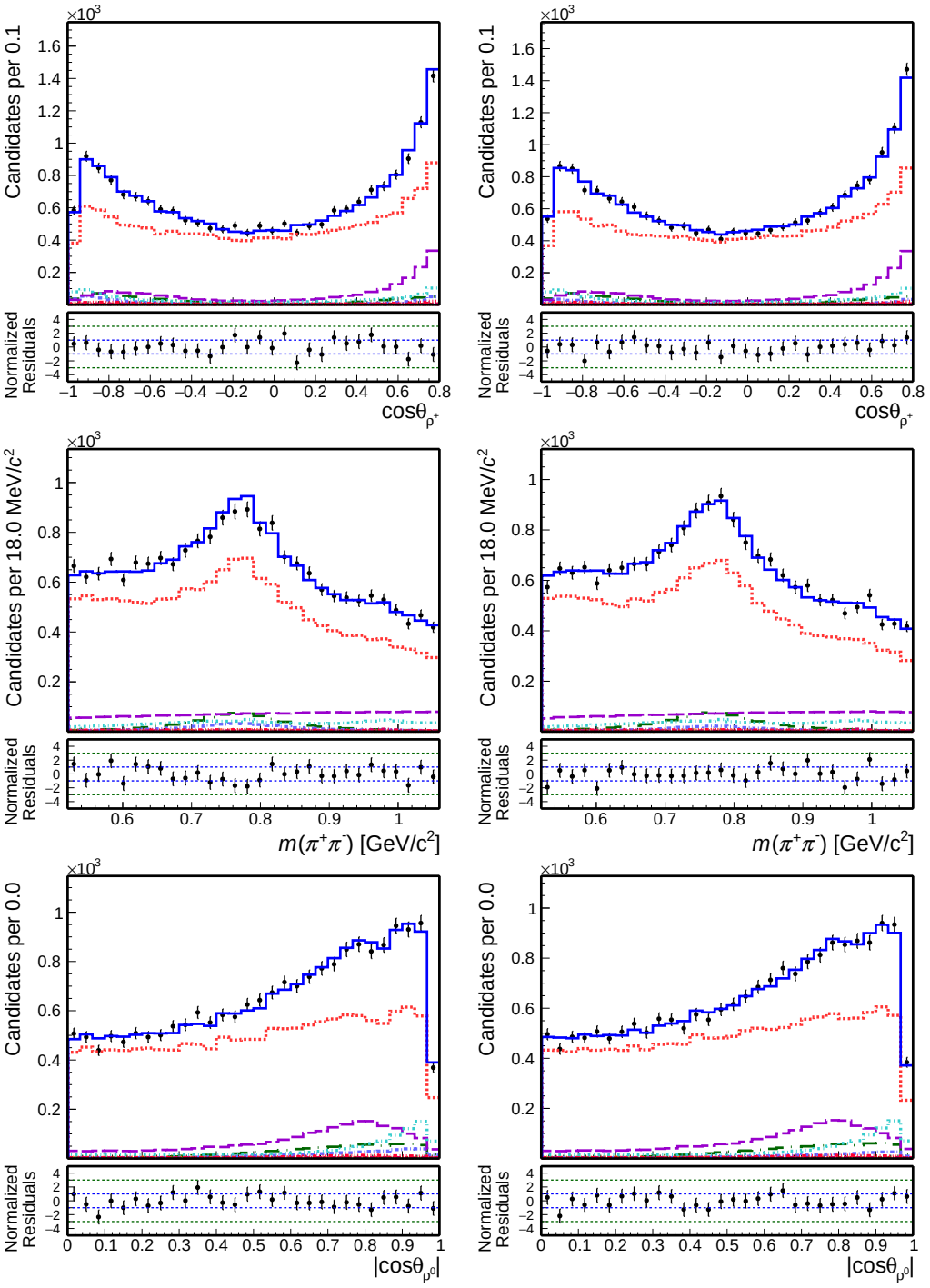


Figure 5.15: Example of fit projections: distributions of (top) $\cos\theta_{\rho^+}$, (middle) $m(\pi^+\pi^-)$, and (bottom)$|\cos\theta_{\rho^0}|$ for a realistic simulated Belle sample of (left) negatively and (right) positively charged $B^+ \to \rho^+\rho^0$ candidates with fit projections are overlaid.

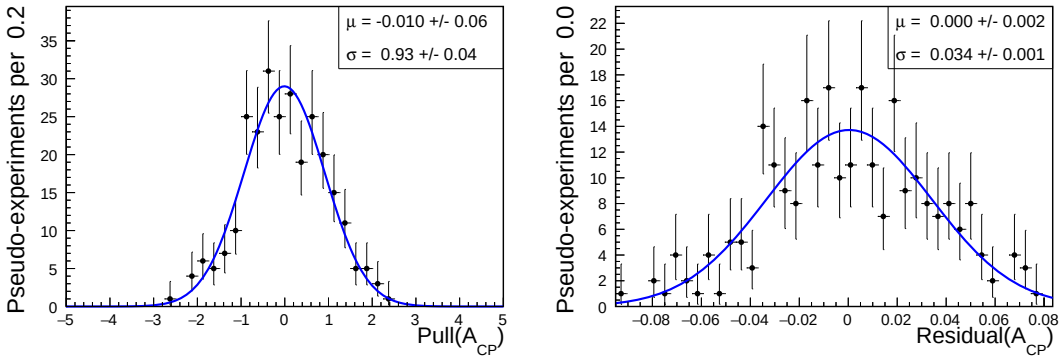


Figure 5.16: Distributions of (left) pull and (right) residual for the charge-parity asymmetry of $B^+ \to \rho^+\rho^0$ decays. Results of fit to Gaussian functions are overlaid (solid).

Chapter 6

Analysis validation on control data

This chapter discusses the validation of the analysis on data.

6.1 Introduction

The sample composition fit is developed based on realistically simulated samples of signal candidates. However, small but potentially harmful discrepancies are known to exist between Belle 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. I compare with known references results of fits similar to the $B^+ \to \rho^+\rho^0$ fit in well-known decays that share relevant features with the signal decay, in search for possible undetected issues or unmodeled effects. In addition, I identify appropriate control regions ("sidebands") in the signal-candidate sample to further compare simulated and observed data.

6.2 Analysis of control samples

I reconstruct in the full Belle data set, and in simulation, abundant control samples of the following decays that feature topologies and final states similar to those of signal.

$B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$. This is a relatively abundant (5×10^4 events expected) known process whose final state matches our final state. This channel allows validating the ΔE and C'_{FBDT} model. The additional requirement imposed on kaon identification is not expected to spoil the validation information.

$B^+ \to \bar{D}^0(\to K^0_S\pi^0)\pi^+$. A total of about 500 events are expected in the full Belle data set of this known process. The final state is the same as for signal. This validates the ΔE and C'_{FBDT} descriptions of the fit with less statistical power than in the previous channel, but freely from the additional requirement on PID.

$B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+(\to \pi^+\pi^0)$. A total of 3×10^4 events is expected. The presence of a ρ meson allows for validating the $m(\pi^+\pi^0)$ and $\cos\theta_{\rho^+}$ treatment, in addition to ΔE and C_{FBDT}. The limitation here is that the reference branching fraction is likely to be affected by unknown systematic uncertainties [10], spoiling the constraining value.

I measure the branching fractions of these channels using the same analysis as for the signal sample. I then compare results with known values to detect possible inconsistencies and data-driven corrections.

6.2.1 Selection

I apply the same selection criteria as for signal except for charm vetoes. In addition, I impose a requirement on the PID likelihood to suppress combinatorial background from pions in $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ and $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+(\to \pi^+\pi^0)$ candidates, which have a final state kaon. I then enrich the sample in D-meson decays by requiring the $K^+\pi^-$ and $K^+\pi^-\pi^0$ masses to meet $1.82 < m(K^+\pi^-, K^+\pi^-\pi^0) < 1.91$ GeV/c^2.

For the $B^+ \to \bar{D}^0(\to K^0_S\pi^0)\pi^+$ channel, I reconstruct K^0_S candidates by combining two opposite-charge particles with $|dr| < 2$ cm to exploit the K^0_S flight distance, and require their dipion mass to be consistent with a K^0_S decay, $0.471 < m(\pi^+\pi^-) < 0.524$ GeV/c^2. I loosen the D mass requirement to $1.79 < m(K^0_S\pi^0) < 1.94$ GeV/c^2 because energetic π^0 smear the D mass resolution.

6.2.2 Likelihood modeling

The composition of each control sample 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 has the same expression for all three control channels,

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

where $p_{\rm sig}$ describes the probability density function of the observables for the signal events, $p_{\rm sxf}$ for the self cross-feed events, $p_{\rm cont}$ for the continuum events, and $p_{B\bar{B}}$ for the non-peaking $B\bar{B}$ background events; N_i corresponds to the yield of generic component i, 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.

6.2.2.1 $B^+ \to \bar{D}^0\pi^+$ model

The likelihood that describes each component i of both $B^+ \to \bar{D}^0\pi^+$ decays is factorized as

$$\mathcal{L}_i = p_i(\Delta E)p_i(C'), \tag{6.2}$$

where index i runs over the signal, self cross-feed, $B\bar{B}$ background, or continuum background component; $p_i(\Delta E)$ is a p.d.f. that describes the ΔE distribution; and $p_i(C')$ is a p.d.f. describing the C' distribution.

The p.d.f.'s of the ΔE term for the signal and self cross-feed components is an empirical combination of three Gaussian functions,

$$p(\Delta E) \propto f_1\frac{e^{-\frac{1}{2}(\frac{\Delta E-\mu_1}{\sigma_1})}}{\sigma_1\sqrt{2\pi}} + f_2\frac{e^{-\frac{1}{2}(\frac{\Delta E-\mu_2}{\sigma_2})}}{\sigma_2\sqrt{2\pi}} + (1-f_1-f_2)\frac{e^{-\frac{1}{2}(\frac{\Delta E-\mu_3}{\sigma_3})}}{\sigma_3\sqrt{2\pi}}, \tag{6.3}$$

where the weight f_i, mean μ_i, and width σ_i of each Gaussian are determined by fits in simulation. The shapes are derived separately for each of the components. All fit projections related to the modeling of control channel distributions are collected in Appendix B. Figures B.1–B.2 show that the model is adequate.

The ΔE shape of the $B\bar{B}$ background is described by a sum of an exponential and a third-order polynomial function,

$$p(\Delta E) \propto f e^{-c\Delta E} + (1-f)(a\Delta E^3 + b\Delta E^2 + d\Delta E + 1), \tag{6.4}$$

where the weight of the exponential function f, and coefficients a, b, c, and d are determined by the fit in simulation. Figure B.3 illustrates that this function is adequate.

The continuum ΔE distribution is described by a straight line where the slope is a parameter determined in simulation. Figure B.4 shows that the model is adequate.

The p.d.f.'s of the continuum suppression term for all components are described by the sum of three Gaussian functions (Eq. 5.9). Figures B.5–B.6 show the results of the modeling fit of the C' distribution for all components of both $B^+ \to \bar{D}^0\pi^+$ control channels.

The functions used in the $B^+ \to \bar{D}^0\pi^+$ fit model are listed in Table 6.1. As in the $B^+ \to \rho^+\rho^0$ fit, I fix from simulation the ratio of the self cross-feed events to the signal events. Figures 6.1–6.2 show example projections obtained from the fit to realistic simulated $B^+ \to \bar{D}^0\pi^+$samples. The chosen model describes the simulated samples accurately.

Component	ΔE	C'
$B^+ \to \bar{D}^0\pi^+$	3G	3G
$B^+ \to \bar{D}^0\pi^+$ sxf	3G	3G
$B\bar{B}$	EXP+Pol3	3G
Continuum	Pol1	3G

Table 6.1: Model summary for $B^+ \to \bar{D}^0\pi^+$ control channels: "3G" indicates the sum of three Gaussian functions, "Poln" indicates a polynomial of nth order, "EXP" indicates an exponential function.

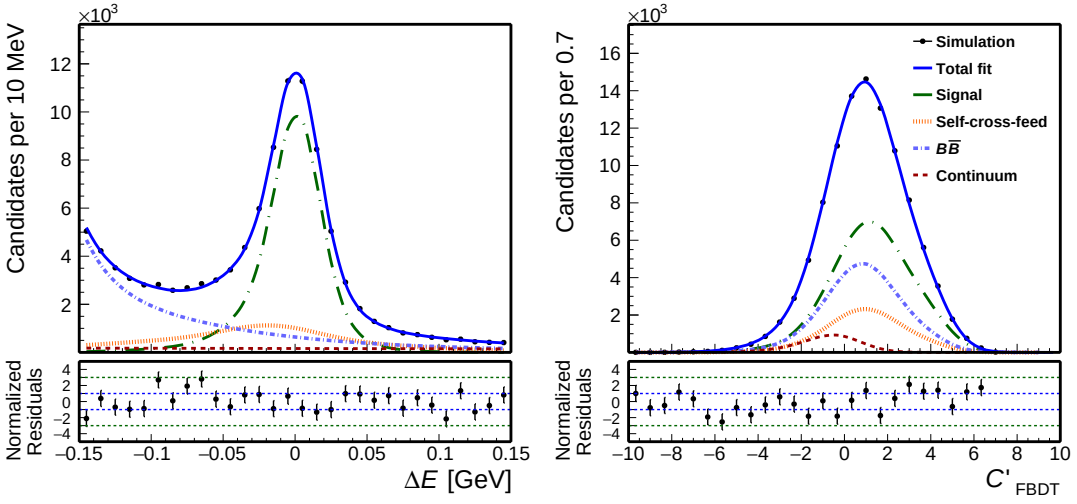


Figure 6.1: Distributions of the (left) ΔE and (right) C' distributions for a realistically simulated $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ sample. Fit projections are overlaid (blue solid line).

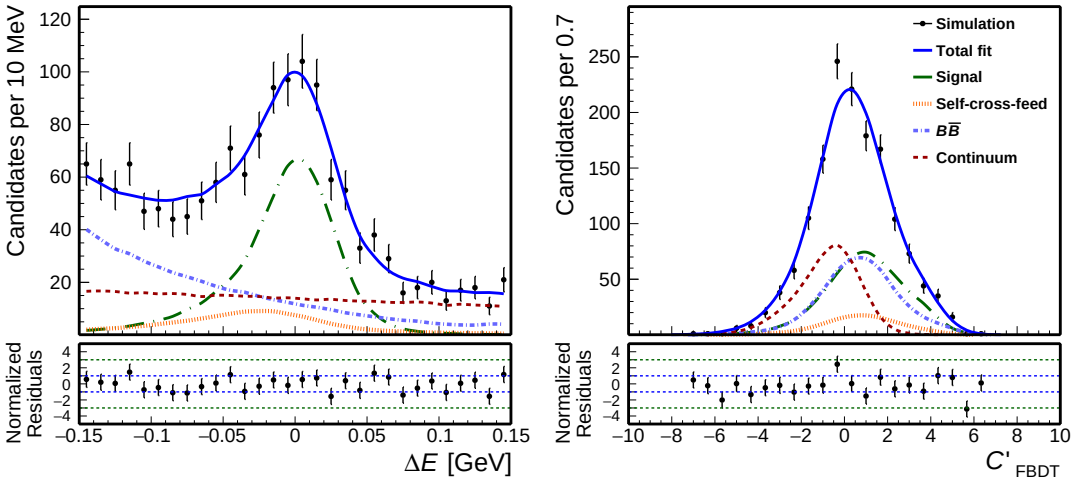


Figure 6.2: Distributions of the (left) ΔE and (right) C' for a realistically simulated $B^+ \to \bar{D}^0(\to K^0_S\pi^0)\pi^+$ sample. Fit projections are overlaid (blue solid line).

6.2.2.2 $B^+ \to \bar{D}^0\rho^+$ model

For $B^+ \to \bar{D}^0\rho^+$ events, the single event likelihood for signal and self cross-feed is factorized as

$$\mathcal{L}_i = p_i(\Delta E, m(\pi^+\pi^0), \cos\theta_{\rho^+}) p_i(C'), \tag{6.5}$$

where the joint $p_i(\Delta E, m(\pi^+\pi^0), \cos\theta_{\rho^+})$ p.d.f., which accounts for the three-dimensional joint distribution of ΔE, $m(\pi^+\pi^0)$, and $\cos\theta_{\rho^+}$ observables, is a histogram template. The p.d.f. $p_i(C')$, which describes the C' distribution, is a sum of three Gaussians (5.9). As shown in Figs. B.7–B.8, the model is adequate.

For the $B^+ \to \bar{D}^0\rho^+$ candidates reconstructed from $B\overline{B}$ background events, the single-event likelihood is factorized as

$$\mathcal{L}_i = p_i(\Delta E, m(\pi^+\pi^0), \cos\theta_{\rho^+}) p_i(C'|\Delta E). \tag{6.6}$$

The p.d.f. describing the C' distribution is a conditional p.d.f. that includes the dependence on ΔE,

$$p(C'|\Delta E) \propto \frac{f_1 + \alpha|\Delta E|}{\sigma_1\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{C'-\mu_1-\beta|\Delta E|}{\sigma_1}\right)} + \frac{f_2 + \alpha|\Delta E|}{\sigma_2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{C'-\mu_2-\beta|\Delta E|}{\sigma_2}\right)} + \frac{1 - (f_1 + f_2 + \alpha|\Delta E|)}{\sigma_3\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{C'-\mu_3-\beta|\Delta E|}{\sigma_3}\right)},$$

where the relative weight of the main Gaussian f, correlation parameters α and β, and Gaussian means μ_i and widths σ_i are determined by a fit in simulation. Figure B.9 shows that the $B\overline{B}$ model is adequate.

The single-event likelihood that describes $B^+ \to \bar{D}^0\rho^+$ candidates reconstructed from continuum events is factorized into three terms,

$$\mathcal{L}_{cont} = p_i(\Delta E) p_i(C') p_i(m(\pi^+\pi^0), \cos\theta_{\rho^+}), \tag{6.7}$$

where $p_i(\Delta E)$ is a p.d.f. that describes the ΔE distribution; $p_i(C')$ describes the C' distribution; and $p_i(m(\pi^+\pi^0), \cos\theta_{\rho^+})$ is a joint p.d.f. that describes the two-dimensional $m(\pi^+\pi^0)$ and $\cos\theta_{\rho^+}$ distribution.

The ΔE term of the continuum single-event likelihood is modeled with a straight line where the slope is a free parameter. Figure B.10 shows that the model is adequate.

The C' p.d.f. is a sum of three Gaussian functions, as represented by Eq. (5.9). The function describes adequately the C' distribution drawn from the simulated continuum sample, as shown in Fig. B.10.

Analytical functions and binned templates used to describe the single-event $B^+ \to \bar{D}^0\rho^+$ likelihood are summarized in Table 6.2. The projections obtained from a fit of the realistically simulated $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ sample are shown in Figure 6.3.

Component	ΔE	C'	$m(\pi^+\pi^0)$	$\cos\theta_{\rho^+}$
$B^+ \to \bar{D}^0\rho^+$	3D (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	3G	3D (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	3D (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)
$B^+ \to \bar{D}^0\rho^+$ sxf	3D (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	3G	3D (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	3D (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)
$B\bar{B}$	3D (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	3G	3D (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	3D (ΔE, $m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)
Continuum	Pol1	3G	2D ($m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)	2D ($m(\pi^+\pi^0)$, $\cos\theta_{\rho^+}$)

Table 6.2: $B^+ \to \bar{D}^0\rho^+$ model summary. "nD" indicates three-, two-dimensional histograms. "3G" indicates the sum of three Gaussian functions, "Poln" indicates a polynomial function of nth order.

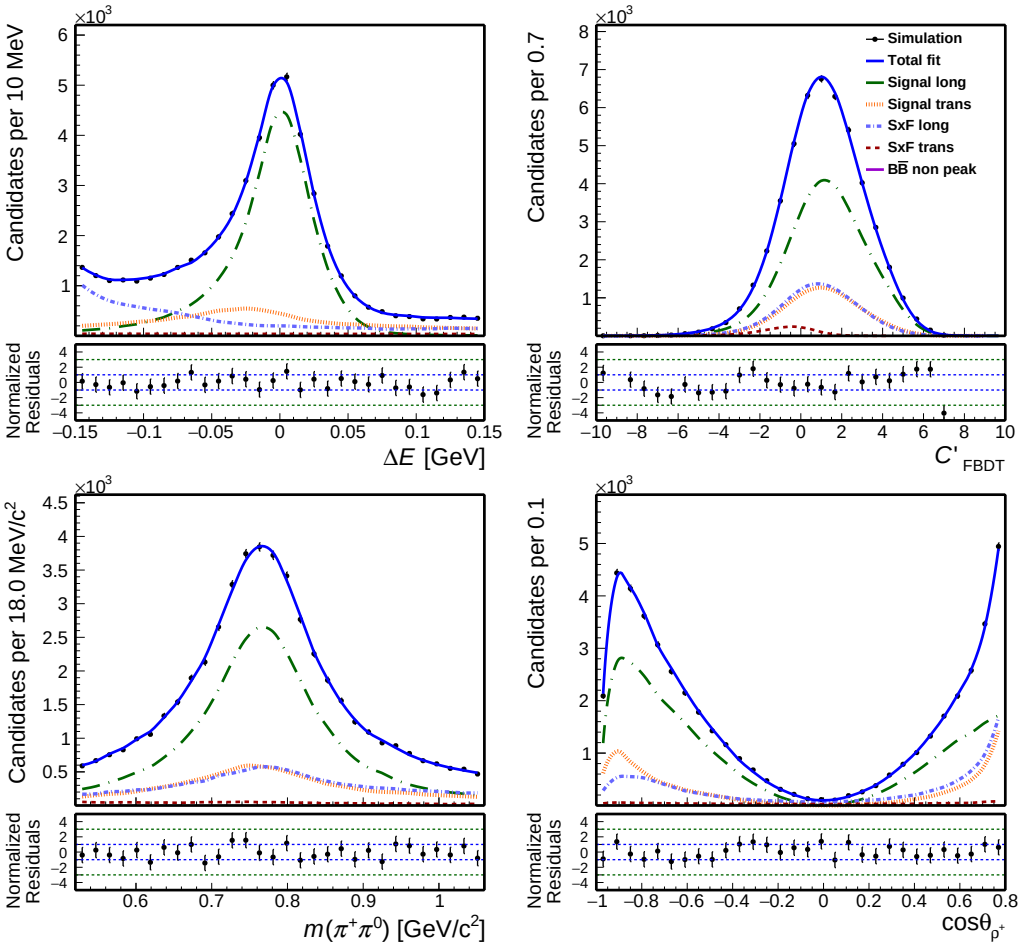


Figure 6.3: Distributions of (top left) ΔE, (top right) C', (bottom left) $m(\pi^+\pi^0)$, and (bottom right) $\cos\theta_{\rho^+}$) for realistically simulated $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ sample. Fit projections are overlaid (blue solid line).

6.2.3 Control channel fits in Belle data

6.2.3.1 $B^+ \to \bar{D}^0\pi^+$

The total likelihood of the $B^+ \to \bar{D}^0\pi^+$ control channel takes the form

$$\mathcal{L} = \prod_{i=1}^{n} \sum_{j} N_j \mathcal{L}_i(\Delta E, C'), \tag{6.8}$$

where $\mathcal{L}_i$ is the ith event likelihood, and N_j, the yield of each jth component, is a free parameter. Other free parameters are the calibration factors μ_C and σ_C. These allows for correcting for the differences known to be possible between data and simulation in the shapes of the core ΔE distributions. I incorporate them in the signal and self cross-feed ΔE p.d.f.'s by transforming each mean and width as $\mu \longrightarrow \mu + \mu_C$ and $\sigma \longrightarrow \sigma \cdot \sigma_C$.

I perform a two-dimensional fit to the same full Belle data set used for the signal-sample fit. A Gaussian constraint is set on the self-cross-feed-to-signal-yield-ratio extracted from simulation. Figure 6.4 shows $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ data with fit projections overlaid. A signal peak is observed in the $\Delta E \approx 0$ region, overlapping smooth continuum and non-peaking $B\overline{B}$ background distributions. In the continuum suppression output distribution, continuum tends to peak at $C' \approx 0$, while the peak corresponding to $B\overline{B}$ events (including signal and background) is shifted toward $C' \approx 1$. All distributions are adequately reproduced by the fit. In the ΔE distribution, the fit reproduces accurately the data both in the central part of the distribution, where the signal dominates, and in the tails where the background dominates. The fit projections for the continuum suppression output distribution show that the fit effectively distinguishes continuum events, which tend to peak at 0, and $B\overline{B}$ events.

I observe 49477±990 $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ decays. Dividing this yield by the known number of $B\overline{B}$ pairs (see Sec. 5.1.4), selection efficiency $\varepsilon = 0.1004 \pm 0.0001$ extracted from simulation, particle-identification correction $\varepsilon_{\rm PID} = 0.9711 \pm 0.0489$ (see Sec. 8.3), and neutral pion reconstruction correction $\varepsilon_{\pi^0} = 0.957 \pm 0.015$ (see Sec. 8.4) results in

$$\mathcal{B}(B^+ \to \bar{D}^0\pi^+) \times \mathcal{B}(\bar{D}^0 \to K^+\pi^-\pi^0) = (6.87 \pm 0.14(\text{stat})) \times 10^{-4}. \tag{6.9}$$

The result agrees with the known value $\mathcal{B}(B^+ \to \bar{D}^0\pi^+) \times \mathcal{B}(\bar{D}^0 \to K^+\pi^-\pi^0)_{\rm WA} = (6.65 \pm 0.3) \times 10^{-4}$, assuming $\mathcal{B}(\bar{D}^0 \to K^+\pi^-\pi^0) = (14.4 \pm 0.5)\%$ [10]. The ΔE calibration factors resulting from the fit are $\mu_C = -1.3 \pm 0.2$ MeV and $\sigma_C = 1.153 \pm 0.017$.

Figure 6.5 shows the $B^+ \to \bar{D}^0(\to K^0_S\pi^0)\pi^+$ data with fit projections overlaid. All distributions are adequately reproduced by the fit. In the ΔE distribution, a signal peak centered at $\Delta E \approx 0$ overlaps a uniform background. The observed yield of 495±56 $B^+ \to \bar{D}^0(\to K^0_S\pi^0)\pi^+$ decays corrected by the number of $B\overline{B}$ pairs, selection efficiency $\varepsilon = 0.01586 \pm 0.00071$ extracted from simulation, particle-identification correction $\varepsilon_{\rm PID} = 0.9711 \pm 0.0489$, and neutral pion reconstruction correction $\varepsilon_{\pi^0} = 0.957 \pm 0.015$ results in

$$\mathcal{B}(B^+ \to \bar{D}^0\pi^+) \times \mathcal{B}(\bar{D}^0 \to K^0_S\pi^0) = (43.5 \pm 4.9(\text{stat})) \times 10^{-6}, \tag{6.10}$$

which agrees with the known value, $\mathcal{B}(B^+ \to \bar{D}^0\pi^+) \times \mathcal{B}(\bar{D}^0 \to K^0_S\pi^0)_{\rm WA} = (38.50 \pm 1.68) \times 10^{-6}$ [10].

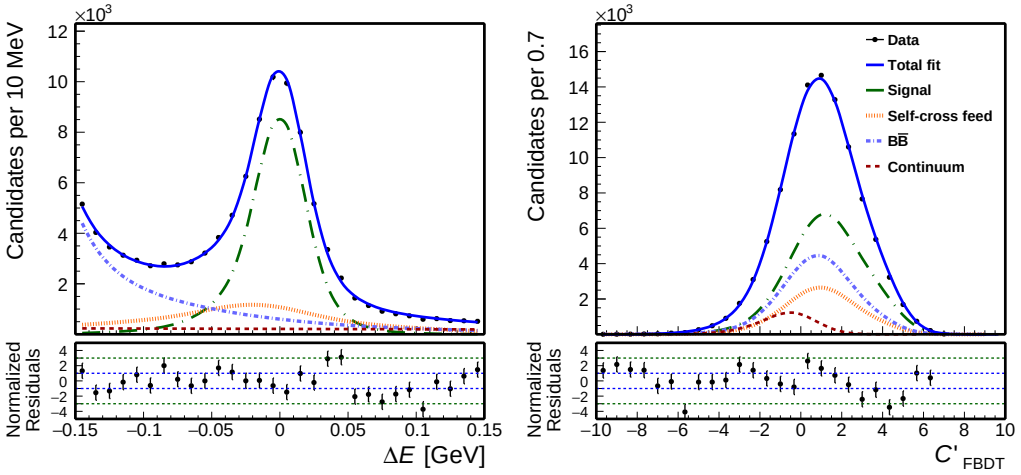


Figure 6.4: Distributions of (left) ΔE and (right) C' for $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ reconstructed in the full Belle data set. Fit projections are overlaid (blue solid line).

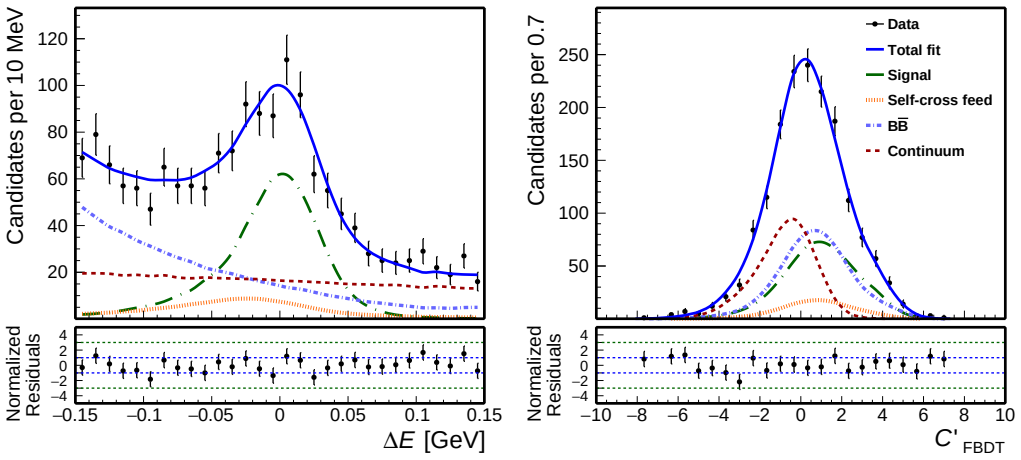


Figure 6.5: Distributions of (left) ΔE and (right) C' for $B^+ \to \bar{D}^0(\to K^0_S\pi^0)\pi^+$ reconstructed in the full Belle data set. Fit projections are overlaid (blue solid line).

6.2.3.2 $B^+ \to \bar{D}^0\rho^+$

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

$$\mathcal{L} = \prod_{i=1}^{n} \sum_{j} N_j \mathcal{L}_i(\Delta E, C', m(\pi^+\pi^0), \cos\theta_{\rho^+}). \tag{6.11}$$

I use histogram templates to model the ΔE signal and self cross-feed shapes. In order to allow for the fit to capture differences between data and simulation in ΔE shapes, I add flexibility in the signal and self cross-feed model, similarly to what done in the $B^+ \to \bar{D}^0\pi^+$ data fit. We introduce flexibility in the histogram templates by generating alternative templates in which each entry is summed to a random number drawn from a Gaussian distribution whose mean and width are correction terms determined on data. The set of corrections that describe more accurately the data as indicated by likelihood ratio tests are chosen as our default corrections.

I perform a four-dimensional fit to the same full Belle data set used for the signal-sample fit. Figure 6.6 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 $C' \approx 1$, 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. All distributions are adequately reproduced by the fit. I observe 19610±230 $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ decays reconstructed in the full Belle data set. Dividing this yield by the number of $B\overline{B}$ pairs, selection efficiency $\varepsilon = 0.0732 \pm 0.0001$ extracted from simulation, particle-identification correction $\varepsilon_{\rm PID} = 0.9711 \pm 0.0489$, and neutral pion reconstruction correction $\varepsilon_{\pi^0} = 0.957 \pm 0.015$ results in

$$\mathcal{B}(B^+ \to \bar{D}^0\rho^+) \times \mathcal{B}(\bar{D}^0 \to K^+\pi^-) \times \mathcal{B}(\rho^+ \to \pi^+\pi^0) = (3.73 \pm 0.04) \times 10^{-4}. \quad (6.12)$$

This disagrees with the known value of $(5.12 \pm 0.71) \times 10^{-4}$ [10]. After various consistency checks aimed at searching for peaking background or wrongly estimated efficiencies, I accumulate growing evidence that the reason for the discrepancy might be attributed to the reference. The fact that the reference is based on CLEO results from 1994 [86] based on a small sample and no results were reported by BaBar or Belle, which had the data to achieve much better precision, suggests caution on the reliability of the reference. The fact that Belle reported a preliminary value [87] compatible with mine builds further confidence on my result. Calibration factors $\mu_C = 0.0003$ GeV and $\delta_C = 0.0135$ GeV (where δ_C is an additive term to the width $\sigma \to \sigma + \delta_C$) are those most favored by the data.

6.2.4 Summary

The results of studies of control samples show no large discrepancy between simulation and data for observables of interest. Our understanding of the relevant distributions in simulation, appropriately modified by the μ_C and σ_C calibration factors, is adequate to provide a robust fit in $B^+ \to \rho^+\rho^0$ data.

6.3 Sideband and off-resonance studies

I study the distributions of $B^+ \to \rho^+\rho^0$ candidates reconstructed in control regions of data devoid of signal and enriched in background and compare them with simulation to ensure that no residual mismodelings are present. Such a sideband study allows further validating the analysis further before unblinding.

I apply the final selection to the $B^+ \to \rho^+\rho^0$ candidates reconstructed in the full Belle data set and in realistic simulated data normalized to the same luminosity. Then, I restrict the samples to adequately chosen sideband regions. Figure 6.7 (left) shows that 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 $M'_{\rm bc} < 5.26\,$GeV/c^2, where continuum and $B\overline{B}$ background dominate and no signal is expected to contribute, for this study.

I further divide the sideband region into five independent regions (Fig. 6.7, left panel) which have different proportions of continuum and $B\overline{B}$ background candidates in order 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 3 and region 4, respectively. Figure 6.7 (right) shows the ratio between the two-dimensional ($M'_{\rm bc}$, ΔE) distributions in data and simulation, normalized to the data luminosity. Good agreement is observed in all regions in terms of total yields.

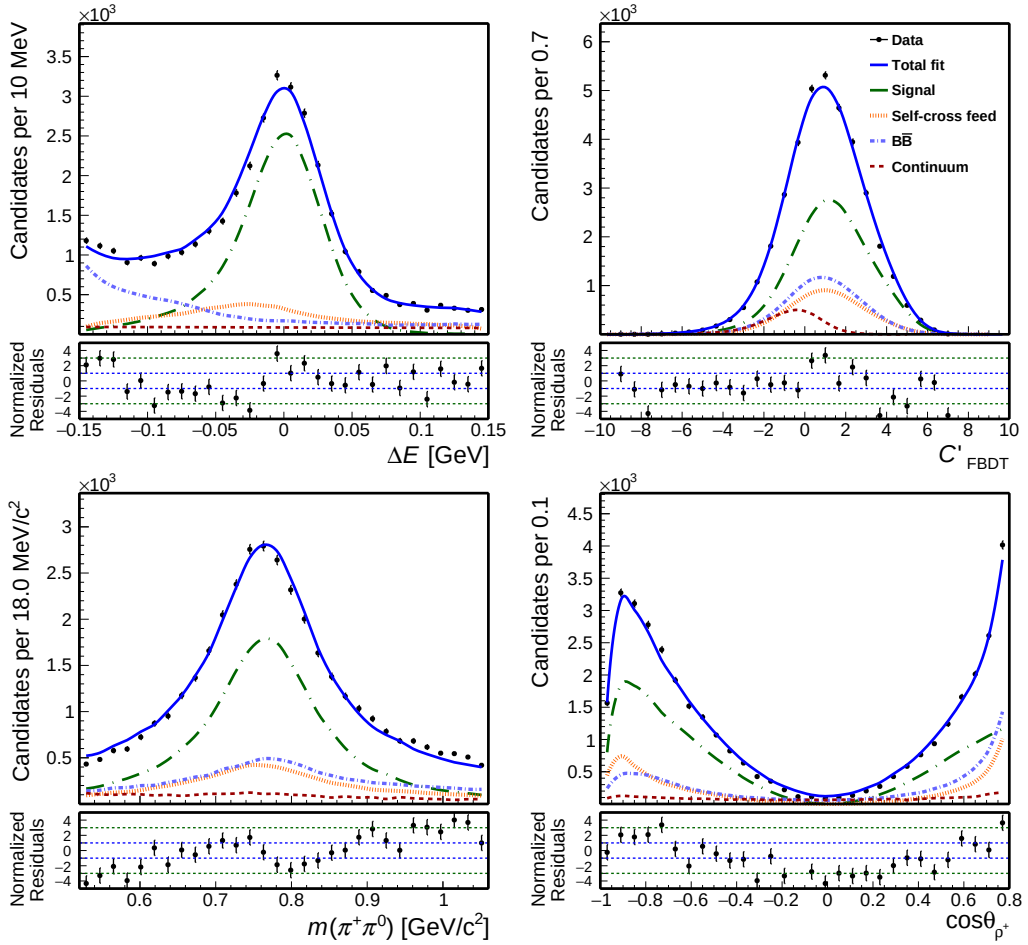


Figure 6.6: Distributions of (top left) ΔE, (top right) C', (bottom left) $m(\pi^+\pi^0)$, and (bottom right) $\cos\theta_{\rho^+}$ for $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ decays reconstructed in the full Belle data set. Fit projections are overlaid (blue solid line).

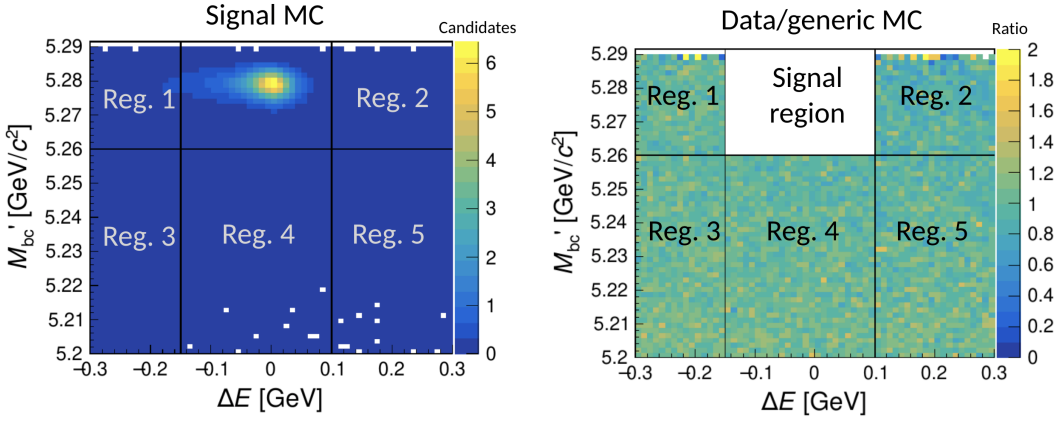


Figure 6.7: Distribution of (left) $M'_{\rm bc}$ as a function of ΔE for events reconstructed in a simulated signal sample. Distribution of (right) ratio between the $M'_{\rm bc}$ vs. ΔE distributions for events reconstructed in collision and simulated data.

I then compare the experimental data and simulated distributions of relevant basic observables associated with the reconstruction of neutral pions, momentum, diphoton mass,

cosine of the polar angle, cosine of the helicity angle, and χ^2 of the kinematic fit (Fig. 6.8). Data and simulation distributions are in good agreement for all the observables but π^0 momentum and diphoton mass.

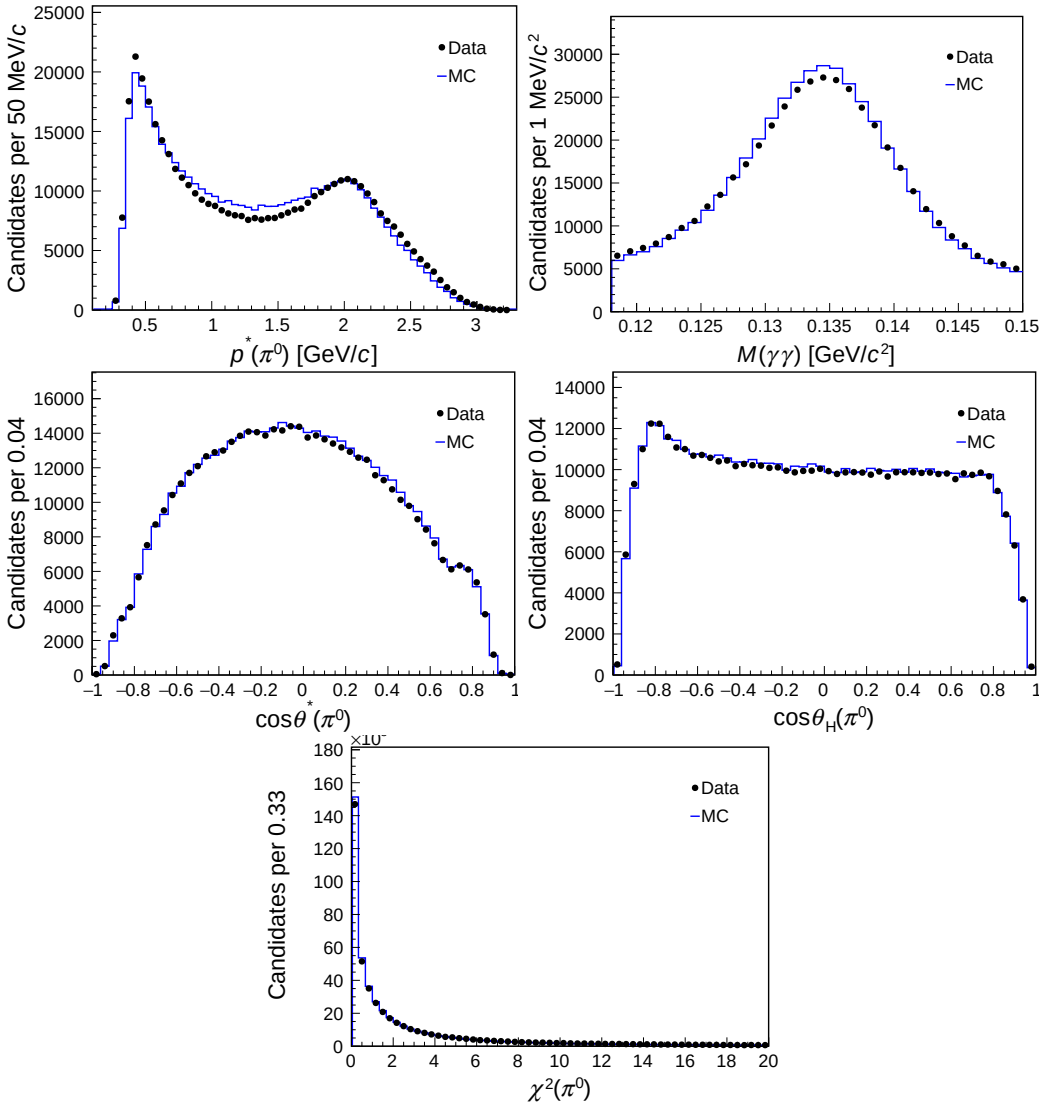


Figure 6.8: Distributions of (top left) momentum, (top right) diphoton mass, (middle left) cosine of the polar angle, (middle right) cosine of the helicity angle, and (bottom) χ^2 of the kinematic fit of neutral pion candidates reconstructed in (black dots) collision and (blue line) simulated data outside the signal box ($M_{bc} > 5.26$ GeV/c^2 and $-0.15 < \Delta E < 0.10$ GeV).

Figure 6.9 shows π^0 momentum distributions obtained in data and simulation in different sideband regions. Since the discrepancies are larger in the regions dominated by continuum background (regions 2 and 5), I inspect the distributions of $B^+ \to \rho^+\rho^0$ candidates reconstructed in off-resonance data (Fig. 6.10) which are only populated by continuum. The distribution shows a data-simulation inconsistency similar to the inconsistency observed in the sideband data. We deduce that the discrepancy is likely to be originated from the continuum simulation. Momentum discrepancies are also observed for final-state charged pions (Fig. 6.11).

To correct for those discrepancies, I weight the momentum distributions of $B^+ \to \rho^+\rho^0$

candidates reconstructed in the simulated off-resonance sample normalized to the Belle luminosity in such a way that simulation mirrors the data.

6.3.1 Continuum background weighting

Figure 6.12 shows that the momentum distributions in off- and on-resonance continuum simulation are compatible. Hence the corrections determined using off-resonance data and off-resonance simulation can be applied to on-resonance simulation. Care is needed when weighting distributions, since modifications in one quantity will propagate to all other quantities that correlate with the probe quantity. If not identified and controlled, such modifications may introduce larger discrepancies than those one tries to suppress.

Hence I study the dependences between the momentum magnitudes of the different final-state particles. Table 6.3 shows the correlation coefficients between momenta; Figure 6.13 shows the corresponding two-dimensional distributions. Large dependences between particle momenta are observed only for the pions that are paired to form ρ meson candidates. I consider those dependences in the extraction of the momentum weights by using two-dimensional histograms.

Figure 6.14 shows the momentum weights determined by dividing the data distributions by the simulation distributions. I then apply the momentum weights to the $B^+ \to \rho^+\rho^0$ candidates reconstructed in simulated continuum. The total weight for each B candidate is

$$w = w_{\rho^+}\left(p\left(\pi_1^+\right), p\left(\pi^0\right)\right) \cdot w_{\rho^0}\left(p\left(\pi_2^+\right), p\left(\pi_3^-\right)\right),$$

where p is the momentum magnitude of the final-state particles, and the labeling convention $B^+ \to \rho^+(\to \pi_1\pi^0)\rho^0(\to \pi_2\pi_3)$ is used.

To test the procedure, I apply the weights to $B^+ \to \rho^+\rho^0$ candidates reconstructed in the sideband of on-resonance simulation and compare with the distributions in sideband on-resonance data. Figures 6.15–6.16 show the $\cos\theta_{\rho^+}$ and $\cos\theta_{\rho^0}$ distributions before and after applying the weights. Reweighted distributions show significantly improved data-simulation agreement in the $\cos\theta_{\rho^+}$ and $\cos\theta_{\rho^0}$ distributions.

	$p\left(\pi_1^+\right)$	$p\left(\pi^0\right)$	$p\left(\pi_2^+\right)$	$p\left(\pi_3^-\right)$
$p\left(\pi_1^+\right)$	1.00	−0.69	−0.02	−0.01
$p\left(\pi^0\right)$		1.00	0.01	−0.01
$p\left(\pi_2^+\right)$			1.00	−0.70
$p\left(\pi_3^-\right)$				1.00

Table 6.3: Correlation coefficients between final-state particle momenta for $B^+ \to \rho^+\rho^0$ candidates reconstructed in continuum simulation. The coefficients are calculated in the range $0 < p < 2.5$ GeV/c.

6.3.2 Continuum suppression output in off-resonance data

Figure 6.17 compares the C'_{FBDT} distribution of candidates reconstructed in off-resonance data and simulation. The small shift observed between the two distributions is accounted for in the systematic uncertainties (see Sec. 8.10.4).

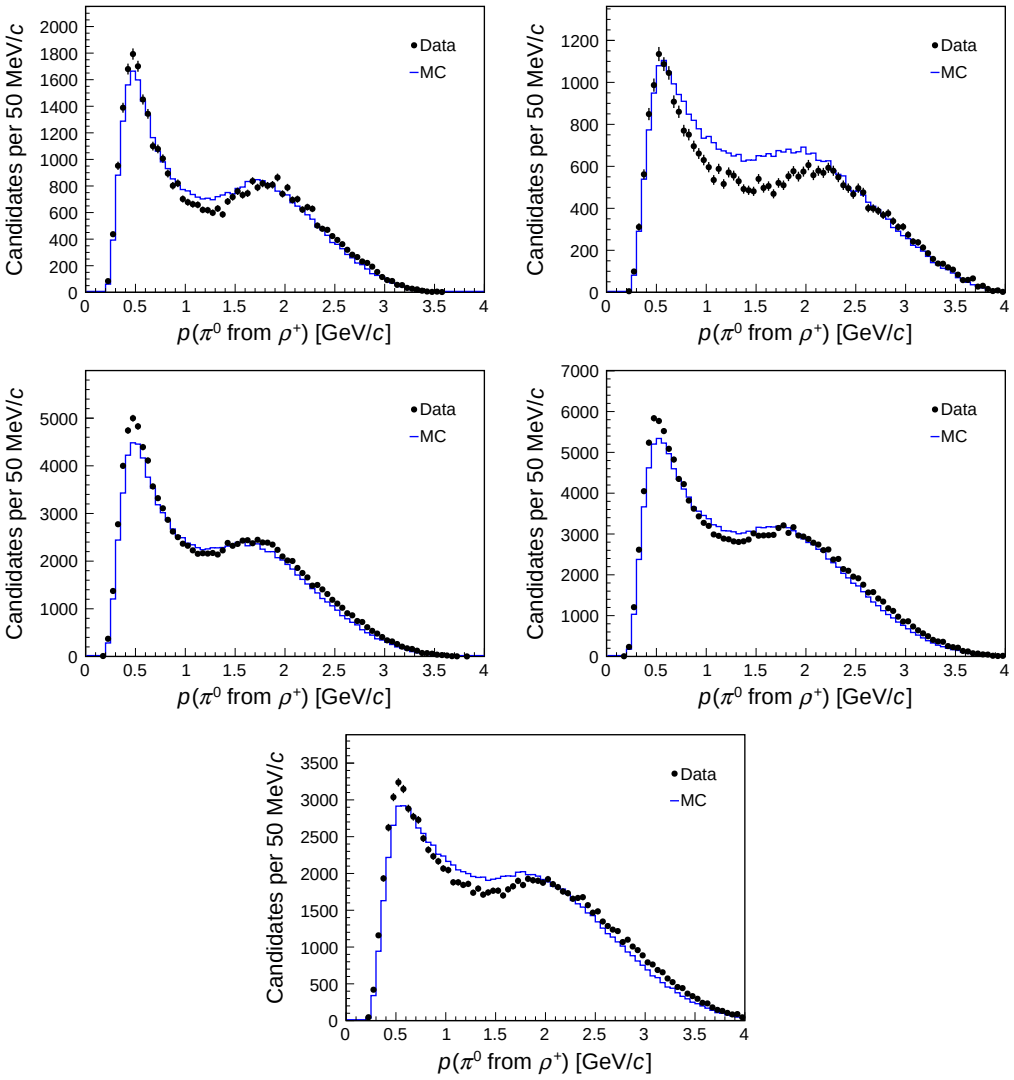


Figure 6.9: Distributions of neutral pion momentum from $B^+ \to \rho^+\rho^0$ candidates reconstructed in (solid) simulation and in (dots) data for the (top left) region 1, (top right) region 2, (middle left) region 3, (middle right) region 4, and (bottom) region 5 of sidebands defined in Figure 6.7.

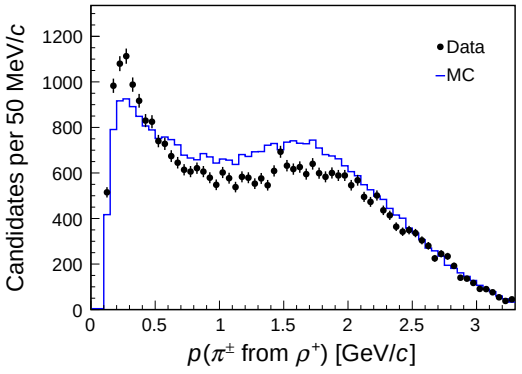


Figure 6.10: Distribution of charged-pion momentum for $B^+ \to \rho^+\rho^0$ candidates reconstructed in luminosity-normalized off-resonance (solid) simulation and (dots) data.

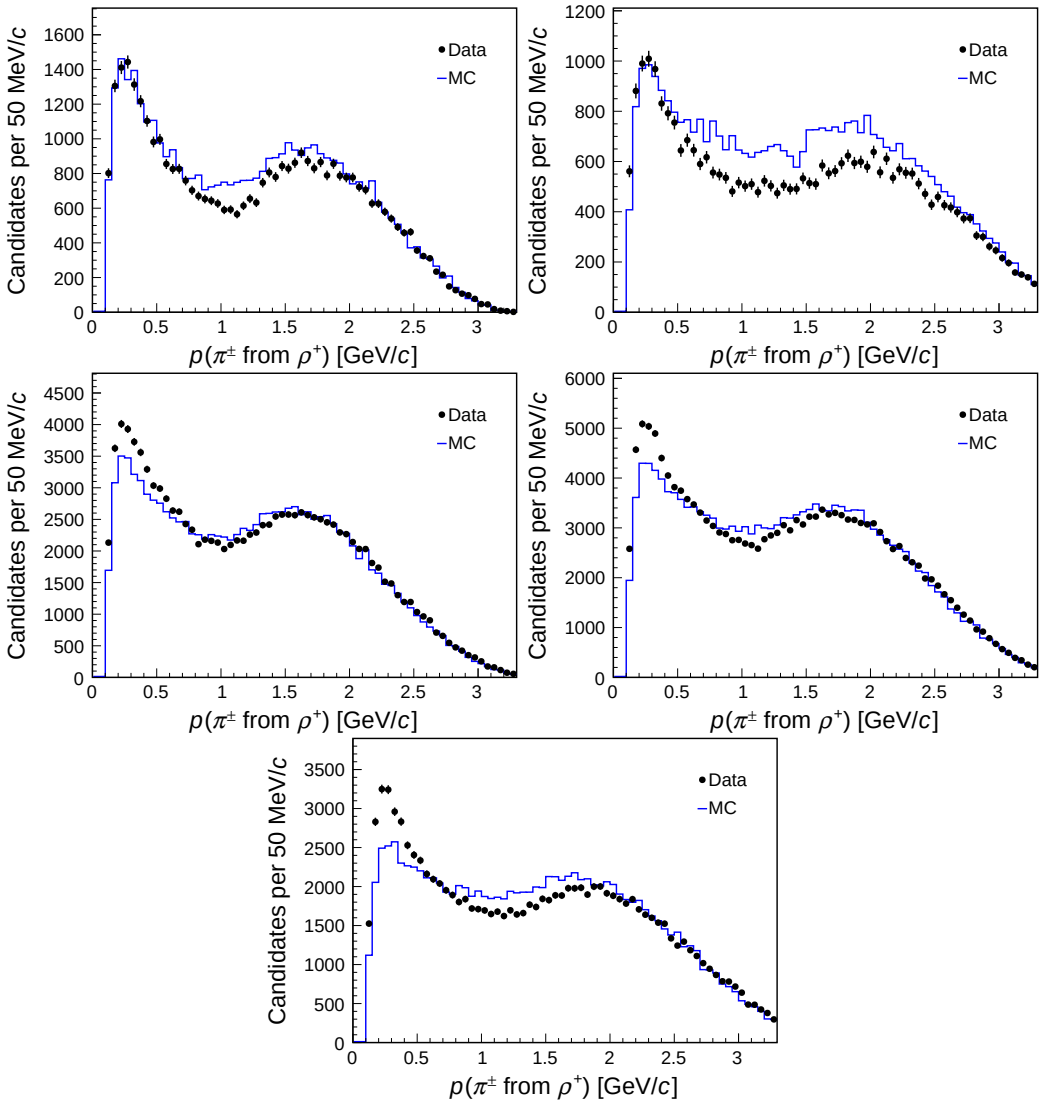


Figure 6.11: Distributions of charged-pion momentum from $B^+ \to \rho^+\rho^0$ candidates reconstructed in (solid) simulation and in (dots) data for the (top left) region 1, (top right) region 2, (middle left) region 3, (middle right) region 4, and (bottom) region 5 of sidebands defined in Figure 6.7.

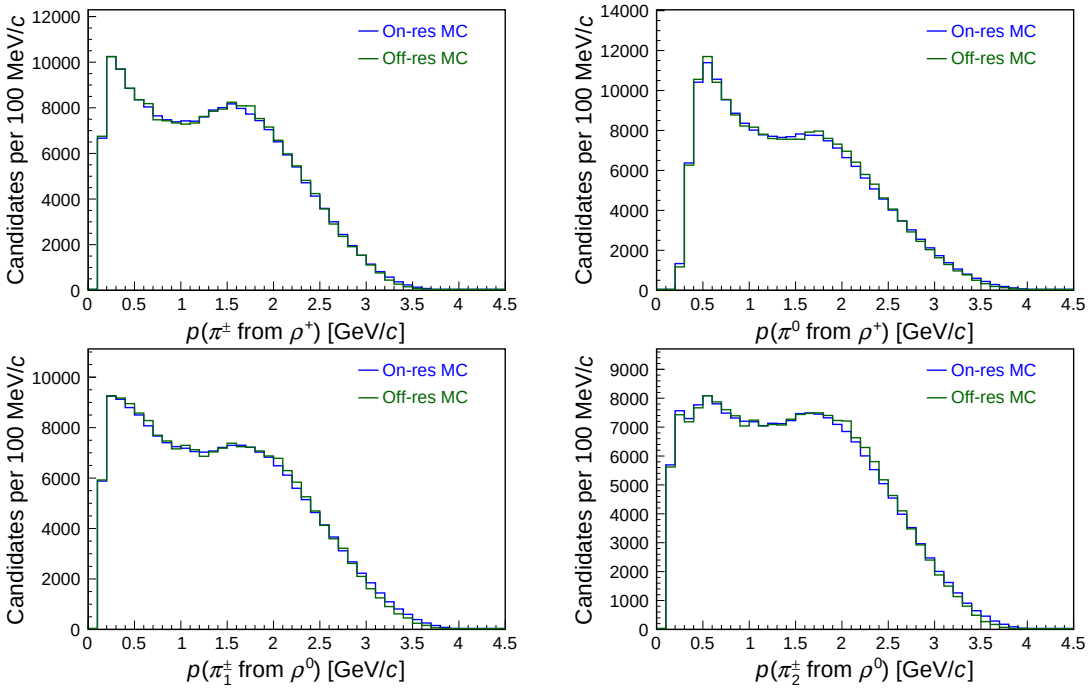


Figure 6.12: Momentum distributions of charged pions from $B^+ \to \rho^+\rho^0$ candidates reconstructed in on- and off-resonance continuum simulation.

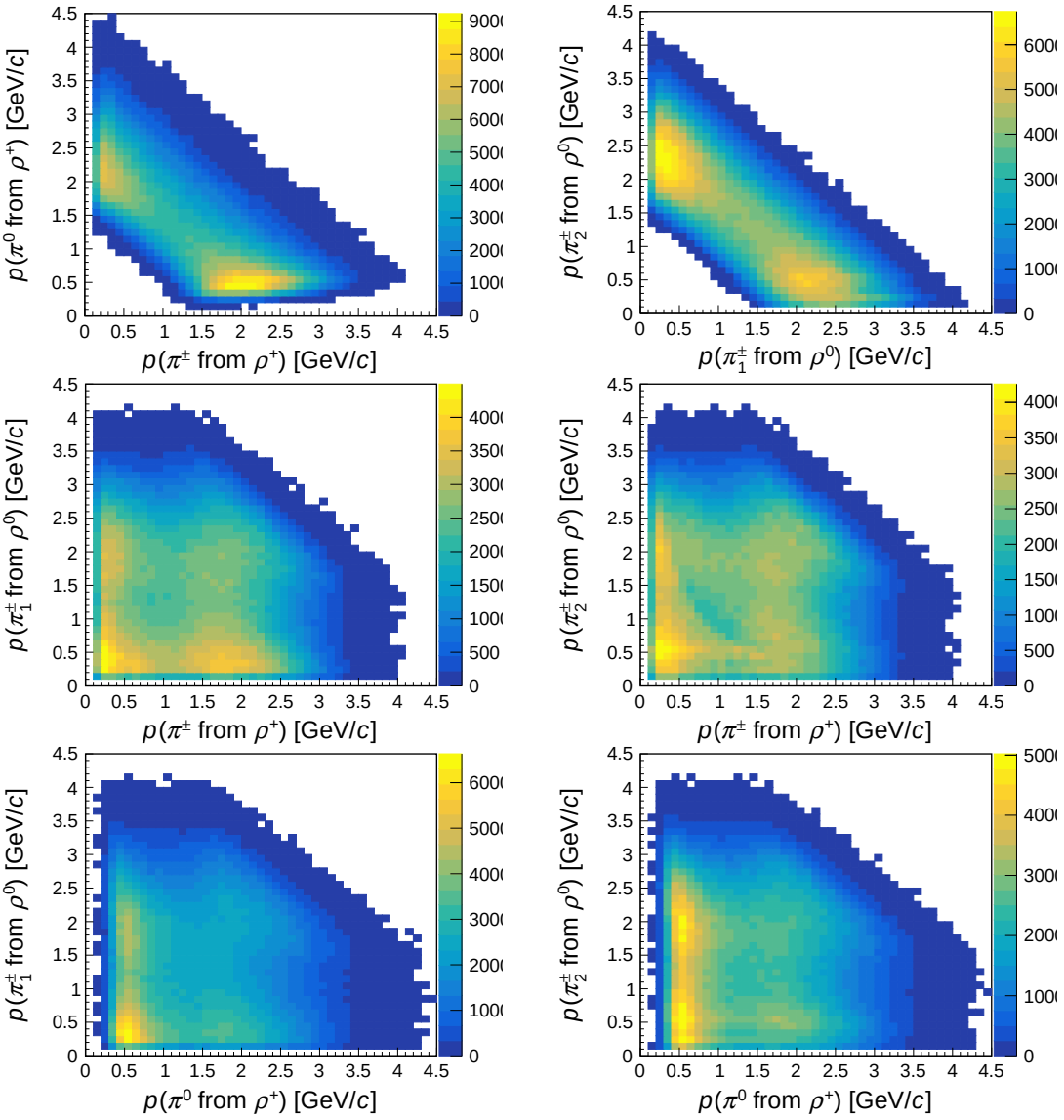


Figure 6.13: Two dimensional distributions of (top left) momentum of π^0 versus the momentum of π^+ from the ρ^+ decay, (top right) momentum of π^+ from ρ^0 versus momentum of π^- from ρ^0, (middle left) momentum of π^+ from ρ^0 versus momentum of π^+ from ρ^+, (middle right) momentum of π^- from ρ^0 versus momentum of π^+ from ρ^+, (bottom left) momentum of π^+ from ρ^0 versus momentum of π^0, and (bottom right) momentum of π^- from ρ^0 versus momentum of π^0 for $B^+ \to \rho^+\rho^0$ candidates reconstructed in continuum simulation.

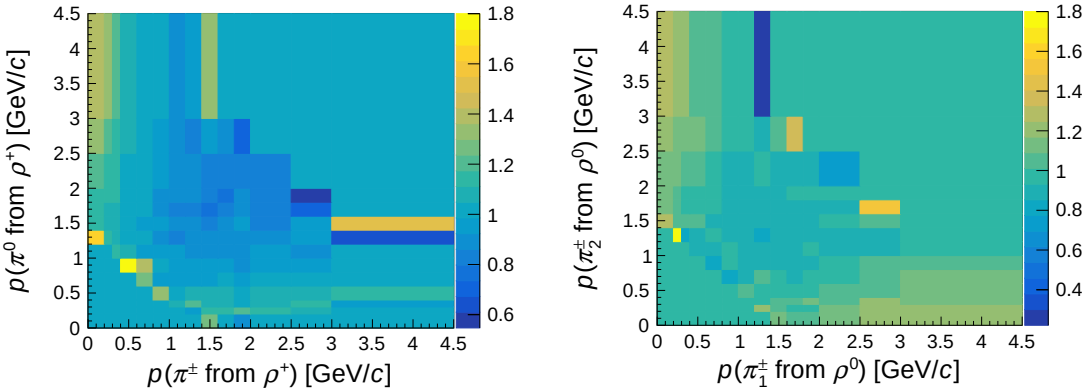


Figure 6.14: Two-dimensional momentum weights obtained as ratio of yields observed in data over those observed in simulation for $B^+ \to \rho^+\rho^0$ candidates reconstructed in off-resonance simulation and data.

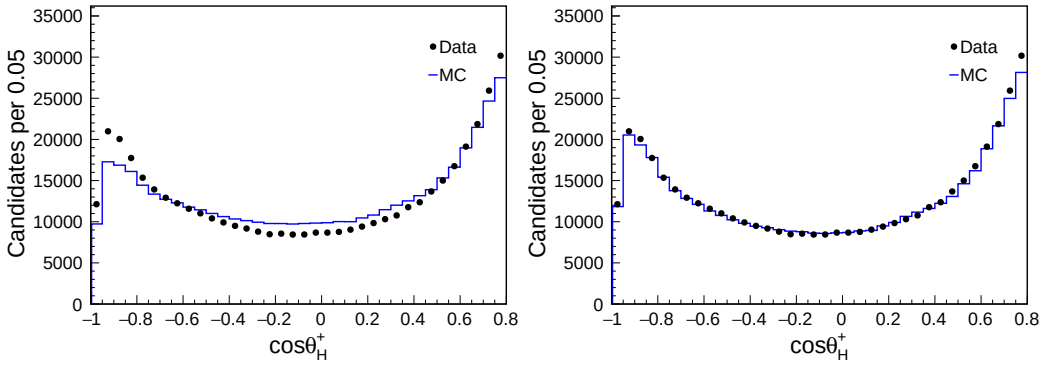


Figure 6.15: Distributions of $\cos\theta_{\rho^+}$ for $B^+ \to \rho^+\rho^0$ candidates reconstructed in (solid) simulation and (dots) data outside the signal box ($M_{bc} > 5.26$ GeV/c^2 and $-0.15 < \Delta E < 0.10$ GeV) (left) before and (right) after weighting the continuum simulation.

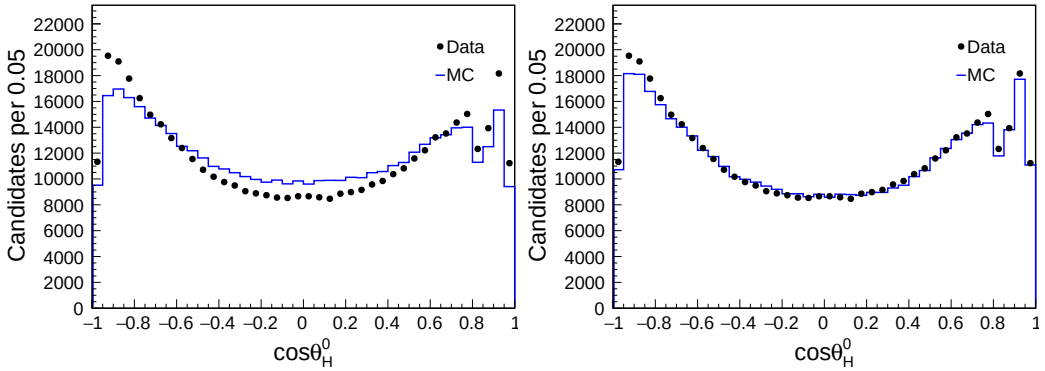


Figure 6.16: Distributions of $\cos\theta_{\rho^0}$ for $B^+ \to \rho^+\rho^0$ candidates reconstructed in (solid) simulation and (dots) data outside the signal box ($M_{bc} > 5.26$ GeV/c^2 and $-0.15 < \Delta E < 0.10$ GeV) (left) before and (right) after weighting the continuum simulation.

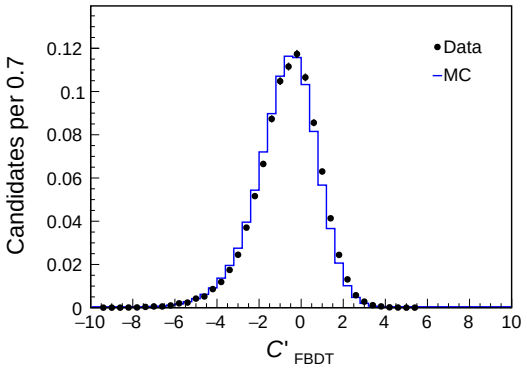


Figure 6.17: Distribution of C'_{FBDT} for $B^+ \to \rho^+\rho^0$ candidates reconstructed in off-resonance (solid) simulation and (dots) data. The distributions are normalized to unity.

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's full data set corresponding to 772 million of $B\overline{B}$ pairs yielding 33572 $B^+ \to \rho^+\rho^0$ candidates after the selection. In the fit to data, I hide the central values because the analysis is still under internal Belle review. Before the fit I apply the reweighting procedure to get continuum p.d.f. in data.

It takes twenty minutes to run the fit with `HESSE` and about five hours with `MINOS`.

Full projections onto the fit observables are shown in Figs. 7.1–7.2 and in logarithmic scale in Figs. 7.3–7.4. As expected, a signal peak is observed in the $\Delta E \approx 0$ region and in the dipion mass distributions. All distributions seem reasonably well modeled by the total fit function with the possible exception of the higher-tails of the $m(\pi^+\pi^-)$ which suggest a partial mismodeling of the $f_0(980)$ component. A possible mismodeling also affects the trailing edge of the $C'_{\rm FBDT}$ distribution ($0 \lesssim C'_{\rm FBDT} \lesssim 2$). A possible concern is the unrealistically high contribution of the $a_1\rho$ component suggested by the projections at lower ΔE values.

The results of the fit with statistical uncertainties are

$$\begin{aligned} \mathcal{B}(B^+ \to \rho^+\rho^0) &= ([\text{xx}]^{+1.20}_{-1.19}) \times 10^{-6}, \\ f_L &= [\text{xx}] \pm 0.025, \\ A_{CP} &= [\text{xx}] \pm 0.052. \end{aligned}$$

Table 7.1 summarizes the estimated correlations. A correlation of -33% is observed between $\mathcal{B}(B^+ \to \rho^+\rho^0)$ and f_L, while their correlation coefficients with A_{CP} are below $|10\%|$. A large correlation is observed between $\mathcal{B}(B^+ \to \rho^+\rho^0)$, f_L, and the estimated fractions of $B^+ \to \rho\pi\pi$ decays, as expected from the pull studies (see Sec. 5.3.1).

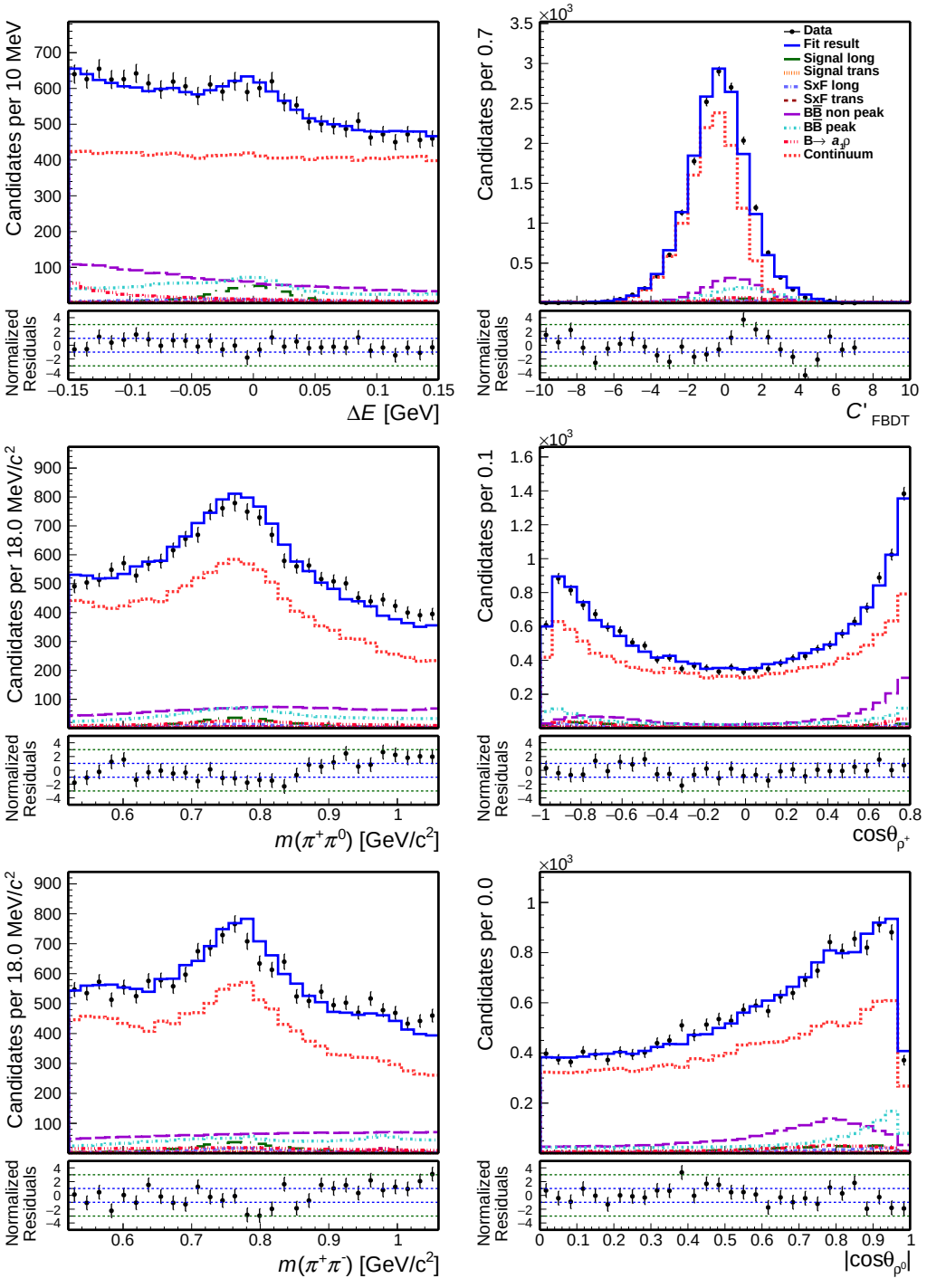


Figure 7.1: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^-\pi^0)$, (middle right) $\cos\theta_{\rho^-}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for $B^- \to \rho^-\rho^0$ candidates reconstructed in the full Belle data set collected at the $\Upsilon(4S)$ resonance and corresponding to 772 million $B\overline{B}$ pairs. Fit projections are overlaid.

	$\mathcal{B}(B^+ \to \rho^+\rho^0)$	f_L	A_{CP}	$r_{\text{sxf }(\rho^+\rho^0)_L}$	$r_{\text{sxf }(\rho^+\rho^0)_T}$	$f_{B\overline{B}}$	$f_{f_0\rho^+}$	$f_{a_1^+\pi^0}$	$f_{a_1^+\pi^-}$	$f_{a_1^0\pi^+}$	$f_{K^{*0}\rho^+}$	$f_{\rho^+\rho^-}$	$f_{a_1^+\rho^-}$	$f_{\rho^0\pi^+\pi^0}$	$f_{\rho^+\pi^+\pi^-}$	$f_{\rho^-\pi^+\pi^+}$	$f_{f_0\pi^+\pi^0}$	$f_{\pi^+\pi^-\pi^+\pi^0}$	$f_{\omega\rho^+}$	$f_{\eta'\rho^+}$	$f_{\text{comb. rare }B}$
$\mathcal{B}(B^+ \to \rho^+\rho^0)$	1	-0.330	-0.087	-0.061	-0.004	-0.006	0.014	-0.057	0.000	0.004	-0.044	-0.029	-0.132	-0.388	-0.430	-0.010	0.056	0.255	-0.025	-0.005	-0.073
f_L		1	0.078	-0.013	0.017	0.059	-0.129	-0.079	0.004	-0.010	0.032	-0.007	-0.074	0.397	0.537	-0.196	-0.027	-0.233	0.001	0.000	-0.006
A_{CP}			1	0.000	-0.003	0.008	-0.009	-0.001	0.001	0.001	0.002	0.002	0.004	0.074	0.067	-0.016	-0.004	-0.044	0.002	0.001	0.004
$r_{\text{sxf }(\rho^+\rho^0)_L}$				1	0.000	-0.003	0.002	-0.013	0.000	0.000	-0.005	-0.003	-0.017	-0.003	0.005	0.005	-0.001	0.005	0.000	-0.001	-0.035
$r_{\text{sxf }(\rho^+\rho^0)_T}$					1	-0.003	0.003	0.002	0.000	0.001	-0.014	0.001	-0.001	-0.008	-0.011	0.006	-0.001	0.005	0.000	-0.002	-0.004
$f_{B\overline{B}}$						1	-0.042	0.031	0.004	0.033	-0.012	0.004	-0.512	0.053	0.078	-0.225	-0.010	-0.085	0.002	0.000	-0.035
$f_{f_0\rho^+}$							1	-0.004	0.001	0.008	0.045	0.000	-0.029	-0.023	-0.256	0.132	-0.404	0.131	0.001	-0.015	-0.074
$f_{a_1^+\pi^0}$								1	0.004	0.006	0.005	-0.001	0.006	-0.068	-0.005	0.112	-0.015	0.007	0.001	-0.003	-0.169
$f_{a_1^+\pi^-}$									1	-0.012	0.002	-0.004	0.005	0.000	0.007	-0.057	0.002	0.007	0.001	-0.003	-0.047
$f_{a_1^0\pi^+}$										1	0.004	-0.009	0.020	-0.004	-0.004	-0.137	-0.003	0.012	0.000	0.001	-0.085
$f_{K^{*0}\rho^+}$											1	0.005	-0.084	0.005	-0.081	0.076	0.048	-0.022	0.002	-0.006	-0.058
$f_{\rho^+\rho^-}$												1	0.004	0.007	0.003	-0.020	0.002	0.005	0.000	0.002	-0.069
$f_{a_1^+\rho^-}$													1	-0.025	-0.004	0.010	0.050	0.009	-0.001	-0.025	-0.230
$f_{\rho^0\pi^+\pi^0}$														1	0.288	0.000	0.202	-0.679	-0.006	0.012	-0.027
$f_{\rho^+\pi^+\pi^-}$															1	-0.220	0.030	-0.549	0.003	-0.012	0.022
$f_{\rho^-\pi^+\pi^+}$																1	-0.057	-0.032	0.000	0.019	-0.362
$f_{f_0\pi^+\pi^0}$																	1	-0.0445	-0.002	0.003	-0.017
$f_{\pi^+\pi^-\pi^+\pi^0}$																		1	0.003	0.005	-0.013
$f_{\omega\rho^+}$																			1	-0.001	0.000
$f_{\eta'\rho^+}$																				1	-0.013
$f_{\text{comb. rare }B}$																					1

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

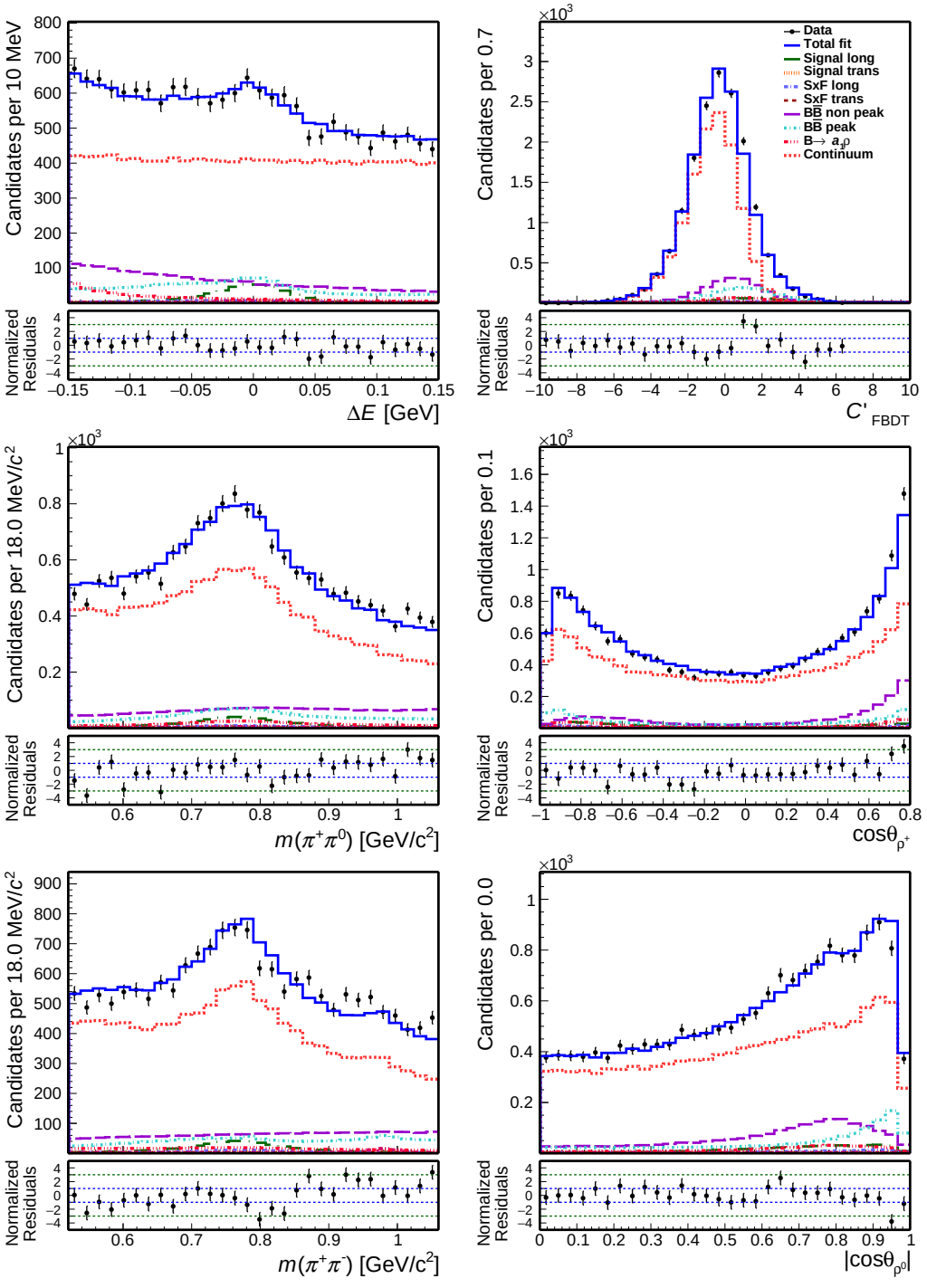


Figure 7.2: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for $B^+ \to \rho^+\rho^0$ candidates reconstructed in the full Belle data set collected at the $\Upsilon(4S)$ resonance and corresponding to 772 million $B\overline{B}$ pairs. Fit projections are overlaid.

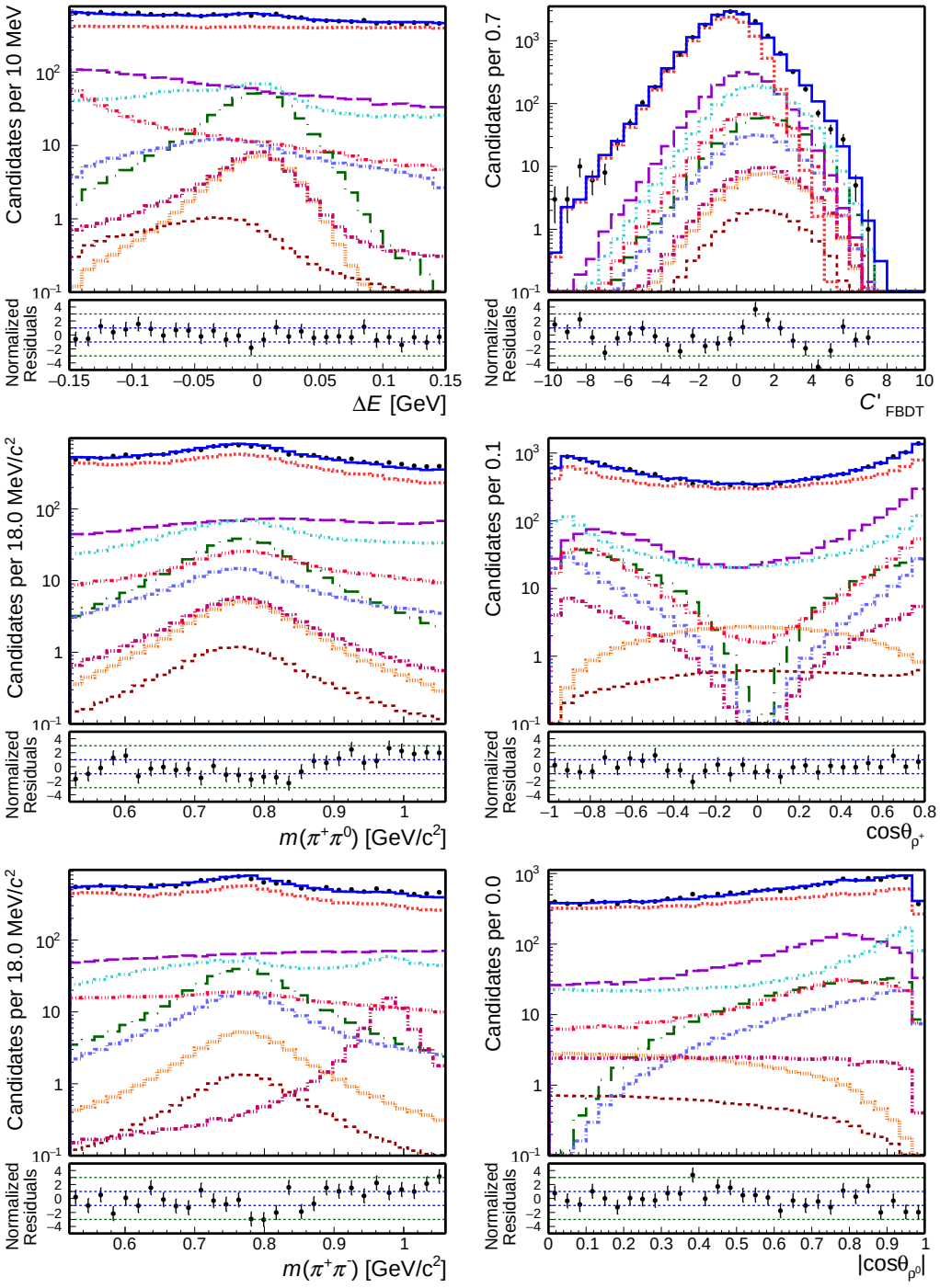


Figure 7.3: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^-\pi^0)$, (middle right) $\cos\theta_{\rho^-}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for $B^- \to \rho^-\rho^0$ candidates reconstructed in the full Belle data set collected at the $\Upsilon(4S)$ resonance and corresponding to 772 million $B\overline{B}$ pairs. The vertical axis is log-scaled. Fit projections are overlaid.

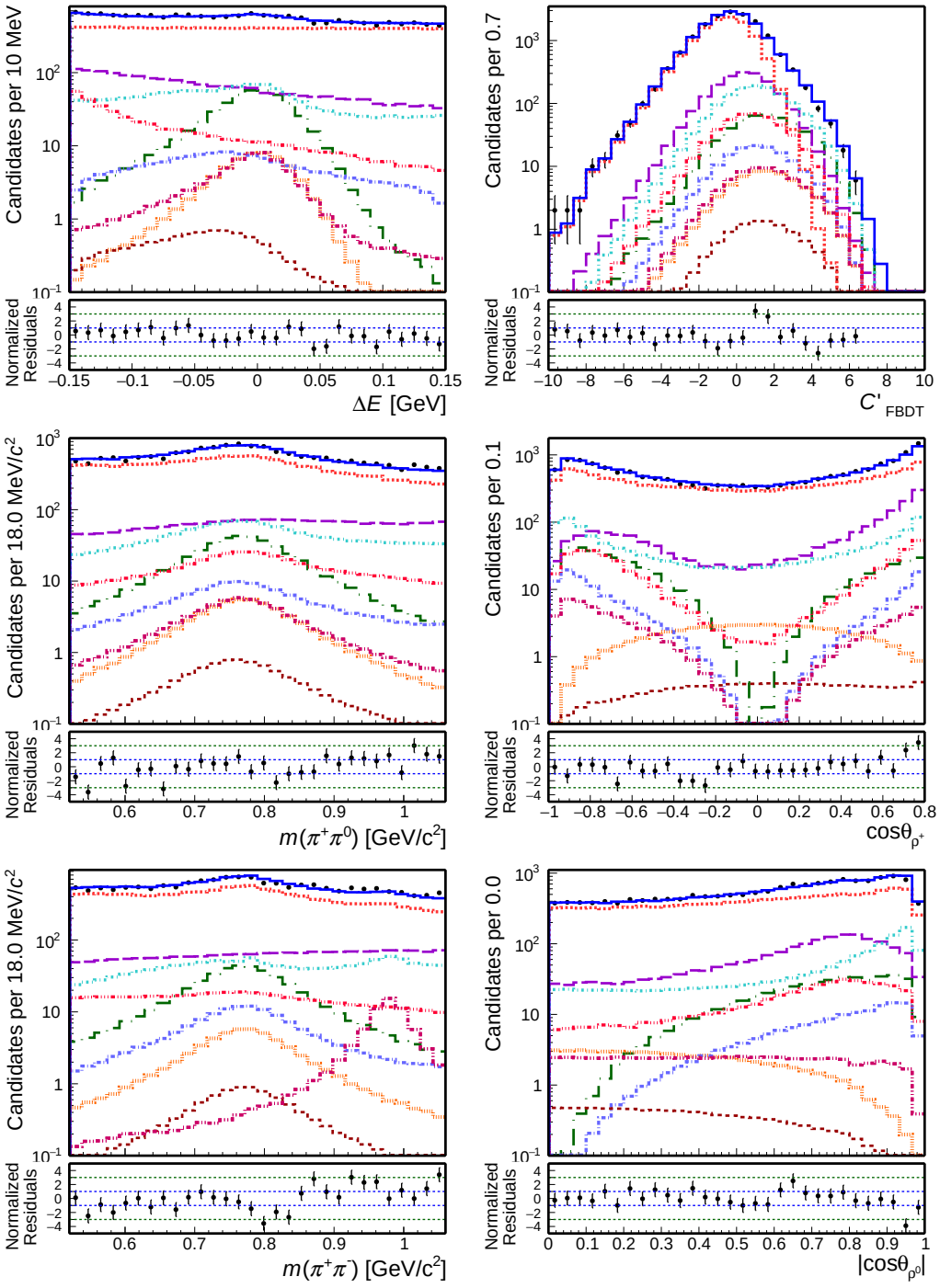


Figure 7.4: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for $B^+ \to \rho^+\rho^0$ candidates reconstructed in the full Belle data set collected at the $\Upsilon(4S)$ resonance and corresponding to 772 million $B\overline{B}$ pairs. The vertical axis is log-scaled. Fit projections are overlaid.

Chapter 8

Systematic uncertainties

This chapter describes the determination of the systematic uncertainties.

8.1 Introduction

We consider various sources of systematic uncertainties associated with the assumptions made in the analysis or the intrinsic uncertainties of external analysis inputs. Whenever a systematic uncertainty is associated to an arbitrary choice in the analysis, we 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.

8.2 Tracking

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 centrally in Belle by using unbiased samples of tracks from $D^{*+} \to D^0(\to K^0_S\pi^+\pi^-)\pi^+$ identified without using some selected tracking information [88].

One of the pions from the K^0_S decay is ignored in the reconstruction thus becoming an unbiased probe of track-finding efficiency. In fact, a signal distinctive enough from background can be reconstructed because the momentum of the unreconstructed pion cancels at first order in the distribution of the $D^{*+} - D^0 - \pi^+$ mass difference, which shows a prominent signal peak over smooth background owing to the small mass difference of 6 MeV/c^2. The momentum of the unreconstructed pion is inferred by using kinematic conservation laws along with the three constraints provided by the D^{*+}, D^0, and K^0_S masses. Additionally, the relatively long lifetime of K^0_S mesons allows determining the flight direction of the unreconstructed pion by using the K^0_S and D^0 decay vertices. The K^0_S vertex is not reconstructed itself but the segment (line) where the vertex can be located. The segment is defined by the direction of its reconstructed daughter and requirement that the K^0_S decays within 3 cm from the beam.

Track-finding efficiency in data and in simulation are consistent. Therefore, 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 0.35% per track contributes to the systematic uncertainty. The uncertainties are added linearly resulting in a total systematic uncertainty of 1.05%.

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 centrally in Belle by using unbiased samples of tracks from $D^{*+} \to D^0(\to K^0_S\pi^+\pi^-)\pi^+$. The asymmetry is compatible with zero. The statistical uncertainty of the asymmetry 1.34% is included as systematic uncertainty.

8.3 Charged particle identification

To reduce backgrounds where kaons are misidentified as pions, I apply particle identification. This introduces an efficiency that is estimated from simulation. Differences between the performance of particle identification in data and simulation exist and need to be accounted for as systematic uncertainty.

The $D^{*+} \to D^0\pi^+$, $D^0 \to K^-\pi^+$ decay is used to study centrally in Belle the particle-identification performance [89]. This channel has a sufficiently high branching ratio of (2.6 %) to ensure availability of large samples and offers striking experimental advantages. The D^{*+} meson decays into a neutral D and a charged pion through the strong interaction. Flavor and charge are conserved; hence, a positively charged D^{*+}, which contains a valence c quark, always decays in a D^0 meson and a π^+, and a D^{*-} always decays into a $\overline{D}^0$ meson and a π^-. Determination of the charge of the pion produced directly in the D^{*+} decay offers therefore a straightforward identification of the flavor of the neutral D. In addition, the large magnitude of V_{cs} (≈ 0.97) causes the D^0 meson to decay into a $K^-\pi^+$ pair much more frequently than into a $K^+\pi^-$ pair, which results from a doubly Cabibbo-suppressed transition with a branching fraction one hundred times smaller. Hence, identification of the D flavor allows identifying the pion and the kaon based on their charges independently from any dedicated particle identification information. Identification of the $D^{*\pm}$ pion is straightforward because its momentum is usually lower than the momenta of the neutral D products, due to the mass difference between the D^{*+} and the D^0, which is just slightly larger than the pion mass. We therefore have a sample where the identities of pions and kaons are known with high accuracy solely from kinematics and charge information, with no need for any dedicated particle identification. This offers unbiased pion and kaon samples to study particle identification performance.

The correction factor $R = \varepsilon_{\text{data}}/\varepsilon_{\text{MC}}$, where ε denotes the PID selection efficiency, and its uncertainty δR (Eqs. (5) and (6) in [89]) are available as functions of track polar-angle and momentum, and particle-identification selection. I use the tabulated values adequate for our sample to determine the correction and its uncertainty for each of the three final state tracks, and then calculate the total correction factor as $R_{total} = R_1 \cdot R_2 \cdot R_3$; the uncertainties are propagated linearly assuming fully correlated corrections. Table 8.1 summarizes the results. I include the correction factor in the branching fraction determination which results in a systematic uncertainty of 2.5%.

Polarization	R_1	R_2	R_3	$R_{\rm total}$	ε	$R_{\rm total}\cdot\varepsilon$
Longitudinal ($f_{\rm L}=1$)	98.68 ± 0.88	98.45 ± 0.85	98.37 ± 0.84	95.57 ± 2.49	7.99	7.64 ± 0.20
Transverse ($f_{\rm L}=0$)	96.98 ± 0.80	97.00 ± 0.80	96.97 ± 0.80	91.21 ± 2.26	14.92	13.61 ± 0.34

Table 8.1: Particle identification corrections (in percent) for each of the charged pions in the decay $B^+\to\rho^+(\pi_1^+\pi^0)\rho^0(\pi_2^+\pi_3^-)$ and for each polarization, total correction $R_{total}=R_1\cdot R_2\cdot R_3$, signal efficiency (from Table 5.2) and corrected signal efficiency. The uncertainties include statistical and systematic uncertainties determined following Ref. [89].

8.4 Neutral pion reconstruction efficiency

The reconstruction and selection of the neutral pion is associated to an efficiency, which is determined from simulation. Belle studies data-simulation discrepancies using $\tau^-\to\pi^-\pi^0\nu_\tau$ decays [90]. This channel provides a low-background environment to reconstruct π^0 and has a large branching ratio of 25% known with high precision ($\delta\mathcal{B}/\mathcal{B}\approx0.3\%$).

The correction factor $R=0.957\pm0.015$, due to the discrepancy in π^0 reconstruction efficiency in data and simulation, is included in the calculation of the branching fraction. The 1.49% uncertainty of this correction is included in the systematic uncertainty.

8.5 $B\overline{B}$ pair counting

For the calculation of the branching fraction, the number of $B\overline{B}$ pairs $(771.6\pm10.6)\times10^6$ collected by Belle is used [85]. This number is obtained by subtracting off-resonance hadronic contributions, which are measured by the number of events meeting the requirements described in Sec. 4.1.4, from the total number of on-resonance hadronic events. The uncertainty of this number is included in the systematic uncertainty.

8.6 Simulated sample size

We determine selection efficiencies from simulation. The 0.2% uncertainties due to the finite size of the simulated samples are included in the systematic uncertainty.

8.7 Continuum suppression selection efficiency

The continuum suppression selection is an essential part of the analysis. This introduces an efficiency which is determined from simulation and could differ from the actual value in data. Since I apply the same continuum suppression criteria to control modes, I use the results of the $B^+\to\bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ branching fraction measurement (see Sec. 6.2.3) to assign a systematic uncertainty. The measured and reference branching fractions are consistent within 5% precision. This uncertainty includes not only any effect due to continuum-suppression selection efficiency mismodeling but also statistical uncertainties and additional systematic uncertainties on track-finding efficiency, π^0 reconstruction efficiency, number of $B\overline{B}$, and PID efficiency. By subtracting these effects the uncertainty decreases to 3.7%, which is attributed as the systematic uncertainty associated with the continuum-suppression selection efficiency. We assign a conservative 1.85% systematic uncertainty for f_L assuming an additional 50% mismodeling difference between the efficiency for reconstructing transversely and longitudinally polarized signal decays.

8.8 Best candidate selection

In the default analysis, I select one candidate per event by choosing the candidate 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 generate 300 simulated $B^+ \to \rho^+\rho^0$ samples where one candidate per event is chosen randomly and fit them with the default model. Pull distributions show biases of -12% for f_L, -33% for $\mathcal{B}(B^+ \to \rho^+\rho^0)$, and 27% for A_{CP}, which are attributed as uncertainties after rescaling them using the world-average central values as references.

8.9 Fit bias

We study the estimator properties, as described in Sec. 5.3. The pull distributions show biases of 20% for $\mathcal{B}(B^+ \to \rho^+\rho^0)$ and -30% for f_L, which are included in the systematic uncertainty after rescaling them using the world-average central values as references.

8.10 Fit model

8.10.1 ΔE signal and self-cross feed model

The signal and self-cross feed ΔE distributions are described by histogram templates extracted from simulation. The simulation is known not to reproduce accurately ΔE distribution in data, as shown in control-channel studies (Sec. 6.2.3). In the default analysis, I modify the signal and self-cross feed ΔE model by using correction factors extracted from fits of $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ signal in data, which is the similar to our signal sample. In order to include systematic uncertainties associated with any sample-dependence of such factors, I evaluate the impact on the results from using correction factors extracted from other control channels, as $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$. I generate 300 toy samples that incorporate the alternative ΔE model and fit them with the default model. The pull distributions show biases of 11% for $\mathcal{B}(B^+ \to \rho^+\rho^0)$, -12% for f_L, and 4% for A_{CP}, which are included in the systematic uncertainties after rescaling them using the world-average central values as references.

8.10.2 ΔE continuum model

The continuum ΔE distribution is modeled from simulation. The distributions in data and in simulation might differ due to possible simulation mismodelings. I compare the results of the fit of $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ data obtained using the continuum ΔE model extracted from simulation and an alternative model based on off-resonance data. The observed difference is $< 0.1\%$ and is included in the systematic uncertainty.

8.10.3 ΔE $B\overline{B}$ background model

The $B\overline{B}$ ΔE distribution is modeled from a finite-size simulated sample. I estimate a systematic uncertainty by repeating the fit in toy samples generated with alternative models based on statistical fluctuations. I obtain 0.2% changes that are attributed as systematic uncertainties.

8.10.4 C'_{FBDT} model

I estimate a systematic uncertainty due to a possible discrepancy in C'_{FBDT} shapes in data and simulation. In the fit of $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ data, I allow for a two-degrees of freedom flexibility for the C'_{FBDT} model to mirror data. By comparing $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ signal yields obtained with the flexible and fixed model, I evaluate a 0.5% difference which is included in the systematic uncertainty.

8.10.5 Signal helicity angle model

The signal helicity angle distributions are described by histogram templates extracted from simulation. The simulation might not reproduce the data accurately. I quote a systematic uncertainty due to possible mismodeling.

I restrict $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ data and simulation samples to $0 < \Delta E < 0.05$ GeV range, where large signal-to-background ratio is expected (Figs. 6.3 and 6.6). By dividing $\cos\theta_{\rho^+}$ distribution obtained in data by that in signal simulated sample, I produce a set of weights that represents the typical size of data-simulation discrepancies. I use these weights to modify the default model and generate toy samples, which are fit with the default model. Pull distributions show biases of 50% for $\mathcal{B}(B^+ \to \rho^+\rho^0)$, -5% for f_L, and 5% for A_{CP}. These values are included as systematic uncertainties after rescaling them using the world-average central values as references.

8.11 Helicity angle distributions

I use histograms obtained from off-resonance data to weight continuum simulation to reduce data-simulation discrepancies in helicity-angle distributions (see Sec. 6.3.1). Since only a small sample of off-resonance data is available, I quote a systematic uncertainty due to statistical fluctuations.

The bin contents of data histograms are sampled according to Poisson distributions providing new weights. I produce twenty sets of weights and use them to obtain twenty weighted continuum models. By using those models, I fit the realistically simulated sample where continuum is weighted with the original set of weights. I study the difference of fit results obtained with the default continuum model and with the fluctuated continuum model. The widths of the distributions are 0.1549×10^{-6} for $\mathcal{B}(B^+ \to \rho^+\rho^0)$, 0.001 for f_L, and negligible for A_{CP}. These values are included as systematic uncertainties after rescaling them using the world-average central values as references.

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

In the default fit, I neglect possible interferences between B^+ decays into four-pion final states such as $B^+ \to a_1^+\pi^0$, $B^+ \to a_1^0\pi^+$, $B^+ \to \rho^0\pi^+\pi^0$, $B^+ \to f_0\rho^+$, etc. I study the possible impact of this assumption following previous Belle measurements in $B^0 \to \rho^+\rho^-$ and $B^0 \to \rho^0\rho^0$ decays [40, 63]. 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 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 so as to identify the value that best reproduces the expected proportion of signal and $B^+ \to a_1^+\pi^0$ decays in the Belle sample. 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 8.2 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.

C, ϕ	$X_{\rm sig}$	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 8.2: 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, $X_{\rm sig} = N_L/\varepsilon_L + N_T/\varepsilon_T$ 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.

8.13 Intermediate resonance branching fractions

The uncertainties on intermediate resonance branching fractions give negligible contributions to the systematic uncertainty [10].

8.14 Summary of systematic uncertainties

Table 8.3 summarizes the 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].

Source	$\mathcal{B}$	$f_{\rm L}$	A_{CP}
Tracking	1.05%	-	-
Instr. tracking asymmetry	-	-	< 0.001
Particle identification efficiency	2.5%	-	-
π^0 efficiency	1.49%	-	-
Number of $B\bar{B}$ pairs	1.37%	-	-
Simulated sample size	0.2%	$<$0.001	-
Continuum suppression selection efficiency	3.7%	0.0185	-
Best candidate selection	1.6%	0.003	0.014
Fit bias	1%	0.008	< 0.001
Fit model	2.5%	0.006	0.006
Helicity angle distributions	0.6%	0.001	-
Interference with 4π final-states	1.2%	< 0.001	-
Branching fractions of intermediate resonances	< 0.1%	-	-
Total	1.46×10^{-6} (6.1%)	0.021	0.015
Statistical uncertainty	$^{+1.20}_{-1.19}\times10^{-6}$ (5%)	0.025	0.052

Table 8.3: 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,

$$\mathcal{B}(B^+ \to \rho^+\rho^0) \;=\; ([\mathrm{xx}]^{+1.20}_{-1.19}\ (\mathrm{stat}) \pm 1.46\ (\mathrm{syst})) \times 10^{-6},$$

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

$$f_L \;=\; [\mathrm{xx}] \pm 0.025\ (\mathrm{stat}) \pm 0.021\ (\mathrm{syst}),$$

and the CP-violating charge-asymmetry,

$$A_{CP} \;=\; [\mathrm{xx}] \pm 0.052\ (\mathrm{stat}) \pm 0.015\ (\mathrm{syst}).$$

The central values are hidden as the analysis is still under internal Belle review.

The systematic contribution to the uncertainty is larger (marginally) than the statistical uncertainty only for the $\mathcal{B}(B^+ \to \rho^+\rho^0)$ result, and dominated by the uncertainty associated with the continuum-suppression selection efficiency. The dominant systematic uncertainties for the f_L and A_{CP} measurements are those associated with the continuum-suppression and best-candidate selection efficiency, respectively.

Table 9.1 shows comparison of these results with previous measurements. I project the 2003 Belle results [41] to the full Belle luminosity by applying a $\sqrt{85/772}$ factor to the 2003 statistical uncertainties to compare the analysis sensitivity independently of sample size.

The A_{CP} precision is 20% better than the projection of the previous Belle measurement and comparable with the current world-best result (by the BaBar collaboration [42]). The achieved statistical uncertainty on f_L is 10% better than that (projected) of the previous Belle measurement, but is slightly worse with respect to the current world-best result. The achieved $\mathcal{B}(B^+ \to \rho^+\rho^0)$ results improve over (the projection of) the previous Belle determination by a factor of two and improve by 10% over current best results. Systematic uncertainties are comparable with those of BaBar, thus making the reported $B^+ \to \rho^+\rho^0$ branching fraction measurement world leading.

I then perform an isospin analysis to determine α (see Sec. 1.6.2) by comparing the current global determination with a determination that includes our results, assumed to have the same central value as the current world average. No S^{00}_{CP} constraint is used.

The theoretical expressions for branching fractions and the CP-violating asymmetries (Eqs. 1.20) are functions of α, α_{eff}, δ, A^{+-}, and A^{+0}. A χ^2-like quantity is constructed based on the difference between observed and theoretical branching fractions divided by the experimental uncertainties. The χ^2 is evaluated at each point of a finely grained scan of the α space, while all other parameters are free. The resulting χ^2 values as functions of α

	Data fit	Belle 2003 [41]	Belle 2003 scaled	BaBar 2009 [42]
$\mathcal{B}$, 10^{-6}	$[\mathrm{xx}]^{+1.20}_{-1.19} \pm 1.46$	$31.7 \pm 7.1^{+3.8}_{-6.7}$	31.7 ± 2.4	$23.7 \pm 1.4 \pm 1.4$
f_L	$[\mathrm{xx}] \pm 0.025 \pm 0.021$	$0.95 \pm 0.11 \pm 0.02$	0.950 ± 0.036	$0.950 \pm 0.015 \pm 0.006$
A_{CP}	$[\mathrm{xx}] \pm 0.052 \pm 0.015$	$0.00 \pm 0.22 \pm 0.03$	0.000 ± 0.073	$-0.054 \pm 0.055 \pm 0.010$

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.

are mapped into confidence levels assuming the asymptotic χ^2 distribution with one degree of freedom. The resulting uncertainties for α correspond to the 68% CL. Figure 9.1 shows the results. By choosing the solution compatible with the SM between the two solutions, we obtain $\alpha = 94^\circ \pm 6.8^\circ$, to be compared with the current value $\alpha = 94^\circ \pm 7.8^\circ$. Our measurement achieves a 13% improvement on the global knowledge of α.

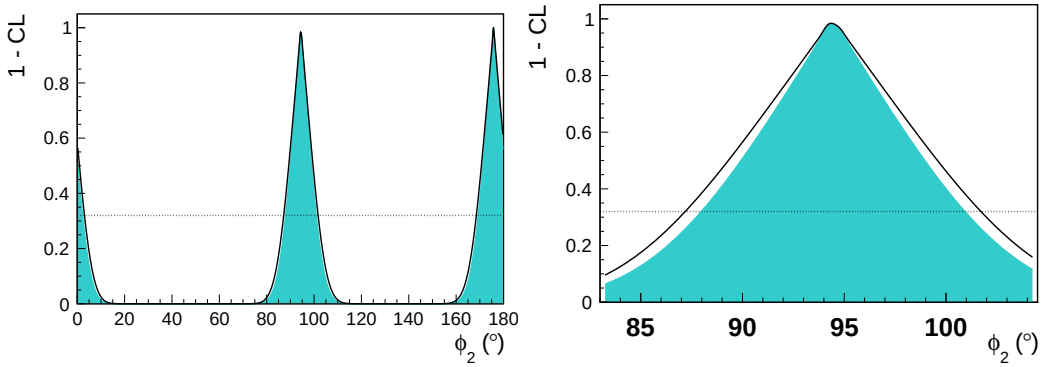


Figure 9.1: P-value as a function of α as resulting from an isospin analysis of the $B \to \rho\rho$ system without including S^{00}_{CP} constraint. The left panel shows the full range; the right panel shows a magnification of the solution compatible with the SM. The black line shows the result of the scan using current world-average results. The cyan shaded area corresponds to a scan where the precision of the current average for $B^+ \to \rho^+\rho^0$ is updated with the uncertainties resulting from this analysis. The dotted horizontal line corresponds to a CL of 68%.

Summary

I report on world-leading measurement of the decay properties of $B^+ \to \rho^+\rho^0$ decays, which are those that mostly limit the precision on the determination of $\alpha = \arg(-(V_{td}V_{tb}^*)/(V_{ud}V_{ub}^*))$ and, consequently, 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 data set, the largest and lower-background sample currently available to achieve these measurements.

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 an angular analysis. This results in two 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 address each of the 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. This is the single most difficult challenge in the analysis, which prevented the Belle collaboration to update it since 2003.

I achieve the reconstruction of approximately 1000 $B^+ \to \rho^+\rho^0$ decays in the full Belle data set and measure

$$\begin{aligned} \mathcal{B}(B^+ \to \rho^+\rho^0) &= ([\mathrm{xx}]^{+1.20}_{-1.19}\ (\mathrm{stat}) \pm 1.46\ (\mathrm{syst})) \times 10^{-6}, \\ f_L &= [\mathrm{xx}] \pm 0.025\ (\mathrm{stat}) \pm 0.021\ (\mathrm{syst}), \\ A_{CP} &= [\mathrm{xx}] \pm 0.052\ (\mathrm{stat}) \pm 0.015\ (\mathrm{syst}). \end{aligned}$$

The central values are hidden as the analysis is still under internal collaboration review. The A_{CP} precision is improved by 20% with respect to the previous Belle result (scaled to the full luminosity) and is comparable with that of the current world-best result. The f_L uncertainty improves over the previous Belle result by 10%, and is slightly worse with respect to the current world best result. The uncertainty on $\mathcal{B}(B^+ \to \rho^+\rho^0)$ is a factor of

two better than the previous Belle result and 10% better than the current world-best result. Our results provide a 13% improvement on the global knowledge of the α parameter.

Appendices

Appendix A

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

To determine the $B^+ \to \rho^+\rho^0$ sample composition in Belle data, I perform a fit over the six-dimensional space of fit observables (see Sec. 5.1). The challenge is to obtain a model that properly accounts for the multidimensional dependences between the observables. For each fit component, I study distributions of each observable conditional on each of the others to identify such 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.26 show modeling fit projections for some of the fit components.

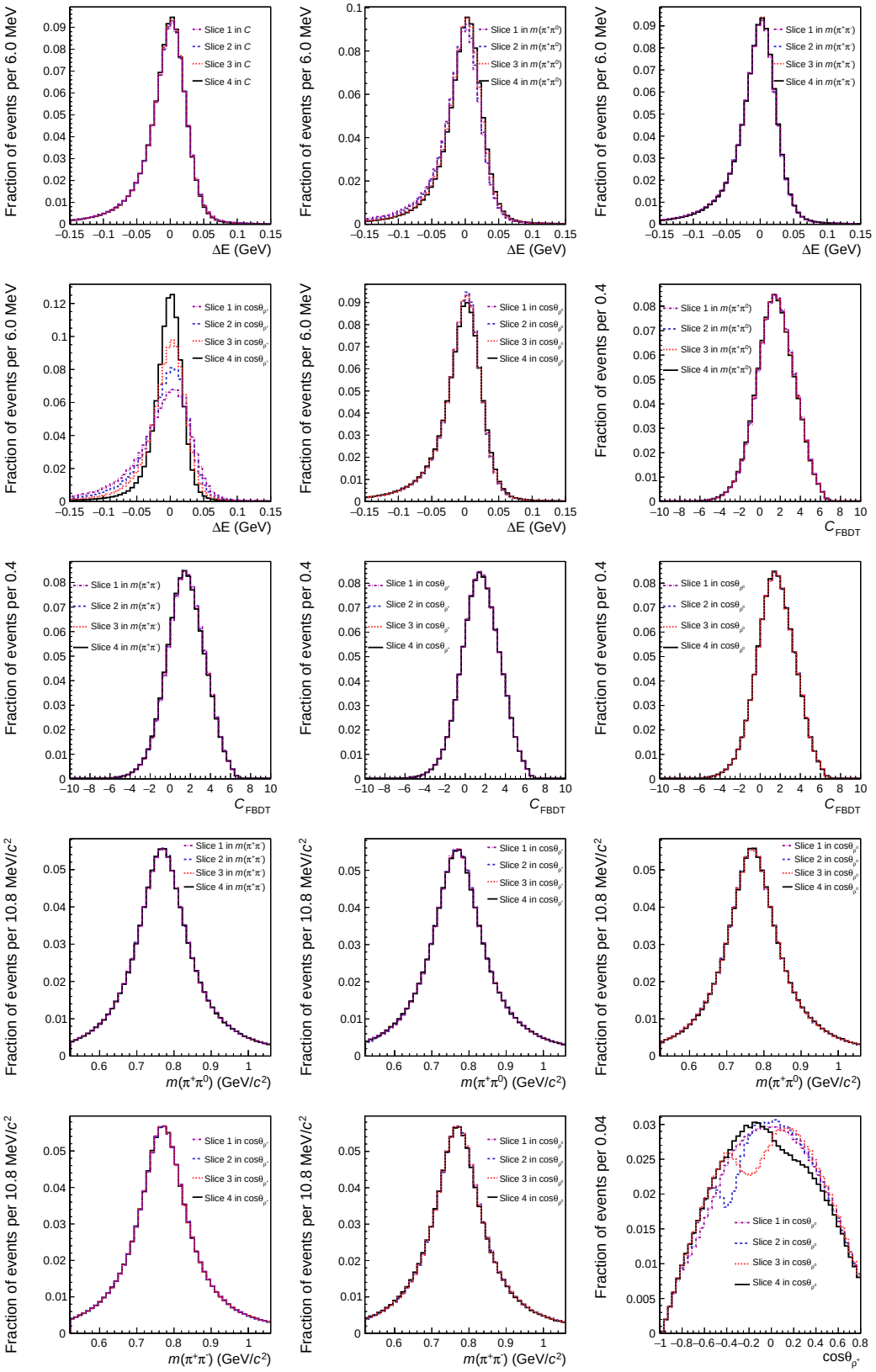


Figure A.1: Distributions of fit observables in slices of other fit observables for simulated transversely polarized signal decays. Distributions normalized to unity.

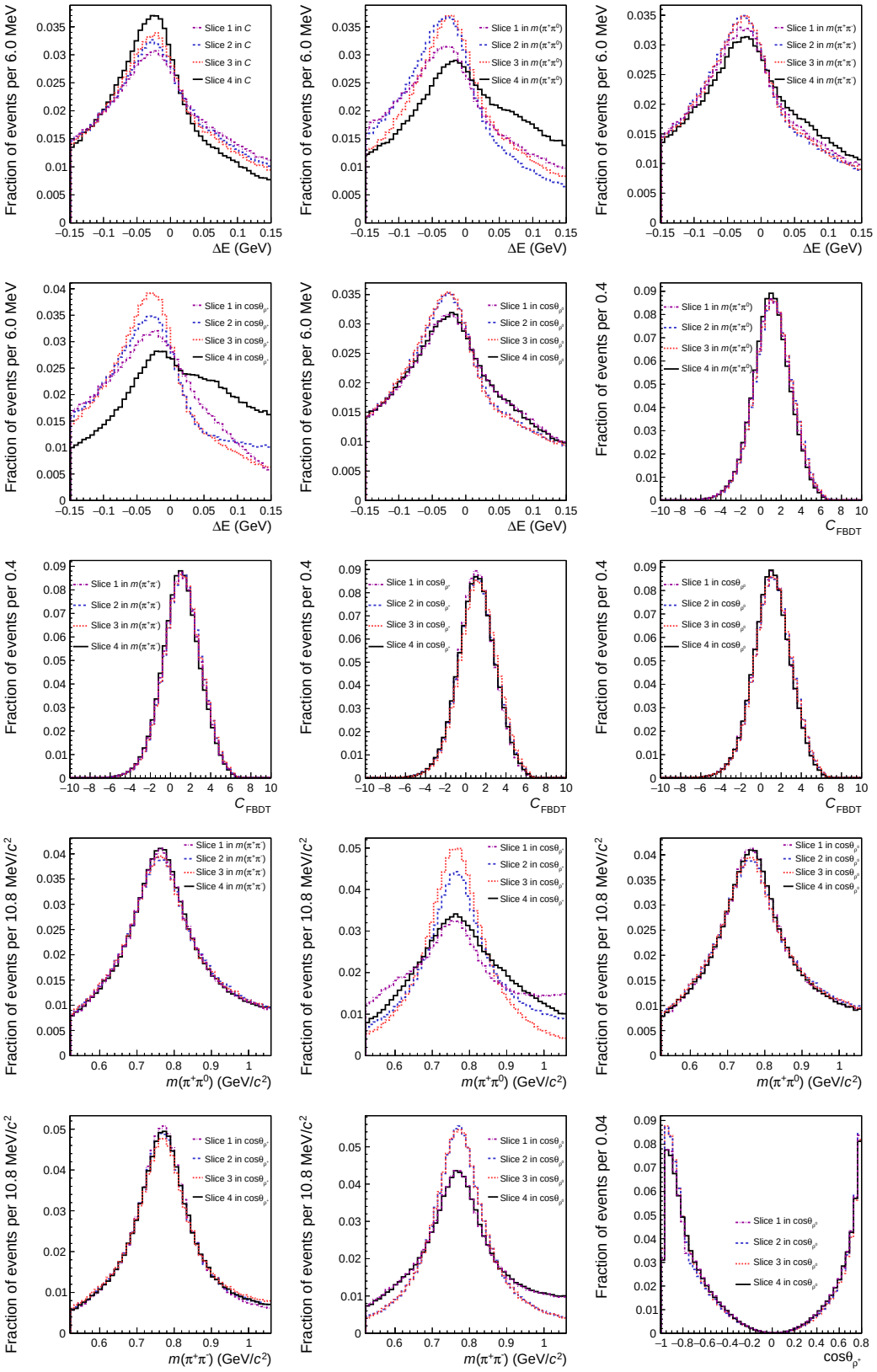


Figure A.2: Distributions of fit observables in slices of other fit observables for simulated longitudinally polarized self-cross-feed events. Distributions normalized to unity.

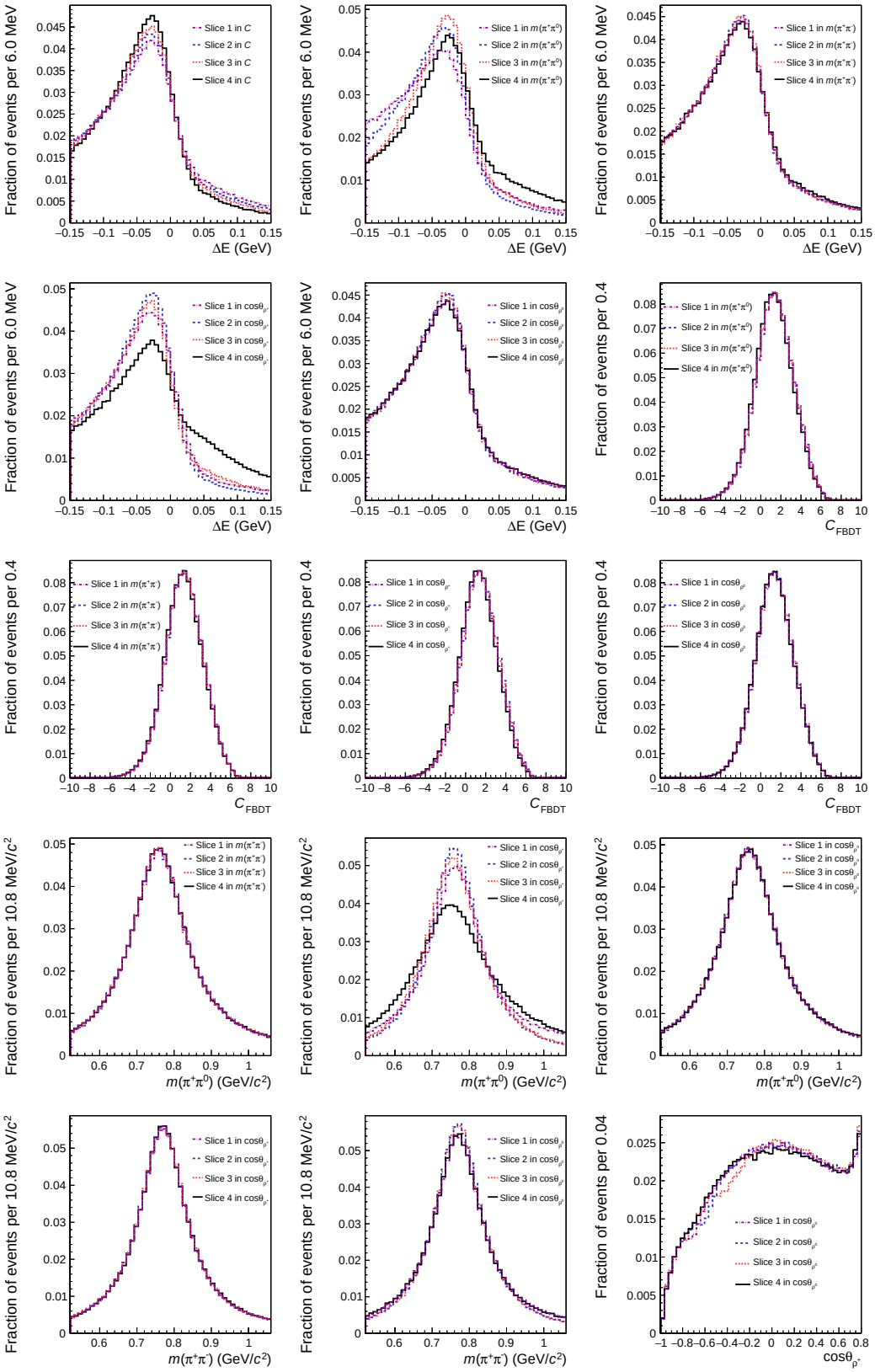


Figure A.3: Distributions of fit observables in slices of other fit observables for simulated transversely polarized self-cross-feed events. Distributions normalized to unity.

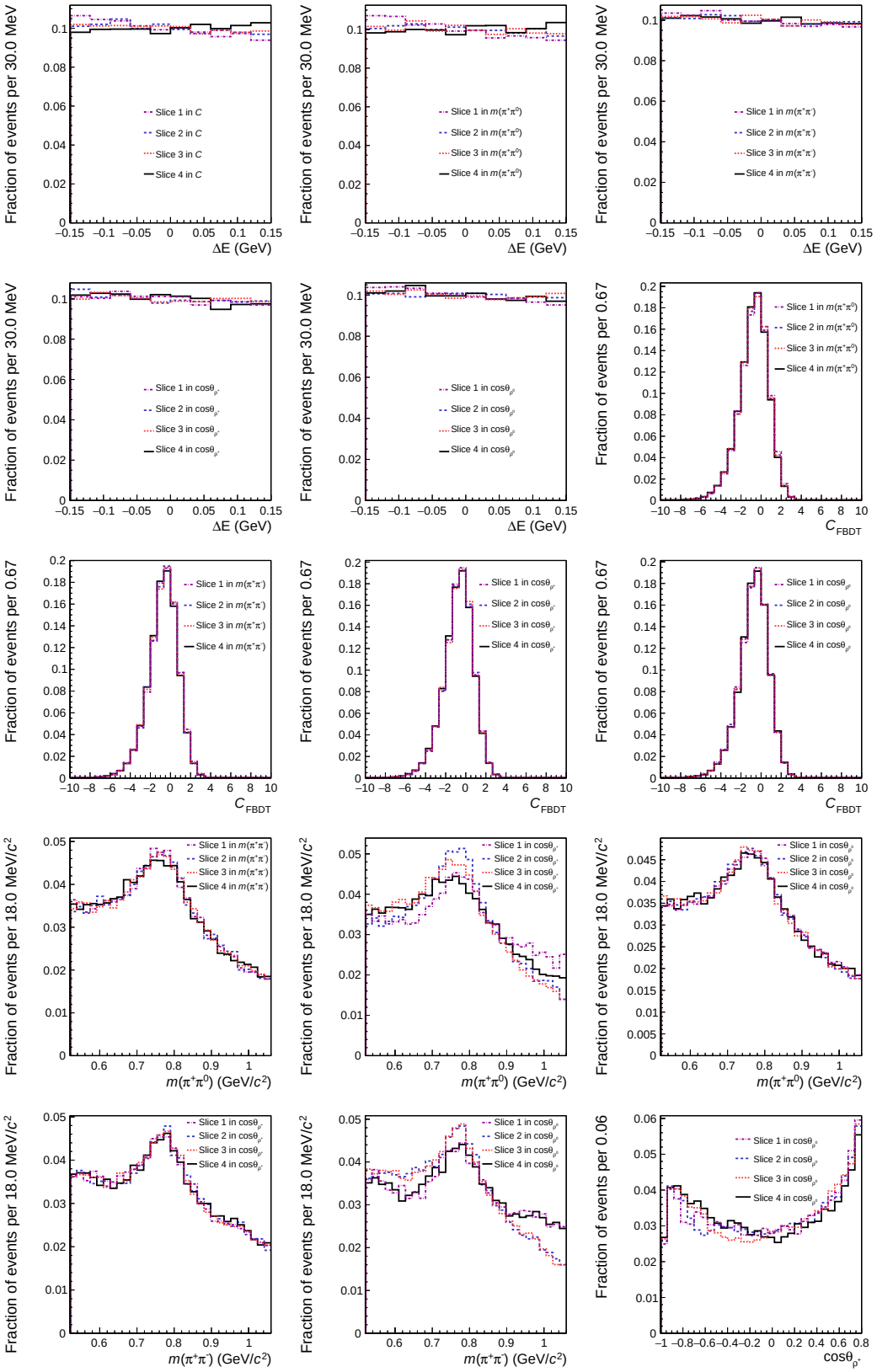


Figure A.4: Distributions of fit observables in slices of other fit observables for simulated continuum events. Distributions normalized to unity.

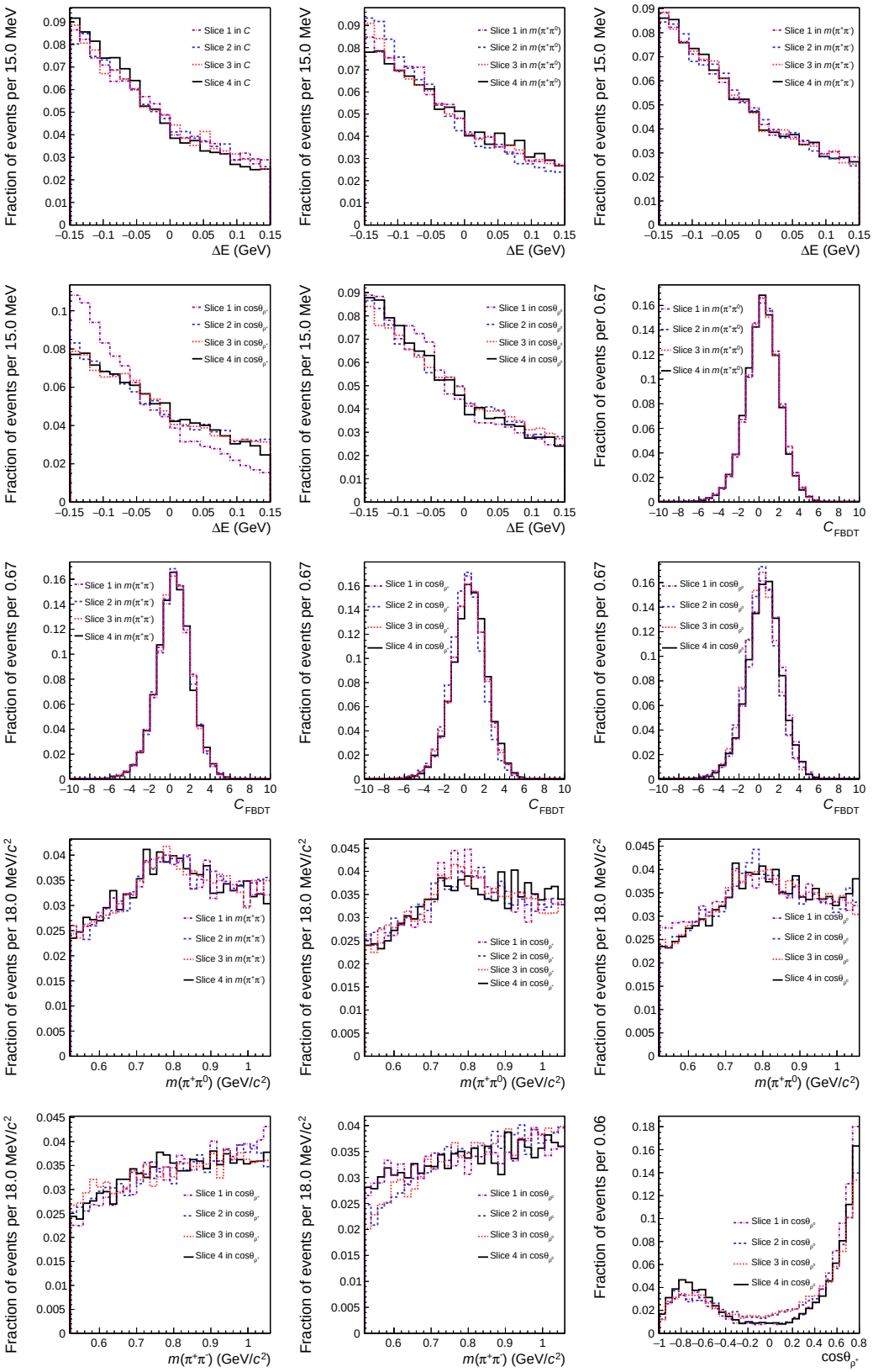


Figure A.5: Distributions of fit observables in slices of other fit observables for simulated non-peaking $B\bar{B}$ events. Distributions normalized to unity.

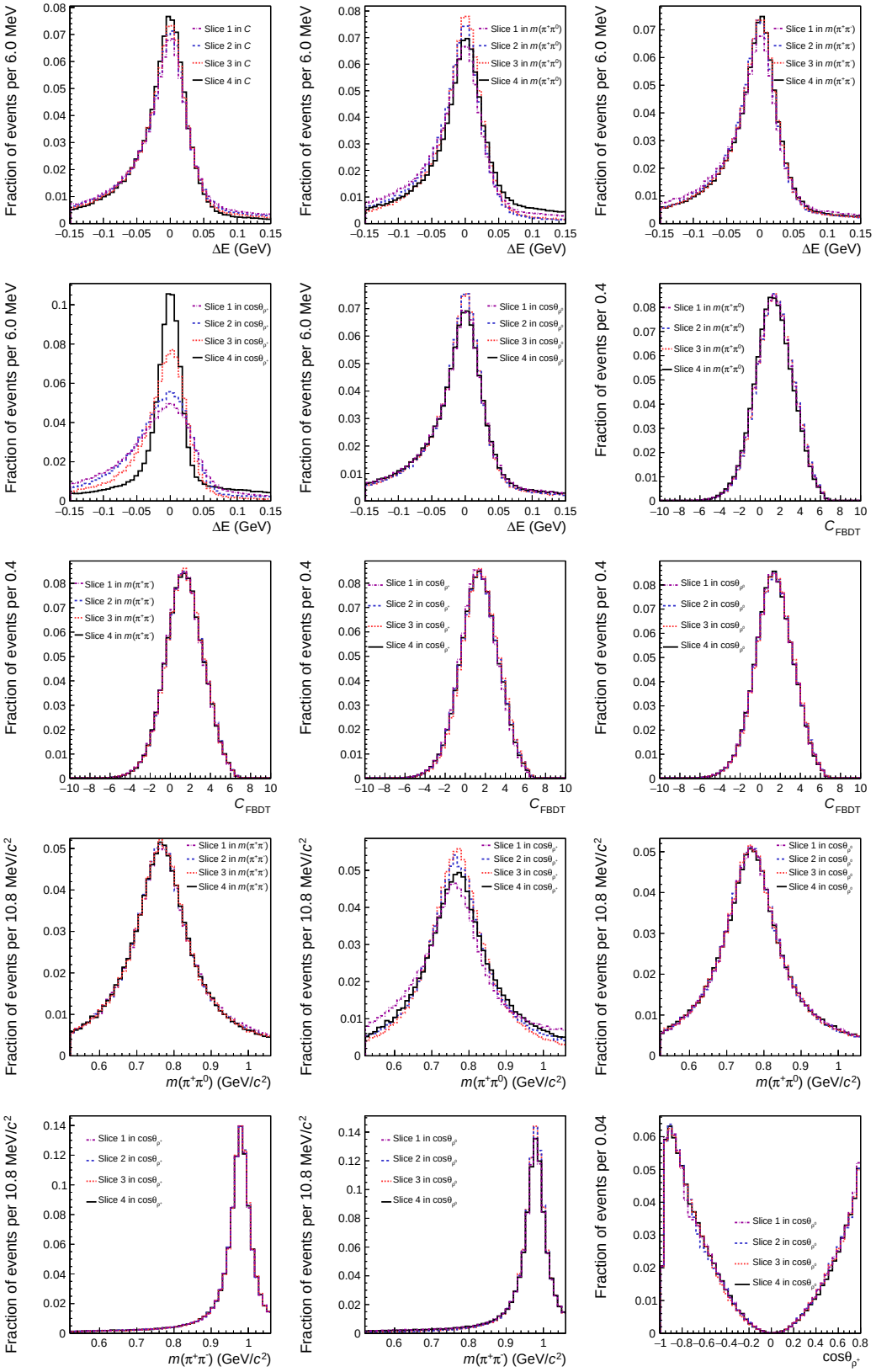


Figure A.6: Distributions of fit observables in slices of other fit observables for simulated $B^+ \to f_0\rho^+$ decays. Distributions normalized to unity.

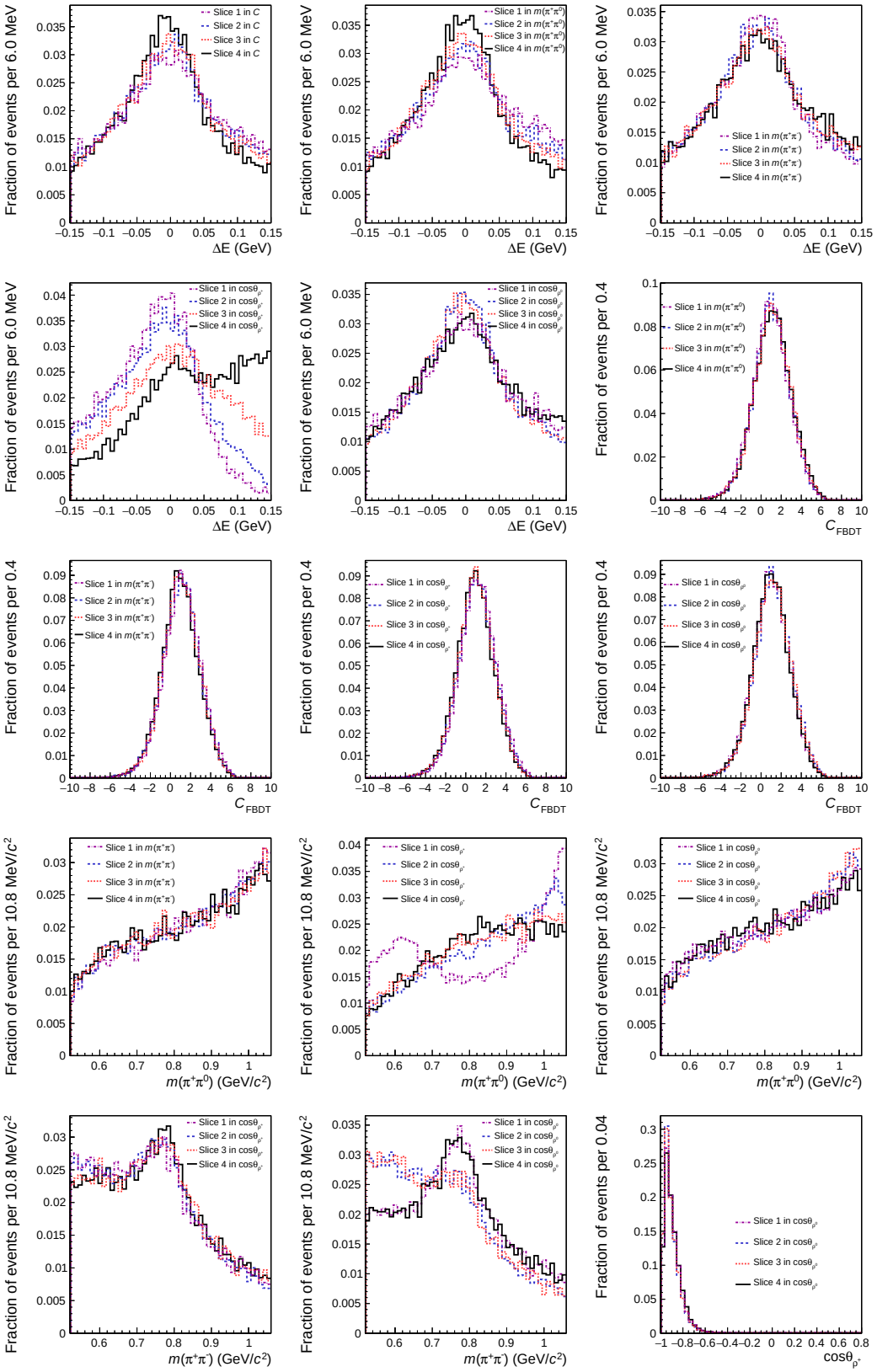


Figure A.7: Distributions of fit observables in slices of other fit observables for simulated $B^+ \to a_1^+\pi^0$ decays. Distributions normalized to unity.

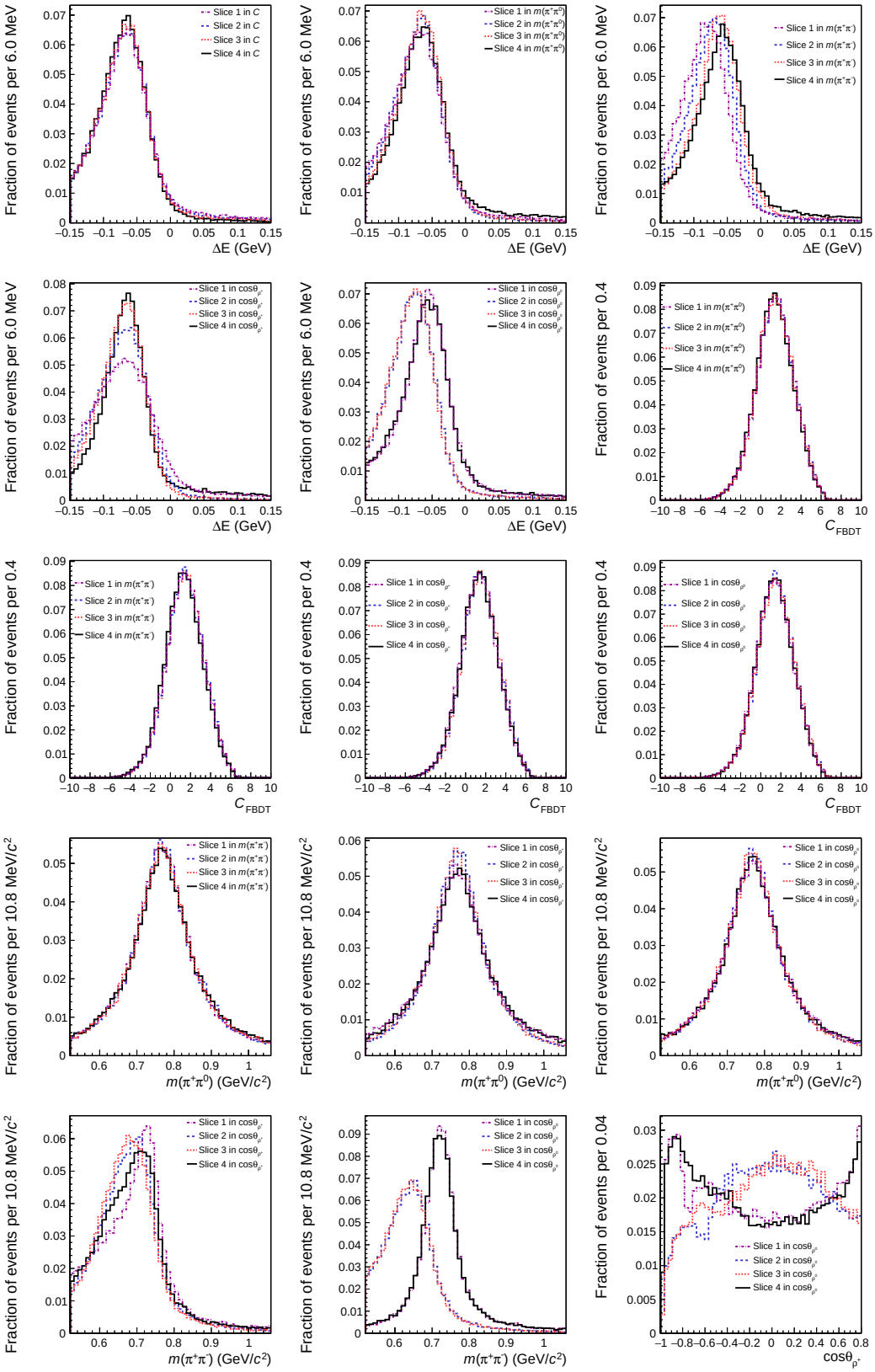


Figure A.8: Distributions of fit observables in slices of other fit observables for simulated $B^+ \to K^{*0}\rho^+$ decays. Distributions normalized to unity.

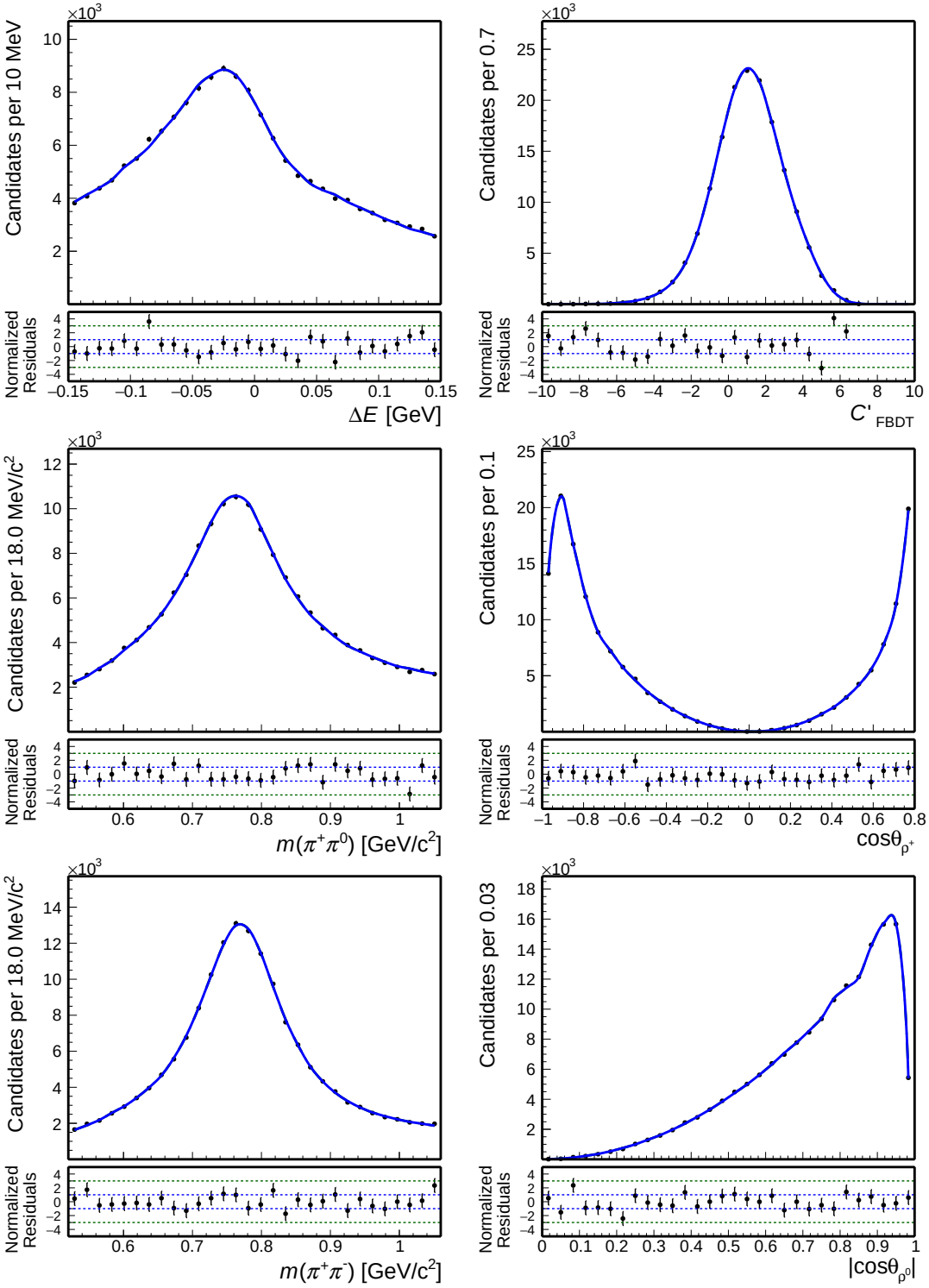


Figure A.9: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated longitudinally polarized self cross feed. The corresponding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

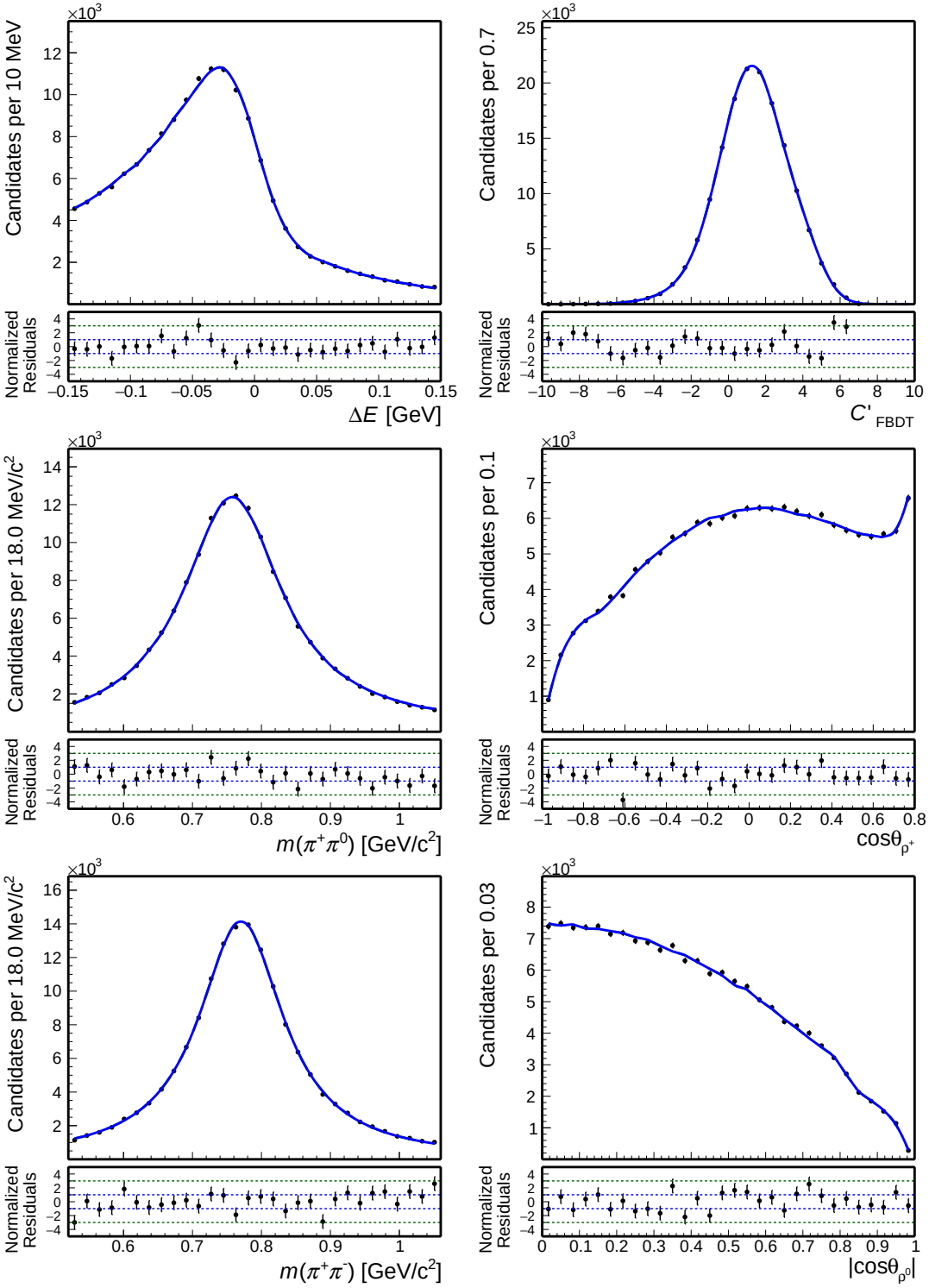


Figure A.10: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated transversely polarized self cross feed. The corresponding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

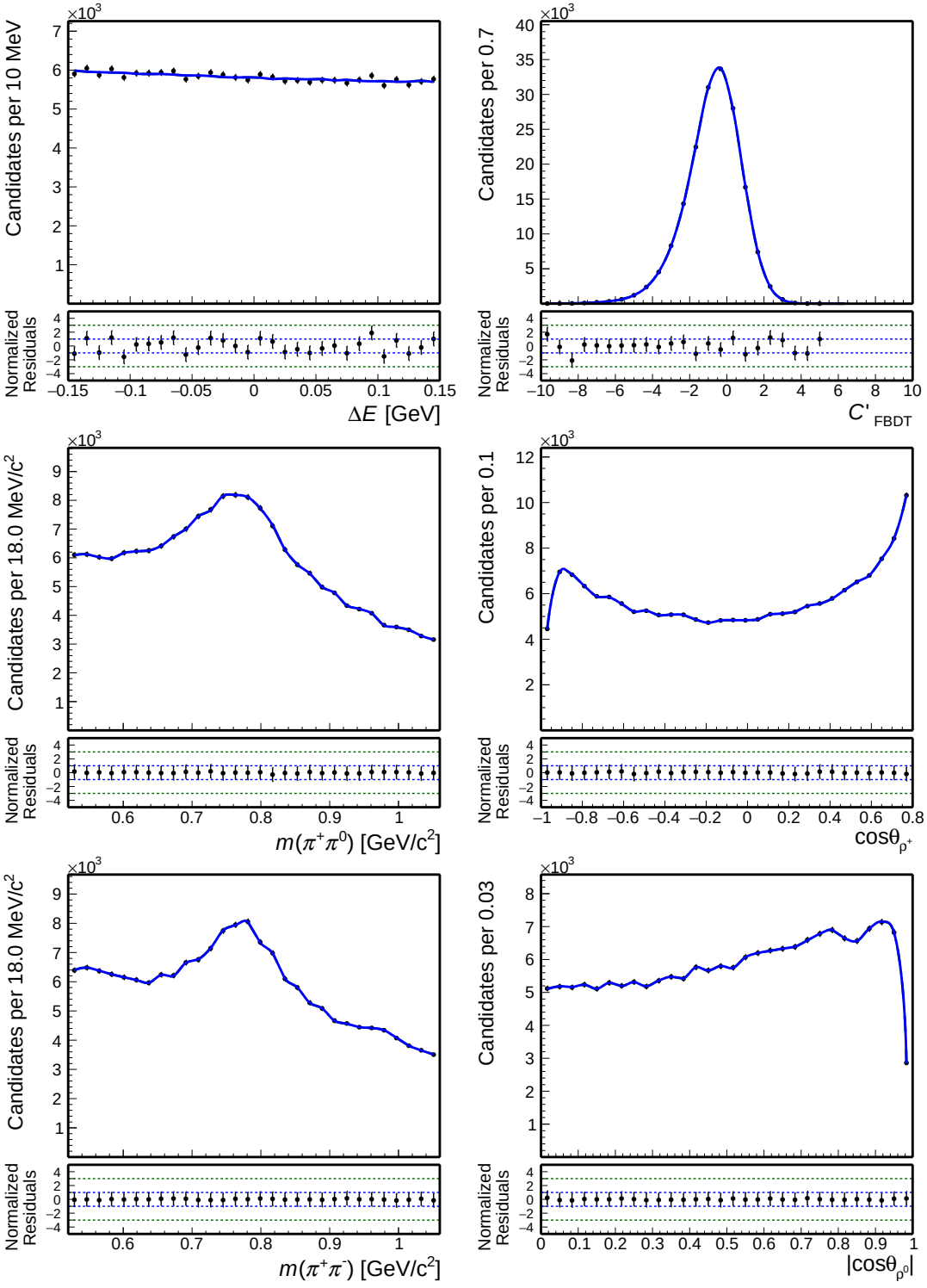


Figure A.11: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated continuum. The corresponding model shapes are overlaid (solid).

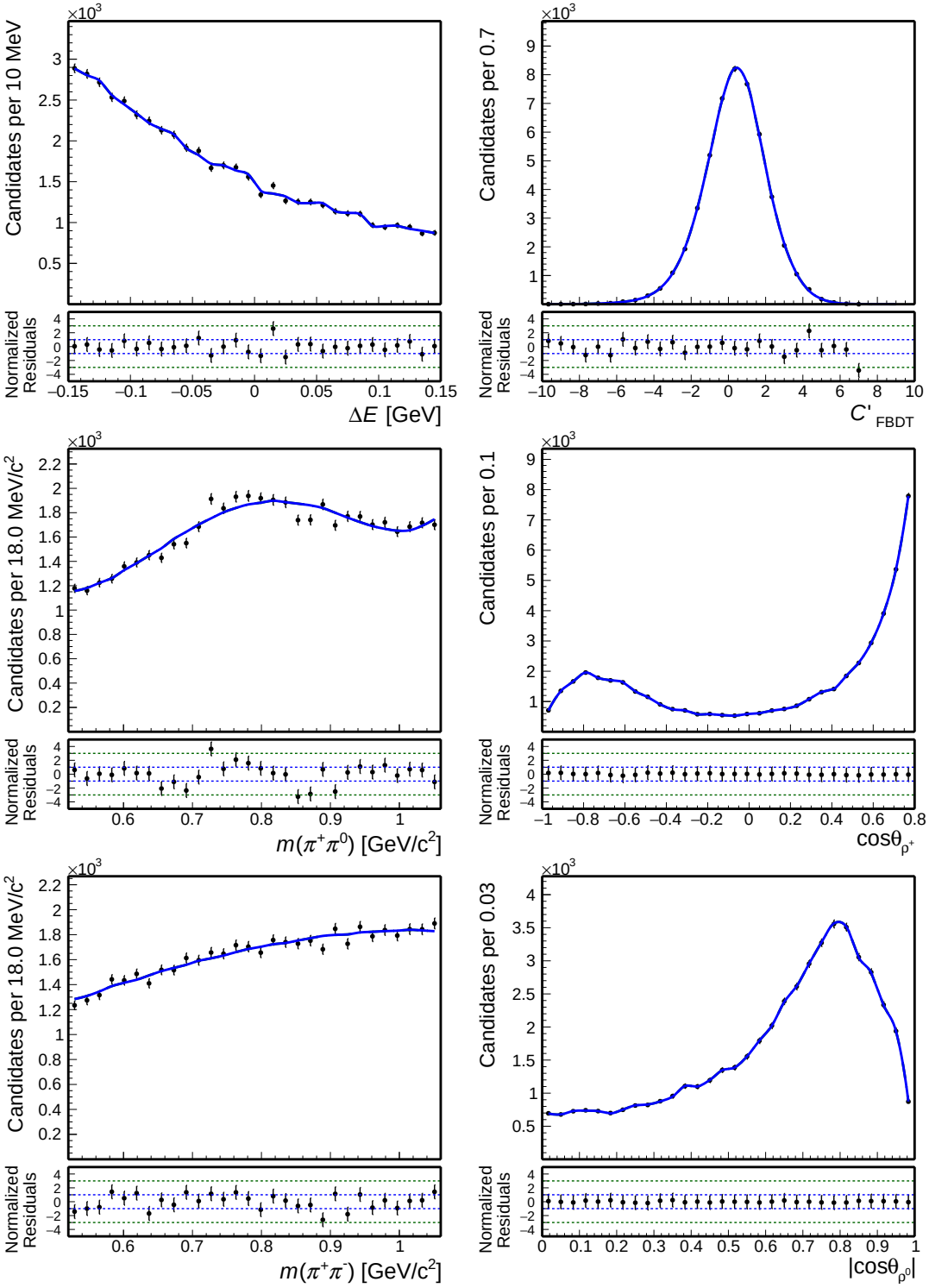


Figure A.12: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated non-peaking $B\overline{B}$ background. The corresponding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

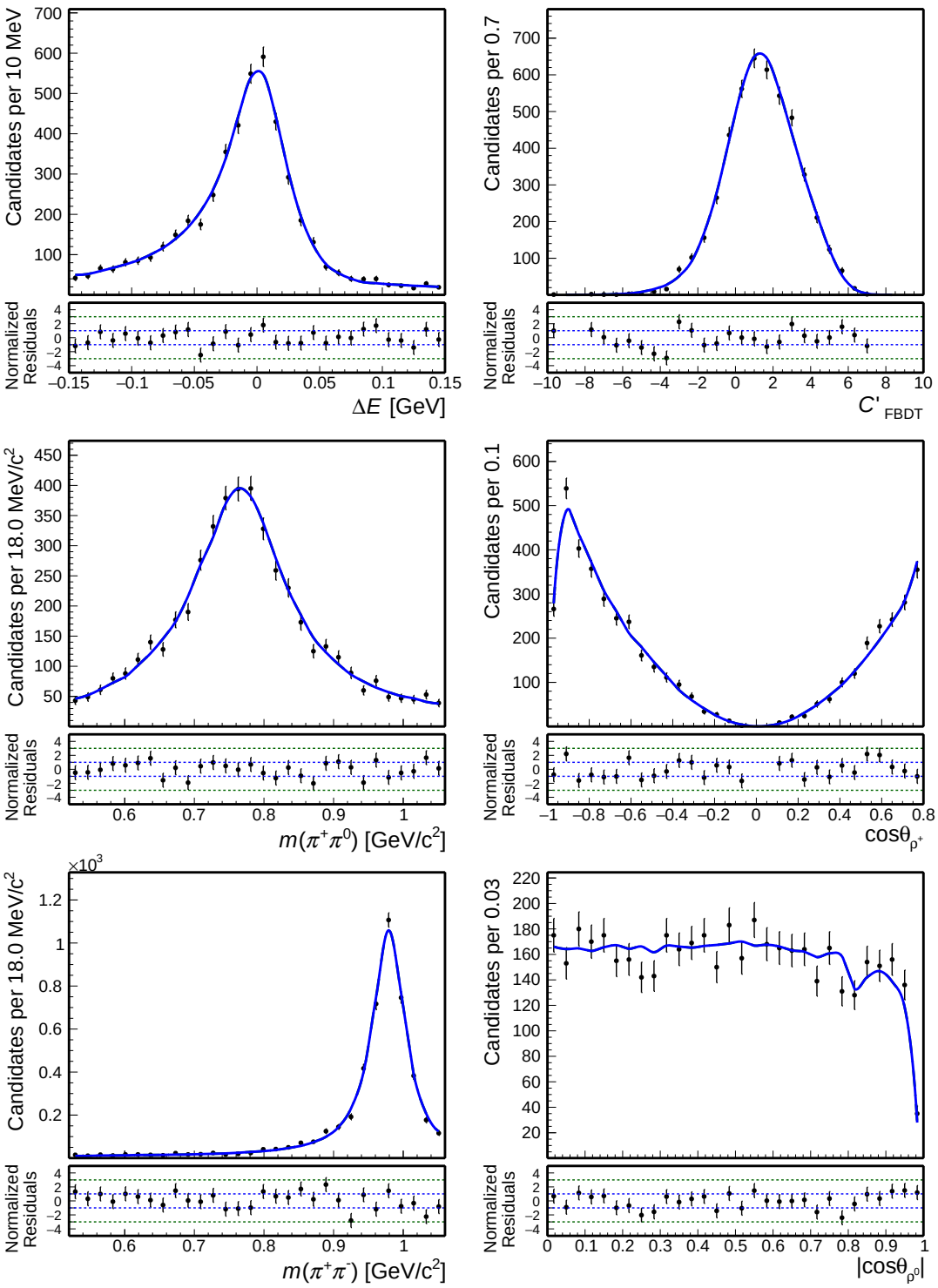


Figure A.13: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated $B^+ \to f_0\rho^+$ decays. The corresponding model shapes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

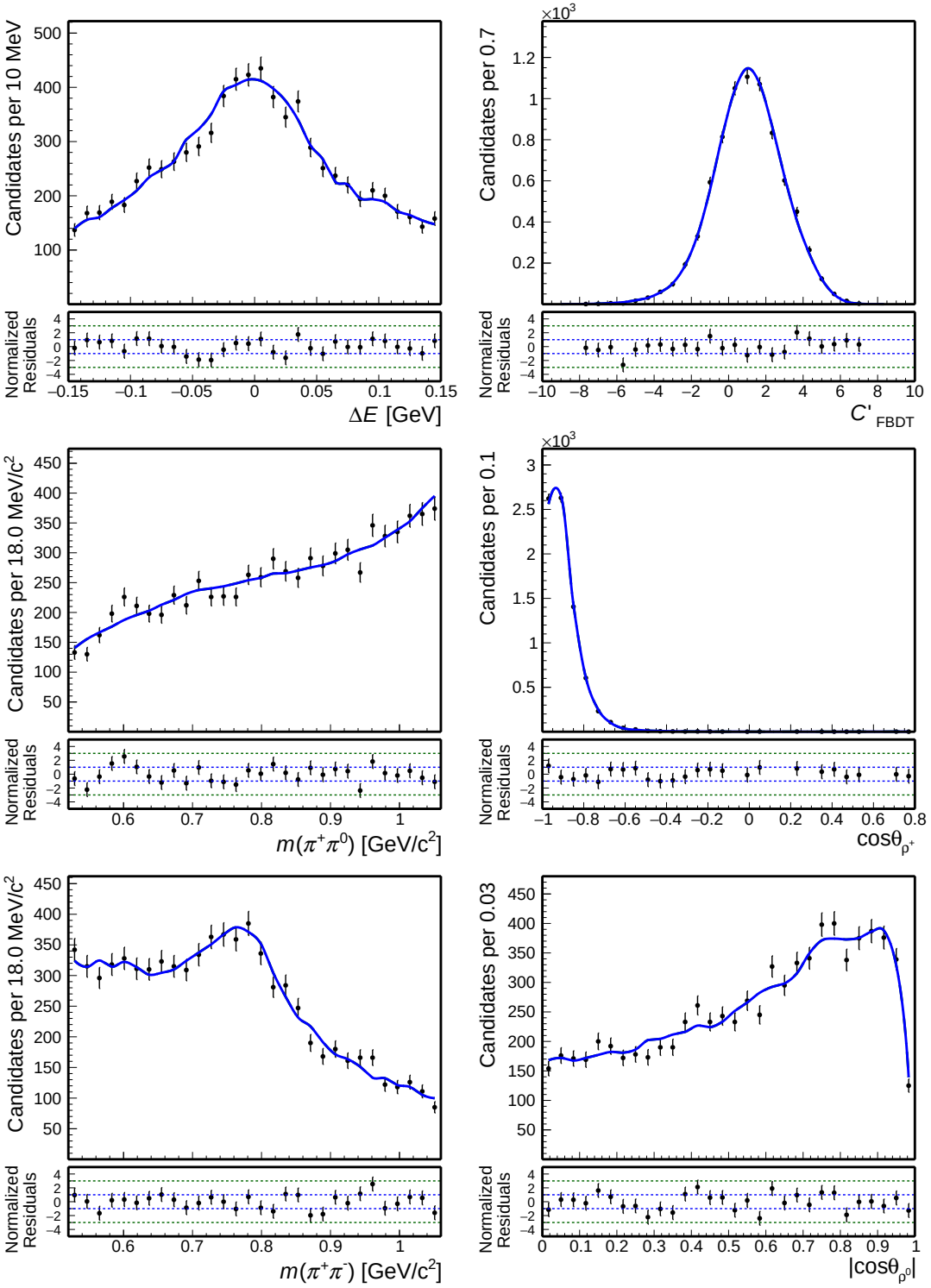


Figure A.14: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated $B^+ \to a_1^+\pi^0$ decays. The corresponding model shapes are overlaid (solid).

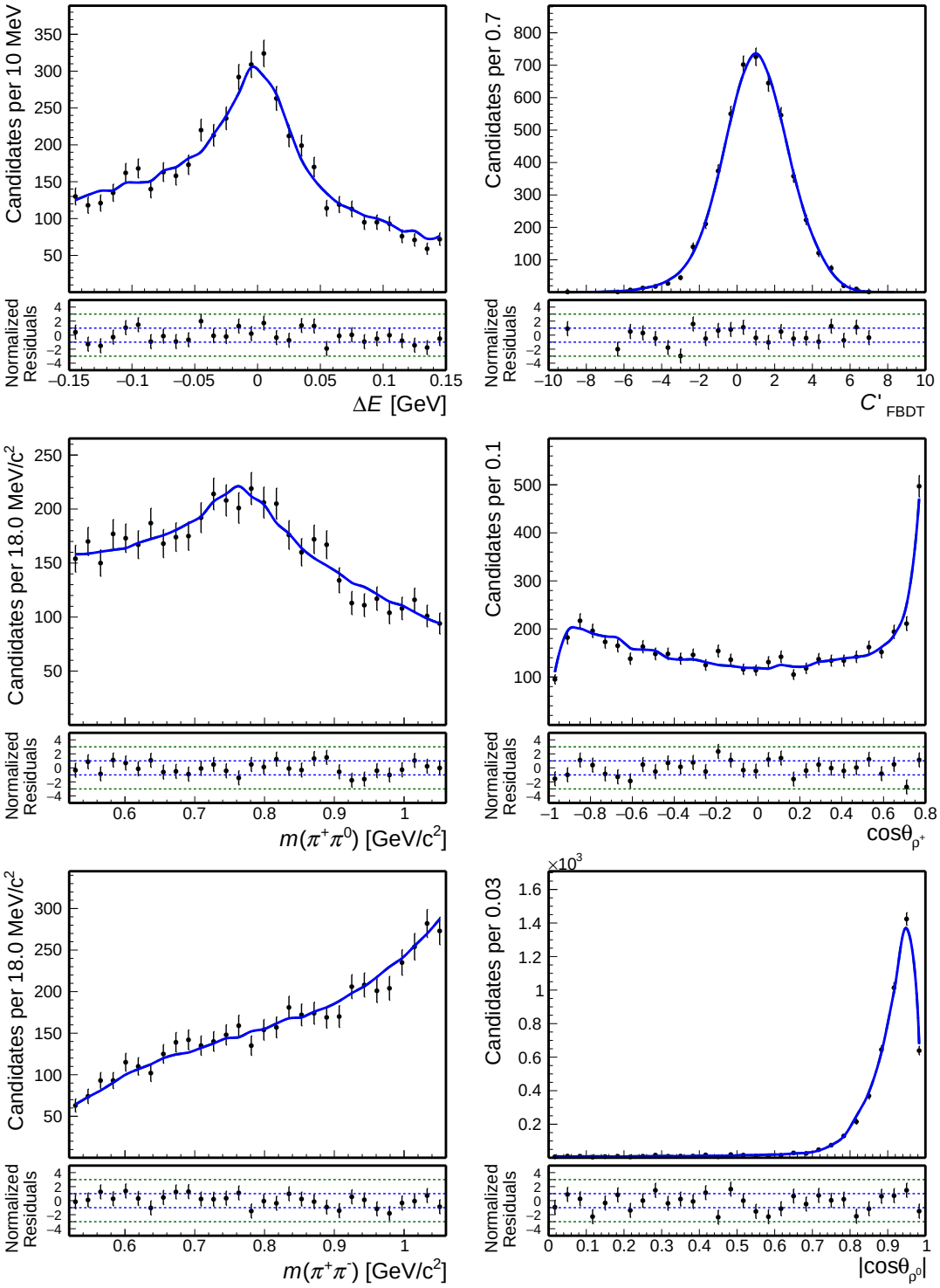


Figure A.15: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated $B^+ \to a_1^+\pi^0$ decays. The corresponding model shapes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

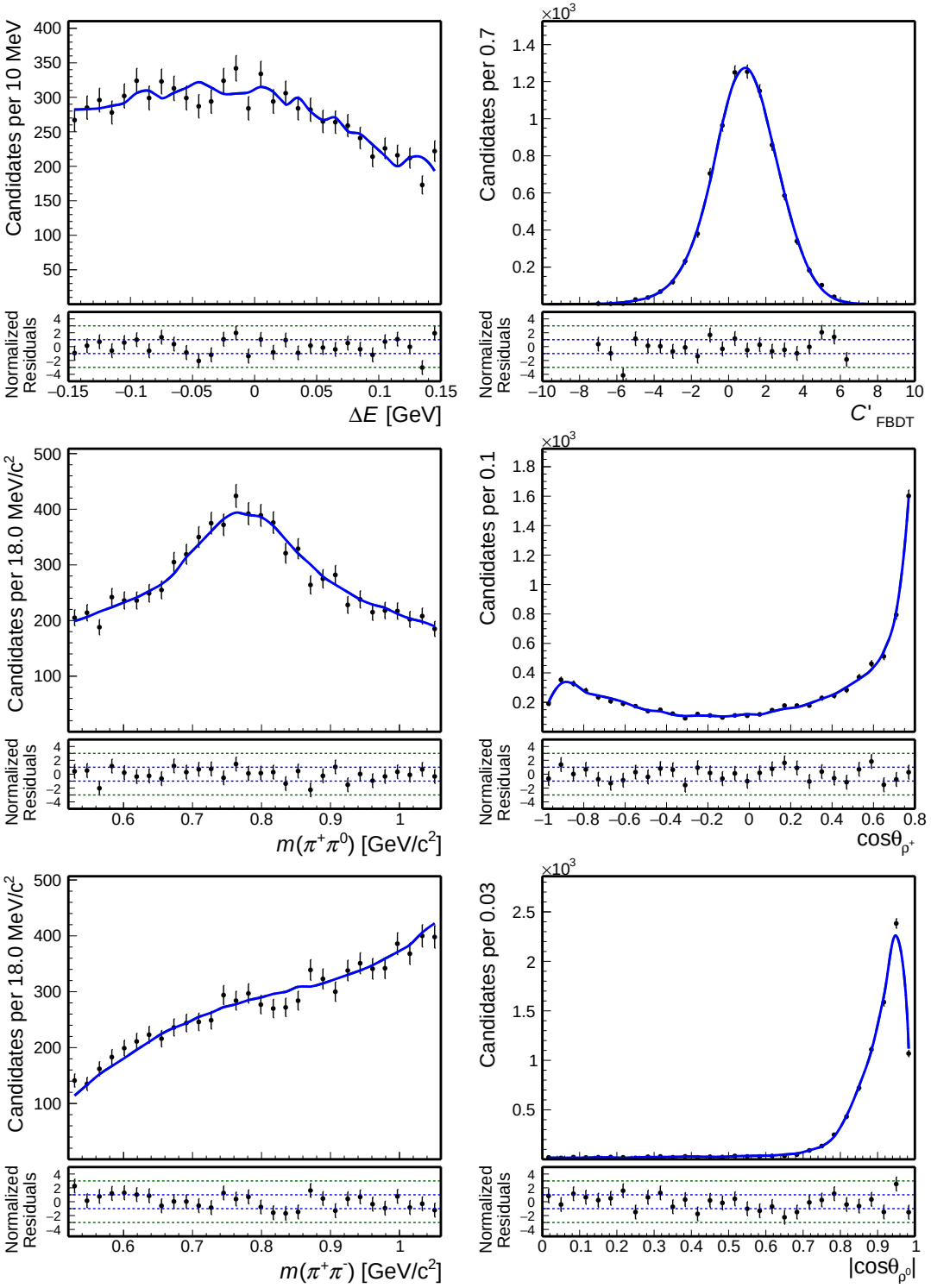


Figure A.16: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated $B^0 \to a_1^\pm\pi^\mp$ decays. The corresponding model shapes are overlaid (solid).

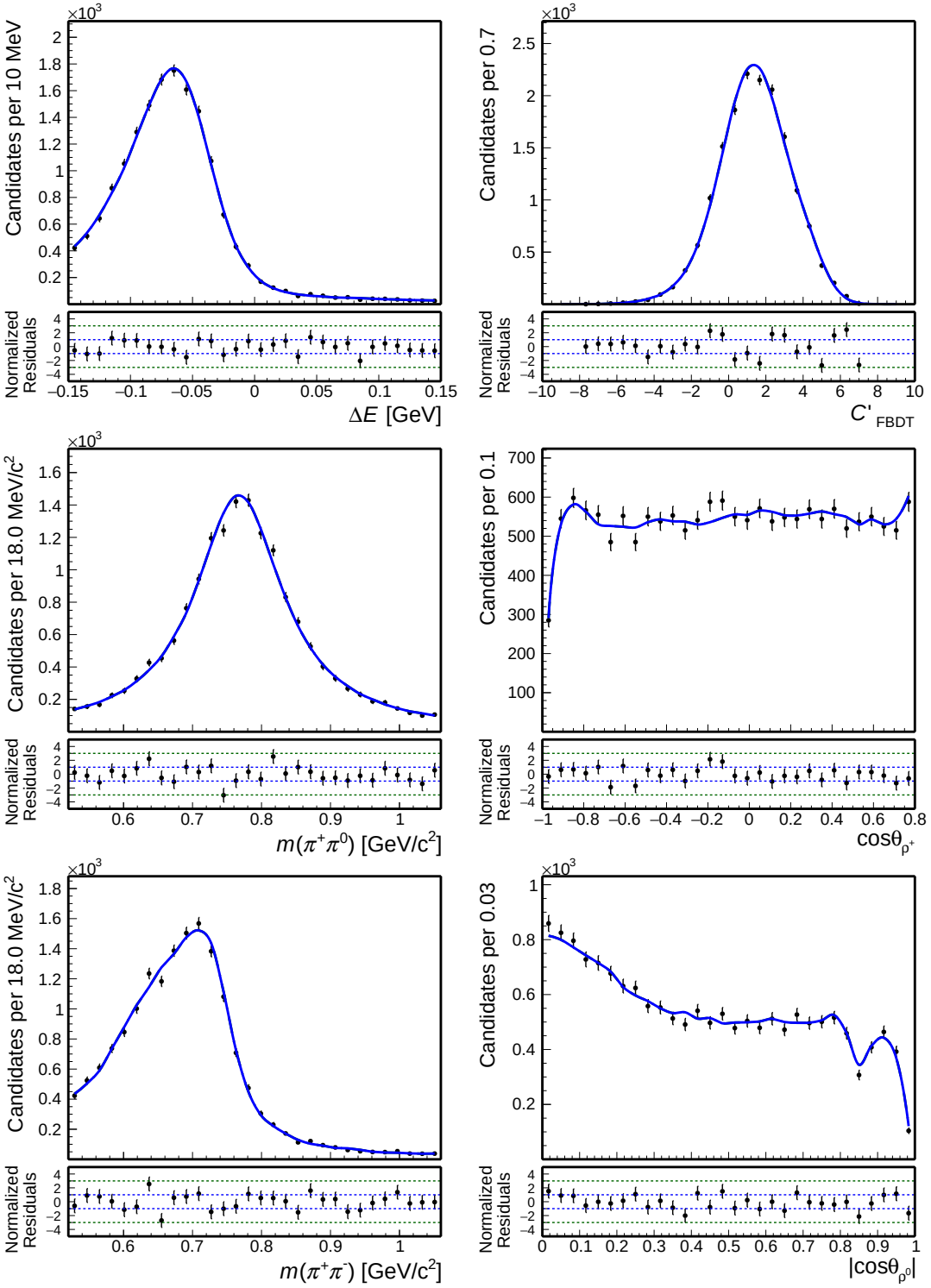


Figure A.17: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated $B^+ \to K^{*0}\rho^+$ decays. The corresponding model shapes are overlaid (solid).

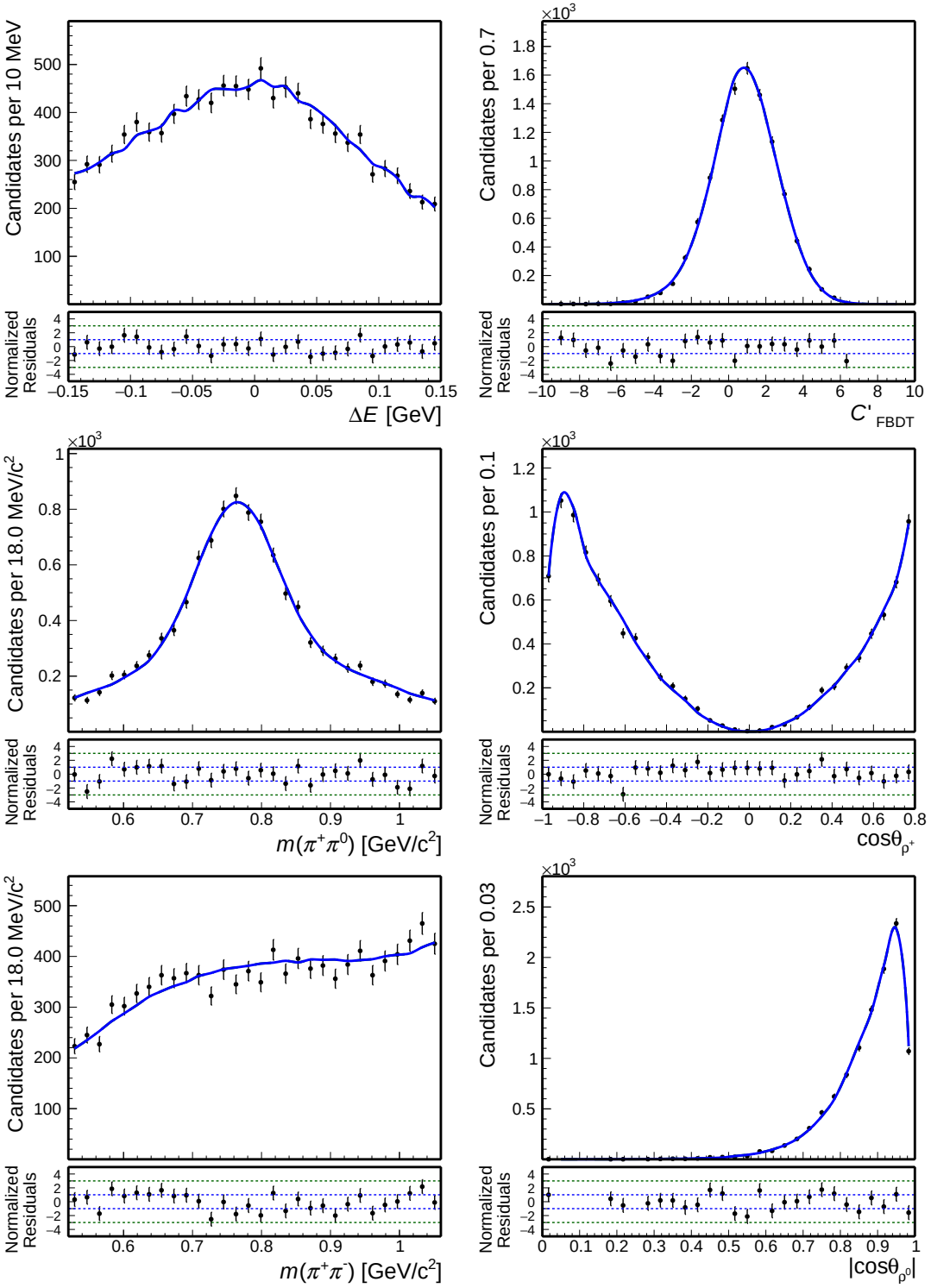


Figure A.18: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated $B^0 \to \rho^+\rho^-$ decays. The corresponding model shapes are overlaid (solid).

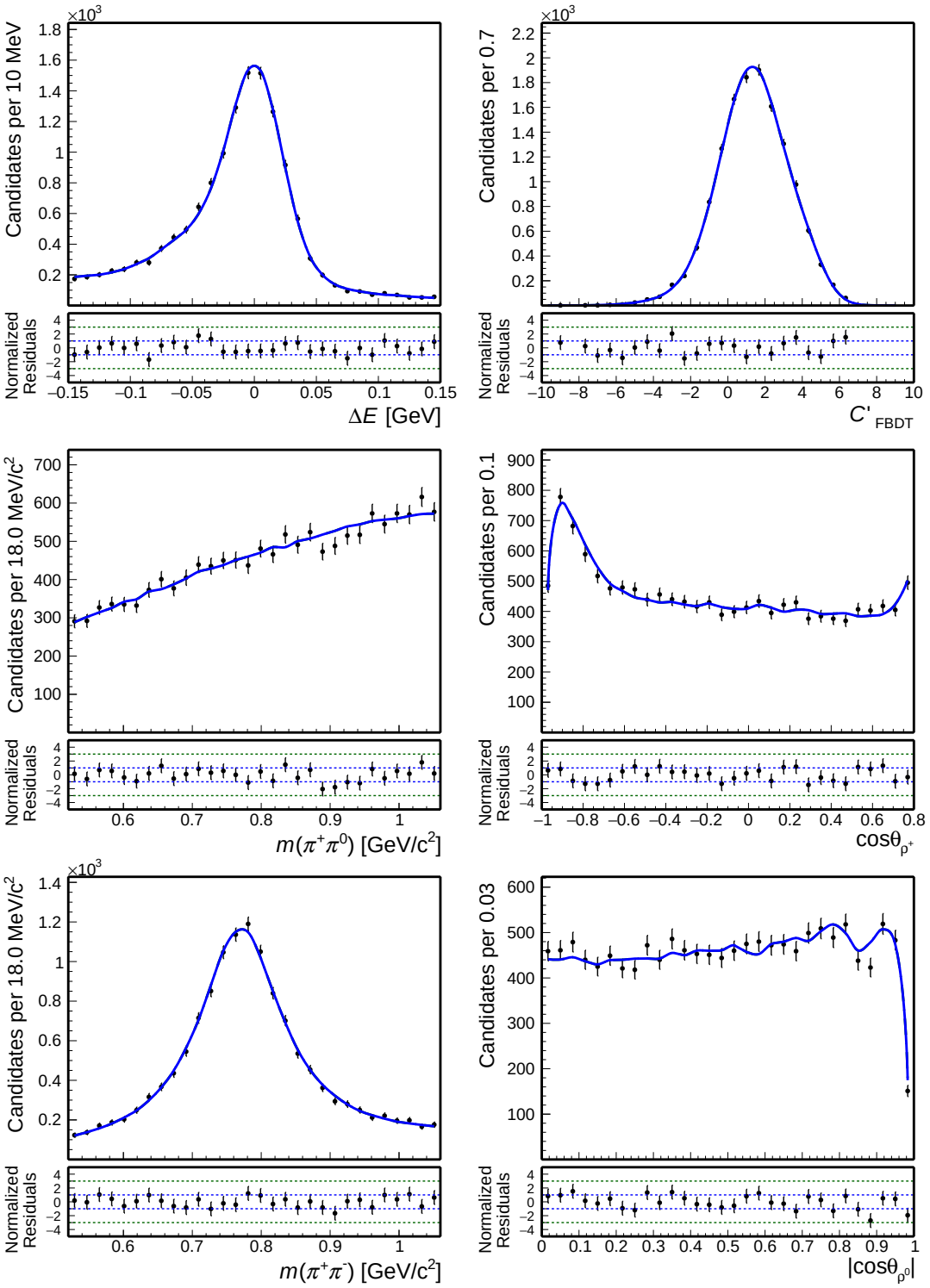


Figure A.19: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated $B^+ \to \rho^0\pi^+\pi^0$ decays. The corresponding model shapes are overlaid (solid).

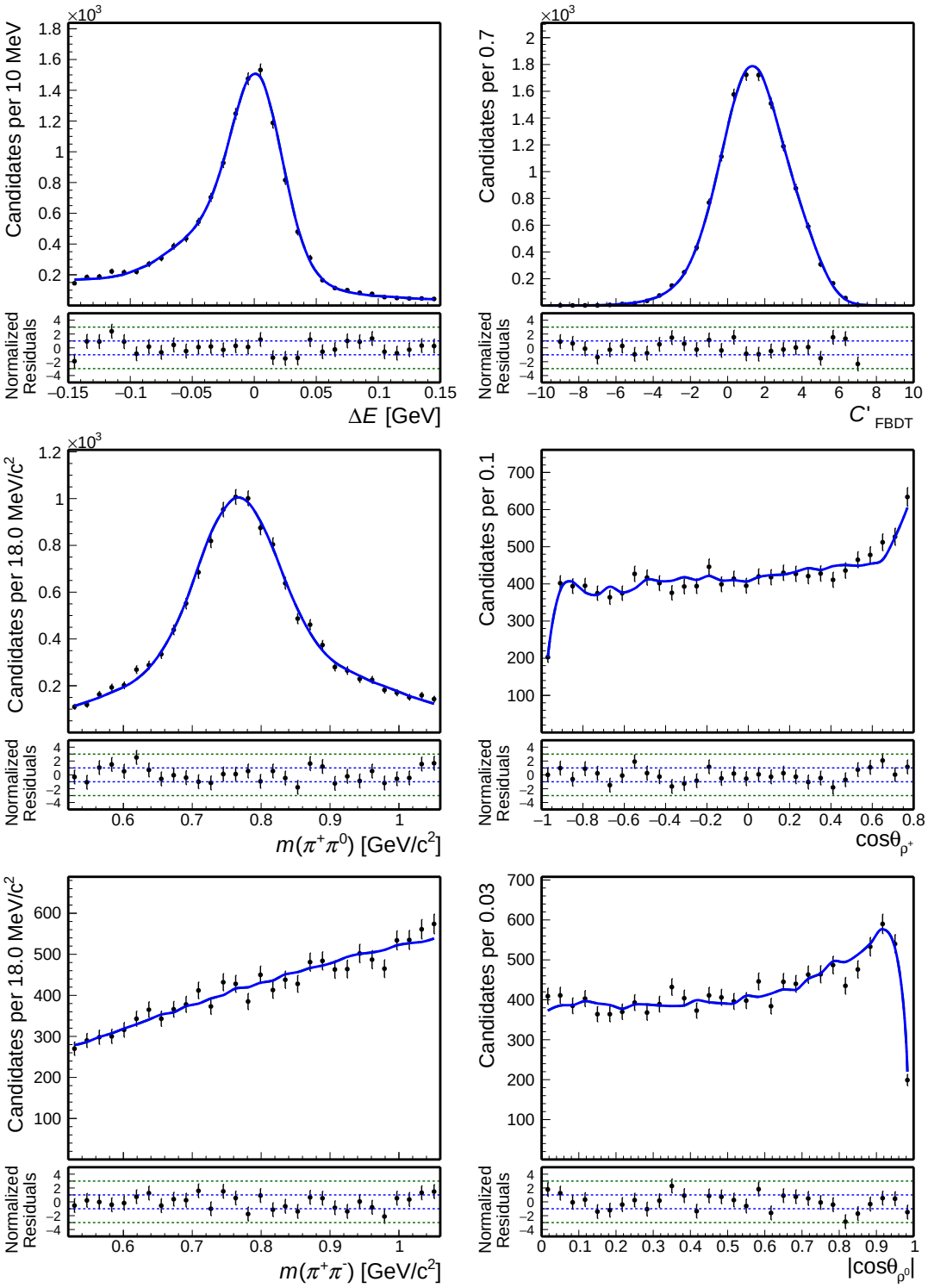


Figure A.20: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated $B^+ \to \rho^+\pi^+\pi^-$ decays. The corresponding model shapes are overlaid (solid).

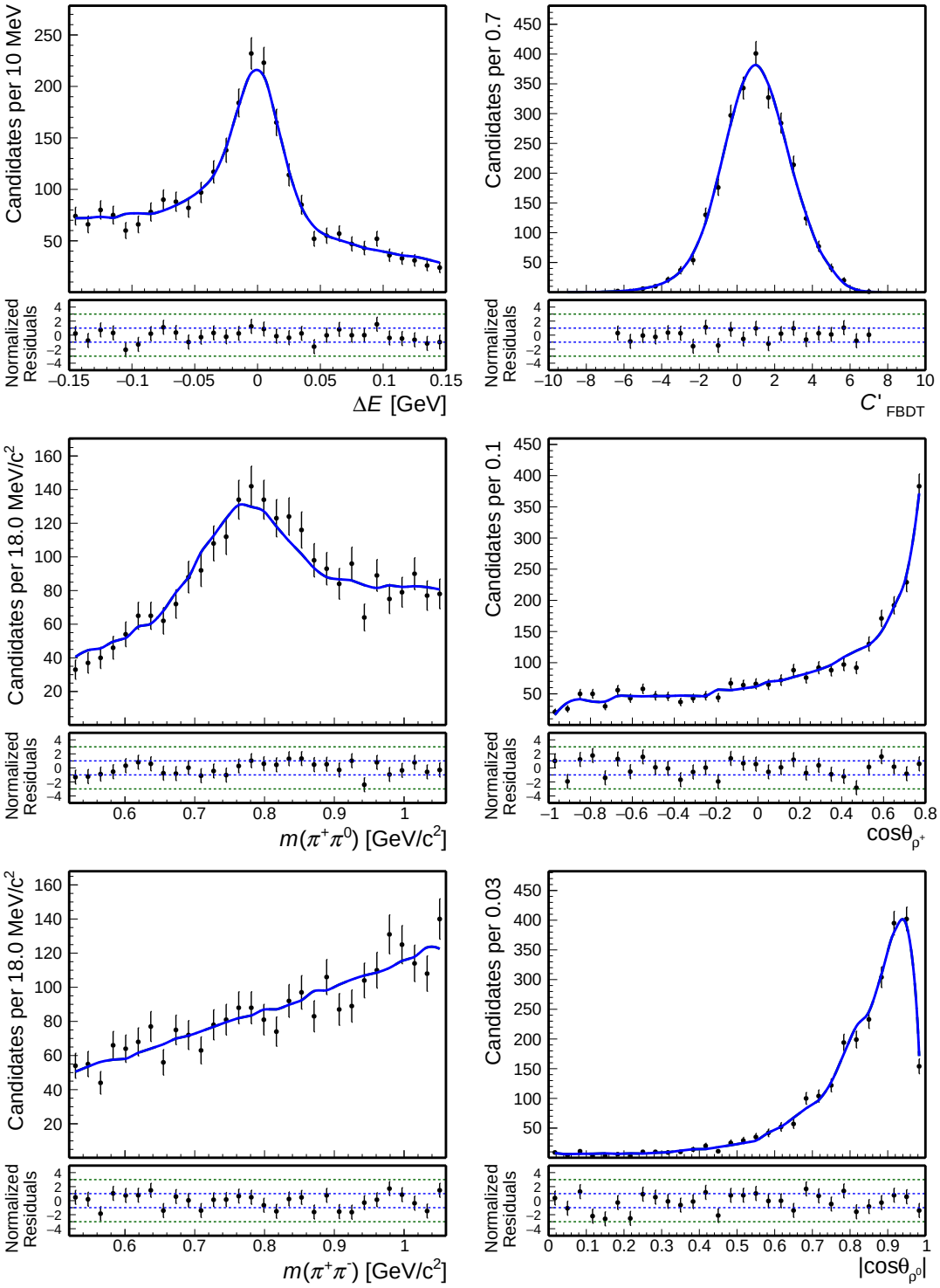


Figure A.21: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated $B^+ \to \rho^-\pi^+\pi^+$ decays. The corresponding model shapes are overlaid (solid).

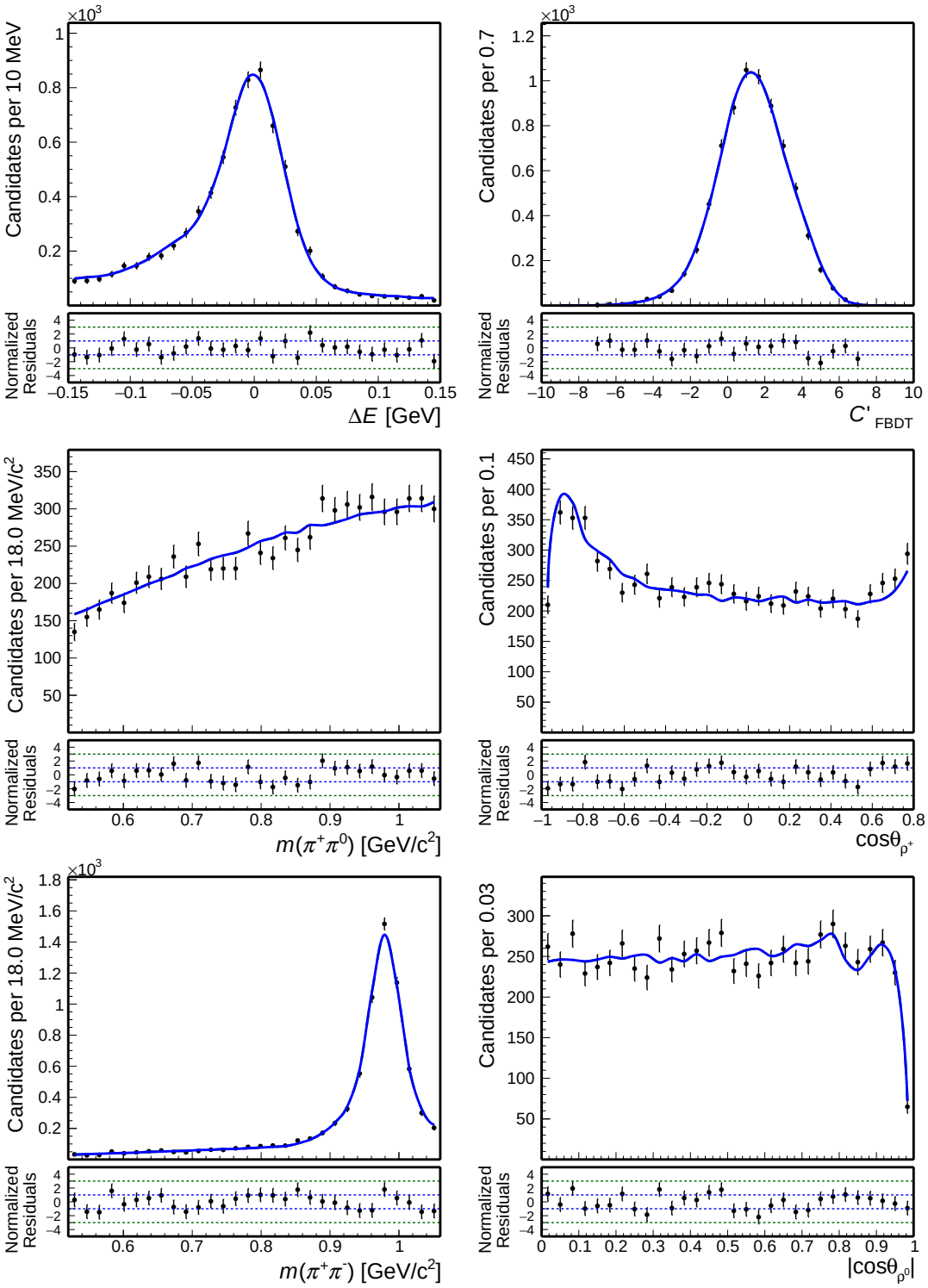


Figure A.22: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated $B^+ \to f_0\pi^+\pi^0$ decays. The corresponding model shapes are overlaid (solid).

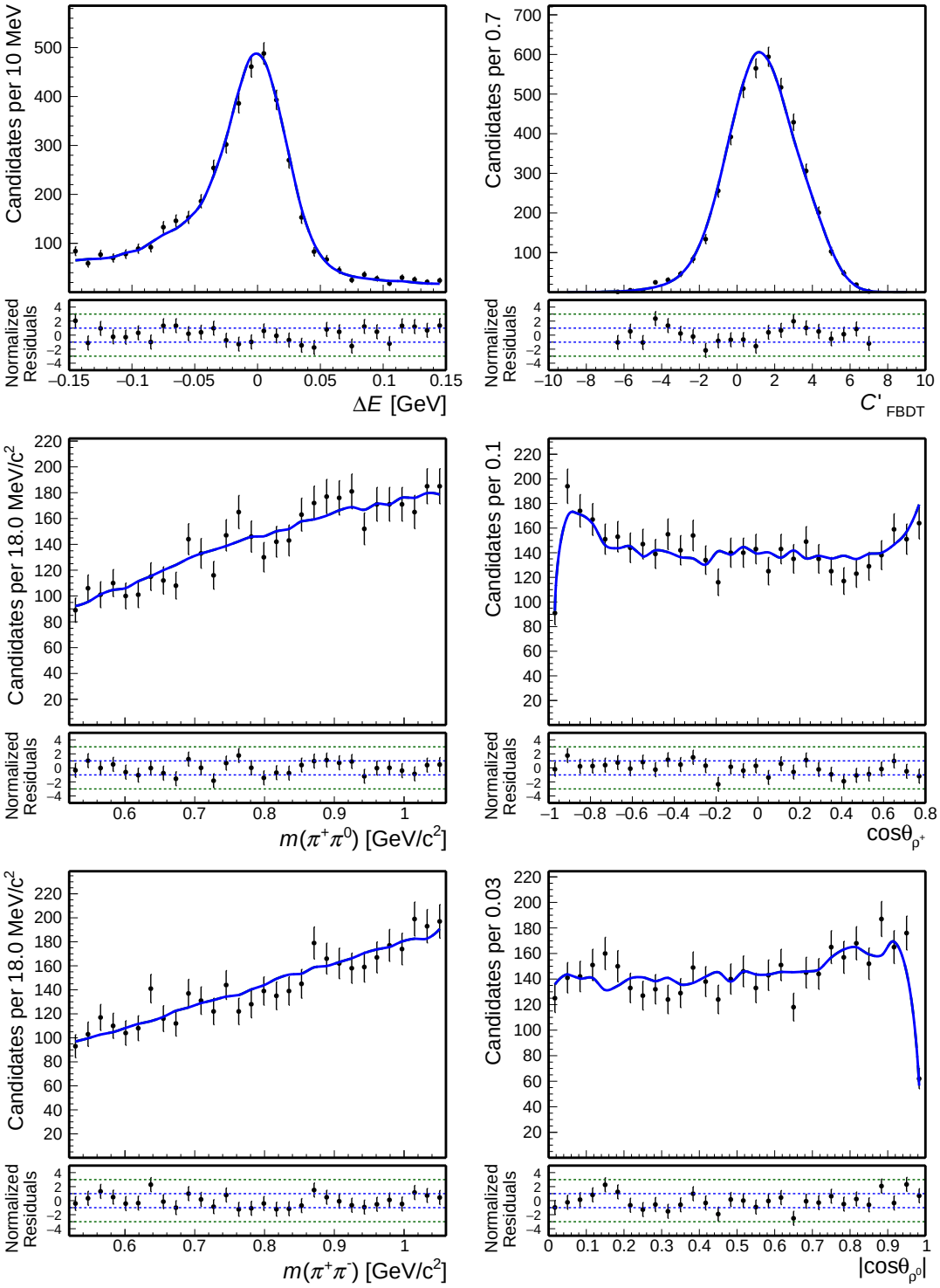


Figure A.23: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated $B^+ \to \pi^+\pi^-\pi^+\pi^0$ decays. The corresponding model shapes are overlaid (solid).

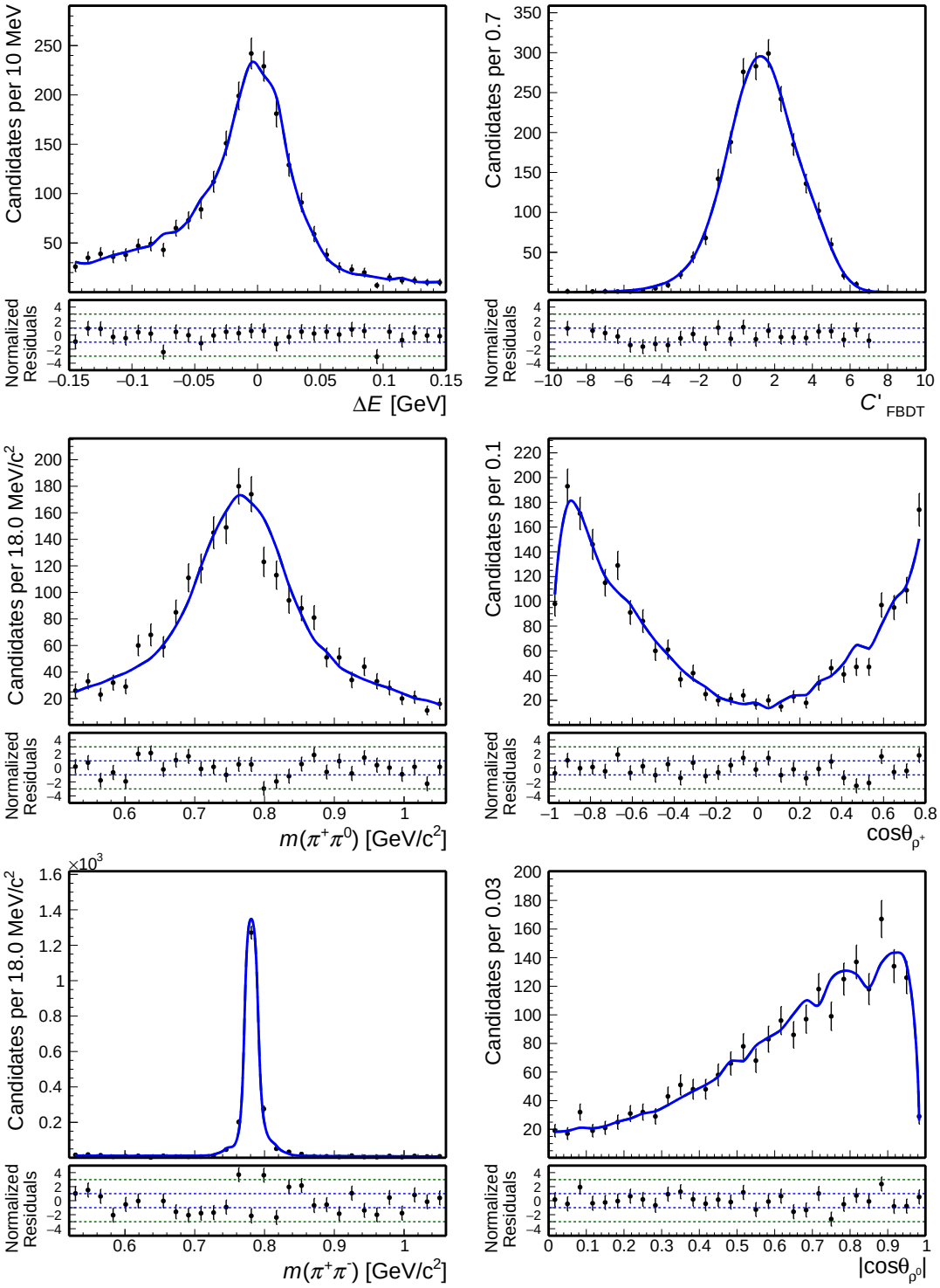


Figure A.24: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated $B^+ \to \omega\rho^+$ decays. The corresponding model shapes are overlaid (solid).

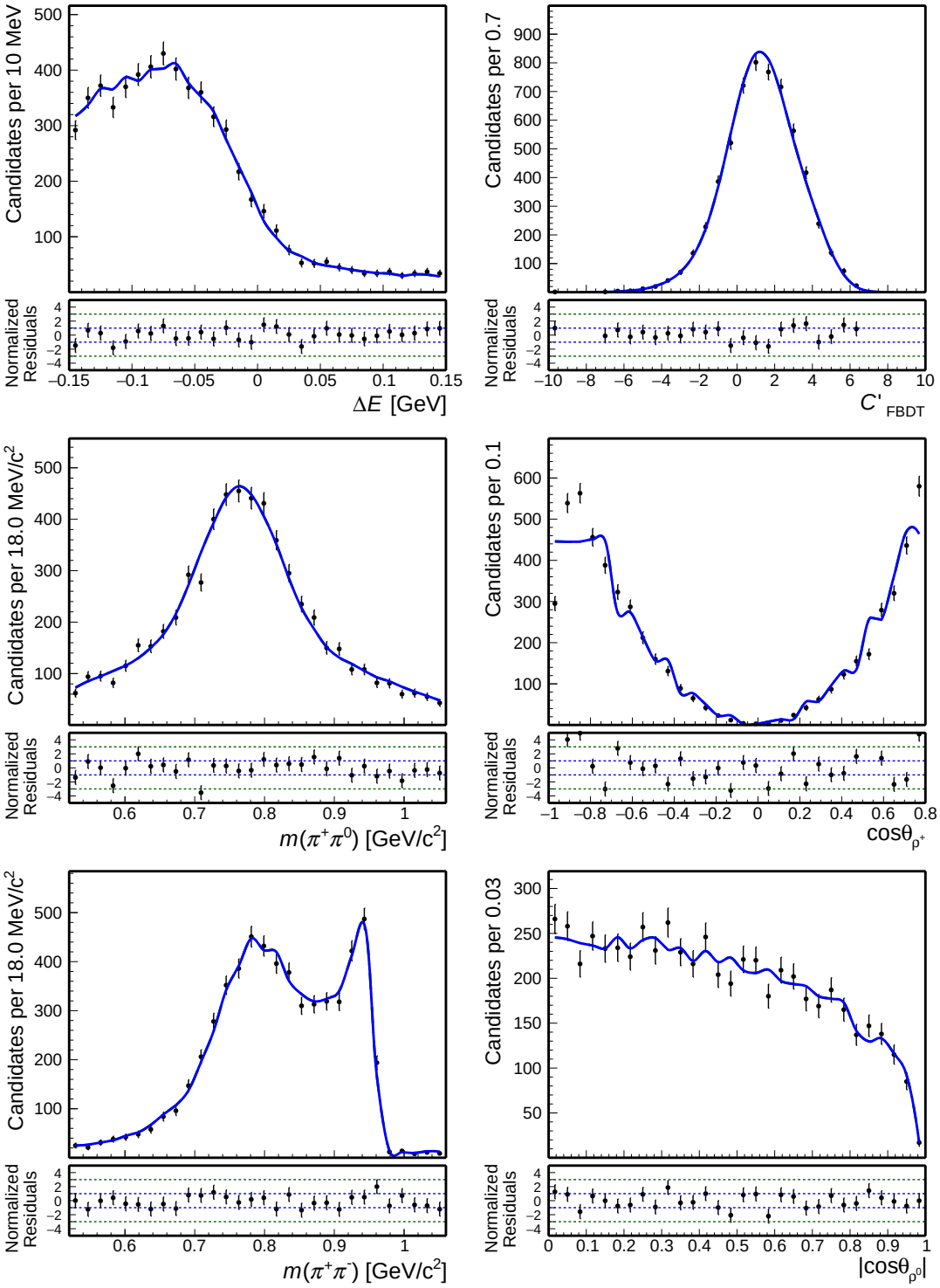


Figure A.25: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated $B^+ \to \eta'\rho^+$ decays. The corresponding model shapes are overlaid (solid).

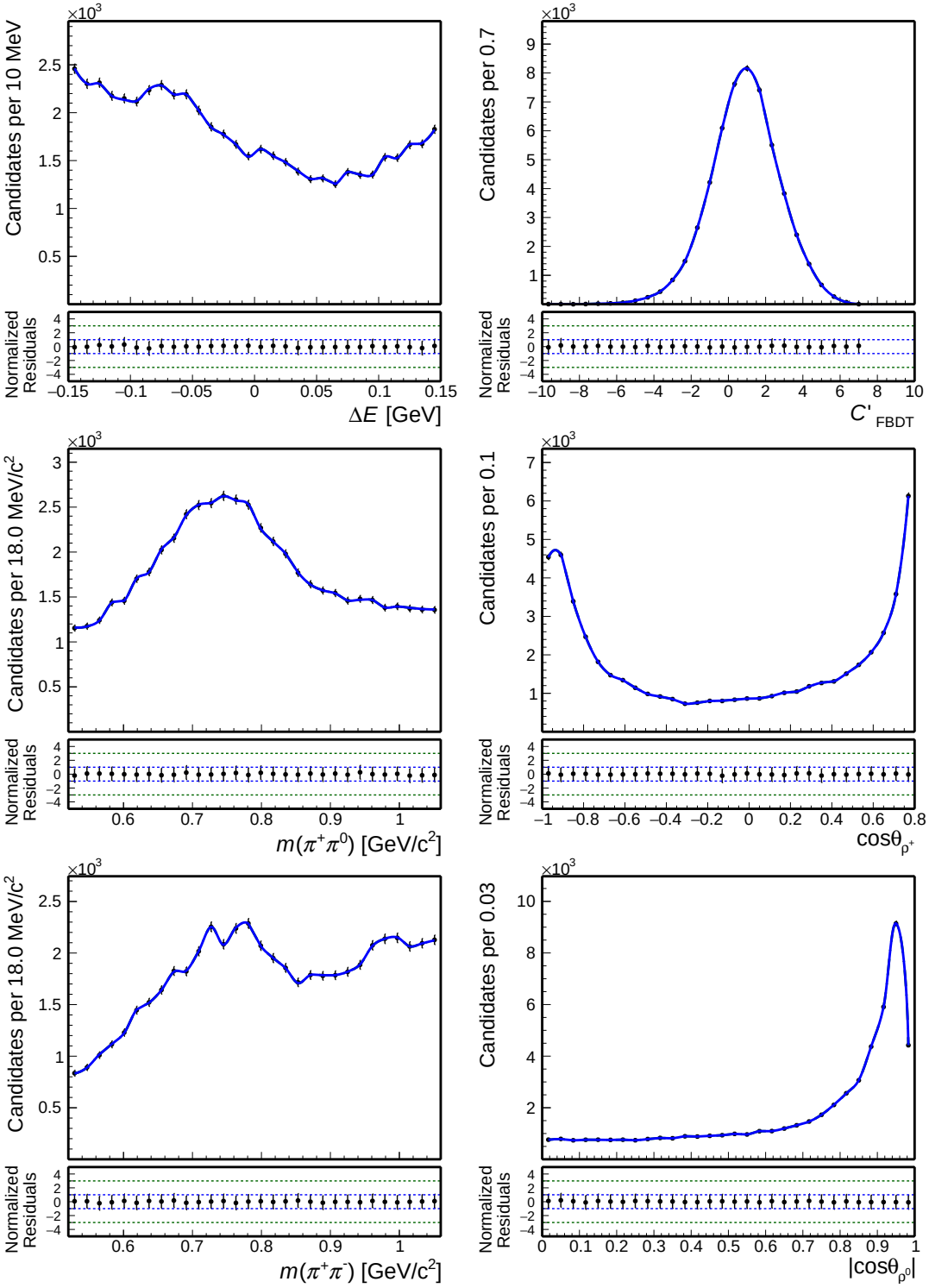


Figure A.26: Distributions of (top left) ΔE, (top right) continuum suppression output, (middle left) $m(\pi^+\pi^0)$, (middle right) $\cos\theta_{\rho^+}$, (bottom left) $m(\pi^+\pi^-)$, and (bottom right) $|\cos\theta_{\rho^0}|$ for simulated combinatorial rare B background. The corresponding model 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 well-known abundant control samples that share relevant features with my signal sample, $B^+ \to \overline{D}^0(\to K^+\pi^-\pi^0)\pi^+$, $B^+ \to \overline{D}^0(\to K^0_S\pi^0)\pi^+$, and $B^+ \to \overline{D}^0(\to K^+\pi^-)\rho^+(\to \pi^+\pi^0)$ decays. By performing a fit, I determine their sample composition in the Belle data. The fit shapes are extracted from simulation. Figures B.1-B.10 show modeling fit projections.

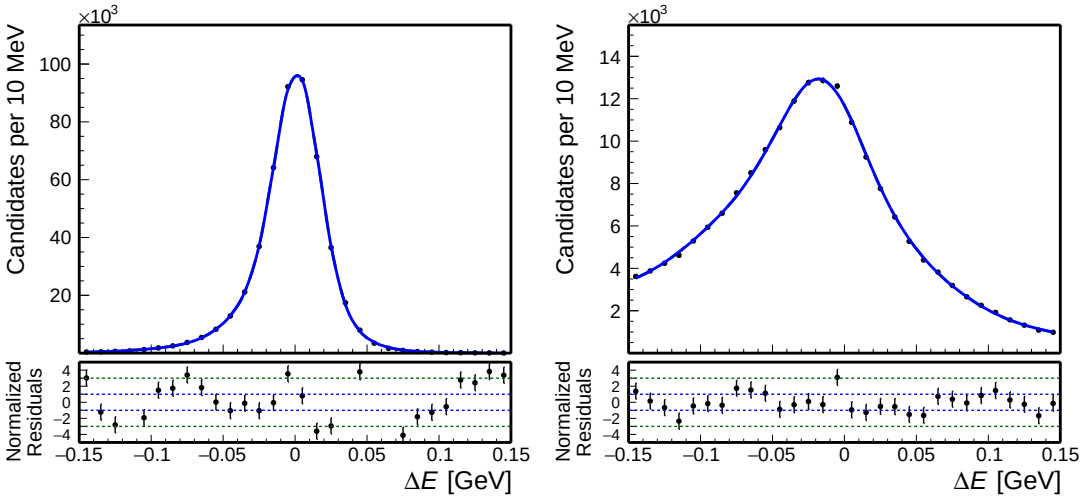


Figure B.1: Distributions of ΔE for simulated $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ (left) signal and (right) self cross-feed events. Modeling fit projections are overlaid.

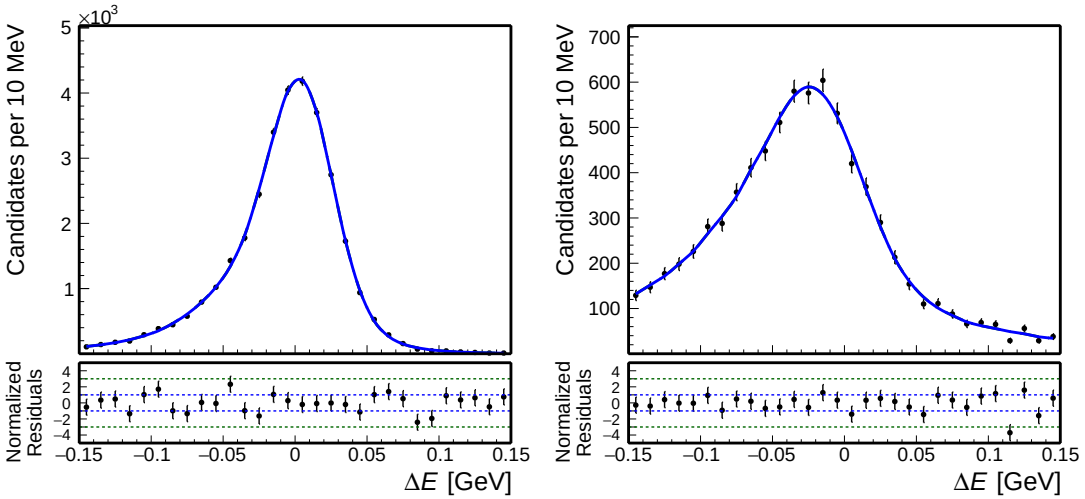


Figure B.2: Distributions of ΔE for simulated $B^+ \to \bar{D}^0(\to K^0_S\pi^0)\pi^+$ (left) signal and (right) self cross-feed. Modeling fit projections are overlaid.

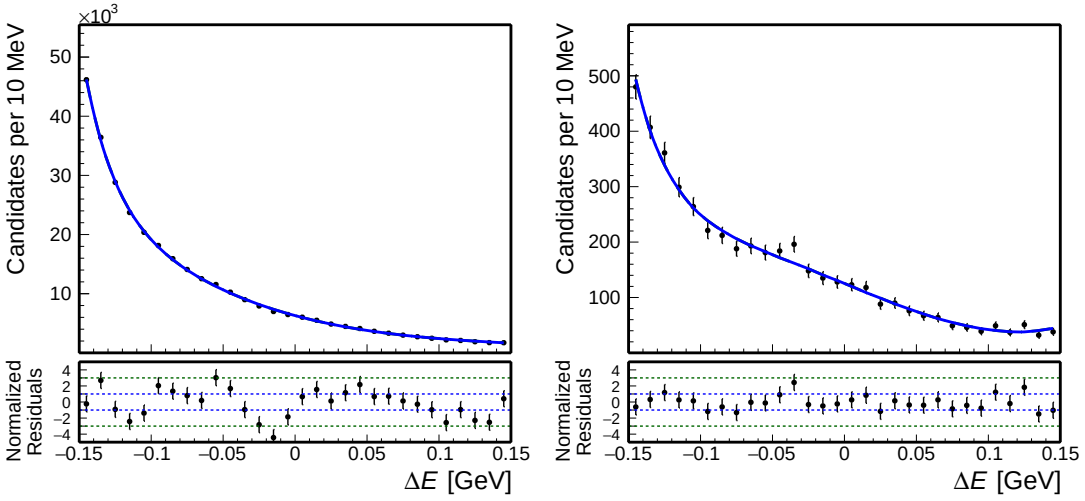


Figure B.3: Distributions of ΔE for $B\bar{B}$ background in simulated (left) $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ and (right) $B^+ \to \bar{D}^0(\to K^0_S\pi^0)\pi^+$ data. Modeling fit projections are overlaid.

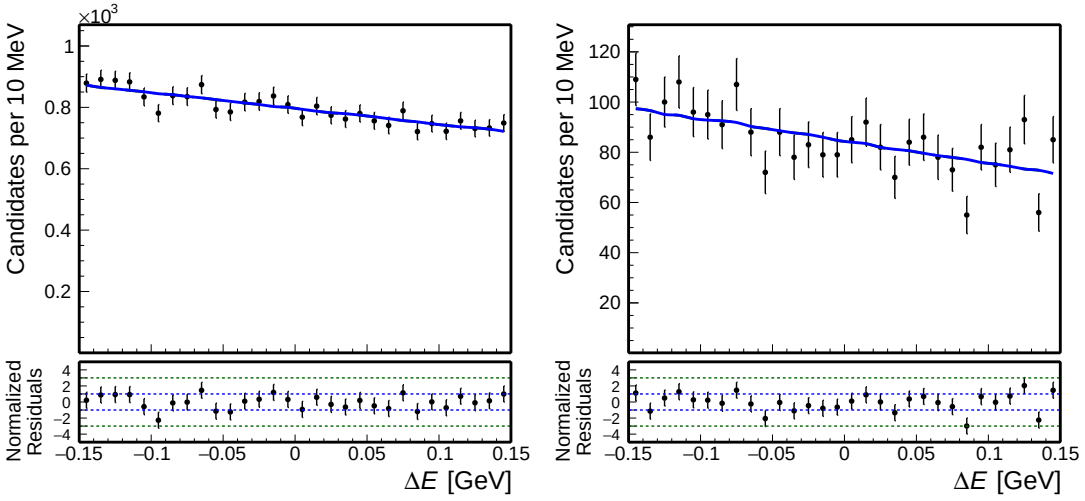


Figure B.4: Distributions of ΔE for continuum in simulated $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ (left) and $B^+ \to \bar{D}^0(\to K^0_S\pi^0)\pi^+$ (right) data. Modeling fit projections are overlaid.

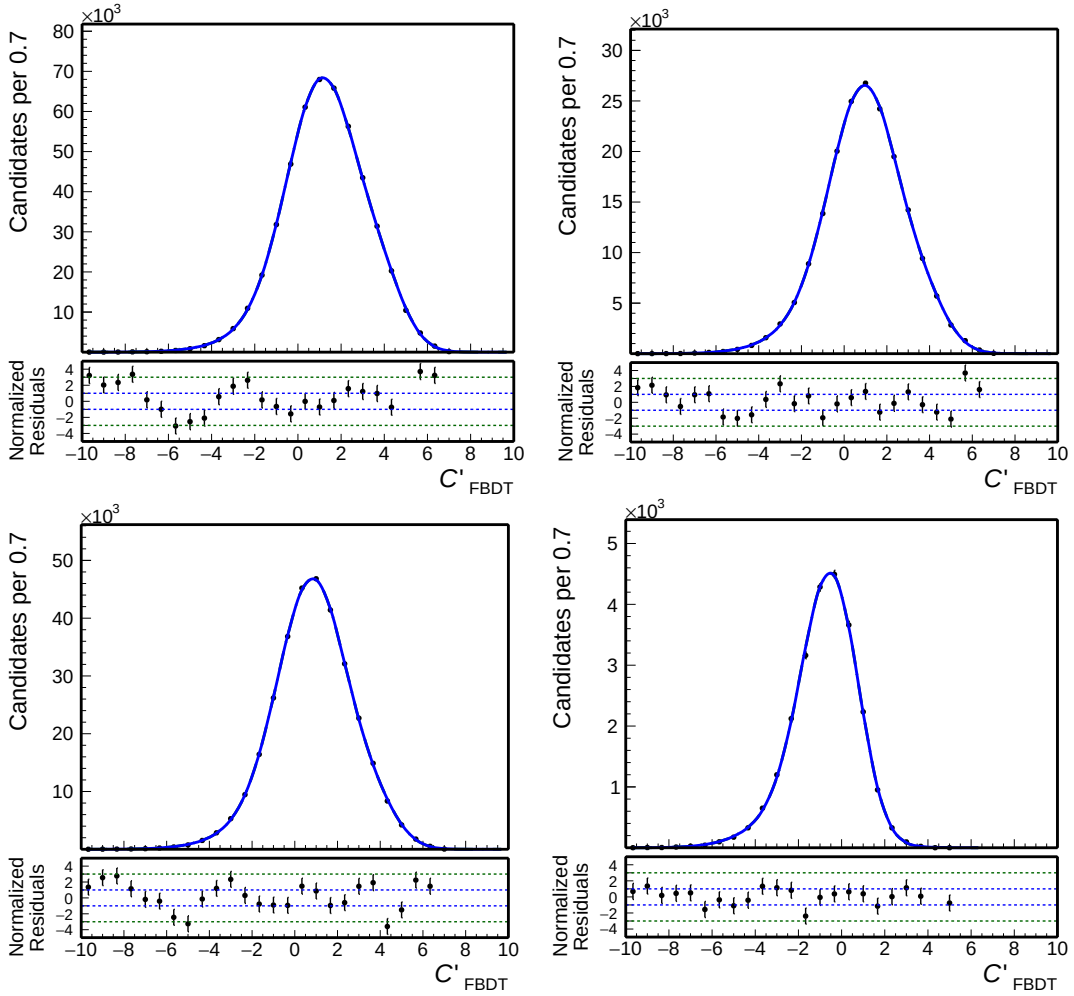


Figure B.5: Distributions of C' for simulated (top left) signal, (top right) self cross-feed, (bottom left) $B\bar{B}$ background, and (bottom right) continuum $B^+ \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^+$ events. Modeling fit projections are overlaid.

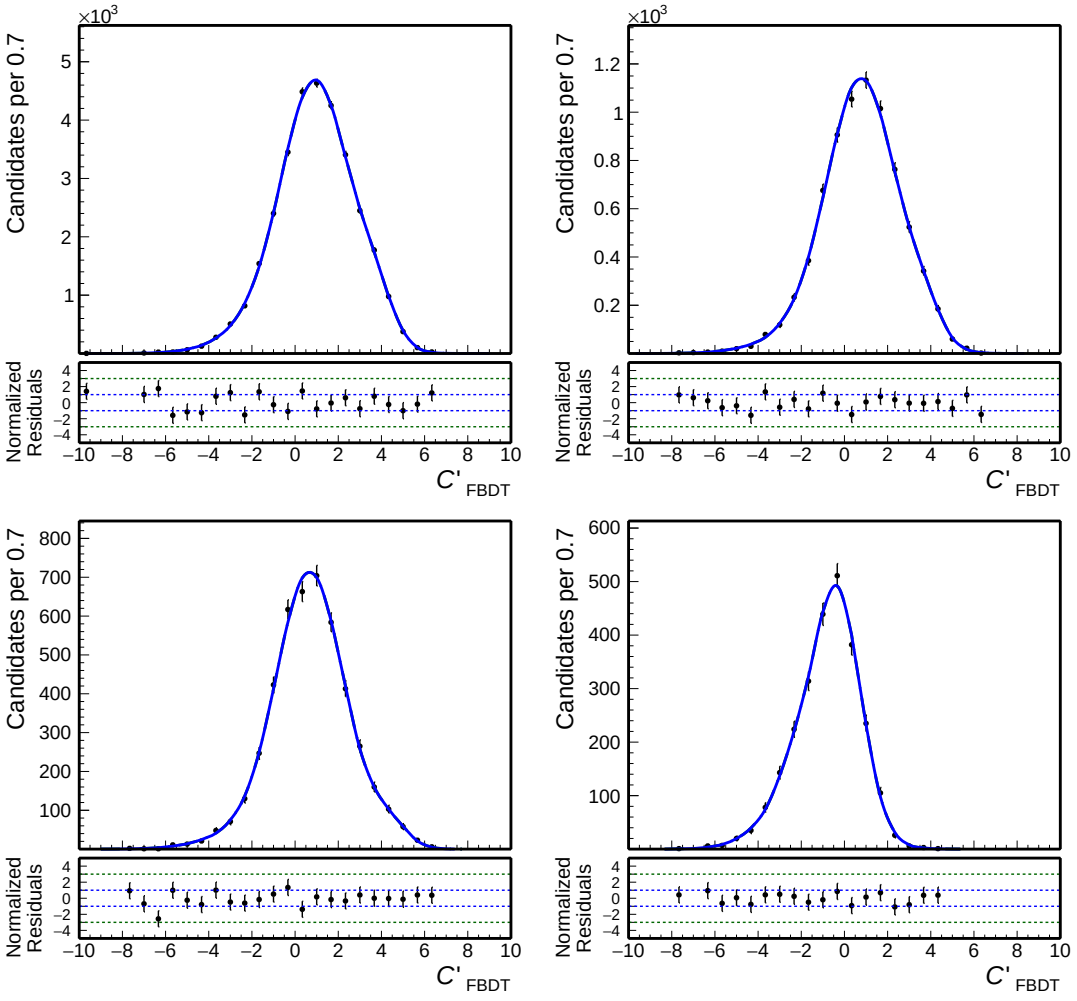


Figure B.6: Distributions of C' for simulated (top left) signal, (top right) self cross-feed, (bottom left) $B\bar{B}$ background, and (bottom right) continuum $B^+ \to \bar{D}^0(\to K^0_S\pi^0)\pi^+$ events. Modeling fit projections are overlaid.

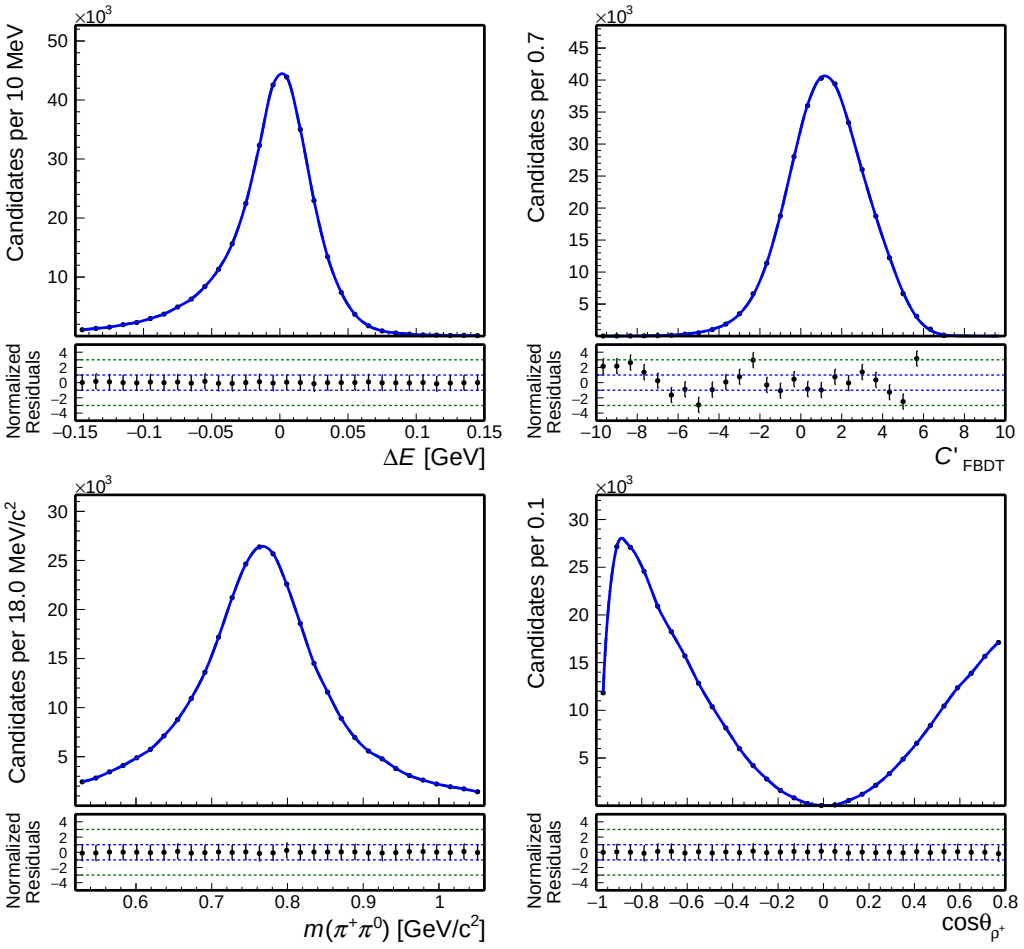


Figure B.7: Distributions of (top left) ΔE, (top right) C', (bottom left) $m(\pi^+\pi^0)$, and (bottom right) $\cos\theta_{\rho^+}$) for simulated signal $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ decays. Fit projections are overlaid (blue solid line).

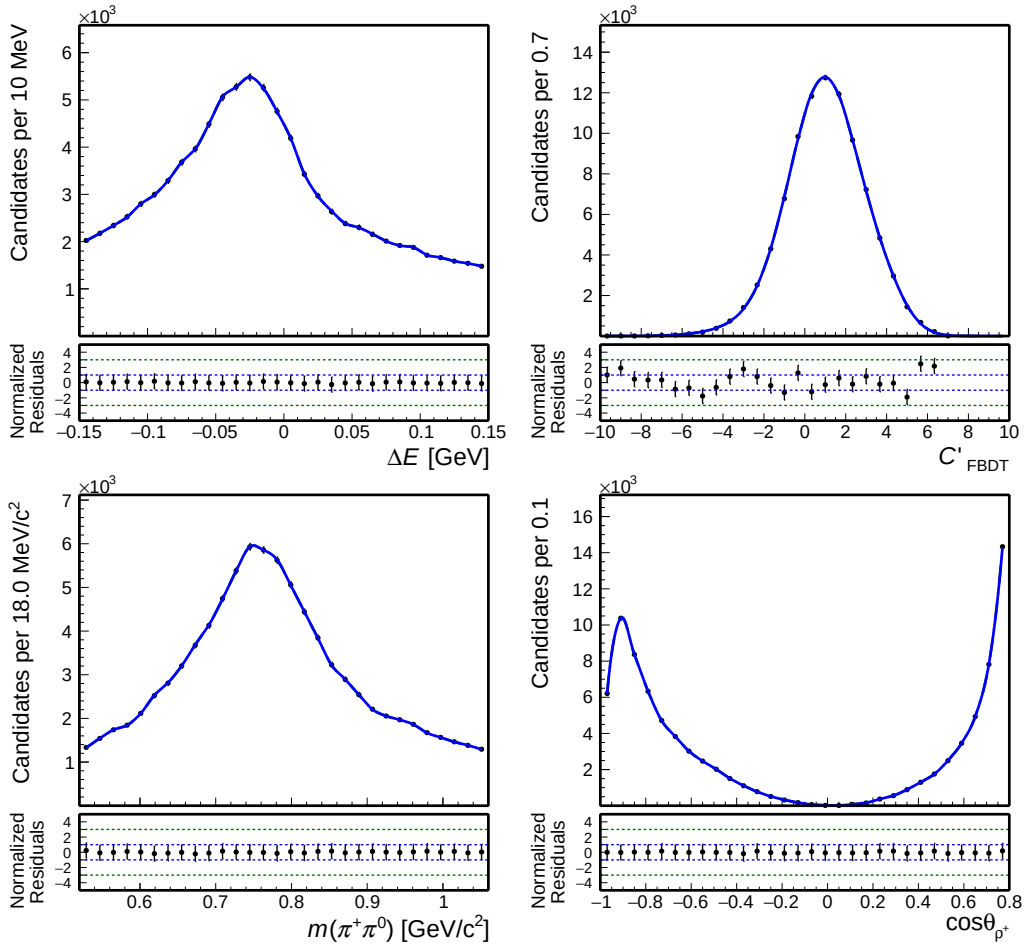


Figure B.8: Distributions of (top left) ΔE, (top right) C', (bottom left) $m(\pi^+\pi^0)$, and (bottom right) $\cos\theta_{\rho^+}$) for simulated self cross-feed $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ candidates. Fit projections are overlaid (blue solid line).

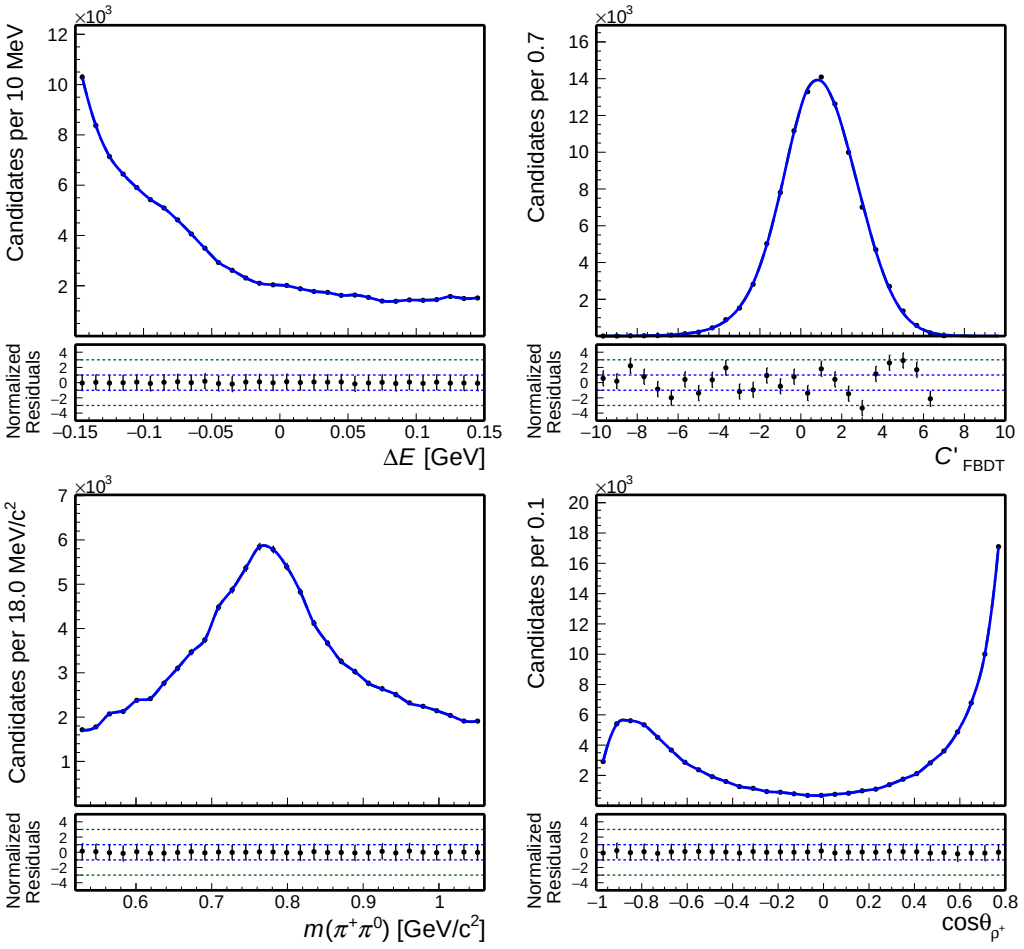


Figure B.9: Distributions of (top left) ΔE, (top right) C', (bottom left) $m(\pi^+\pi^0)$, and (bottom right) $\cos\theta_{\rho^+}$) for $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ candidates reconstructed in simulated $B\overline{B}$ background sample. Fit projections are overlaid (blue solid line).

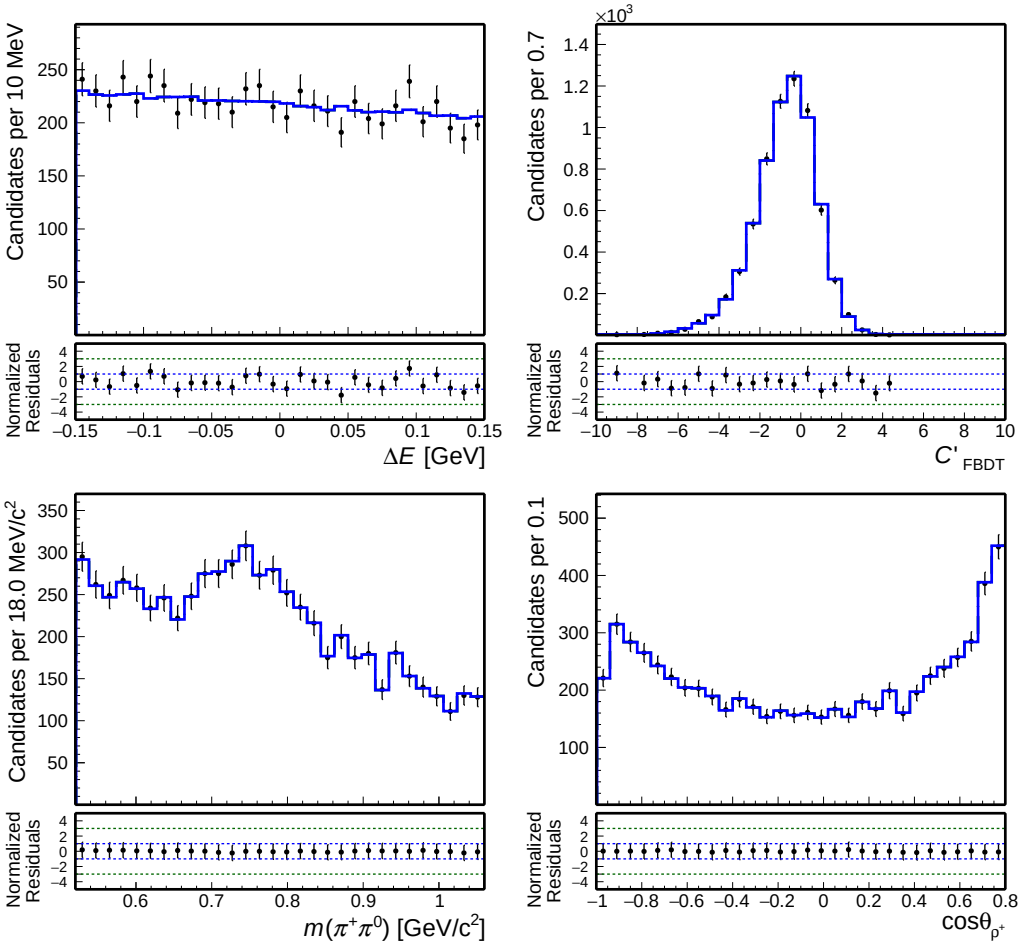


Figure B.10: Distributions of (top left) ΔE, (top right) C', (bottom left) $m(\pi^+\pi^0)$, and (bottom right) $\cos\theta_{\rho^+}$) for $B^+ \to \bar{D}^0(\to K^+\pi^-)\rho^+$ candidates reconstructed in simulated continuum sample. Fit projections are overlaid (blue solid line).

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] *Overview of the KEKB accelerators*, Nucl. Instrum. Methods Phys. Res. A **499** (2003) no. 1, 1 – 7. KEK-B: The KEK B-factory.

[44] A. Tetsuo *et al.*, *Achievements of KEKB*, Progress of Theoretical and Experimental Physics **2013** (2013) no. 3, . https://doi.org/10.1093/ptep/pts102. 03A001.

[45] 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.

[46] A. Abashian *et al.*, *The Belle Detector*, Nucl. Instrum. Methods Phys. Res. A **479** (2002) 117–232.

[47] A. Bevan *et al.*, BaBar, Belle, *The Physics of the B Factories*, Eur. Phys. J. C **74** (2014) 3026, arXiv:1406.6311 [hep-ex].

[48] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444–450.

[49] Z. Natkaniec *et al.*, *Status of the Belle silicon vertex detector*, Nucl. Instrum. Meth. A **560** (2006) 1–4.

[50] 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.

[51] K. Miyabayashi, *Belle electromagnetic calorimeter*, Nucl. Instrum. Meth. A **494** (2002) 298–302.

[52] H. Kichimi *et al.*, *The BELLE TOF system*, Nucl. Instrum. Meth. A **453** (2000) 315–320.

[53] T. Iijima *et al.*, *Aerogel Cherenkov counter for the BELLE detector*, Nucl. Instrum. Meth. A **453** (2000) 321–325.

[54] J. Brodzicka *et al.*, *Physics achievements from the Belle experiment*, Progress of Theoretical and Experimental Physics **2012** (12, 2012) . https://doi.org/10.1093/ptep/pts072. 04D001.

[55] Y. Ushiroda, A. Mohapatra, H. Sakamoto, Y. Sakai, M. Nakao, Q. An, and Y. Wang, Belle, *Development of the central trigger system for the BELLE detector at the KEK B factory*, Nucl. Instrum. Meth. A **438** (1999) 460–471.

[56] Belle, H. Ozaki *et al.*, *Mini-DST Tables V0.0.*, 1996. Belle note 146 (internal) (1996) .

[57] Belle, B. Casey *et al.*, *HadronB*, Belle note 390 (internal) .

[58] Belle, R. Itoh *et al.*, *BASF User's Manual*, 1996. Belle note 161 (internal) (1996) .

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

[60] M. Gelb *et al.*, *B2BII: Data Conversion from Belle to Belle II*, Comput. Softw. Big Sci. **2** (2018) no. 1, 9, arXiv:1810.00019 [hep-ex].

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

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

[63] 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.

[64] 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.

[65] 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.

[66] W. Waltenberger, *RAVE: A detector-independent toolkit to reconstruct vertices*, IEEE Trans. Nucl. Sci. **58** (2011) 434–444.

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

[68] 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.

[69] 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.

[70] 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.

[71] 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.

[72] 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].

[73] 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.

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

[75] 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.

[76] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini, *GEANT 3: user's guide Geant 3.10, Geant 3.11; rev. version.* CERN, Geneva, 1987. https://cds.cern.ch/record/1119728.

[77] 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].

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

[79] 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.

[80] G. Calderón, J. H. Muñoz, and C. E. Vera, *Nonleptonic two-body B decays including axial-vector mesons in the final state*, Physical Review D **76** (Nov, 2007) . http://dx.doi.org/10.1103/PhysRevD.76.094019.

[81] H.-Y. Cheng and K.-C. Yang, *Hadronic charmless B decays $B \to AP$*, Phys. Rev. D **76** (Dec, 2007) 114020. https://link.aps.org/doi/10.1103/PhysRevD.76.114020.

[82] V. Laporta, G. Nardulli, and T. N. Pham, *Nonleptonic B decays to axial-vector mesons and factorization [Phys. Rev. D 74, 054035 (2006)]*, Phys. Rev. D **76** (Oct, 2007) 079903. https://link.aps.org/doi/10.1103/PhysRevD.76.079903.

[83] B. S. M. Bauer and M. Wirbel, *Exclusive non-leptonic decays of D^-, D_s^- and B^- mesons*, Z. Phys. C - Particles and Fields **34** (1987) 103–115.

[84] 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.

[85] Belle, *The Number of B events in the Belle Hadronic Skims"*., https://belle.kek.jp/secured/nbb/nbb_old.html#details.

[86] 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.

[87] Belle, H. M. et al, *Dalitz analysis of the decays $B^- \to D^0\pi^-\pi^0$ and $B^0 \to D^-\pi^+\pi^0$*, 2015. Belle note 1371 (internal) (2015) .

[88] Belle, P. Koppenburg, *A measurement of the track finding efficiency using partially reconstructed D* decays*, 2003. Belle note 621 (internal) (2003) .

[89] Belle, S. Nishida, *Study of kaon and pion identification using inclusive D* sample*, 2005. Belle note 779 (internal) (2005) .

[90] Belle, S. Ryu, *Study of π^0 efficiency using $\tau^- \to \pi^-\pi^0\nu_\tau$*, 2012. Belle note 1224 (internal) (2012) .