



UNIVERSITÀ DEGLI STUDI DI TRIESTE

DIPARTIMENTO DI FISICA

Corso di Laurea Magistrale in Fisica

PRIMA RICOSTRUZIONE DI DECADIMENTI
CHARMLESS DEL B A BELLE II

Candidato:
Sebastiano Raiz

Relatore:
Dott. Diego Tonelli
Correlatore:
Dott. Fernando J. Abudinén G.

ANNO ACCADEMICO 2019–2020



UNIVERSITÀ DEGLI STUDI DI TRIESTE

DIPARTIMENTO DI FISICA

Corso di Laurea Magistrale in Fisica

**FIRST CHARMLESS B DECAY
RECONSTRUCTION IN BELLE II DATA**

Candidato:
Sebastiano Raiz

Relatore:
Dott. Diego Tonelli
Correlatore:
Dott. Fernando J. Abudinén G.

ANNO ACCADEMICO 2019–2020

The risk I took was calculated,

but man,

I am bad at math.

Sunto

Questo è un lavoro di fisica sperimentale delle particelle finalizzato alla prima ricostruzione di decadimenti *charmless* dei mesoni B nei dati dell'esperimento Belle II. Belle II è un rivelatore ottimizzato per la ricostruzione di miliardi di decadimenti di mesoni pesanti e leptoni τ prodotti in collisioni elettrone-positrone a 10 GeV. Il suo obiettivo scientifico è la ricerca indiretta di estensioni del Modello Standard, la teoria attualmente accettata.

I decadimenti dei quark b in quark diversi dal c (charmless) offrono opportunità uniche di misurare alcuni parametri fondamentali delle interazioni deboli sensibili a fisica oltre il Modello Standard, ed hanno pertanto un ruolo centrale nel programma scientifico di Belle II.

Il lavoro consiste di tre parti sinergiche. Il filo conduttore è usare le sfide sperimentali associate alla ricostruzione di decadimenti charmless per studiare, comprendere maggiormente e possibilmente migliorare le prestazioni di Belle II affinché siano ottimali in vista dei campioni di dati aspettati nel prossimo futuro.

Nella prima parte discuto la ricostruzione di decadimenti $B^+ \rightarrow K^+\pi^0$ nel campione di dati raccolto da Belle II nel 2019 e corrispondente a 8.7 fb^{-1} di luminosità integrata. Il canale $B^+ \rightarrow K^+\pi^0$ sonda diverse capacità sperimentali cruciali del rivelatore, come la soppressione del fondo da coppie di quark leggeri, la ricostruzione di fotoni ad alta energia e l'identificazione di adroni carichi. Il risultato è la prima ricostruzione di decadimenti $B^+ \rightarrow K^+\pi^0$ nei dati di Belle II, con 27 ± 8 eventi di segnale. Il segnale evidenzia anomalie nella calibrazione dell'energia dei fotoni, rimaste finora inosservate, che caratterizzo in dettaglio portando la collaborazione ad adottare un nuovo approccio alla calibrazione del calorimetro. Nella seconda parte studio le prestazioni di identificazione di particelle cariche, un elemento chiave nell'analisi di decadimenti di mesoni B . Il risultato è l'osservazione di dipendenze tra le variabili di identificazione misurate dai vari sottorivelatori, finora trascurate nella ricostruzione ma potenzialmente dannose per le prestazioni di lungo periodo. Nell'ultima parte ottengo la prima misura di asimmetrie tra i tassi di decadimento di mesoni B^+ e B^- dovute alla non conservazione della simmetria carica-parità (\mathcal{CP}) nei decadimenti $B^+ \rightarrow K^+K^-K^+$ e $B^+ \rightarrow K^+\pi^-\pi^+$ ricostruiti nei dati raccolti nel 2019–2020 e corrispondenti a 34.6 fb^{-1} di luminosità integrata. I risultati, $\mathcal{A}_{\mathcal{CP}}(B^+ \rightarrow K^+K^-K^+) = -0.049 \pm 0.063(\text{stat}) \pm 0.022(\text{syst})$ e $\mathcal{A}_{\mathcal{CP}}(B^+ \rightarrow K^+\pi^-\pi^+) = -0.063 \pm 0.081(\text{stat}) \pm 0.023(\text{syst})$, hanno richiesto il primo studio approfondito sulle asimmetrie di carica strumentali in Belle II, uno strumento fondamentale per tutte le future misure di violazione di \mathcal{CP} .

Tutti i risultati sono compatibili con quanto aspettato dalla simulazione e dimostrano prestazioni già competitive con le migliori degli esperimenti precedenti. L'impatto principale è una validazione dettagliata e significativi miglioramenti di molti aspetti cruciali delle prestazioni di Belle II. Questo lavoro getta solide basi per le analisi del prossimo futuro, quando le dimensioni dei campioni permetteranno risultati competitivi con le migliori misure odierne.

Abstract

This is an experimental particle physics thesis aimed at the first reconstruction of charmless decays of B mesons in Belle II data. Belle II is a detector designed to reconstruct billions of decays of heavy mesons and τ leptons from 10 GeV electron-positron collisions in search for indirect indications of non-Standard-Model physics. Decays of b quarks in quarks other than c (charmless B decays) provide unique access to fundamental parameters of the weak interactions that are particularly sensitive to non-SM physics and play an important role in the Belle II scientific program.

The work consists in three synergic parts. The common thread is the use of challenging physics to study, understand, and possibly optimize the most relevant experimental capabilities of the detector and reconstruction to reach optimal performance when the data set size will be competitive for world-leading results.

I first focus on the reconstruction of $B^+ \rightarrow K^+\pi^0$ decays using the sample collected by Belle II in 2019 and corresponding to 8.7 fb^{-1} of integrated luminosity. The $B^+ \rightarrow K^+\pi^0$ channel probes several key aspects of the detector and reconstruction performance such as background suppression, reconstruction of energetic photons, and charged-hadron identification. I achieve the first reconstruction of $B^+ \rightarrow K^+\pi^0$ decays in Belle II data, with a signal yield of 27 ± 8 events, and expose a previously undetected energy-miscalibration in photon reconstruction, prompting the collaboration to adopt a novel approach to monitor the calorimeter energy-scale. A study of the charged-particle identification performance – a key capability in B physics – follows. I observe subtle instrumental dependences between particle-identification variables measured by different sub-detectors, so far neglected in Belle II reconstruction, and assess their impact on long-term performance. Finally, I achieve the first measurement of charge-parity (\mathcal{CP}) violating asymmetries in three-body decays $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$ reconstructed in 2019–2020 Belle II data corresponding to 34.6 fb^{-1} of integrated luminosity. The results, $\mathcal{A}_{\mathcal{CP}}(B^+ \rightarrow K^+K^-K^+) = -0.049 \pm 0.063(\text{stat}) \pm 0.022(\text{syst})$ and $\mathcal{A}_{\mathcal{CP}}(B^+ \rightarrow K^+\pi^-\pi^+) = -0.063 \pm 0.081(\text{stat}) \pm 0.023(\text{syst})$, demand a thorough data-based investigation of instrumental charge-asymmetries, which is a fundamental input for all future \mathcal{CP} -violation measurements.

All results are compatible with expectations from simulation and show performance on par with the best performance of the predecessor experiments.

This work provides solid ground for upcoming analyses of B decays at Belle II and is expected to have significant impact on the quality of Belle II physics throughout the next few years, when the expected data set will enable competitive precisions with current world-best results.

Contents

Introduction	1
1 Flavor physics to overcome the Standard Model	3
1.1 The Standard Model of particle physics	3
1.2 Where do we stand?	5
1.3 Flavor physics in the Standard Model	6
1.4 Flavor physics to overcome the Standard Model	8
1.4.1 Flavor-changing neutral currents	8
1.4.2 Violation of charge-parity symmetry	8
1.5 Role of charmless B decays	10
1.5.1 Principal amplitudes and relevant measurements	11
1.5.1.1 Isospin sum-rules to address the $K\pi$ puzzle	12
1.5.1.2 β/ϕ_1 from $B^0 \rightarrow \eta' K_S^0$ and $B^0 \rightarrow \phi K_S^0$ decays	12
1.5.1.3 α/ϕ_2 from $B \rightarrow \pi\pi$, $B \rightarrow \rho\pi$, and $B \rightarrow \rho\rho$ decays	12
1.5.1.4 Multibody decays	13
1.6 Current experimental status	13
2 The Belle II detector at the SuperKEKB collider	16
2.1 The SuperKEKB collider at the KEK laboratory	16
2.2 Overview of the Belle II detector	20
2.3 Tracking detectors	23
2.3.1 Silicon-pixel vertexing detector	23
2.3.2 Silicon-microstrip vertexing detector	24
2.3.3 Central drift chamber	24
2.4 Particle-identification detectors	25
2.4.1 Time of propagation counter	25
2.4.2 Aerogel ring-imaging Cherenkov	26
2.4.3 Electromagnetic calorimeter	26
2.4.4 K_L^0 and muon detection system	27
2.5 Online event selection	27
2.6 Status of Belle II operations	28
2.7 B decay reconstruction at Belle II	28
2.7.1 Particle-level variables	30
2.7.2 Candidate-level variables	30
2.7.3 Event-level variables	31

3	First reconstruction of the charmless decay $B^+ \rightarrow K^+\pi^0$ at Belle II	34
3.1	Motivation and strategy	34
3.2	Data and tools	35
3.3	Baseline optimization	35
3.3.1	Selection performance metric	36
3.3.2	Skimming selection quantities	36
3.3.3	Charged-particle classes	36
3.3.4	Charged-particle selection optimization	37
3.3.5	π^0 selection optimization	38
3.4	$B^+ \rightarrow K^+\pi^0$ reconstruction	40
3.5	Continuum suppression	41
3.6	Offline selection optimization	46
3.7	Sample composition	46
3.8	Determination of signal yields in simulation	47
3.9	Results	48
3.10	Summary and overview of other charmless channels	49
4	Observation of photon-energy miscalibration	51
4.1	Motivation	51
4.2	Studies on control channel	53
4.3	Analytical derivation	54
4.4	Simulation	57
4.5	Summary	57
5	Studies of charged-hadron identification	58
5.1	Motivation and strategy	58
5.2	Sample and tools	58
5.3	Particle identification at Belle II	59
5.3.1	Detector-specific likelihoods and PID	60
5.3.2	Detector combination	62
5.3.3	Performance metrics	63
5.4	Detector-specific and detector-combined performances	64
5.5	Exploring likelihood dependences	68
5.6	Impact of neglected likelihood dependences on performance	70
5.6.1	Simplified simulation	71
5.6.2	Results and impact on physics	72
5.7	Summary	75
6	First measurement of \mathcal{CP}-violating asymmetries in three-body charmless decays at Belle II	76
6.1	Motivation and strategy	76
6.2	Raw charge-asymmetry determination	77
6.2.1	Tools, selection, and reconstruction	77
6.2.2	Selection optimization	77
6.2.3	Sample composition	78
6.2.4	Asymmetry determination	82
6.2.4.1	Fit model	82
6.2.4.2	Results using simulated data	82
6.2.4.3	Results using Belle II data	83
6.3	Determination of instrumental asymmetries	84

CONTENTS

6.3.1	Samples, tools and selection	87
6.3.2	$K^-\pi^+$ instrumental asymmetry	88
6.3.3	$\bar{K}^0\pi^+$ instrumental asymmetry	89
6.3.4	Determination of the final K^+/K^- instrumental asymmetry	90
6.4	Systematic uncertainties	90
6.5	Consistency checks	92
6.6	Results and summary	92
	Summary	94
	Bibliography	96

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 less than 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 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 main 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 available 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 theory 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 II experiment is an international collaboration of more than 1000 physicists that aims at *indirectly* testing the Standard Model by studying billions 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 GeV. Belle II started taking data in March 2019, and has now collected collision samples corresponding to 70 fb^{-1} of integrated luminosity, or, equivalently, 80 million pairs of B mesons (bound states of a b quark and a lighter quark). As it is a recent experiment, its chief priority in this initial stage is the data-driven validation of the physics performance. This is an essential task where *known* physics processes are used as references to understand, calibrate, and optimize the performance of *not yet fully optimized* detector and reconstruction algorithms.

My work focuses on the first reconstruction of *charmless* B decays in Belle II data and on their use as sensitive probes of Belle II performance. Charmless B decays are transitions where a b quark decays in quarks other than c . Charmless B decays provide unique access to fundamental parameters of the weak interactions that are particularly sensitive to non-SM physics. They thus play a key role in the scientific program of Belle II. However, studying charmless B decays poses experimental and theoretical challenges. They are difficult to reconstruct due to their rarity (branching fractions $< 2 \times 10^{-5}$) and to the similarity of final states with backgrounds 10^5 times more abundant. Moreover, interpretation of measurements suffers from phenomenological challenges associated with the presence of non-perturbative quantum chromodynamics effects, which spoil the precision of predictions.

Thanks to the abundance and variety of available decay channels, charmless decays allow for detailed characterization of multiple and diverse experimental aspects relevant for flavor physics analysis at Belle II.

I join this effort with three synergic contributions. In the first part of my thesis I target the first reconstruction in Belle II data of $B^+ \rightarrow K^+\pi^0$ decays, and an ensuing dedicated investigation of photon reconstruction. The $B^+ \rightarrow K^+\pi^0$ decay is particularly challenging even among the charmless channels because it is rare, it is contaminated by an abundant background, and it poses the additional experimental challenges associated with π^0 reconstruction. Reconstructing $B^+ \rightarrow K^+\pi^0$ decays imposes therefore a harsh strain on many key experimental capabilities, such as background suppression algorithms, π^0 reconstruction, and charged-hadron identification, allowing to benchmark the physics performance in depth.

The central part of my thesis is devoted to charged-hadron identification, as correct identification of particles is key to suppress backgrounds to charmless decays due to misidentifications. I focus on the combination of identification information from the various subdetectors, assumed independent in Belle II, and investigate the validity of this approximation and its impact on physics.

In the final part of my thesis I use a larger data sample to target the first measurement of \mathcal{CP} -violating asymmetries in $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$ decays. These offer a new set of peculiar challenges due to the larger number of final-state tracks and insidious peaking backgrounds that need to be suppressed or modeled. This work involves developing a thorough data-based investigation of instrumental charge-asymmetries, which could bias the \mathcal{CP} -violating asymmetries if left unconstrained.

This document is structured as follows: Chapter 1 outlines the flavor sector of the Standard Model; Chapter 2 describes the Belle II experiment and provides an outline of the general principles of B meson reconstruction; Chapter 3, where the description of my direct, original contributions begins, reports on the first reconstruction of the $B^+ \rightarrow K^+\pi^0$ decay channel; Chapter 4 shows an ensuing study on photon-energy miscalibrations; Chapter 5 presents my studies on the charged-hadron identification in Belle II; Chapter 6 describes the first Belle II measurement of charge-parity asymmetries in three-body charmless B decays; and the final chapter summarizes the salient results.

Charge-conjugate processes are implied throughout the document unless specified otherwise.

The work reported in this thesis is documented in two published Belle II preprints [1,2] and in five internal Belle II documents [3–7].

Chapter 1

Flavor physics to overcome the Standard Model

This is a concise introduction to the weak interactions of quarks and how they are incorporated in the Standard Model of particle physics. I outline the main concepts subtending 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, which are the subject of this thesis.

1.1 The Standard Model of particle physics

The Standard Model (SM) is a quantum field theory that describes three of the four fundamental interactions in nature (gravity is not included) [8–13].

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.

In physics, 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. 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, and 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 (quan-

tum 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 The group $SU_C(3)$ has eight generators, $SU_L(2)$ has three generators $W_{1,2,3}$, and $U_Y(1)$ has one, B . All elementary particles are excitations of fields.

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 [14] 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.

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

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

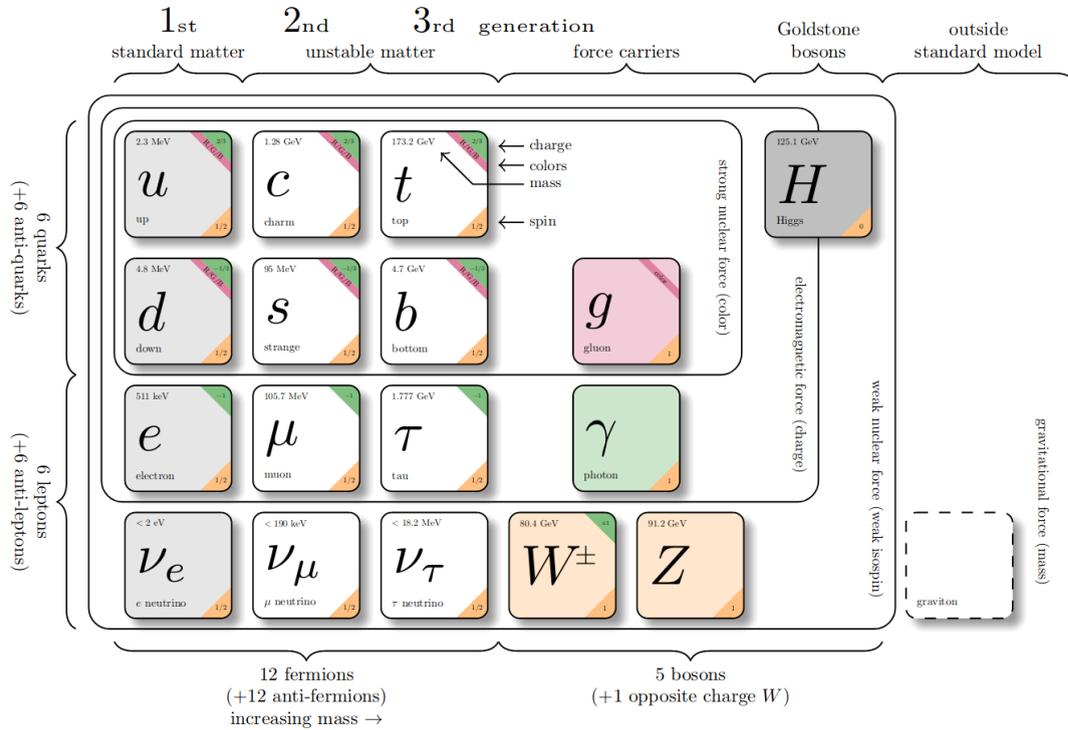


Figure 1.1: Scheme of particles and interactions in the Standard Model.

In addition to gauge symmetry, discrete symmetries are important too in constraining the dynamics. Parity (\mathcal{P}) is a transformation that inverts all spatial coordinates; charge conjugation (\mathcal{C}) is the exchange of every particle with its own antiparticle; and time reversal (\mathcal{T}) inverts the time axis. The product of the three discrete symmetry transformations is found to be conserved in all interactions, but they are not conserved individually [15, 16]. Parity symmetry is maximally violated in the weak interactions, while the combined \mathcal{CP} symmetry is violated in the weak interactions at the 0.1% level. In principle, the strong interaction too could violate \mathcal{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 [17]. 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 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, 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. The 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 $\Delta E \Delta t > \frac{\hbar}{2}$ where energy conservation can be violated for short enough times. 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 will restrict the scope by referring solely to the quark interactions in this work.

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 [18] and then Cabibbo [19], who postulated differing mass (d) and weak (d') eigenstates for down-type quarks. This was realized 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 $K^0 \rightarrow \mu^+ \mu^-$ 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 $2 \text{ GeV}/c^2$ mass, whose contribution in the $K^0 \rightarrow \mu^+ \mu^-$ decay amplitude would cancel the u -quark contribution, suppressing the branching fraction down to values consistent with experimental limits [20]. In the early 1970's, 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 \mathcal{CP} violation observed in 1964 [21]. They introduced a 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 quark-mixing matrix or V_{CKM} , a complex matrix that is unitary because the sum of all the transition probabilities must equal one. The $N \times N$ CKM matrix has $(N - 1)^2$ free parameters [22], where N is the number of quarks families. If $N = 2$, the only free parameter is the Cabibbo angle θ_C , whereas if $N = 3$, the free parameters are three Euler angles (θ_{12} , θ_{13} , and θ_{23}) and a complex phase (δ), which allows for \mathcal{CP} -violating couplings.

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.

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 quarks indicate the weak-interaction eigenstates corresponding to the unprimed 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* [23], 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 \mathcal{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, prompting a convenient geometric representation in terms of so-called *unitarity triangles* in the complex plane. A \mathcal{CP} conserving theory would yield null-area triangles or, equivalently, a vanishing Jarlskog invariant $J = \Im(V_{us} V_{cb} V_{ub}^* V_{cs}^*)$ [24–26]. All elements of the second equation in the middle column have similar magnitudes, yielding a non-degenerate triangle. That equation is therefore 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 .

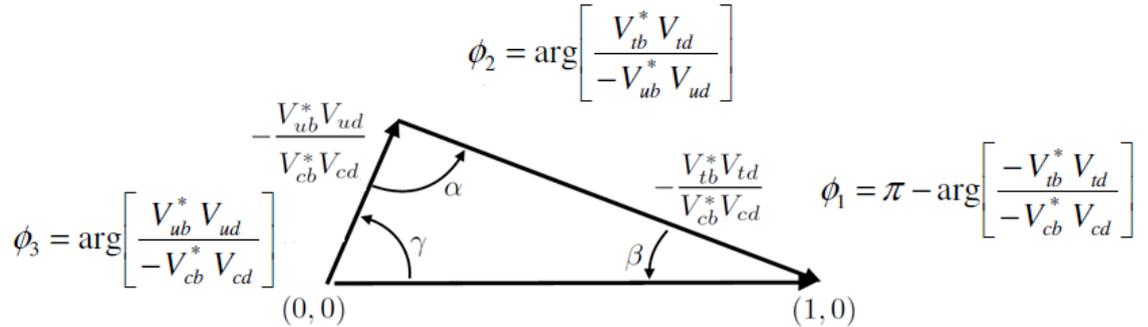


Figure 1.2: Graphical representation of the Unitarity Triangle.

1.4 Flavor physics to overcome the Standard Model

Many physicists find the current understanding of flavor dynamics and \mathcal{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 with a possibly reduced set of parameters. In addition, while the CKM mechanism offers a framework to include \mathcal{CP} violation in the SM, it does not really enlightens 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 \mathcal{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 and promising 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 \mathcal{CP} -violating processes.

1.4.1 Flavor-changing neutral currents

Flavor-changing neutral currents (FCNC) are processes in which particle flavor changes between initial and final states but 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. 1.3. Such amplitudes are naturally sensitive to non-SM contributions, since any particle with proper quantum numbers and nearly arbitrary mass can replace the SM-quarks closed-line in these diagrams thus altering the rate. FCNC can therefore be used to identify contribution from non-SM particles by measuring rate enhancements or suppressions with respect to Standard Model expectations.

1.4.2 Violation of charge-parity symmetry

Non-SM contributions may exist that do not change appreciably the rate of a process, but could largely modify the amplitude *phases*. Such contributions can be identified by studying processes sensitive to \mathcal{CP} violation, which allow comparing measurements of amplitude's phases with their theoretical predictions. In classical physics we measure the phase difference between two waves through their interference. In the same way, when a particle evolves transitioning from an initial to a final state, various possible paths are open, corresponding to various quark-level processes: if the amplitudes corresponding to each process have different \mathcal{CP} -violating and \mathcal{CP} -conserving phases, the ensuing interference allows measuring phase differences. The phenomenology of \mathcal{CP} violation is further enriched by the phenomenon of neutral meson mixing [27]. This consists in particle-antiparticle oscillations as functions of time that occur because the mass and the flavor eigenstate of neutral flavored mesons differ, and contributes further quark-level processes to the decay amplitudes.

The experimental phenomenology of \mathcal{CP} violation in hadrons is conventionally classified into three classes of phenomena:

- Direct \mathcal{CP} violation (or \mathcal{CP} violation in the decay amplitudes) occurs when the pure decay rate of the X hadron ($X \rightarrow f$) and the rate of the \mathcal{CP} -conjugate process

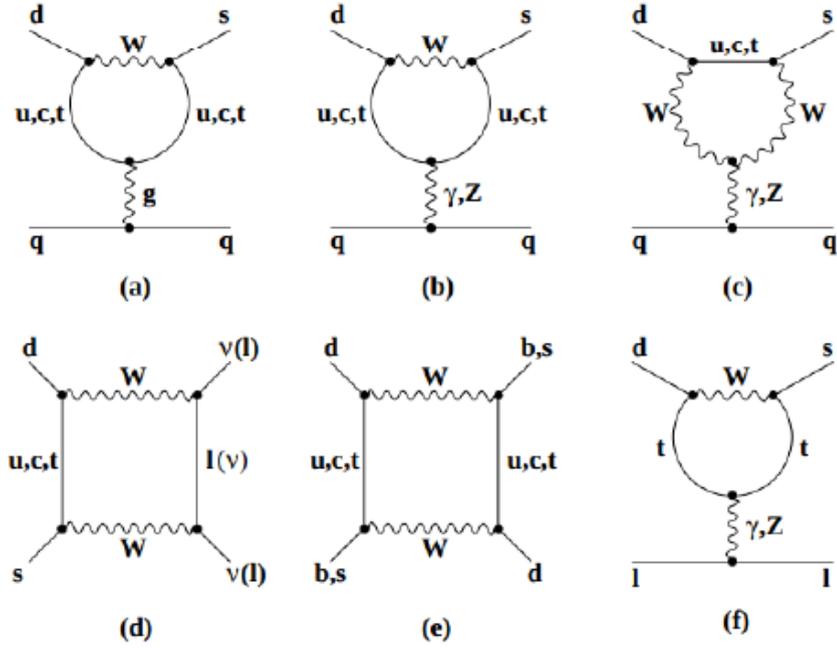


Figure 1.3: Examples of leading FCNC diagrams.

$(\bar{X} \rightarrow \bar{f})$ differ. The observable typically used to measure direct \mathcal{CP} violation is the partial-width asymmetry

$$\mathcal{A}_f = \frac{\Gamma(X \rightarrow f) - \Gamma(\bar{X} \rightarrow \bar{f})}{\Gamma(X \rightarrow f) + \Gamma(\bar{X} \rightarrow \bar{f})}. \quad (1.1)$$

This is the only \mathcal{CP} -violating phenomenon possible for charged mesons and baryons and it is the object of measurement discussed in Chap. 6.

- Indirect \mathcal{CP} violation in mixing occurs when the probability of neutral-meson mixing from one state to its \mathcal{CP} conjugate ($X^0 \rightarrow \bar{X}^0$) differs from the \mathcal{CP} -conjugate process ($\bar{X}^0 \rightarrow X^0$). The observable typically used is the semileptonic rate asymmetry,

$$\mathcal{A}_{\text{SL}} = \frac{d\Gamma(X^0(t) \rightarrow \bar{f})/dt - d\Gamma(\bar{X}^0(t) \rightarrow f)/dt}{d\Gamma(X^0(t) \rightarrow \bar{f})/dt + d\Gamma(\bar{X}^0(t) \rightarrow f)/dt}, \quad (1.2)$$

where the final state \bar{f} is chosen such that X^0 cannot directly decay into f while its conjugate can, and viceversa.

- If both particle and antiparticle decay into the same final state, indirect \mathcal{CP} violation may occur through the interference between direct decays and decays following mixing, since the final state can be reached either from direct decay or from decay following oscillations. The observable used is

$$\mathcal{A}_{f_{\mathcal{CP}}} = \frac{d\Gamma(X^0(t) \rightarrow f_{\mathcal{CP}})/dt - d\Gamma(\bar{X}^0(t) \rightarrow f_{\mathcal{CP}})/dt}{d\Gamma(X^0(t) \rightarrow f_{\mathcal{CP}})/dt + d\Gamma(\bar{X}^0(t) \rightarrow f_{\mathcal{CP}})/dt}. \quad (1.3)$$

1.5 Role of charmless B decays

Decays governed by transitions of b quarks into quarks other than c , so-called charmless B decays, are the topic of this thesis. We therefore discuss them in more detail here.

Charmless B decays proceed through the quark-level processes $b \rightarrow u$, $b \rightarrow d$, and $b \rightarrow s$. These processes yield final-state hadrons without net charm, such as kaons, pions, ρ mesons, and ϕ mesons. Typical examples of charmless decays include $B^0 \rightarrow K^+\pi^-$, $B^0 \rightarrow \pi^+\pi^-$, $B^+ \rightarrow \rho^+\rho^0$ and many others.

Because the magnitude of the CKM element V_{ub} ($\approx 4 \times 10^{-3}$) is much smaller than $|V_{cb}|$ ($\approx 4 \times 10^{-2}$), the tree-level amplitudes are suppressed.³ The resulting branching fractions are typically small (less than 2×10^{-5}), which makes charmless decays difficult to observe. In fact, the first reconstruction of charmless B decays was reported in 1993 by the CLEO collaboration [28], 16 years after the discovery of the *bottom* quark [29]. On the other hand, contributions from loop amplitudes turn out to be comparable to, or even larger than, those from tree amplitudes, offering multiple opportunities to probe efficiently the Standard Model and its extensions.

Charmless decays play a key role in flavor physics. They offer multiple avenues for refining our understanding of \mathcal{CP} violation and for probing the weak and strong interactions. Transitions $b \rightarrow du\bar{u}$ provide, for example, the only reliable determination of the α/ϕ_2 parameter. Transitions $b \rightarrow q\bar{q}s$ enable measurements of the β/ϕ_1 CKM phase through loop-level processes. Moreover, charmless B decays offer multiple accesses to direct \mathcal{CP} violation: all neutral and charged B meson decays involving both tree and penguin amplitudes are possible sources of direct \mathcal{CP} violation, as observed for example in $B^0 \rightarrow K^+\pi^-$, $B^0 \rightarrow \pi^+\pi^-$, and $B^+ \rightarrow \rho^0 K^+$ decays.⁴

Experimental and phenomenological challenges need to be addressed for a successful charmless program. The rarity of these modes and the similarity of their final states with the background dominated by light-quark pairs (which combine to form mainly pions and kaons) make their observation difficult: one has to separate signal from irreducible backgrounds that are at least 10^5 times more abundant. The lack of narrow intermediate resonances offering invariant mass restrictions to reduce combinatorial background makes discrimination even more challenging. In addition, interpretation of measurements of charmless decays suffers from phenomenological challenges associated with the presence of non-perturbative QCD effects. In hadronic decays, quarks interactions through low-energy gluons exchanges are relevant compared to the weak interactions. The perturbative QCD approach, for which higher-orders terms in the strong-coupling-constant α_s expansion tend to vanish, is valid only when α_s is small (that is, when the energies involved in the process are larger than ≈ 1 GeV). In low-energy gluon exchanges, the sum of the contributions from the expansion amplitudes diverges. The perturbative expansion is therefore unable to approximate the total amplitude of a process. The implication is that the poor precision of predictions limits strongly the sensitivity of comparisons with precise measurements.

Theorists proposed various approaches to overcome the issue. Currently the most promising strategy is to exploit exactly (or approximately) valid flavor symmetries existing between quark-level processes to reduce, with appropriate approximations, the QCD unknowns. An approximation often used is based on $SU(3)$ symmetry, where $m_u = m_d = m_s$ is assumed. The comparison between precise measurements and theoretical predictions

³Tree-level amplitudes are first-order amplitudes that involve the emission of a W boson. Higher-order processes, like loop amplitudes, involve the emission and the reabsorption of the W boson.

⁴A penguin amplitude involves the emission and reabsorption of a W boson and typically the presence of a gluon.

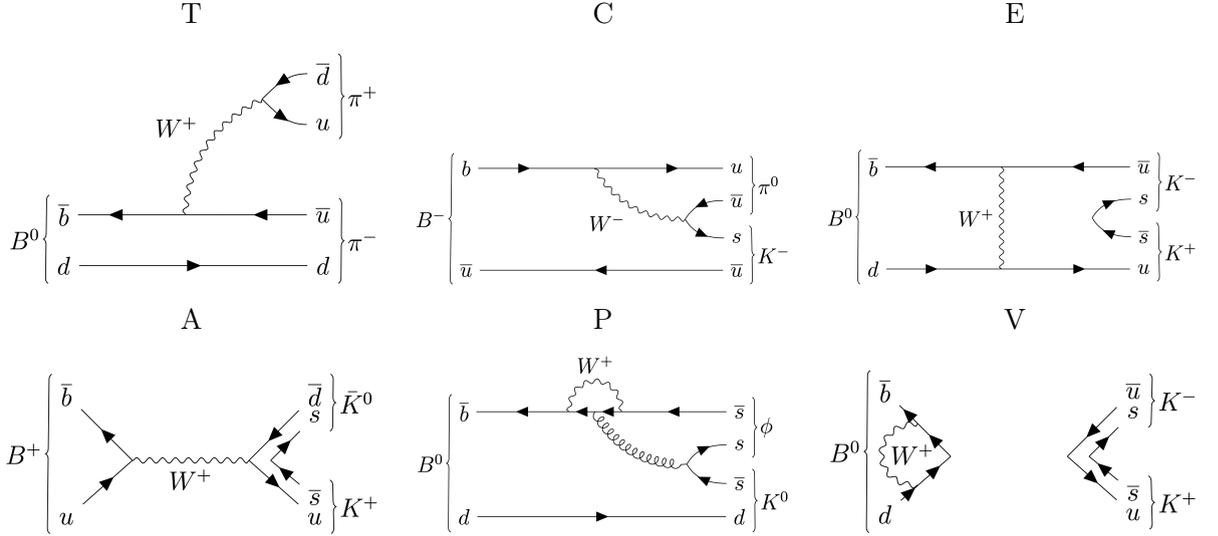


Figure 1.4: Amplitudes relevant in charmless B decays: T) color-allowed external W -emission tree diagram; C) color-suppressed internal W -emission tree diagram E) W -exchange diagram A) W -annihilation diagram P) penguin diagram with gluon exchange V) W -loop diagram.

is therefore performed on combinations of experimental results based on channels related by $SU(3)$. This way, for example, the ratio between the branching fractions $\mathcal{B}(B^0 \rightarrow K^+\pi^-)/\mathcal{B}(B^0 \rightarrow \pi^+\pi^-)$ is more accurately predicted than the individual branching fractions. Theorist also suggest experimental tests to check the extent of violations of the assumed symmetries, which allow achieving a fairly reliable, and sometimes precise, description of the dynamics.

1.5.1 Principal amplitudes and relevant measurements

Figure 1.4 shows examples of amplitudes relevant for charmless B decays. A detailed discussion of these amplitudes is beyond the scope of this thesis, but some intuitive features can be identified.

The first two examples represent a $B^0 \rightarrow \pi^+\pi^-$ decay and a $B^- \rightarrow K^-\pi^0$ decay proceeding through tree-level processes. The second amplitude (color-suppressed tree-amplitude) is suppressed with respect to the first by the color factor $N_C=3$, since the two internal quarks coming from the W boson need to match the color of the B -meson quarks to combine in a colorless final state.

Tree-level amplitudes enjoy accurate predictions, but in charmless decays they are suppressed to rates typically comparable to higher-order amplitudes. An example of the latter is the penguin amplitude (fifth picture), where the presence of an internal quark line opens up sensitivity for non-SM contributions, but makes the calculations difficult.

These and many other amplitudes offer a plethora of hadron-level processes, which are conventionally divided in ‘two-body decays’ and ‘multibody decays’: the two-body decays include decays in long-lived final states ($\pi\pi$, $K\pi$, KK , etc.) and decays where one of the decay products is a resonance (η , ρ , ω , etc.), so called ‘quasi-two-body decays’. Multibody decays include final states with π , K , but also resonances or protons.

These decays offer hundreds of possibilities for measurements of decay rates and \mathcal{CP} violation. Those currently considered as the most promising toward furthering the under-

standing of flavor dynamics follow.

1.5.1.1 Isospin sum-rules to address the $K\pi$ puzzle

Two-body charmless B decays in kaon-pion pairs have been attracting great interest in the past decade. The key dynamical feature is that a simple additive relation based on constraints dictated by isospin,

$$\sqrt{2}A(K^0\pi^0) + A(K^+\pi^-) = \sqrt{2}A(K^+\pi^0) + A(K^0\pi^+)$$

connects the amplitudes (A) of the decays $B^0 \rightarrow K^+\pi^-$, $B^+ \rightarrow K^0\pi^+$, $B^+ \rightarrow K^+\pi^0$, and $B^0 \rightarrow K^0\pi^0$ [30]. The relevant amplitudes have been determined by the Belle, BaBar, and LHCb collaborations based on measurements of branching fractions and \mathcal{CP} -violating asymmetries, and an apparent inconsistency emerges [17]. Once the experimental values are input, the equality is violated at a level in excess of three standard deviations. This could hint at an indirect indication of non-SM physics, but could also be reconciled by admitting unexpected differences in strong-interaction effects or a more prominent role of subdominant amplitudes not included in the original derivation of the relation. This so-called ‘ $K\pi$ puzzle’ is considered a relevant open problem in flavor physics. More precise measurements are expected to provide a conclusive outcome.

Belle II is particularly apt to perform these measurements thanks to the precision expected on channels with final-state neutral particles and to the unique capability of measuring, in the same experiment and in a consistent fashion, all four decays, thus suppressing common systematic uncertainties.

In the first part of my work (Chap. 3) I target the first reconstruction in Belle II data of $B^+ \rightarrow K^+\pi^0$ decays, an essential ingredient to prepare the study of the $K\pi$ puzzle at Belle II.

1.5.1.2 β/ϕ_1 from $B^0 \rightarrow \eta'K_S^0$ and $B^0 \rightarrow \phi K_S^0$ decays

The measurement of the β/ϕ_1 angle was the original motivation for B Factories, where it has been measured with high precision using $B^0 \rightarrow J/\psi K_S^0$ decays, which are dominated by tree-level amplitudes. The angle β/ϕ_1 is today the best known parameter of the Unitarity Triangle. Such precisely known reference offers therefore the opportunity of a stringent SM test through the comparison with alternate determinations of the same angle from processes dominated by loop amplitudes. Non-SM particles could indeed contribute additional amplitudes and alter the value of the observed phase, thus indicating non-SM effects.

Belle II plans to use the charmless decays $B^0 \rightarrow \eta'K_S^0$ and $B^0 \rightarrow \phi K_S^0$ to measure β/ϕ_1 using loop processes, because they are affected by the least theoretical uncertainties. The advantages of Belle II over competing experiments like LHCb are the more straightforward reconstruction of decays with final-state neutral particles as $B^0 \rightarrow \eta'(\rightarrow \gamma\gamma)K_S^0$ and the higher precision on the B^0/\bar{B}^0 flavor determination, an essential element in the measurement of β/ϕ_1 .

1.5.1.3 α/ϕ_2 from $B \rightarrow \pi\pi$, $B \rightarrow \rho\pi$, and $B \rightarrow \rho\rho$ decays

With an uncertainty of about 5%, the CKM angle α/ϕ_2 is currently among the major limiting factors on the global precision of the Unitarity Triangle. The most reliable approach to measure α is to reconstruct large samples of $B^0 \rightarrow \pi^+\pi^-$, $B^+ \rightarrow \rho^+\pi^0$, and $B^0 \rightarrow \rho^0\rho^0$

decays and combine appropriately branching fractions, \mathcal{CP} -violating asymmetries, and polarization fractions in a global fit that determines α by exploiting the existing symmetries between the channels.

Competing experiments like LHCb, where reconstruction of decays with final-state neutral particles like $B^+ \rightarrow \pi^0\pi^+$ and $B^0 \rightarrow \pi^0\pi^0$ is difficult if at all possible, need inputs from other experiments to perform this fit. Belle II enjoys the advantage of the access to precise measurements of the full set of inputs in an experimentally consistent way, which ensures uniformity of analysis across channels with reduced systematic uncertainties.

1.5.1.4 Multibody decays

In recent years, multibody charmless decays have been attracting significant attention. Their rich dynamical structure due to the presence of many possible intermediate states allows for a plethora of measurements of \mathcal{CP} -violating asymmetries localized in different regions of the Dalitz plot. These provide more information with respect to measurements based on two-body decays, where only integrated \mathcal{CP} violation is available. Recent measurements by LHCb show \mathcal{CP} -violating asymmetries interestingly large (50–60%) that certainly need a more thorough exploration [31].

In the analysis of multibody decays, Belle II has the advantage of a straightforward amplitude analysis of the Dalitz plot. In LHCb, stringent trigger selections sculpt the kinematic distributions of the B mesons in ways that are not trivially reproduced in simulation, resulting in significant experimental challenges and large systematic uncertainties.

In the last chapter of this work I present a first measurement of the \mathcal{CP} -violation in the three-body charmless decays $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$ at Belle II.

1.6 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 [32]. The resulting global picture is that the CKM interpretation of quark-flavor phenomenology is the dominant mechanism at play in the dynamics.

The current status of constraints on sides and angles of the Unitarity Triangle is shown in Fig. 1.5 [33]. Measurements of $\sin 2\beta$ reached a precision of 2.5%, mainly due to the availability of large samples of $B^0 \rightarrow J/\psi K_S^0$ decays in e^+e^- collisions, while the angle α is known down to a 5% precision from $B \rightarrow hh$ 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 7% precision using combinations of several measurements involving $B \rightarrow DK$ decays (B , D , K charged or neutral) reconstructed in e^+e^- and pp collisions. Mild discrepancies in the determinations of $|V_{cb}|$ and $|V_{ub}|$ are found between values measured in different 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 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. One example is the recent observation of direct \mathcal{CP} violation in charm decays at LHCb [34].

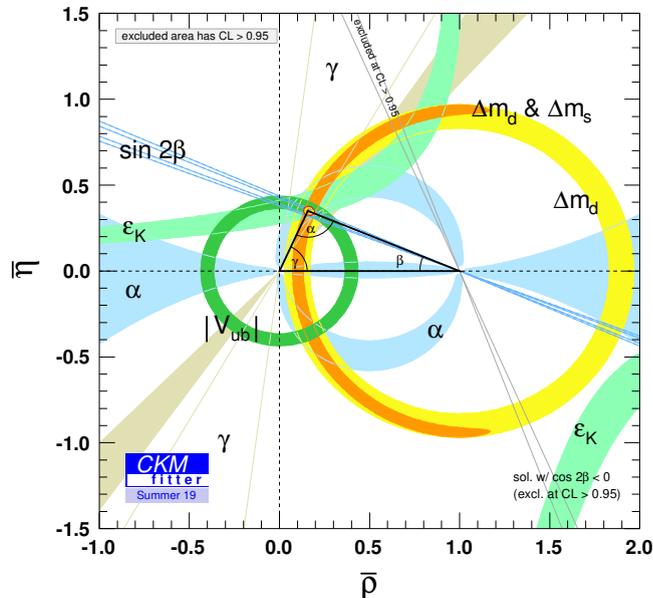


Figure 1.5: Current constraints on sides and angles of the Unitarity Triangle.

Recent direct searches for non-SM physics, mainly in pp collisions at the Large Hadron Collider (LHC) at CERN, have 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 when 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.

LHCb is a forward spectrometer that operates at the LHC. It studies proton-proton collisions at energies of up to 14 TeV, where incoherent QCD production yields large samples of $b\bar{b}$ pairs (≈ 45 kHz) and $c\bar{c}$ pairs (≈ 1 MHz). The high production rate enables low statistical uncertainties, and TeV collision energy allows the production of all species of *bottom* hadrons, and therefore to the possibility of investigating a broader set of processes. LHCb drawbacks are associated with the fact that LHC collisions occur between composite particles. These lead to large backgrounds, which require challenging online selections and offer looser constraints on the production kinematics. This impacts specially the reconstruction of neutral final-state particles, including photons and neutrinos.

Belle II is an experiment operating at the energy asymmetric e^+e^- collider SuperKEKB at KEK, which produces $B\bar{B}$ pairs from $\Upsilon(4S)$ decays at ≈ 1 kHz rate. Compared to LHCb, Belle II has a lower production rate but a better sensitivity to B and D final states that include neutral particles, because of the much lower background and additional kinematic constraints provided by the point-like nature of colliding particles. Figure 1.6 shows the display of two typical events in these two experiments. Moreover, the coherent evolution in time of the $B\bar{B}$ pairs produced in e^+e^- collisions make the determination of the flavor

of the signal neutral B meson more efficient than in LHCb. Belle II aims at operating until 2030 at the intensity frontier of the flavor sector to explore indirectly of non-SM physics, in competition and synergy with LHCb [35].

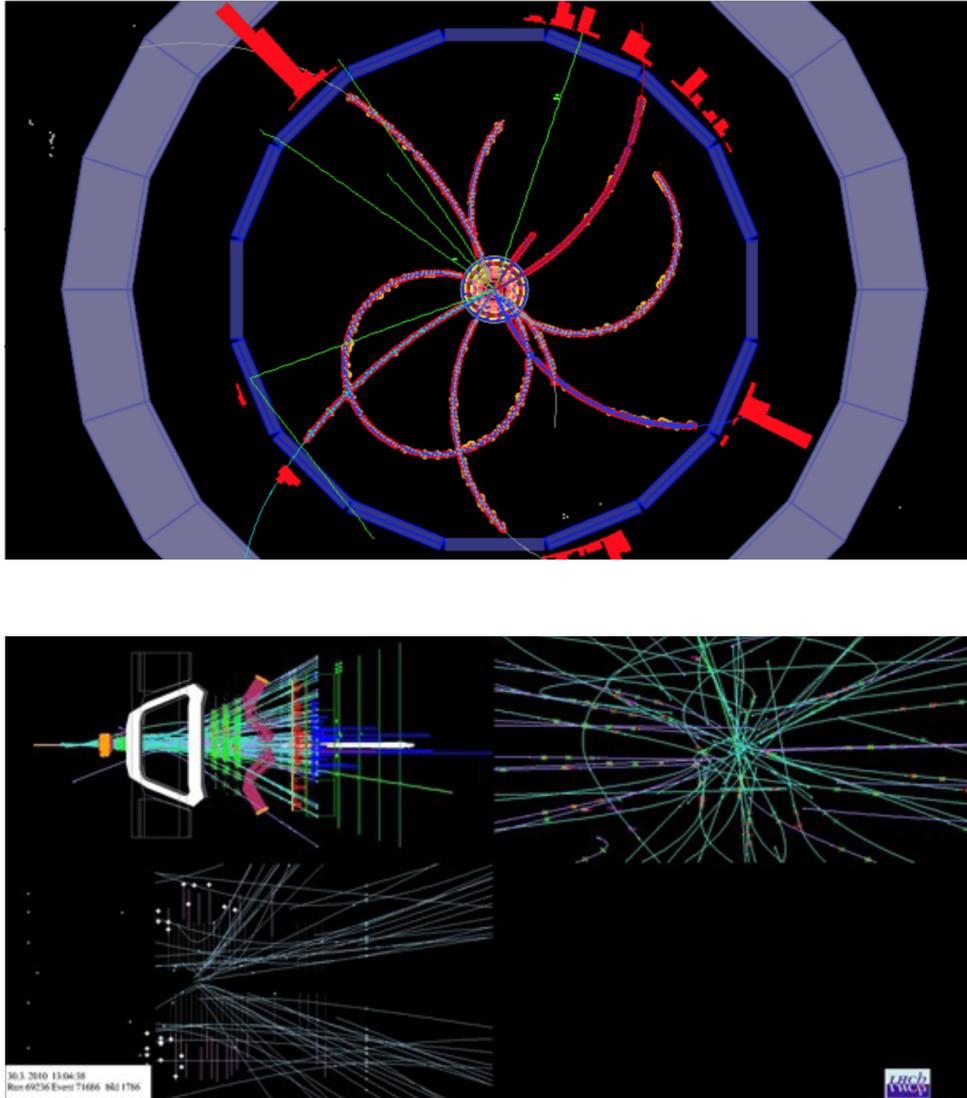


Figure 1.6: Typical event display for Belle II (top) and LHCb (bottom).

Chapter 2

The Belle II detector at the SuperKEKB collider

This chapter introduces the Belle II experiment, whose data I used for my thesis work.

2.1 The SuperKEKB collider at the KEK laboratory

SuperKEKB is a high-luminosity electron-positron (e^+e^-) energy-asymmetric collider, designed to produce nearly 1000 $B\bar{B}$ pairs ($B^0\bar{B}^0$ and B^+B^- in approximately equal proportions) per second via decays of $\Upsilon(4S)$ mesons produced at threshold [36]. Such colliders are called ‘ B -factories’, and were proposed in the 1990’s for the dedicated exploration of \mathcal{CP} -violation in B mesons. The main goal of B -factories is to produce low-background quantum-correlated $B\bar{B}$ 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 $B\bar{B}$ production kinematic threshold. Such finely tuned collision energy is key. The ensuing production of $\Upsilon(4S)$ mesons, which decay in $B\bar{B}$ pairs 96% of the times with little available energy to produce additional particles, suppresses backgrounds, which are mainly due to competing nonresonant hadron production. In addition, usage of beams of point-like particles allows for knowing precisely the collision energy, which sets stringent constraints on the final-state 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 $B\bar{B}$ pair; even though B^0 and \bar{B}^0 undergo flavor oscillations before decaying, their time-evolution is quantum-correlated in such a way that no B^0B^0 or $\bar{B}^0\bar{B}^0$ pairs are present at any time. In fact, angular-momentum conservation implies that the decay of a spin-1 particle in two spin-0 particles yields total angular momentum $L = 1$. Because the simultaneous presence of two identical particles in an antisymmetric state would violate Bose statistics, the system evolves coherently as an oscillating $B^0\bar{B}^0$ 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 \mathcal{CP} -violating quantities.

Not just $\Upsilon(4S)$ mesons are produced in 10 GeV e^+e^- collisions; Figure 2.1 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 and the nearly uniform baseline at about 4 nb represents the so-called continuum of lighter-quark pair production ($e^+e^- \rightarrow q\bar{q}$, where q identifies u, d, c, s), which exceeds $\Upsilon(4S)$ production in rate. In addition, the most frequent outcomes of 10 GeV e^+e^- collisions are electroweak processes of lepton production, such as $e^+e^- \rightarrow e^+e^-(\gamma)$, $e^+e^- \rightarrow e^+e^-e^+e^-$ and $e^+e^- \rightarrow \mu^+\mu^-(\gamma)$. These amount to about 94% of the total cross-section, but can straightforwardly be discriminated online owing to their distinctive final states.

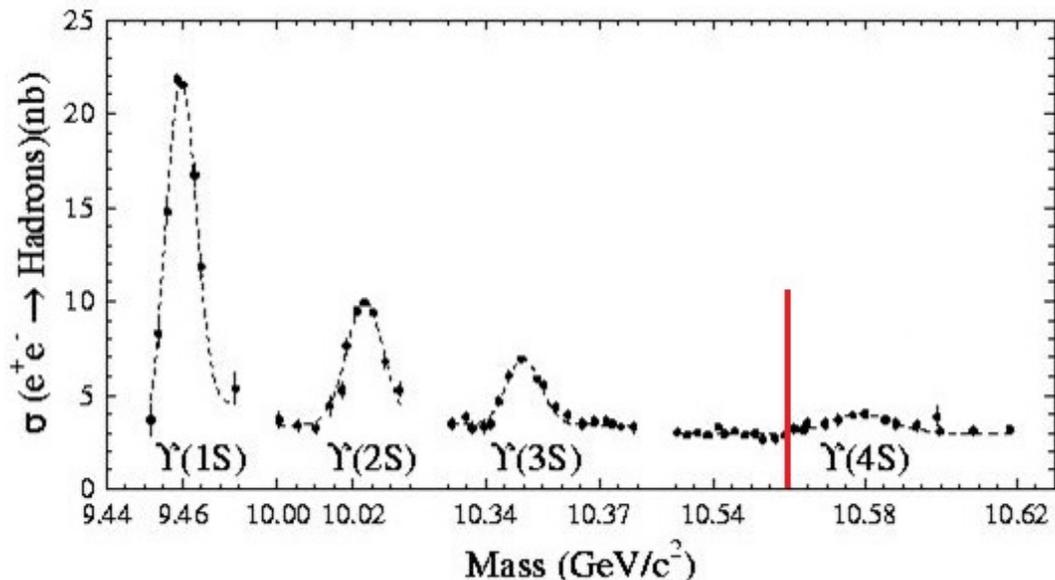


Figure 2.1: Hadron production cross section from e^+e^- collisions as a function of the final-state mass. The vertical red line indicates the $B\bar{B}$ production threshold.

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 ($\approx 10 \text{ MeV}/c$) in the laboratory, because of the $21 \text{ MeV}/c^2$ difference between the $\Upsilon(4S)$ mass and the mass of a $B\bar{B}$ pair. With such low momenta they would only travel approximately $1 \mu\text{m}$ before decaying. The $10 \mu\text{m}$ typical spatial resolution of vertex detectors would not be sufficient to separate B -decay vertices and enable the study of the decay-time evolution for measurements involving mixing. 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 vertices separations resolvable with current vertex detectors [37]. SuperKEKB (Fig. 2.3) implements a 7-4 GeV energy-asymmetric double-ring design, which achieves a vertex displacement of about $150 \mu\text{m}$.

SuperKEKB is designed to deliver collisions corresponding to 50 ab^{-1} of integrated luminosity by 2030, corresponding $\approx 5.3 \times 10^{10}$ $B\bar{B}$ pairs, about 40 times the total amount collected by B Factories to date.

Electrons are produced via photoelectric effect by illuminating a cold cathode with a pulsed laser, then accelerated to 7 GeV with a linear accelerator and injected in the High-Energy Ring (HER). Positrons are produced by colliding electrons on tungsten; they are first injected in a damping ring to reduce their emittance, which is the average spread of particles in position and momentum phase-space (a small emittance corresponds to particles confined in a small region and having a similar momentum), then accelerated to 4 GeV with the linear accelerator and injected in the Low-Energy Ring (LER). When the circulating beams are sufficiently intense, they are brought to collision. Since luminosity degrades

due to various phenomena including beam-beam interactions, beam-gas interactions, and others. SuperKEKB currently operates in a ‘continuous injection’ regime, with refills of the beams every few seconds.

The collision products fly from the interaction point through the volume of the detector, where various features of the final-state particles (Fig. 2.4) are measured. To achieve high luminosities, a nano-beam, large crossing-angle collision scheme is implemented [38]. This is an innovative configuration based on keeping small horizontal and vertical emittance and large crossing angle, as shown in Fig. 2.2. This is obtained with a specially designed final-focus superconducting-quadrupole-magnet system, made of magnets, corrector coils, and compensation solenoids installed at each longitudinal end of the interaction region. Functionally the nano-beam scheme mimics a collision with many short micro-bunches, offering great advantages in luminosity with respect to previous schemes. The reduction of the luminous volume size to about 5% with respect to the predecessor KEKB, combined with doubling of beam currents, is expected to yield a factor 40 gain in intensity.

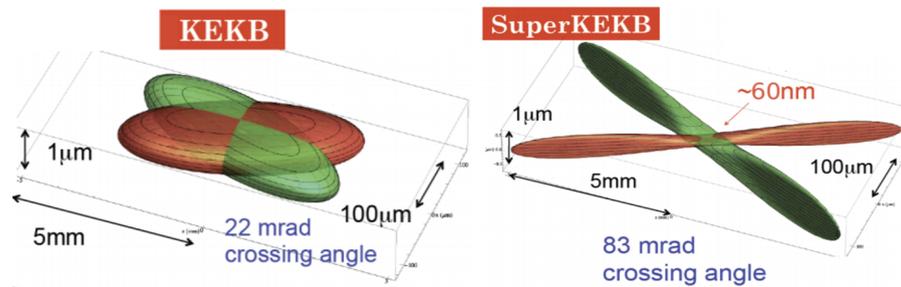


Figure 2.2: Two-dimensional sketch of the nano-beam mechanism implemented in SuperKEKB (right) compared with the previous KEKB collision scheme (left).

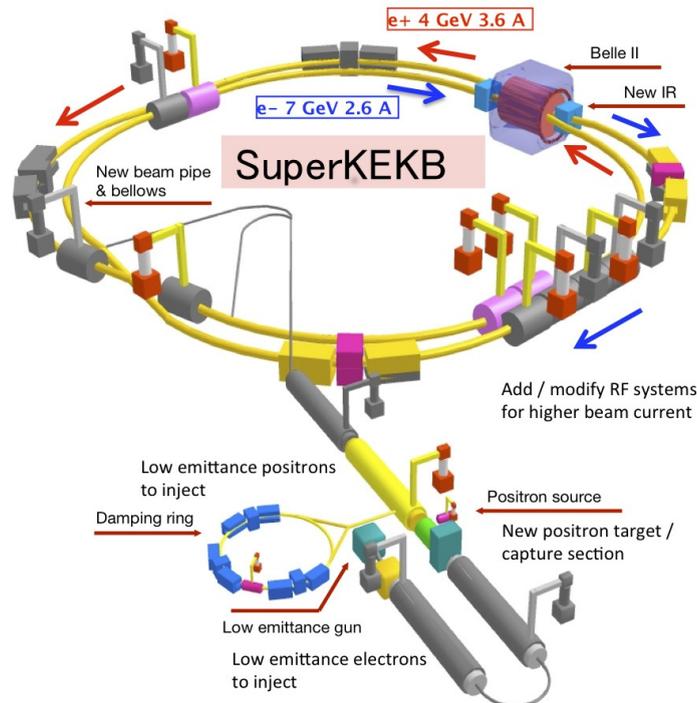


Figure 2.3: Scheme of the SuperKEKB collider.

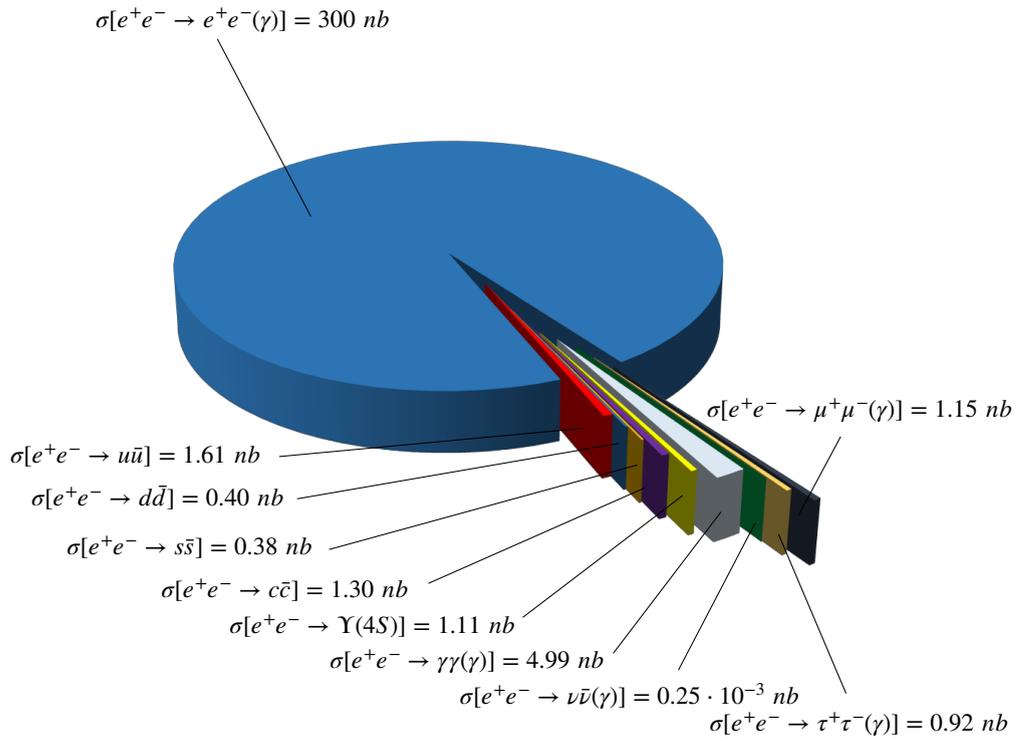


Figure 2.4: Cross sections of the main final states produced in $e^+ e^-$ collision at the $\Upsilon(4S)$ center-of-mass energy.

2.2 Overview of the Belle II detector

The Belle II detector is a major upgrade of its predecessor, Belle, and targets more ambitious physics goals. It is a system of multiple subdetectors, each optimized to reconstruct some specific features of the collision final-states, arranged in concentric layers forming an approximately cylindrical layout around the collision point of the SuperKEKB accelerator. It is about 8 meters in length, width, and height, and weights 1400 tons. While it is located in the same experimental hall and has a similar design to its predecessor, most of its subdetectors are new or considerably upgraded. A more efficient charged hadron identification is implemented for increased separation of final-state charged hadrons, reduction of backgrounds, and improved flavor-tagging; a better resolution in the reconstruction of the decay positions (vertices) of long lived particles is achieved, to enhance background suppression and sensitivity in measurements of decay-time dependent quantities; a more hermetic acceptance is achieved thanks to the smaller boost of the center of mass of the collision. These improvements are achieved through technologies designed to sustain the side-effects of higher SuperKEKB luminosity. These include high occupancies and radiation damages, associated with higher beam-induced backgrounds on the detectors closer to the beam pipe; an increased pile-up noise on the electromagnetic calorimeter; and more neutron hits on the muon detectors.

The main subsystems are summarized in Table 2.1 and shown in Figure 2.5. They are detailed in next sections and can be broadly classified as follows:

- **Detectors for charged particle tracking:** silicon pixel and strip detectors close to the beam pipe and a wire drift chamber, all immersed in a 1.5 T magnetic field parallel to the beam axis, are used for reconstruction of charged-particle trajectories (tracks).
- **Detectors for particle identification:** Cherenkov radiators, an electromagnetic calorimeter, and scintillators for muon and long-lived neutral hadrons achieve particle identification.
- **Data acquisition system:** a two-stage online trigger is designed to acquire interesting events at the high rates expected at design luminosities.

Purpose	Name	Technology	Configuration	Readout channels	Polar coverage (θ)
Enclose beams	Beam pipe	Beryllium	Cylindrical, inner radius 10 mm, 10 μm Au, 0.6 mm Be, 1 mm paraffin, 0.4 mm Be		
Tracking	PXD	Silicon pixel (DEPFET)	Sensor size: $15 \times (L1 \ 136, L2 \ 170)$ mm ² , Pixel size: $50 \times (L1a \ 50, L1b \ 60, L2a \ 75, L2b \ 85)$ μm^2 ; two layers at radii: 14, 22 mm	10M	[17°;150°]
	SVD	Silicon strip	Rectangular and trapezoidal, strip pitch: $50(p)/160(n) - 75(p)/240(n)$ μm , with one floating intermediate strip; four layers at radii: 38, 80, 115, 140 mm	245k	[17°;150°]
	CDC	He-C ₂ H ₆ drift chamber	14336 wires in 56 layers, inner radius of 160mm outer radius of 1130 mm	14k	[17°;150°]
Particle ID	TOP	RICH with quartz radiator	16 segments in ϕ at $r \approx 120$ cm, 275 cm long, 2cm thick quartz bars with 4×4 channel MCP PMTs	8k	[31°;128°]
	ARICH	RICH with aerogel radiator	2×2 cm thick focusing radiators with different n , HAPD photodetectors	78k	[14°;30°]
Calorimetry	ECL	CsI(Tl)	Barrel: $r = 125 - 162\text{cm}$, end-cap: $z = -102 - +196\text{cm}$	6624 (Barrel), 1152 (FWD), 960 (BWD)	[12.4°;31.4°], [32.2°;128.7°], [130.7°;155.1°]
Muon ID	KLM	barrel: RPCs and scintillator strips end-cap: scintillator strips	2 layers with scintillator strips and 12 layers with 2 RPCs	θ 16k, ϕ 16k	[40°;129°]
	KLM	end-cap: scintillator strips	12 layers of $(7-10) \times 40$ mm ² strips	17k	[25°;40°], [129°;155°]

Table 2.1: Summary of the Belle II components and specifics.

2.3 Tracking detectors

The innermost detectors are used for charged-particle reconstruction (tracking). Effective track reconstruction is of great importance since flavor-physics final states are mostly composed of charged particles and analyses rely strongly on precise measurements of their momenta and the decay positions of their long-lived parent particles. Precisely measured momenta and vertices allow separation of signal from backgrounds thanks to invariant-mass signal peaks that are narrower and therefore more distinctive from smoothly distributed backgrounds, and higher mean-valued decay lengths. Moreover, information on the decay time is key for measurements of \mathcal{CP} -violating asymmetries involving mixing. To simplify pattern recognition, tracks are first reconstructed in the outer tracking volume, and are then extrapolated into the innermost detector to define coarse regions of interest around their expected intersection points in the inner active layers. If an actual measurement point is found within the region of interest, the corresponding event is included in the pattern recognition algorithm, otherwise it is discarded.

2.3.1 Silicon-pixel vertexing detector

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

PXD sensors are based on depleted field-effect transistor technology [39]. They are made of p-channel MOSFET integrated on a silicon substrate, which is fully depleted by applying an appropriate voltage. Incident particles generate electron-hole pairs in the depleted region, and thus induce a current passing through the MOSFET. Sensors are $75\ \mu\text{m}$ thick, which allows on-pixel integration of most of the electronics.

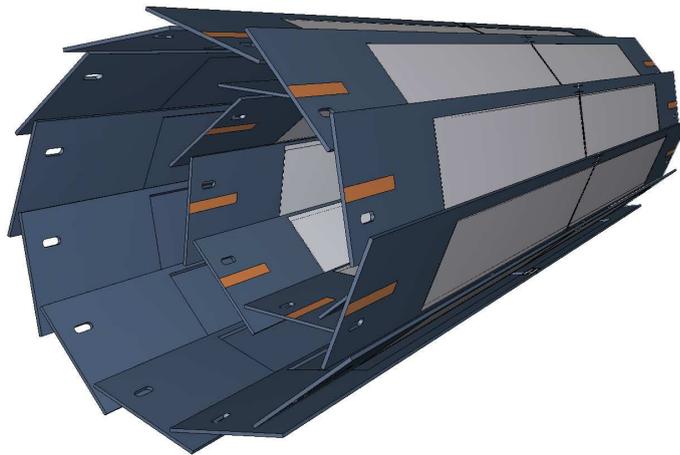


Figure 2.6: Scheme of the PXD detector geometry.

The PXD consists of two sets of rectangular layers arranged around the beam pipe on a cylindrical layout, at 14 and 22 mm radii: the PXD extends longitudinally by 174 mm at the radius of the outer layer. It comprises around 8 million pixels, $50 \times (50 - 55)\mu\text{m}^2$ (inner layer) and $50 \times (70 - 85)\mu\text{m}^2$ (outer layer) each. The polar acceptance ranges from 17° to 150° .

The transverse impact-parameter resolution is $12\ \mu\text{m}$, achieved by weighting the charge deposited in neighbouring pixels. At the moment only the inner PXD layer and part of the second layer is installed; the full outer layer will be added in 2022.

2.3.2 Silicon-microstrip vertexing detector

Around the PXD is the SVD, a silicon detector aimed at reconstructing decay vertices and charged-particle tracks at high spatial resolution [40].

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

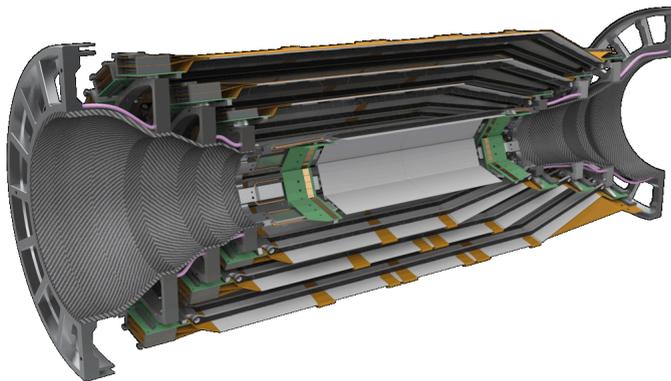


Figure 2.7: Exploded view of a SVD detector half.

As shown in Fig. 2.7, SVD has a polar-asymmetric geometry that mirrors the asymmetry in particle density resulting from the center-of-mass boost. The polar acceptance ranges from 17° to 150° . SVD is radially structured into four concentric layers at 39, 80, 104, and 135 mm, composed by, respectively, 7, 10, 12, and 16 independently readout modules arranged in a cylindrical geometry.

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

2.3.3 Central drift chamber

The CDC [41] is a drift chamber. It samples charged-particle trajectories at large radii, thus providing trigger signals for events containing charged particles, and it contributes to identification of charged-particle species by measuring their specific-ionization energy-loss (dE/dx).

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

The CDC has a hollow cylindrical geometry with an inner radius of 16 cm and an outer radius of 113 cm. The chamber is composed of 14336 $30\text{-}\mu\text{m}$ -diameter sense wires, divided in 56 layers, immersed in a gaseous mixture of 50% He and 50% C_2H_6 , while 42240 $126\text{-}\mu\text{m}$ -diameter aluminum wires shape the electric field. The azimuthal acceptance ranges

from 17° to 180° .

The single hit spatial resolution is about $100 \mu\text{m}$ and the dE/dx resolution is 11.9% for an incident angle of 90° . Figure 2.8 shows a reconstructed cosmic-ray track in the CDC. The typical transverse momentum resolution is $\sigma(p_T)/p_T^2 \approx 0.5\% / [\text{GeV}/c]$.

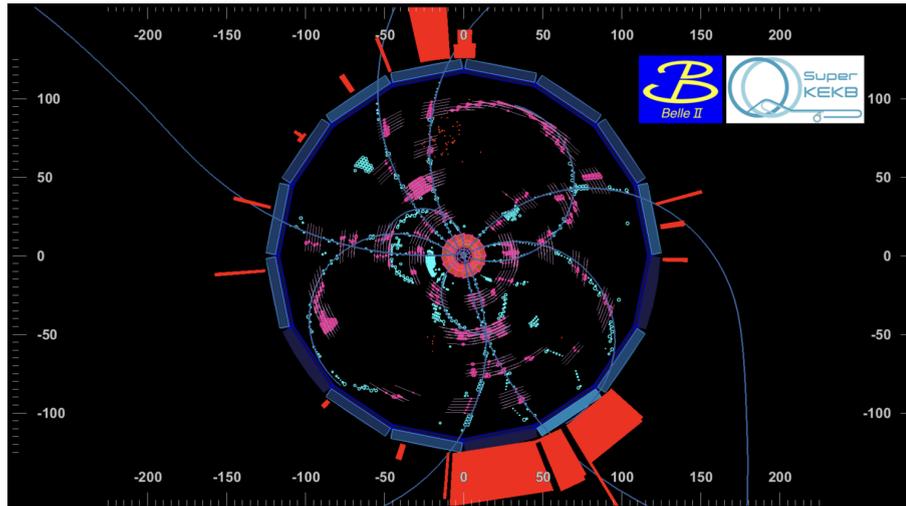


Figure 2.8: Example of event display of a typical hadronic event at Belle II. Curved tracks are reconstructed in the CDC.

2.4 Particle-identification detectors

Charged particle identification in the Belle II experiment is mainly performed by two detectors: the time of propagation counter (TOP) and the aerogel ring-imaging Cherenkov counter (ARICH). Both use Cherenkov light to identify charged particles. Particle-identification information is also provided by the electromagnetic calorimeter (ECL) and the K_L^0 and muon detector (KLM).

2.4.1 Time of propagation counter

The TOP counter is located in the barrel region. It measures the time of propagation of the Cherenkov photons produced by charged particles undergoing internal reflection in its quartz radiator. A three-dimensional image of the photon cone is reconstructed using the correlation between hits positions in the x-y plane and time of propagation. The TOP consists of 16 quartz bars mounted on the barrel at 1.2 m radius from the interaction point. Each bar is a photon radiator and has three main components (Fig.2.9): a long section that acts as a Cherenkov radiator, where photons are generated and propagate towards the bar end; a spherical mirror mounted on the forward end, which focuses the light and reduces the chromatic error; and a prism, mounted on the backward end of the bar, which collects and guides the photons to a photomultiplier.

The polar angular acceptance ranges from 31° to 128° . The single-photon time resolution is about 100 ps, providing a good separation of pions and kaons in the $0.4 - 4 \text{ GeV}/c$ momentum range (kaon identification efficiency is about 95%, pion fake rate is about 10%). This time resolution is achieved with a micro-channel plate photo-multiplier specially developed for this purpose.

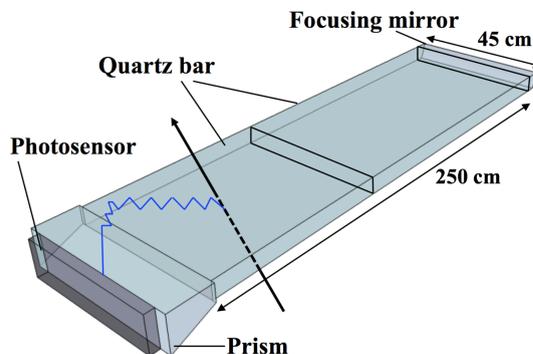


Figure 2.9: Sketch of a TOP quartz bar. A charged particle is shown passing through the radiator and emitting a Cherenkov photon.

2.4.2 Aerogel ring-imaging Cherenkov

Charged-particle identification in the forward end-cap is provided by the ARICH counter, which measures the Cherenkov ring produced by the passage of charged particles through a radiator. The ARICH provides discrimination between pions and kaons in a broad momentum range, and discrimination between pions, muons, and electrons below $1 \text{ GeV}/c$. When charged particles pass through the aerogel radiator, Cherenkov photons are produced; they propagate in an expansion volume where they form a ring on a photon sensitive surface made by position-sensitive photo-diodes. Photocathodes are used to convert photons in photoelectrons and generate electrical signals. As shown in Fig. 2.10, two aerogel radiators with different refraction indexes are used to increase the number of generated photons without degrading the Cherenkov-angle resolution [35].

The ARICH is composed of 420 modules for photon detection arranged in seven layers extending from 0.41 to 1.14 m radii, and by 248 aerogel tiles placed on the detector endcaps. The polar angular acceptance ranges from 14° to 30° . The observed ARICH performances allow for a 5σ separation between kaons and pions of $0.4 - 4 \text{ GeV}/c$ momenta, and a 4σ separation between pions, muons and electrons with momenta smaller than $1 \text{ GeV}/c$.

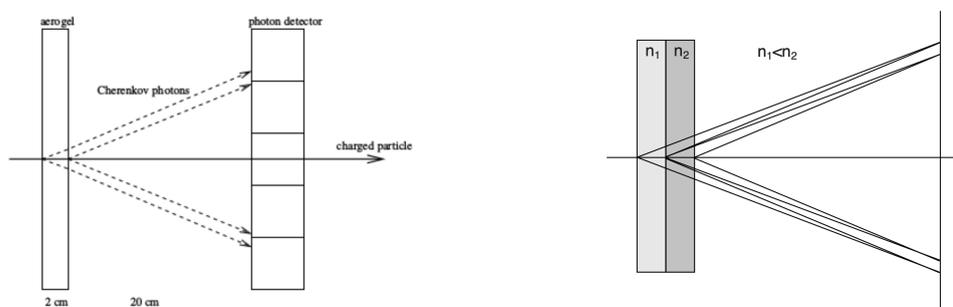


Figure 2.10: Scheme of the ARICH counter.

2.4.3 Electromagnetic calorimeter

The electromagnetic calorimeter (ECL) is used to detect photons over a wide energy range as well as to identify electrons and separate them from hadrons by extrapolation of CDC tracks. The ECL is a highly-segmented array of thallium-doped, cesium iodide crystals

assembled in a projective geometry pointing to the interaction region. Its detection principle is based on scintillation: the energy released by an incident particle causes a molecular excitation in the material with the passage of an electron from the valence band to the conduction band. The deexcitation of the electron to the valence band is associated with the emission of a photon, called scintillation light, that is usually inefficient and low-energetic. To improve the probability to emit a photon in the visible, Thallium impurities are added to create activator sites for the electrons with energy levels in the forbidden zone between the two bands. The scintillation light is then collected by phototubes. If the incident particle stops in the calorimeter, the total energy is measured. The CsI(Tl) crystals offer short scintillation time, which reduces the contamination of beam-background photons, which are often ‘out of time’ with respect to collision products. In the ECL, photons and electrons are identified through their kinematics, shower shapes and timing information, as they have different (in shape and magnitude) energy losses with respect to charged hadrons. To separate electrons from photons, information from tracking detectors is correlated with the ECL signal. The ECL consists of a 3-m-long barrel section with an inner radius of 1.25 m and annular endcaps at $z = 1.96$ m (forward) and $z = -1.02$ m (backward) from the interaction point. The polar angle coverage ranges from 12.4° to 155.1° .

The energy resolution ranges from $\sigma_E/E = 4\%$ at 100 MeV to $\sigma_E/E = 1.6\%$ at 8 GeV. The observed resolution for the reconstructed π^0 mass is $8 \text{ MeV}/c^2$. The ECL also allows for determining luminosity by measuring the Bhabha scattering rate and using its precisely known cross-section.

2.4.4 K_L^0 and muon detection system

The KLM detects muons and neutral particles that do not get absorbed in the inner detectors, such as K_L^0 [42]. It is made of alternating 4.7-cm-thick iron plates and active detector elements. Iron elements act also as magnetic flux returns for the tracking solenoid. In the inner layers, the active material is scintillator, while in the outer layers are glass-electrode resistive-plates chambers, with a gas mixture filling the space between electrodes. When particles traverse the KLM, they produce charges that are collected by applying an appropriate voltage. The barrel section of the detector covers 45° to 125° in polar angle. The endcaps cover 20° to 45° and 125° to 155° . Design reconstruction efficiency exceeds 80% for muons with momentum greater than 1 GeV/ c and K_L^0 with momentum greater than 3 GeV/ c .

2.5 Online event selection

Various processes occur in 10 GeV e^+e^- collisions (Fig. 2.4). Since the events of physical interest for the Belle II program make up for a small fraction of the total cross section, the online event selection (trigger) identifies them in real time while rejecting background events, to reduce the data-writing rate. The trigger must be efficient for recording hadronic event from $\Upsilon(4S) \rightarrow B\bar{B}$ and for the continuum to a manageable level, up to a maximum accept rate of about 30 kHz, due to data-acquisition restrictions. $B\bar{B}$ events have distinctive high-track multiplicity, and therefore are relatively straightforward to select; events containing τ decays are harder, since they have fewer tracks in the final state and can therefore be misclassified as electrodynamics backgrounds as $e^+e^- \rightarrow e^+e^-$ or $e^+e^- \rightarrow \mu^+\mu^-$ processes, that are not interesting for Belle II physics.

The trigger is composed by a hardware stage called Level1 (L1) [43] followed by a software high-level stage (HLT) [43]. The L1 decision is mainly based on informations

from CDC and ECL, but also TOP and KLM informations are available. The L1 decision feeds the global decision logic [43], that sends out the final trigger based on the informations it receives from the detector. The HLT is based on a more complete software reconstruction of the event similar to the offline reconstruction, that uses charged particles from the CDC and energy deposits in the ECL. The HLT selects events on the base of tracking multiplicity, vertex position, and total energy deposition, achieving a 30% event rate reduction without efficiency loss for $e^+e^- \rightarrow \text{hadron}$ processes.

2.6 Status of Belle II operations

An overview of the SuperKEKB and Belle II operations is in Table 2.2.

Phase	Period	Main goals	Detectors
Phase 1	February- July 2016	Accelerator tuning and background tests w/o final focusing	w/o Belle II, other detectors for background studies
Phase 2	March-July 2018	Verification of nano-beam, background study and physics data	w/ partial Belle II (partial PXD and SVD)
Phase 3	From March 2019	Complete physics program	w/ Belle II

Table 2.2: Summary of the three operation phases of SuperKEKB and Belle II.

First SuperKEKB operations started in 2016 with non-colliding low-intensity beams. The first collisions were achieved in April 2018 and reconstructed with a partially incomplete Belle II detector, with the goal of validating the nano-beam scheme while learning how to operate the detector with controlled backgrounds.

Then Belle II was completed with its silicon vertex detector and, since March 2019, the goal is to take physics data, steadily increasing the instantaneous luminosity. By June 2020, instantaneous luminosity has reached $2.2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, which is a record for a colliding-beam accelerator. In the coming years, the luminosity of SuperKEKB will be increased to reach approximately 40 times this value.

By summer 2020 the Belle II experiment has collected an integrated luminosity of 70 fb^{-1} . This sample, and a subset of it, are used for the analyses discussed in Chapters 3–6.

2.7 B decay reconstruction at Belle II

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

Figure 2.11 shows a sketch of a $B^+ \rightarrow K^+\pi^0$ decay. 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 $\Upsilon(4S)$ particle 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. 2.11. These events produce about ten tracks in

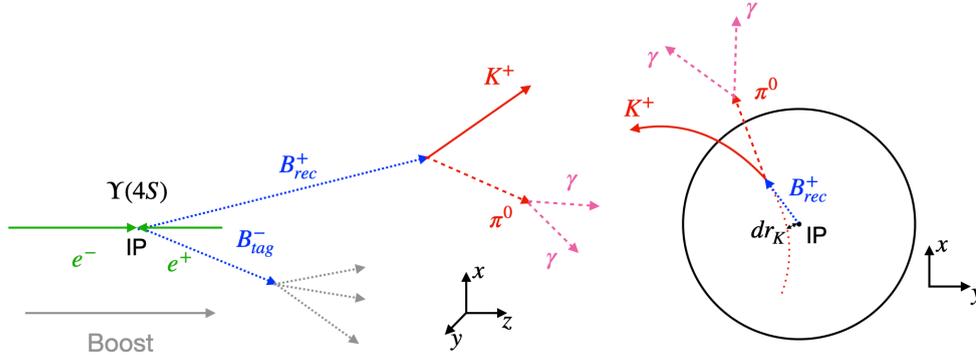


Figure 2.11: Sketch of the charmless decay $B^+ \rightarrow K^+ \pi^0 (\rightarrow \gamma\gamma)$ in (left) longitudinal and (right) transverse views. The particle labeled as B_{rec}^+ represents the signal B meson, while B_{tag}^- is the partner B -meson.

acceptance on average, which are used by the trigger to recognize hadronic events and save them onto permanent memory.

The B meson lifetime is 1.5 ps, which, combined with a typical momentum of 1.5 GeV/ c in the laboratory frame, results in a flight length of about 100–150 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 kaon and a neutral pion that have typical momenta of 2–3 GeV/ c in the laboratory frame. The trajectory of the kaon is curved by the Lorentz force produced by the axial magnetic field that pervades the tracking volume and it is reconstructed by a geometric fit of the 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 is reconstructed by summing the four-momenta of the two photons from its decay, which are in turn identified and reconstructed by using information from the electromagnetic calorimeter where their energy is recorded. The trajectory is inferred from the position of their impact on the calorimeter.

Reconstructed final-state particles are then used to determine the kinematic properties of the signal candidate of interest B_{rec}^+ . The four momenta of the final-state particles, along with the space points where they originate, are combined using constraints from momentum and energy conservation in a kinematic fit of the decay. Every combination of them that meets the reconstruction quality-requirements is a possible decay candidate.

However, decay candidates can either originate from genuine signal events or background events. Background events can in turn be originated from two broadly defined sources,

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

continuum, where light mesons such as pions, originated from light quark-antiquark pairs produced in the e^+e^- collision yield, form random combinations that accidentally meet the reconstruction and selection requirements.

To separate signal from background, Belle II exploits the available event information to construct discriminating variables at particle-, candidate-, and event-level.

2.7.1 Particle-level variables

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

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

Hit multiplicity. Informations 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 momentum, depends on the number of hits, but also false tracks from accidental combinations of unrelated hits are typically suppressed by hit-multiplicity requirements;

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 reconstructed track assuming true a mass hypothesis out of six possibilities (kaon, pion, electron, muon, proton and deuteron) and enhances discrimination against background, especially from B decays.

2.7.2 Candidate-level variables

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

χ^2 probability of the vertex fit. Frequently, multiple decay candidates corresponding to the various combinations of particles that meet the quality and reconstruction requirements are reconstructed in an event. Each candidate has an associated decay vertex fit. The corresponding χ^2 probability corresponds to a p -value, which may be used to choose the decay candidate most likely to correspond to the genuine signal.

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

The peculiar kinematic environment of B -Factory colliders provides additional constraints that further background separation. The $\Upsilon(4S)$ is produced at rest and decays in two same-mass particles, B and \bar{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,

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

$$M_{bc} = \sqrt{s/4 - |\vec{p}_B^*|^2}, \quad (2.1)$$

where \vec{p}_B^* is the momentum of the B meson in the center-of-mass frame reconstructed from the momenta of its decay products, and s is the squared center-of-mass energy (that is, the squared collision energy). While conceptually similar to the B invariant mass, M_{bc} is significantly more precise since the beam energy is known with much higher precision from the magnetic lattice of the machine than the reconstructed B energy. In the M_{bc} distribution, all B decays peak at the B meson mass, while non- B events have a smooth distribution that drops to zero at the kinematic limit of half of the collision energy. This makes M_{bc} powerful to separate B -events from $q\bar{q}$ continuum events.

- energy difference $\Delta E = E_B^* - \sqrt{s}/2$. This is the difference between the reconstructed B -candidate energy in the center-of-mass frame 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, signals peak at zero in the ΔE distribution, while continuum background follows a smooth distribution. In addition to discriminating against continuum, ΔE is particularly helpful in suppressing background from misidentified B decays. If a 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, and the candidate ΔE gets shifted from zero.

Figure 2.12 shows the bidimensional ΔE - M_{bc} distribution for simulated $B^0 \rightarrow K^+\pi^-$ decays and their one-dimensional projections. Continuum background is represented in blue, signal $B^0 \rightarrow K^+\pi^-$ events are represented in red, and B -background $B^0 \rightarrow \pi^+\pi^-$ events, where a pion is misidentified as a kaon, are represented in green. In the M_{bc} distribution, both $B^0 \rightarrow K^+\pi^-$ and $B^0 \rightarrow \pi^+\pi^-$ decays peak at the B mass, while in the ΔE distribution the $B^0 \rightarrow \pi^+\pi^-$ component is shifted from the signal peak by about 0.05 GeV.

M_{bc} and ΔE are the principal variables used in fits aimed at statistically separating signal from background at Belle II. Their complementary features are often combined in two-dimensional fits to achieve optimal discrimination.

2.7.3 Event-level variables

At a further level of reconstruction, global-event informations are used too in signal-from-background discrimination. Relevant examples include

Event-shape variables. Hadronic e^+e^- cross-sections are dominated by non- B events, so-called continuum background, consisting in production of light $q\bar{q}$ pairs (where q is u , d , c , s) that mostly yield pions and kaons. Because of the kinematic features associated with at-threshold $B\bar{B}$ 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 $B\bar{B}$ events from continuum. In a $B\bar{B}$ event, both B mesons are nearly at rest in the $\Upsilon(4S)$ frame. The B decay products are therefore emitted isotropically in that frame. In contrast, light quarks are produced with a comparatively large initial momentum due to their small mass compared to the total energy available in the collision. This results in a highly collimated fragmentation into two narrow back-to-back jets of light hadrons. Hence, the spatial distributions of $B\bar{B}$ decay products are approximately spherical, compared to pencil-like shapes for continuum. Information based on these distributions is therefore useful to discriminate

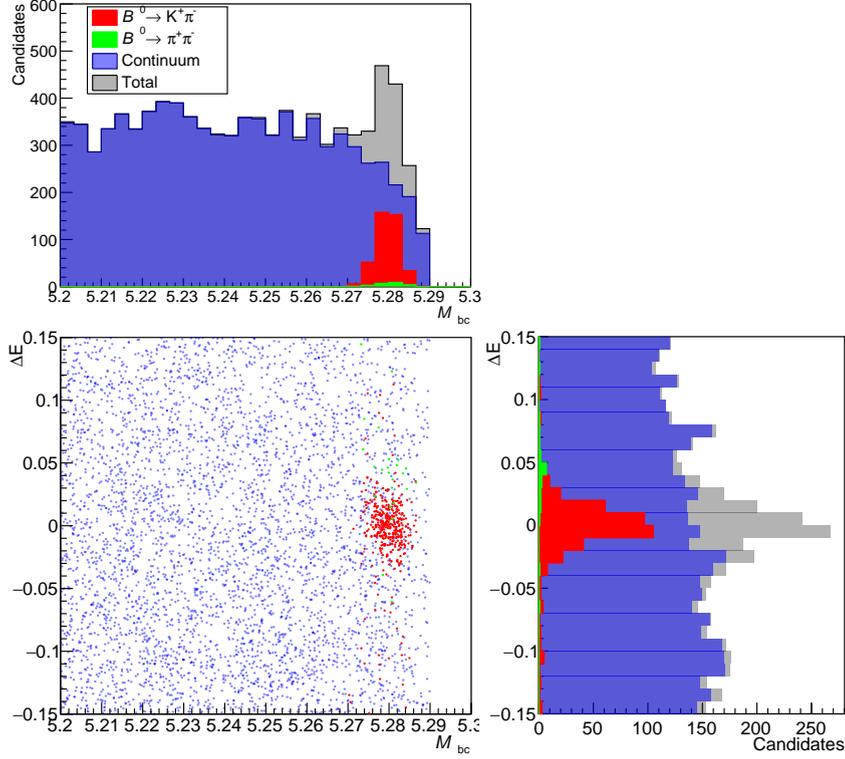


Figure 2.12: Distribution of ΔE as a function of M_{bc} for simulated $B^0 \rightarrow K^+\pi^-$ decays. Continuum background is represented in blue, $B^0 \rightarrow K^+\pi^-$ decays are represented in red, and $B^0 \rightarrow \pi^+\pi^-$ decays where a pion is misidentified as a kaon are represented in green.

signal events from continuum and it is exploited through several variables. Perhaps the most relevant example is the normalized second Fox-Wolfram moment [44],

$$R_2 = \frac{\sum_{m,n} |\vec{p}_m| |\vec{p}_n| P_2(\cos \theta_{mn})}{\sum_{m,n} |\vec{p}_m| |\vec{p}_n|}, \quad (2.2)$$

where \vec{p}_m and \vec{p}_n are the momenta of any pair of particles m and n , $P_2(\cos \theta_{mn})$ is the Legendre polynomial of second order in the cosine of the angle between \vec{p}_m and \vec{p}_n , and the sum runs over all the particles in the event. This variable captures the spatial distribution and energy flow in that event. Events with collimated jets, and therefore more likely to originate from continuum, show values of R_2 close to ≈ 0.5 , while $B\bar{B}$ events cluster at lower R_2 values (right panel of Fig. 2.13).

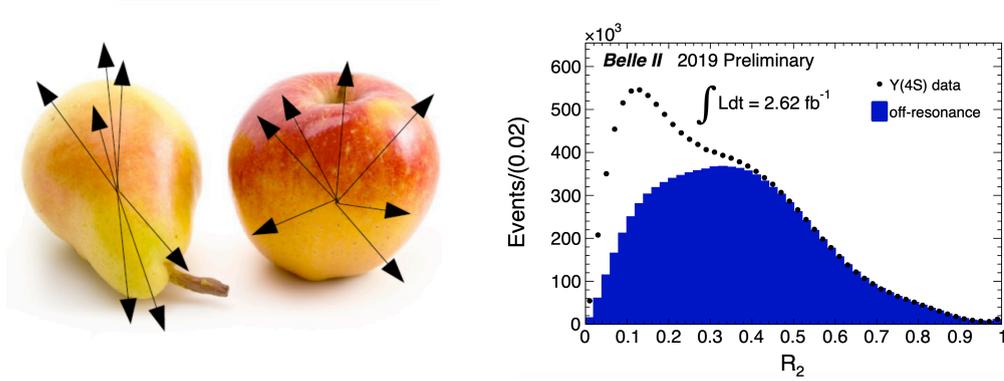


Figure 2.13: (Left panel) Pictorial representation of the different shapes of (left) a continuum background event and (right) a signal event. (Right panel) Distribution of the R_2 variable for (circles) $\Upsilon(4S)$ data and (histogram) off-resonance data. Off-resonance data are taken at an energy about 600 MeV below the $\Upsilon(4S)$ and are therefore composed purely of continuum.

Chapter 3

First reconstruction of the charmless decay $B^+ \rightarrow K^+ \pi^0$ at Belle II

This chapter reports the initial part of my work, which focuses on the first reconstruction of the $B^+ \rightarrow K^+ \pi^0$ decay in Belle II data.

3.1 Motivation and strategy

Charmless B decays play a key role in the Belle II scientific program as they provide essential information on flavor dynamics and its interplay with nonperturbative QCD.

Only recently Belle II has started collecting data, with luminosities that are still moderate; the chief priority in this initial stage is the validation and optimization of detector and reconstruction performance, an essential task where *known physics* processes are used as references to understand and optimize the performance of *not yet fully optimized* detector and reconstruction algorithms. The detector and the reconstruction software are both complex technical infrastructures whose optimal working points and performance depend on a plethora of configuration parameters. It is therefore key to study, understand, and possibly optimize the most relevant capabilities in early data to ensure optimal performance when the data set will enable world-leading results. Such capabilities include the reconstruction of charged and neutral particles, their identification, the reconstruction of vertices, flavor tagging, and others.

In Trieste we choose to use charmless B decays, which are our principal physics interest, to contribute to this collaboration-wide validation effort. The group targets the channels $B^0 \rightarrow K^+ \pi^-$, $B^0 \rightarrow \pi^+ \pi^-$, $B^+ \rightarrow K_S^0 (\rightarrow \pi^+ \pi^-) \pi^+$, $B^+ \rightarrow K^+ \pi^0 (\rightarrow \gamma \gamma)$, $B^+ \rightarrow K^+ K^- K^+$, and $B^+ \rightarrow K^+ \pi^- \pi^+$. These were chosen because they offer a comprehensive validation of most of the salient features associated with charmless physics, including reconstruction of charged particles, neutral particles, and K_S^0 mesons; continuum suppression; charged-hadron identification; beam-background suppression; and multibody B -decay modeling. In addition, they feature branching fractions sufficiently high ($\mathcal{O}(10^{-6})$ or larger) and final states sufficiently simple to allow straightforward reconstruction in the relatively small data set available by spring 2020.

My direct personal contribution focuses on the reconstruction of the $B^+ \rightarrow K^+ \pi^0$ decay, which is an interesting channel physics-wise as the current precision on its rate is among the limiting factors in the verification of the isospin sum rule (see Sec. 1.5.1.1) [30]. The $B^+ \rightarrow K^+ \pi^0$ channel is also important for validation purposes as it probes the details of high-energy π^0 reconstruction and charged-particle identification.

The $B^+ \rightarrow K^+\pi^0$ branching fraction is $(1.29 \pm 0.05) \times 10^{-5}$ [17] and its signature is a high-energy charged kaon and a pair of high-energy photons all originated by a common space point.

Multiple factors compound to make the $B^+ \rightarrow K^+\pi^0$ reconstruction challenging. The signal is rare and overwhelmed by combinatorial background from continuum 10^5 times more abundant. In addition, signal efficiency is affected by the challenges associated with π^0 reconstruction. These include the uncertainty in determining precisely the π^0 direction due to the finite size of the ECL cells that collect the final-state photons, and the uncertainty in determining the π^0 energy due to overlapping low-energy background photons, energy leakage across ECL cells and others.

An efficient discriminating selection is key. I first optimize a baseline (‘skimming’) selection on low-level quantities, capable of reducing the data set size while maintaining nearly full signal efficiency. Then I optimize the offline signal selection using a combination of higher-level discriminating variables. I finally extract the signal yield by fitting the distribution of the difference ΔE between expected and observed candidate energy.

I first develop and finalize all analysis choices on simulated data to avoid any experimenter’s bias in the final results. Then I apply the analysis on the Belle II data sample available in winter 2020, corresponding to 8.7 fb^{-1} , and compare the results with known references to gain insight on detector and reconstruction performance.

3.2 Data and tools

The experimental data consists in the whole Belle II data set collected at the $\Upsilon(4S)$ resonance in 2019 and corresponding to an integrated luminosity of 8.7 fb^{-1} , or equivalently, 10 million $B\bar{B}$ pairs.

I use generic simulated data to develop and finalize the analysis choices and to optimize the offline signal selection. These are Monte Carlo samples that include $e^+e^- \rightarrow B^0\bar{B}^0$ and $e^+e^- \rightarrow B^+B^-$ events generated using EvtGen [45], and $e^+e^- \rightarrow u\bar{u}$, $e^+e^- \rightarrow d\bar{d}$, $e^+e^- \rightarrow c\bar{c}$, and $e^+e^- \rightarrow s\bar{s}$ events generated using Pythia [46], passed through the standard Belle II simulation [47]. Events are combined in proportions according to their individual production cross-sections, for a total sample corresponding to an integrated luminosity of 50 fb^{-1} . In addition I simulate two million signal-only events for modeling the signal shape in the fit.

3.3 Baseline optimization

The key feature of B factories is high luminosity, a necessary condition to achieve the desired precision on the final measurements. The final Belle II sample is expected to contain about 50 billion $B\bar{B}$ events. The typical $B\bar{B}$ event size is about 100 KB, for a projected size of the raw data set of about 5 PB. 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 dimension and contain higher-level information related to physics primitives, including four-momenta, vertices, particle identification informations, and others. A second centralized step consists in applying loose selection criteria on the mdst data to obtain analysis-specific subsets reduced in dimension (5% of the mdst) that can be released to collaborators for detailed studies in each particular channel. This is called *skimming*. An

optimal skimming is important because any signal inefficiency incurred at this stage is laborious to recover in downstream analysis.

At the start of my work, the skimming selection for charmless decays was highly inefficient. It achieved significant sample-size reduction due to stringent selection criteria, at the price of a substantial degradation in signal efficiency, which was a factor of two lower than in Belle, the predecessor experiment. The fundamental issue was that, up to mid-2019, the skimming performance for charmless channels was configured based on the metric of *retention* only, with no regard for signal efficiency. The retention is the fraction of events where at least one signal candidate is reconstructed and corresponds to the fraction of mdst data that passes the skimming selection.

In principle, a proper skimming demands an effective compromise between keeping fully efficient signal reconstruction for downstream analysis and reducing effectively the data set size. To achieve this trade-off, I develop a new skimming concept that accounts for both efficiency and retention simultaneously.

3.3.1 Selection performance metric

My first contribution to Belle II was therefore to propose a new, more complete metric for charmless skimming performance. This is the correlation between signal efficiency and retention rate as functions of selection configurations, as displayed by the receiver operating characteristic curve (ROC). I then set on to choose the optimal quantities to use for an effective skimming.

3.3.2 Skimming selection quantities

The basic quantities available for skimming in an event are grouped into

- inclusive event variables (total number of tracks, total visible energy, etc.);
- variables associated with physics primitives (selections of tracks, selections of neutral particles, etc.);
- high-level variables (invariant masses, energies, event shape, etc.).

Mid-2019 skimming selections were based on stringent requirements on both physics primitives and high-level variables. This reduced retention down to $< 2\%$ but greatly degraded the signal reconstruction efficiency.

My strategy for the 2020 results is to focus on variables associated with physics primitives only, and attempt at achieving high efficiency and reasonable retention using those alone. I keep the possibility of using a mild requirement on R_2 (see Sec. 2.7.3) to control accept-rate as a safety measure should retention become an issue.

3.3.3 Charged-particle classes

Belle II central reconstruction provides three classes of tracks, based on the selection criteria they meet, *loose* tracks, *standard* tracks, and *tight* tracks.

Loose tracks undergo no selection. They are basically all results of a track fit, subtracted of duplications.

Standard tracks are selected by requiring full polar acceptance in the CDC, a minimum number of CDC hits, and a maximum displacement from the e^+e^- interaction-point

to suppress beam-backgrounds. In addition, particle misidentification is reduced by a loose PID criterion.

Tight tracks are standard tracks further restricted with a more stringent PID selection.

Details of the selections are shown in Table 3.1.

Variable	Loose	Standard	Tight
Inside CDC acceptance	-	Yes	Yes
dr	-	< 0.5 cm	< 0.5 cm
$ dz $	-	< 2 cm	< 2 cm
Number of hits in CDC	-	> 20	> 20
PID	-	> 0.002	> 0.1

Table 3.1: Selection criteria for each class of tracks. Variables are explained in Sec. 2.7.1

3.3.4 Charged-particle selection optimization

The skimming-optimization strategy is straightforward: I test various skimming configurations by varying track selections and R_2 requirements; then I determine the signal efficiency and background retention for each configuration using, respectively, 2×10^6 simulated signal-only events and generic simulation; I display the results in ROC curves to compare the performances of the various choices; and I finally choose the optimal skim configuration as that achieving the highest signal efficiency while maintaining a retention rate compatible with data-processing constraints ($\leq 10\%$).

The ROC curve in Figure 3.1 shows the results. The blue curve is obtained using loose tracks, the green curve corresponds to standard tracks, and the red curve to tight tracks. The yellow curve is obtained using the optimized selection proposed with this work. Every point represents the signal efficiency of the tested track choice as a function of its retention for a different selection on R_2 between 0.3 and 1.0 in 0.025 steps. Points to the (right) left correspond to (loose) stringent R_2 criteria.

Prior to this work, the default skimming configuration required tight tracks and a 0.5 upper limit on the event-shape variable R_2 . This corresponded to 74% $B^+ \rightarrow K^+\pi^0$ signal efficiency at 2% retention. Figure 3.1 shows that this was suboptimal. The curve corresponding to the tight tracks shows consistently equal or lower efficiency than at least one of the other curves at any given retention. Loose tracks achieve a nearly-full signal efficiency at the cost of a high retention rate ($\approx 10\%$). To reduce retention while maintaining approximately the same signal efficiency I explore some additional baseline criteria to be applied on loose tracks, and converge to the following requirements,

- polar angle in CDC acceptance, to ensure high-quality tracks fully reconstructed in the drift chamber;
- transverse (longitudinal) distance from the interaction point $dr < 0.5$ cm ($|dz| < 3$ cm), to remove tracks produced outside the interaction region, which are typically associated with particles from beam-background events that interact with the material surrounding the interaction region.

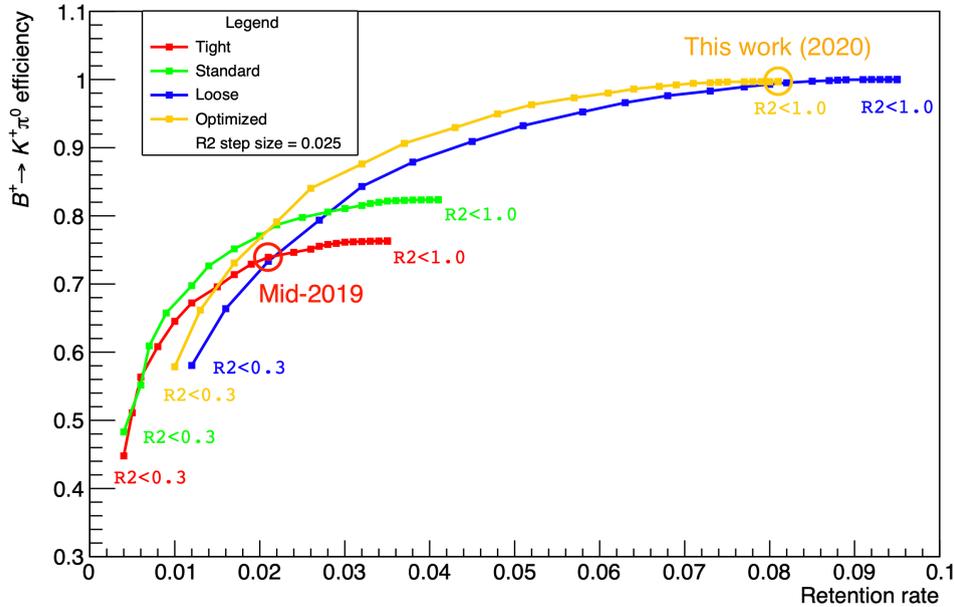


Figure 3.1: Signal efficiency as a function of retention rate for simulated $B^+ \rightarrow K^+\pi^0$ decays selected through various charged-particle selections. Loose tracks are displayed in blue, standard tracks in green, and tight tracks in red. Tracks with the optimized selection are displayed in yellow. Every square corresponds to an upper limit on the variable R_2 between 0.3 and 1.0 in 0.025 steps.

- χ^2 probability of the track fit greater than 0 to ensure fit convergence, and therefore that the track is more likely to come from an actual charged particle.

The corresponding optimized tracks are displayed in Fig. 3.1 in yellow: signal efficiency is approximately the same of loose tracks, but the retention rate is reduced.

Using the mid-2019 configuration as a reference (tight and $R_2 < 0.5$), an improvement in signal efficiency of 5% (absolute) at the mid-2019 retention is achieved by using the optimized criteria instead of the tight criteria on tracks. In addition, in this analysis I choose not to apply a selection on the variable R_2 , accepting a more generous retention rate ($\approx 8\%$) but achieving nearly-full signal efficiency, corresponding to a $\approx 30\%$ efficiency improvement with respect to the mid-2019 skimming.

3.3.5 π^0 selection optimization

As the $B^+ \rightarrow K^+\pi^0$ decay features a π^0 in the final state, I also explore an adequate π^0 baseline selection to be included in the skimming.

Prior to this work, the skimming selection for π^0 imposed criteria on low-level ECL variables for which a mismatch between simulated and experimental data was known to exist. For instance, the photon selection in mid-2019 included requirements on the timing of the arrival of photons in the ECL cells, which was mismodeled in simulation. An optimization of the π^0 selection based on simulated data would lead to strongly suboptimal results. I therefore choose to lift any requirement on variables that are not fully understood and perform this optimization using experimental data.

I use an abundant sample of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ decays, restricted in kinematics such that the π^0 momentum mirrors the π^0 momentum in the signal $B^+ \rightarrow K^+\pi^0$ decay.

The π^0 selection efficiency depends on the momentum of the pion as part of the π^0 requirements act on the energy deposited in the electromagnetic calorimeter by the associated photons. It is therefore important to restrict the determination of the optimized selection to neutral pions in a kinematic range similar to the signal π^0 kinematics for robust results.

I reconstruct $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ candidates by selecting only high-quality tracks reconstructed in the full polar acceptance of the central drift chamber through a converged fit. I require that the transverse (longitudinal) distance of the reconstructed tracks from the interaction point is $|dr| < 0.5$ cm ($|dz| < 3$ cm), to suppress tracks pointing outside the interaction region, which are typically associated with beam-background particles that interact with the material. I require the momentum of the neutral pion to meet $2.0 < p(\pi^0) < 3.5$ GeV and its transverse component to meet $1.0 < p_T(\pi^0) < 2.8$ GeV, to mirror the kinematics of signal π^0 . I also require a positive identification of the kaon to suppress misidentified $B^+ \rightarrow \bar{D}^0(\rightarrow \pi^+\pi^-\pi^0)\pi^+$ decays. Then I combine K^+ , π^- , and π^0 candidates to form \bar{D}^0 -meson candidates and suppress combinatorial background by keeping only those consistent with a \bar{D}^0 decay, which meet $1.84 < m(K^-\pi^+\pi^0) < 1.89$ GeV/ c^2 . I finally combine the \bar{D}^0 candidates with a charged pion to form the B -meson candidates, further enriched in signal by requiring the event-shape variable $R_2 < 0.3$ and $M > 5.27$ GeV/ c^2 .

I optimize the π^0 selection by maximizing the figure of merit (FoM) $S/\sqrt{S+B}$, where S is the signal yield and B is the background yield under the signal, as functions of different π^0 selection criteria.¹ These are based on various requirements on the photon energy recorded in the calorimeter and on the opening angles between photon directions. Higher thresholds on the photon-energy typically correspond to higher signal purities but lower efficiencies.

Topology and kinematic differences between $B^+ \rightarrow K^+\pi^0$ and $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ decays yield largely different S and B values in the two channels at any given π^0 selection. If not corrected, such differences would lead to a suboptimal $B^+ \rightarrow K^+\pi^0$ selection. To offset this, I use information from $B^+ \rightarrow K^+\pi^0$ decays and the known proportions between $B^+ \rightarrow K^+\pi^0$ and $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ rates. I determine the background yield B' in a sideband of the M_{bc} distribution of $B^+ \rightarrow K^+\pi^0$ decays, where only background events are expected, and extrapolate it to the expected background B under the M_{bc} signal. This is reliable because the distribution of continuum in M_{bc} has a known shape, represented by the blue contribution in Fig. 3.2. I determine S by scaling the observed $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ signal yield by the ratio of branching fractions $\mathcal{B}(B^+ \rightarrow K^+\pi^0)/\mathcal{B}(B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+) = 1.29 \times 10^{-5}/6.65 \times 10^{-4} \approx 0.02$.

I determine the $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ signal yields through extended maximum-likelihood fits on the unbinned ΔE distributions from multiple samples of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ candidates, each selected with different π^0 requirements (Fig. 3.3).

The observed asymmetric signal shape with a tail at negative ΔE values suggests a fit model based on the sum of a Gaussian distribution to reproduce the bulk, and an asymmetric distribution to reproduce the low- ΔE tail. The latter, so-called Crystal-Ball distribution, corresponds approximately to an asymmetric Gaussian distribution with a single tail [48],

$$f(x; \alpha, n, \bar{x}, \sigma) = N \begin{cases} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right), & \text{for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A\left(B - \frac{x-\bar{x}}{\sigma}\right)^{-n}, & \text{for } \frac{x-\bar{x}}{\sigma} \leq -\alpha \end{cases}, \quad (3.1)$$

¹The figure $S/\sqrt{S+B}$ offers minimum variance on the determination of a signal yield in a counting experiment of a narrow signal peak over a uniform background. Even if such simple model differs from our case, we assume $S/\sqrt{S+B}$ to be sufficiently optimal, as done in many other analyses.

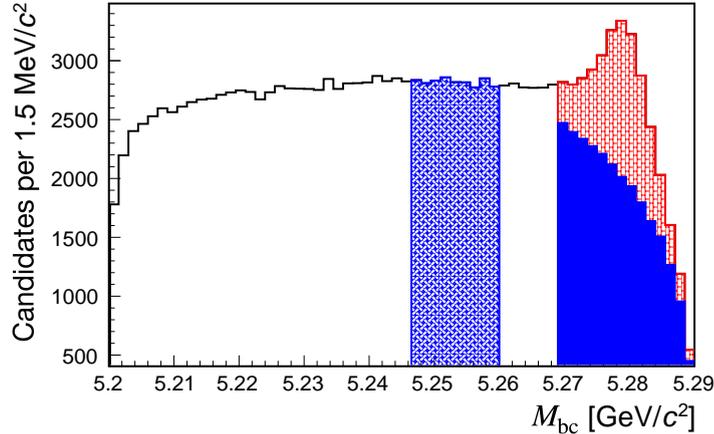


Figure 3.2: Distribution of M_{bc} for simulated $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ decays and definition of the sideband and signal region. Continuum background is represented in blue.

with $A = (n/|\alpha|)^n \exp(-|\alpha|^2/2)$, $B = n/|\alpha| - |\alpha|$, $N = [\sigma(C + D)]^{-1}$,
 $C = n/|\alpha|(1/n - 1) \exp(-|\alpha|^2/2)$, and $D = \sqrt{\pi/2} (1 + \operatorname{erf}(|\alpha|/\sqrt{2}))$. An exponential function models the background.

All distributions feature a large $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ peak at $\Delta E \approx 0$ overlapping an abundant, smoothly distributed background. The signal-to-background ratio at peak ranges from 6 for the most stringent selection to 1.2 for the loosest. The FoM values as functions of π^0 selections are shown in Fig. 3.4. The selection yielding the largest FoM corresponds to the following requirements:

- photon-energy deposited in one of the three calorimeter polar regions (barrel, forward end-cap, and backward end-cap) larger than 20 MeV, to exclude low-energy photons, which are more likely originated from light quark backgrounds and beam-backgrounds;
- diphoton mass $105 < m(\gamma\gamma) < 150 \text{ MeV}/c^2$ to suppress combinatorial background not associated with $\pi^0 \rightarrow \gamma\gamma$ decays.
- cosine of the angle between the π^0 momentum and the momentum of one of the associated photons smaller than 0.98. Combinatorial π^0 background is dominated by photon pairs formed by a high-energy photon (preferentially from genuine signal) and a low-energy photon from background (background photons have typically lower energies than signal photons). To match the π^0 mass, the high-energy photon in those events has to be energetic and therefore its momentum forms preferably a small angle with the π^0 momentum. A requirement that removes extreme values of angles reduces this contribution;
- mass of the π^0 candidate constrained to its known value.

3.4 $B^+ \rightarrow K^+\pi^0$ reconstruction

I reconstruct $B^+ \rightarrow K^+\pi^0$ candidates by selecting only high-quality tracks reconstructed in the full polar acceptance of the central drift chamber through a converged fit. I also require

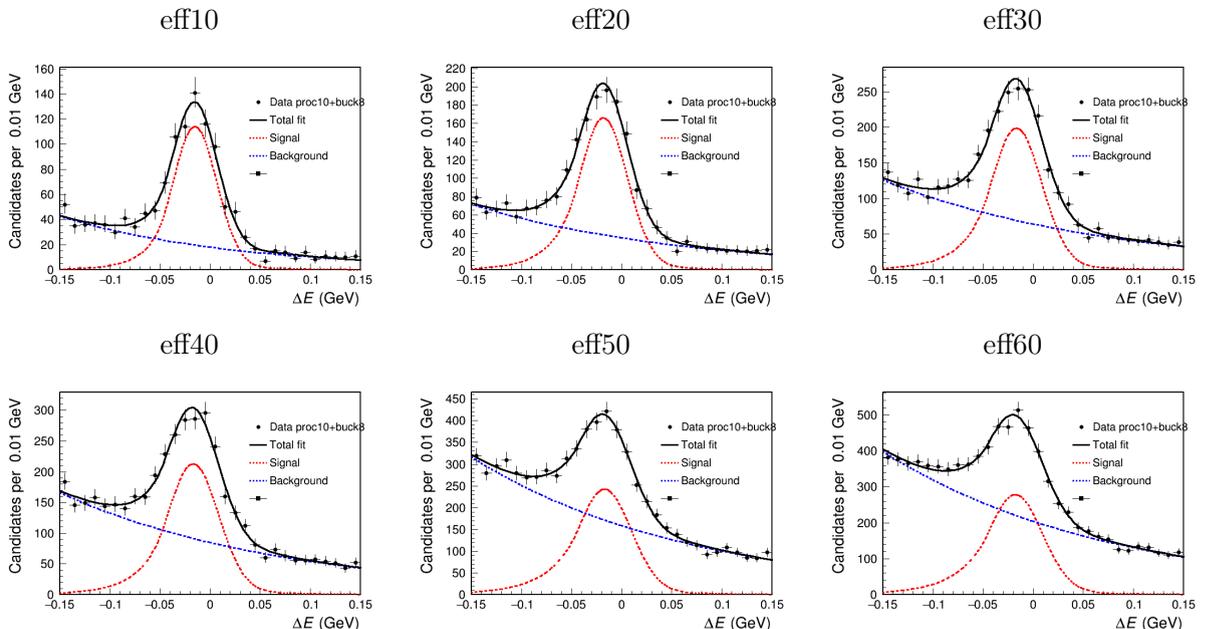


Figure 3.3: ΔE distribution of $B^+ \rightarrow \bar{D}^0 (\rightarrow K^+\pi^-\pi^0)\pi^+$ candidates reconstructed in 2019 Belle II data, corresponding to 8.7 fb^{-1} of integrated luminosity, for various π^0 selections; fits are overlaid.

that the transverse (longitudinal) distance of the reconstructed tracks from the interaction point is $|dr| < 0.5 \text{ cm}$ ($|dz| < 3 \text{ cm}$), to suppress tracks pointing outside the interaction region, which are typically associated with beam-background particles that interact with the material. I apply the optimized selection obtained in the previous section on the neutral pions. Then, I combine K^+ and π^0 candidates to form B -meson candidates, restricted to the signal region by requiring $M_{bc} > 5.27 \text{ GeV}/c^2$.

Figure 3.5 shows the resulting $m(\gamma\gamma)$ distribution for $B^+ \rightarrow K^+\pi^0$ candidates reconstructed in simulated and experimental data after the additional requirement $\Delta E < 0.3 \text{ GeV}$. A prominent signal peak, centered on the π^0 mass value and about 8 MeV wide, is visible over a small and smooth background.

The resulting ΔE distribution of simulated $B^+ \rightarrow K^+\pi^0$ candidates after applying the optimized requirements on the charged kaon and neutral pion is shown in Fig. 3.6. The expected signal is much smaller than the background, dominated by continuum. A further selection on high-level discriminating variables is needed to suppress it.

3.5 Continuum suppression

Many quantities related to the topology of the event, the features of the decay vertex, and the partner B -meson produced in the e^+e^- collision are available to discriminate signal from the predominant continuum background due to light quark-antiquark pairs. In addition, the presence of a kaon in the signal final-state offers further discrimination, since most of continuum is composed by pions.

Up to this point, the selection is based on the logic ‘and’ of unidimensional requirements imposed on individual variables. The challenging task of continuum suppression calls for a refined statistical approach to exploit maximally the available information in its full

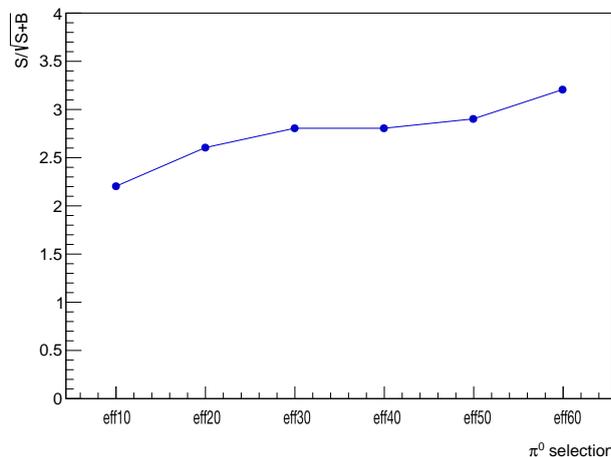


Figure 3.4: $S/\sqrt{S+B}$ for various π^0 selections.

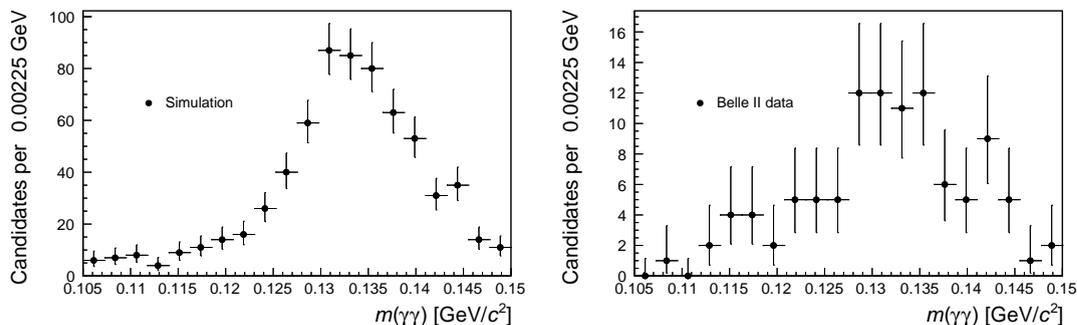


Figure 3.5: Diphoton mass distribution for $B^+ \rightarrow K^+\pi^0$ candidates reconstructed in (left) simulated and (right) 2019 Belle II data corresponding to 8.7 fb^{-1} of integrated luminosity.

multidimensionality. Belle II uses statistical-learning techniques to combine non-linearly multiple discriminating variables and exploit maximally their correlated informations, thus furthering background rejection.

Statistical-learning-based classification allows for calculating the probability of a given datum, defined as a set of discriminating-variable values and a class label (signal or background), to be a signal or background. In supervised statistical learning this involves (i) a training phase, which uses data-points sampled from a known (typically simulated) class (signal or background) to ‘fit’ the internal parameter values of the classifier model, and (ii) a classification phase, where the resulting classifier is applied to new data-points sampled from unknown classes for classification.

In Belle II a centralized software package to build one’s own fast binary boosted decision-tree classifier is available [49]. This combines non-linearly any number of user-defined discriminating input variables to output a scalar, called *continuum suppression* (CS).

A binary decision tree searches recursively for the best binary selections over a set of requirements applied to a set of discriminating input variables. Given a training data set, the classifier iterates recursively on all variables, and all requirements on each, to search for the combination of binary decisions that maximizes separation. The goal is to achieve the most accurate partition of the mixed-class sample in terms of signal-pure and background-

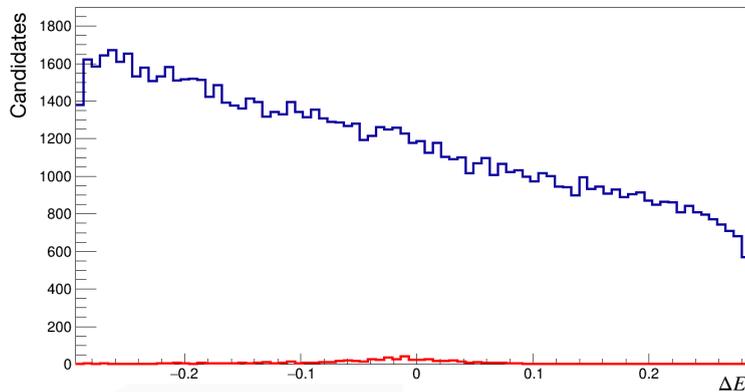


Figure 3.6: ΔE distribution of $B^+ \rightarrow K^+\pi^0$ candidates reconstructed in 50 fb^{-1} of generic simulated data after applying the optimized requirements on the charged kaon and neutral pion (with the additional selection $M_{bc} > 5.27 \text{ GeV}/c^2$). Signal events are represented in red, background events are represented in blue.

pure subsamples. The result is a set of hypercubic regions in the input variable space, each classified as either signal- or background-dominated.

The input variables available for $B^+ \rightarrow K^+\pi^0$ discrimination are related to various salient features of the events, including

Topology. Many input variables are sensitive to the topological shape of the event and provide discrimination of signal from continuum. These include 14 Fox-Wolfram moments [44], which express the energy and momentum flow of the event (see discussion in Sec. 2.7.3); CLEO cones, which include information about the number and momenta of charged particles reconstructed inside conic regions around the decay vertex [50]; thrust axis and thrust, that offer measures of the average momentum of a set of particles coming from a decay;²

Kinematics Global variables related to the kinematics of the event, such as transverse energy and missing mass, also offer discrimination.³ For instance, transverse energies in $q\bar{q}$ events are typically larger than in $B\bar{B}$ events, while the missing mass of an event is particularly important when dealing with decays involving neutrinos, such as semileptonic decays;

Flavor tagging Many variables used by algorithms dedicated to inferring the quark content of the partner B meson offer additional discriminating information: if such algorithms indicate that a partner B meson is reconstructed in an event, that in-

²In a $B\bar{B}$ event, the final states are produced almost at rest, and are therefore isotropically distributed in space. Their thrust axes are randomly distributed. The cosine of the angle between the B thrust axes follows therefore a uniform distribution in the range $[0, 1]$. In contrast, the momenta of particles in $q\bar{q}$ events align with the principal axis of the jet-like structures in the event. The resulting thrusts are collimated, yielding a cosine distribution strongly peaked at large values.

³The transverse energy is defined as the scalar sum of the transverse momentum of each particle, $\sum_{i=1}^N |p_{T,i}|$. The missing mass is defined as

$$M_{\text{miss}}^2 = \left(E_{\Upsilon(4S)} - \sum_{i=1}^N E_n \right)^2 - \sum_{i=1}^N |p_n|^2, \quad (3.2)$$

where E_n and p_n are the energy and momentum of particle n , respectively. The missing mass is a measure of the unreconstructed energy in an event, for example due to the presence of neutrinos.

formation obviously correlates strongly with the presence of a signal B meson and therefore helps suppressing continuum.

Because the classifier is trained using simulated events, if the distributions of the inputs are poorly simulated, the classifier training converges to an incorrect model and the ensuing signal-from-background discrimination is suboptimal. To ensure that only properly modeled inputs are used in the decision tree, I compare the distributions of all input variables in experimental and simulated data from $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ decays. I use $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ decays because they have larger signals than $B^+ \rightarrow K^+\pi^0$ and similar final states, making the input validation reliable for $B^+ \rightarrow K^+\pi^0$ decays too. Differences in input distributions between $B^+ \rightarrow K^+\pi^0$ and $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ are not expected to introduce significant limitations. I reconstruct $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ decays as discussed in Sec. 3.3.5 and compare the distributions for each input variable in data and simulation, separately for signal and background events. Background events are selected in the sideband $5.2464 < M_{bc} < 5.26 \text{ GeV}/c^2$. Signal events are those in the signal-rich region $M_{bc} > 5.27 \text{ GeV}/c^2$ and subtracted of the sideband (Fig. 3.2).

I train the classifier using 10^5 correctly reconstructed simulated signal events as a model for signal and 10^5 simulated continuum events ($u\bar{u}$, $d\bar{d}$, $c\bar{c}$ and $s\bar{s}$ in the expected proportions) surviving the baseline selection as a model for background.

General consistency exists in most inputs, but mild residual discrepancies are present. Figure 3.7 shows an example comparison for two input variables. The modified Fox-Wolfram moment (top panels) shows a typical case of good agreement; R_2 (bottom panels) shows the worse case of disagreement.

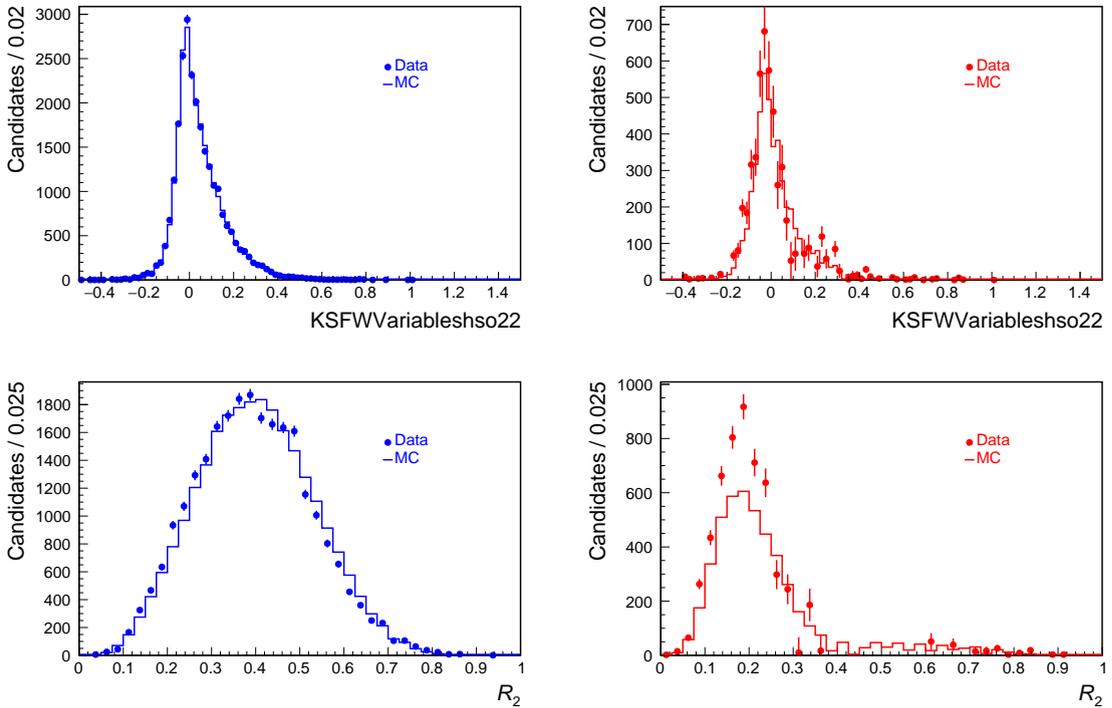


Figure 3.7: Distributions of (top) modified Fox-Wolfram moment and (bottom) R_2 in (circles) data and (solid) simulation for (left) background and (right) signal $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ decays.

Now that a subset of properly modeled inputs is identified, I use ROC curves to determine the final choice of input variables. Figure 3.8 shows the curves correlating signal efficiency and background rejection for various input-variable choices. I converge on a set of 39 variables, chosen for offering (i) reasonable agreement between data and simulation, (ii) significant discriminating power, and (iii) small correlations with the variables ΔE and M_{bc} used in subsequent analysis to fit the signal yield. The latter condition is important as a strong dependence between input variables and ΔE or M_{bc} would bias the final signal yields. The CS developed in this work outperforms the classifiers used in 2019: the red curve, corresponding to the CS used in this work, has the largest area under the curve and therefore the best performance, with a relative improvement of 2%.

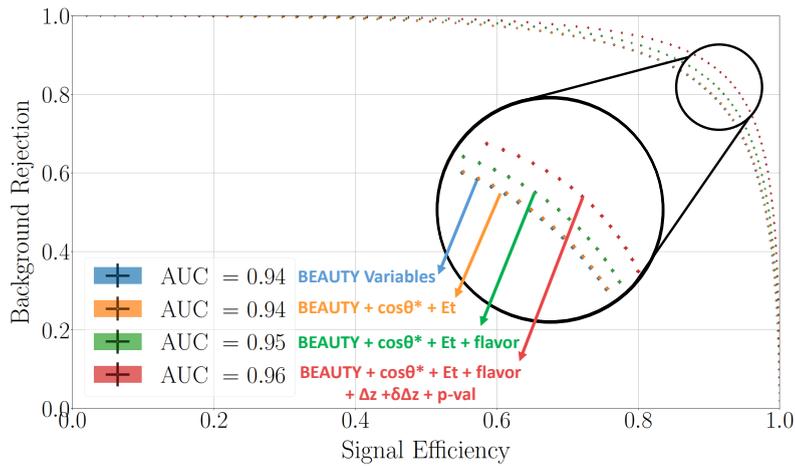


Figure 3.8: ROC curves and area under the curves (AUC) for various choices of input variables. The uppermost red curve corresponds to the CS configuration used in this analysis.

Figure 3.9 shows the CS-output distribution compared in simulated and experimental data for background and signal events. Both distributions show good agreement, thus confirming that the mild discrepancies observed in the input distributions do not invalidate the global picture.

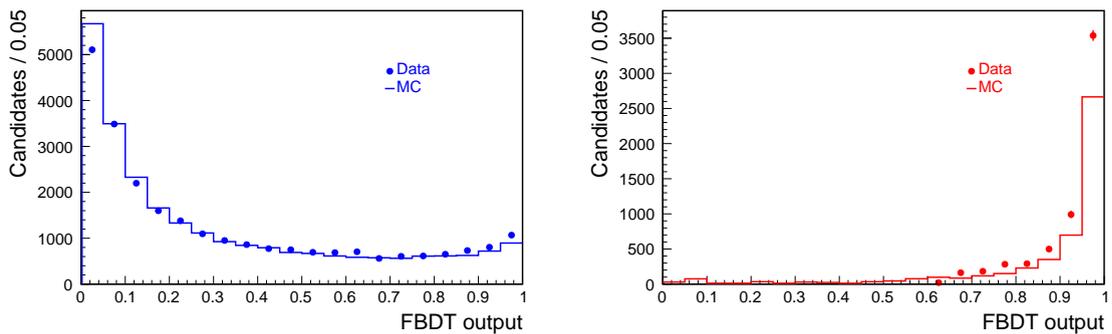


Figure 3.9: Distributions of the output of the CS classifier for (left) background events and (right) signal events in (circles) data and (solid) simulation.

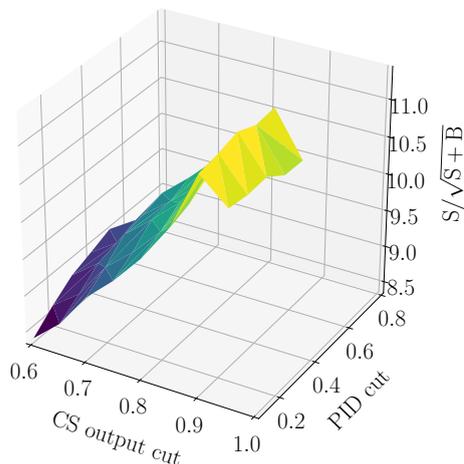


Figure 3.10: $S/\sqrt{S+B}$ as a function of the lower threshold on continuum suppression output and PID requirements for the $B^+ \rightarrow K^+\pi^0$ decay channel.

3.6 Offline selection optimization

Once established the discriminator, a selection needs to be defined.

I determine the offline selection by simultaneously optimizing selection criteria on the continuum suppression output and kaon identification. The latter is important to reduce backgrounds where pions are misidentified. As CS and PID can be correlated, a simultaneous optimization is expected to achieve better performance than independent optimizations.

I perform the optimization on generic simulated data to mimic the expected proportions of signal and background. I vary the selection criteria on both the CS output and kaon PID to maximize $S/\sqrt{S+B}$, where S is the signal yield and B is the background yield under the signal, as obtained from extended maximum likelihood fits of the unbinned ΔE distribution.

The $S/\sqrt{S+B}$ values for the various CS-PID selection configurations are shown in Fig. 3.10. High values of figure of merit are observed at stringent thresholds in continuum suppression (>0.9), independently from the PID selection. The optimized configuration for the $B^+ \rightarrow K^+\pi^0$ channel corresponds to a CS output larger than 0.9 and a kaon PID larger than 0.3.

The resulting ΔE distribution in simulation is shown in Fig. 3.11: a prominent signal peak populates the $\Delta E \approx 0$ region with a signal-to-background ratio of about three at the peak, overlapping a smooth decreasing exponential background.

3.7 Sample composition

The sample of $B^+ \rightarrow K^+\pi^0$ candidates resulting from the final selection receives, in principle, contributions from three main sources: signal, continuum background, and peaking backgrounds. Peaking backgrounds are candidates from $B\bar{B}$ events that peak close or under the signal in the ΔE distribution and are therefore particularly insidious in fits. According to known information about charmless B decay rates, the only possible peaking background is the $B^+ \rightarrow \pi^+\pi^0$ decay where a pion is misidentified as a kaon; however, we expect its contribution to be marginal due to the ten-fold suppression in branching fraction and the additional suppression due to misidentification. Nevertheless, simulation

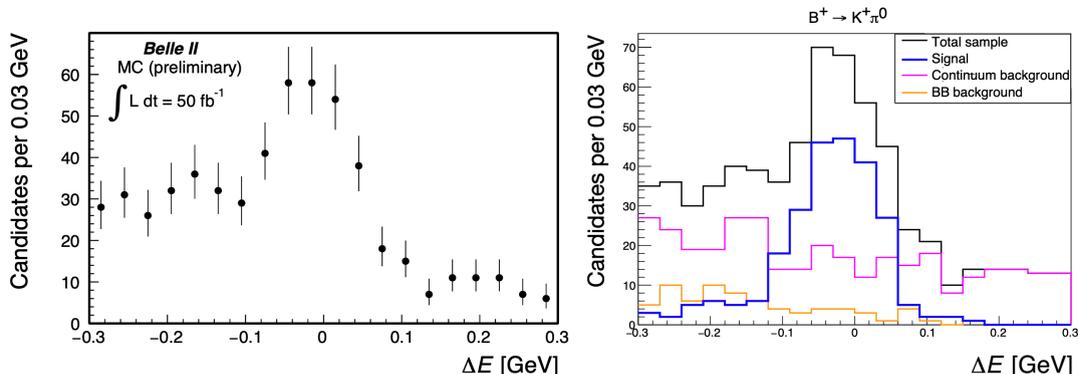


Figure 3.11: Distribution of ΔE for $B^+ \rightarrow K^+\pi^0$ candidates reconstructed in 50 fb^{-1} of simulated data and selected through the baseline and the optimized selection. Colors indicate leading contributions to the sample in the right panel.

offers a comprehensive independent way to check the expected composition of the sample. The left panel in Fig. 3.11 shows the expected individual contributions. The signal peak overlaps a smooth continuum background, which is naturally blended with a smoothly decreasing contribution from $B\bar{B}$ backgrounds. The $B\bar{B}$ background is devoid of any peaking component, and therefore is not a concern in fitting.

Simulation also shows that more than one candidate per event populates the ΔE distribution of $B^+ \rightarrow K^+\pi^0$ candidates. Candidate multiplicity is 1.05 per event. I then restrict the samples to include one candidate per event, by first selecting the π^0 candidate with highest p -value of the diphoton-mass fit, and then by selecting a single B candidate randomly.

3.8 Determination of signal yields in simulation

I determine the signal yields by performing an extended maximum likelihood fit of the unbinned ΔE distribution of the candidates restricted to the signal region $M_{bc} > 5.27 \text{ GeV}/c^2$. A two-dimensional joint ΔE - M_{bc} fit using a multidimensional model would make a more efficient use of the M_{bc} information in correlation with ΔE , with expected 5%–10% improvements in precision. However, a two-dimensional fit would also imply additional systematic uncertainties associated with the choice of the signal modeling in the two-dimensional ΔE - M_{bc} space, which may offset the modest advantage in statistical precision. For this first analysis I therefore choose to perform a simple one-dimensional fit.

I empirically determine the shape of the signal distribution from simulated signal events, using a sum of a Gaussian distribution and a Crystal-Ball distribution (see Sec. 3.3.5). The background model is also empiric and based on an exponential function with slope and yield determined by the fit.

Figure 3.12 shows the resulting ΔE distribution in generic simulation, with fit results overlaid. The signal yield is 191 ± 20 events in a simulated sample corresponding to 50 fb^{-1} , about six times larger than Belle II data sample available in winter 2020 for this analysis.

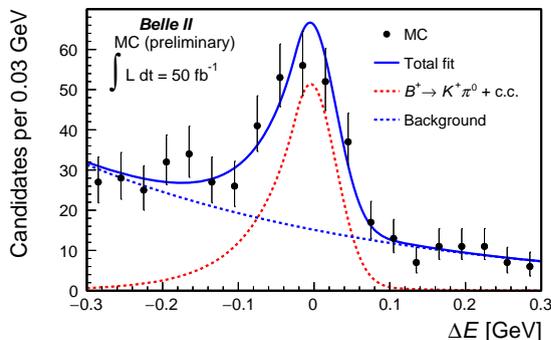


Figure 3.12: Distribution of ΔE for $B^+ \rightarrow K^+\pi^0$ candidates reconstructed in a simulated sample corresponding to 50 fb^{-1} of integrated luminosity and selected through the baseline plus continuum-suppression selection, further restricted to $M_{bc} > 5.27 \text{ GeV}/c^2$ with a kaon-enriching requirement of PID. The fit projection is overlaid.

3.9 Results

After having developed and performed the analysis on simulated data, I finally apply it to the sample of Belle II data available in winter 2020 and corresponding to 8.7 fb^{-1} of integrated luminosity.

To account for possible discrepancies between the simulated and experimental distributions, I check if the signal model may benefit from a few possible additional degrees of freedom that may accommodate expected mismodelings: a signal peak-shift, which accounts for possible miscalibrations in track reconstruction and energy measurements, and a signal width scale-factor, which is related to a possible mismatch of the energy resolution in data and simulation. I calculate the likelihood ratio $2 \ln(\mathcal{L}_1/\mathcal{L}_0)$, where \mathcal{L}_1 is the likelihood corresponding to the test model and \mathcal{L}_0 is the likelihood corresponding to the model fixed from simulation. In the limit of large sample size, the likelihood ratio follows a χ^2 distribution with a number of degrees of freedom corresponding to the number of additional free parameters in \mathcal{L}_1 [51]. Hence, assuming asymptotic regime, I associate p -values to the observed likelihood ratios. This indicates which additional degrees of freedom are favored by data and which would only generate overfitting.

The results, shown in Tab. 3.2, favor the addition of a peak-shift degree of freedom as free parameter in the fit in data.

Degree of freedom	p -value
Signal peak-shift	0.317
Signal peak-shift and width scale factor	0.604

Table 3.2: p -values for the alternate fit models.

Figure 3.13 shows the ΔE distribution for $B^+ \rightarrow K^+\pi^0$ decay candidates reconstructed in a Belle II data sample collected in 2019 and corresponding to 8.7 fb^{-1} of integrated luminosity, with fit results overlaid. A signal peak is observed in the $\Delta E \approx 0$ region, overlapping a smooth, decreasing background, with a signal-to-background ratio of about two at peak. The signal yield is of 27 ± 8 events, corresponding to a statistical significance of 3.95σ , obtained from the p -value associated with the likelihood ratio $2 \ln(\mathcal{L}_1/\mathcal{L}_0)$, where \mathcal{L}_1 is the likelihood corresponding to the fit model and \mathcal{L}_0 is the likelihood corresponding

to the model where the signal yield is fixed to zero. This is the first reconstruction of $B^+ \rightarrow K^+\pi^0$ decays in Belle II data. The observed yield corresponds to 3.1 ± 0.9 signal events per 1 fb^{-1} in data, in good agreement with the 3.8 ± 0.4 figure obtained in Monte Carlo. The signal peak is asymmetric as expected, with a negative- ΔE tail. The position of the peak seems shifted with respect to the position in Monte Carlo by $-21 \pm 12 \text{ MeV}$. This may indicate some mismodeling in the π^0 reconstruction, an hypothesis studied thoroughly in the next chapter.

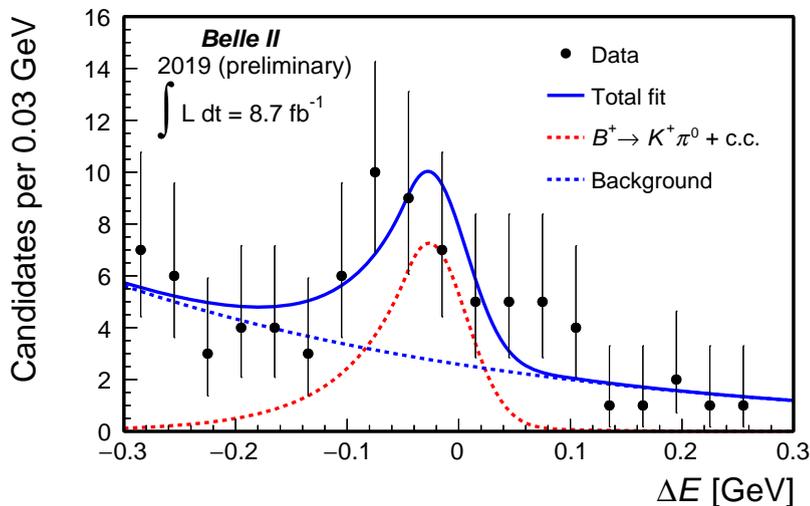


Figure 3.13: Distribution of ΔE for $B^+ \rightarrow K^+\pi^0$ candidates reconstructed in 2019 Belle II data corresponding to 8.7 fb^{-1} of integrated luminosity and selected through the baseline plus continuum-suppression selection, further restricted to $M_{bc} > 5.27 \text{ GeV}/c^2$ with a kaon-enriching requirement of PID. The fit projection is overlaid.

3.10 Summary and overview of other charmless channels

This chapter reports the first reconstruction of the charmless decay $B^+ \rightarrow K^+\pi^0$ in Belle II data. The data set corresponds to an integrated luminosity of 8.7 fb^{-1} and was collected by Belle II in 2019. Observed signal yield and purity are consistent with expectations from simulation showing a remarkable maturity of the detector and reconstruction performances.

This work is part of a broader effort toward the first reconstruction of charmless decays in Belle II data carried out by the Trieste group [6]. Part of the studies I developed for my analysis were also used in the reconstructions of the $B^0 \rightarrow K^+\pi^-$, $B^0 \rightarrow \pi^+\pi^-$, $B^+ \rightarrow K_S^0(\rightarrow \pi^+\pi^-)\pi^+$, $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$ channels. Table 3.3 reports a concise summary of all charmless results. Figure 3.14 shows the summed ΔE distributions of the 400 charmless decays reconstructed in Belle II 2019 data and presented at the Winter 2020 conferences.

In the next chapter I study in detail the observed ΔE shift of the $B^+ \rightarrow K^+\pi^0$ signal peak.

Decay	Yield data	Yield/fb ⁻¹ MC	Yield/fb ⁻¹ Data
$B^0 \rightarrow K^+\pi^-$	79 ± 11	7.4 ± 0.5	9.1 ± 1.3
$B^0 \rightarrow \pi^+\pi^-$	16 ± 5	1.6 ± 0.2	1.8 ± 0.6
$B^+ \rightarrow K_S^0\pi^+$	18 ± 5	1.7 ± 0.2	2.1 ± 0.6
$B^+ \rightarrow K^+K^-K^+$	92 ± 12	11.2 ± 0.6	10.6 ± 1.4
$B^+ \rightarrow K^+\pi^-\pi^+$	160 ± 19	20.2 ± 0.9	18.4 ± 2.2
$B^+ \rightarrow K^+\pi^0$	27 ± 8	3.8 ± 0.4	3.1 ± 0.9

Table 3.3: Summary of charmless yields obtained in 2019 Belle II data corresponding to 8.7 fb^{-1} of integrated luminosity.

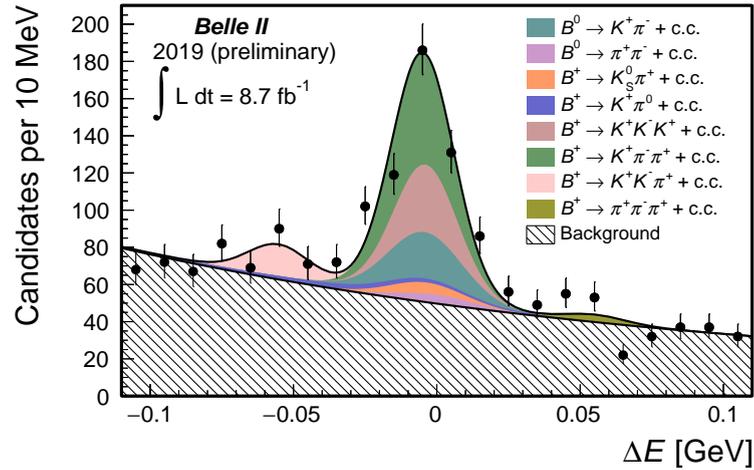


Figure 3.14: Sum of the ΔE distributions of charmless channels reconstructed in 2019 Belle II data, corresponding to 8.7 fb^{-1} of integrated luminosity.

Chapter 4

Observation of photon-energy miscalibration

This chapter reports my studies on the observed energy bias in the reconstruction of the $B^+ \rightarrow K^+\pi^0$ decay channel.

4.1 Motivation

My analysis of the channel $B^+ \rightarrow K^+\pi^0$ not only achieves the first reconstruction of this decay in Belle II data, but also exposes a possible issue in the detector calibration.

A -21 ± 12 MeV shift of the signal peak between experimental and simulated data is observed in the ΔE distribution, as shown in Figure 4.1.

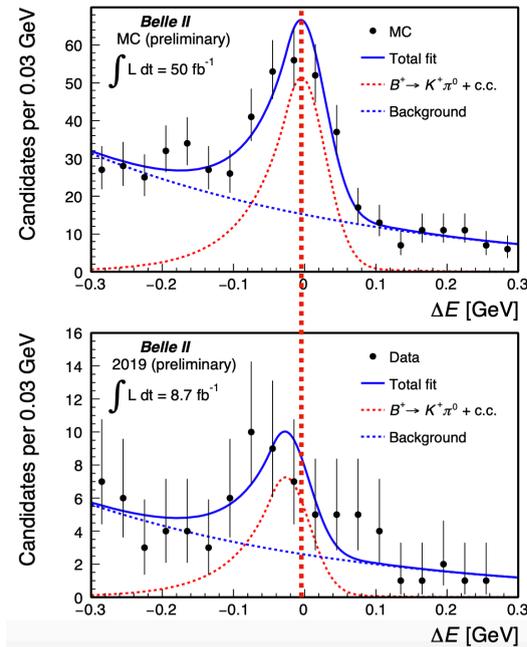


Figure 4.1: Distributions of ΔE for $B^+ \rightarrow K^+\pi^0$ candidates reconstructed in (top) a generic-simulation sample corresponding to 50 fb^{-1} of integrated luminosity and (bottom) the 2019 Belle II data sample, corresponding to 8.7 fb^{-1} of integrated luminosity.

While the statistical significance of the shift is marginal, the consequences of a real

effect would be so significant that a dedicated study is in order.

I therefore perform extensive studies to understand the possible causes of this potential shift. Causes of a mismeasurement of ΔE in $B^+ \rightarrow K^+\pi^0$ decays could be

- a miscalibration of charged-particle momentum, which would enter through the kaon;
- a mismeasurement of the beam energy;
- a miscalibration of the photon energy, which would enter through the π^0 ;
- any combination of the above.

As for the first effect, in fact Belle II reports a small mismatch between true and reconstructed charged-particle momentum in 2019–2020 data and provides an offline correction [52]. However, the lack of a ΔE shift in decays into charged particles only (for example $B^0 \rightarrow K^+\pi^-$ and $B^+ \rightarrow K^+K^-K^+$ [6]) suggests that this is unlikely to be the culprit. This is further confirmed by the permanence of the ΔE shift in $B^+ \rightarrow K^+\pi^0$ decays even after standard Belle II corrections to the momentum mismatch are applied.

Beam-energy mismeasurements have been shown to induce a bias in the M_{bc} measurement [53]. However, the associated shift in the ΔE distribution is ≈ 1 MeV, which is insufficient to explain by itself the 20 MeV effect possibly observed in our data.

Based on these considerations, I explore the hypothesis of a photon-energy miscalibration in the electromagnetic calorimeter.

The calibration of photon-energy at Belle II is achieved by calibrating the individual ECL cells using known-energy photons from $e^+e^- \rightarrow \gamma\gamma$ events (thanks to the precise knowledge of the collision energy) and by determining from simulation cell-specific corrections to compensate possible energy leakages. The calibration is then validated by inspecting the invariant mass in $\pi^0 \rightarrow \gamma\gamma$ and $\eta \rightarrow \gamma\gamma$ decays. The maximum observed discrepancy between experimental and simulated diphoton mass positions is 1% [54].

I therefore study the $m(\gamma\gamma)$ distribution in $B^+ \rightarrow K^+\pi^0$ decays to search for possible visible effects (Fig. 4.2). A large signal peak corresponding to reconstructed π^0 particles is present in both simulated and experimental data, overlapping a small uniform contribution due presumably to random photon pairs that accidentally meet the selection requirements. I determine the peak position through a maximum likelihood fit of the unbinned diphoton mass distribution. The signal shape is determined from simulation by an (empiric) Novosibirsk function [55], which is an asymmetric distribution with a long single tail,

$$P(x) = e^{-0.5[(\ln q_y)^2/\Lambda^2] + \Lambda^2}, \quad (4.1)$$

where $q_y = 1 + \Lambda(x - x_0) \times \frac{\sinh(\Lambda\sqrt{\ln 4})}{\sigma\Lambda\sqrt{\ln 4}}$ and σ and Λ are parameters related to the width of the signal and the shape of the tail. The parameter σ is fixed from Monte Carlo, while the peak shift and the shape of the tail are determined by the fit. I describe the background using a straight line with free slope.

A comparison of fit results in experimental and simulated data shows no obvious anomaly (Fig. 4.2). Both distributions are properly peaking at the π^0 mass, consistently with the findings of the photon-energy calibration group.

To further investigate the observed shift in ΔE , my strategy for exploring the hypothesis of a calorimetric-energy miscalibration is to identify an abundant control channel with a final-state neutral pion and monitor the shift size as a function of various aspects associated with π^0 reconstruction. In addition, I use analytical derivations and simulation to reproduce the observed phenomenology.

I choose $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ decays as a control channel, because they are more abundant than $B^+ \rightarrow K^+\pi^0$ decays, thus offering a more precise characterization of the effects at play.

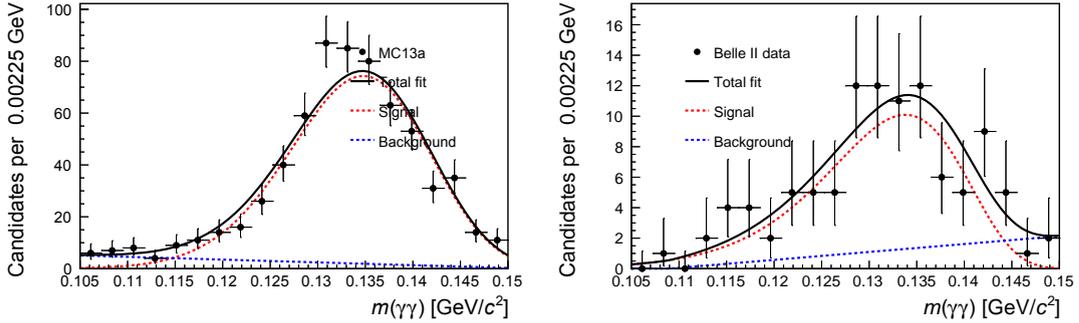


Figure 4.2: Diphoton mass distributions for $B^+ \rightarrow K^+\pi^0$ candidates reconstructed in (left) simulated data and (right) Belle II 2019 data \rightarrow corresponding to 8.7 fb^{-1} of integrated luminosity.

4.2 Studies on control channel

I restrict the kinematic properties of the $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ decays in such a way that the π^0 momentum mirrors the momentum of the π^0 from $B^+ \rightarrow K^+\pi^0$ decays. This allows for studying neutral pions with similar kinematic properties as those of the signal of interest, which may be relevant for a proper characterisation of the effect.

I reconstruct $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ decays as described in Sec. 3.4.4, using various selections to reconstruct π^0 candidates based on different requirements on the photon energy recorded in the calorimeter, ranging from 20 MeV to 200 MeV. When displaying ΔE distributions, I restrict candidates to the signal-rich region $M_{bc} > 5.27 \text{ GeV}/c^2$, while I restrict candidates to $|\Delta E| < 0.05 \text{ GeV}$ when displaying the M_{bc} distributions.

Figure 4.3 shows the ΔE and M_{bc} distributions in data and simulation for $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ candidates. In the ΔE distribution, a prominent peak at $\Delta E \approx 0 \text{ GeV}$ is visible, overlapping a smoothly decreasing background. Similarly, a narrow peak corresponding to $B\bar{B}$ events is present in the M_{bc} distribution at the B meson mass, while a uniform component due to continuum populates the whole range. Both distributions show evidence for a shift between data and Monte Carlo. The ΔE shift is $\mathcal{O}(10)$ MeV and negative, while the shift is smaller and positive for M_{bc} .

These findings confirm and consolidate conclusively the (statistically insignificant) indication from the $B^+ \rightarrow K^+\pi^0$ sample, calling for a systematic study on the nature and magnitude of ΔE biases as a function of various, possibly relevant quantities, such as π^0 selection, π^0 momentum, and R_2 selection. In addition to studying ΔE , where the effect is more evident, I also monitor M_{bc} and $m(\gamma\gamma)$ to form a more complete picture.

I fit the ΔE , M_{bc} , and $m(\gamma\gamma)$ distributions of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ samples selected with various criteria on π^0 selection, momentum, and R_2 selection. I model the shapes from simulation with the addition of an additional peak-position degree of freedom, which is meant to capture the shift under study.

The shifts observed as functions of π^0 selections are shown in Table 4.1. A negative 10 MeV shift in ΔE is observed consistently for all π^0 selections, while the M_{bc} distribution

is affected by a smaller ($0.6 - 0.7$ MeV) positive shift and $m(\gamma\gamma)$ shows no significant shift. These findings establish further strong evidence for an instrumental effect.

Table 4.2 shows the results of a similar study as a function of π^0 momentum, indicating a positive correlation between the average π^0 momentum and the size of the ΔE bias. No indication of a dependence on the R_2 selection used to reconstruct the sample is observed, thus confirming that the effect is unlikely to be an artifact of higher-level selection biases and it is likely to be generated in the calorimeter reconstruction.

	eff20	eff30	eff40	eff50	charmless
ΔE shift (MeV)	-10 ± 1	-10 ± 1	-10 ± 1	-10 ± 1	-11 ± 1
M_{bc} shift (MeV/c^2)	0.66 ± 0.11	0.58 ± 0.10	0.59 ± 0.10	0.61 ± 0.09	0.68 ± 0.10
$m(\gamma\gamma)$ shift (MeV/c^2)	-0.3 ± 0.2	-0.7 ± 0.2	-0.7 ± 0.3	-0.8 ± 0.3	-0.9 ± 0.3

Table 4.1: Observed ΔE , M_{bc} , and $m(\gamma\gamma)$ peak shifts for $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ candidates reconstructed in 2019 Belle II data corresponding to 8.7 fb^{-1} of integrated luminosity for various choices of π^0 selections.

π^0 momentum (GeV/c)	< 0.5	0.5 - 1.5	> 1.5
ΔE shift (MeV)	-8 ± 2	-9 ± 1	-14 ± 3

Table 4.2: ΔE shift observed in $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ candidates reconstructed in 2019 Belle II data corresponding to 8.7 fb^{-1} of integrated luminosity as a function of π^0 momentum.

Now that the evidence for an effect is compelling, I try to corroborate the hypothesis of photon-energy miscalibration with the aid of simple analytical calculations and simulation.

4.3 Analytical derivation

I study the impact on the ΔE , M_{bc} , and $m(\gamma\gamma)$ distributions of a putative instrumental bias of the calorimetric measurement of the photon energy in a decay $B \rightarrow \pi^0 X_{\text{chrg}}$, where

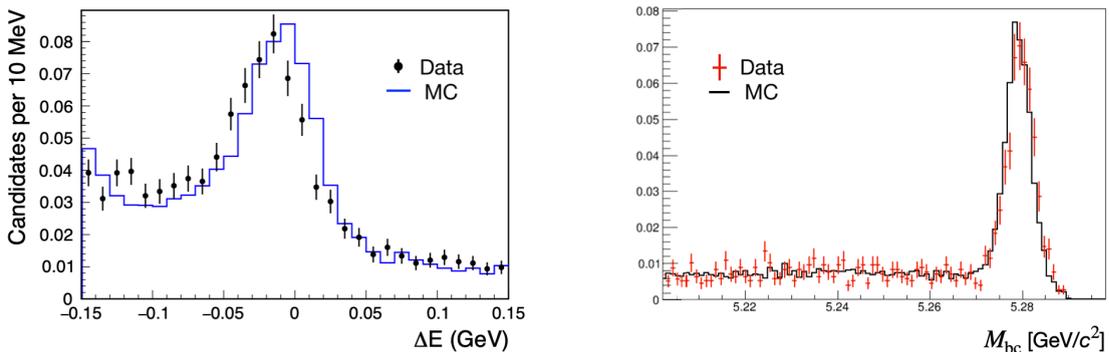


Figure 4.3: Distributions of (left) ΔE and (right) M_{bc} for $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ candidates reconstructed in (histogram) generic simulation and (markers) Belle II 2019 data corresponding to 8.7 fb^{-1} of integrated luminosity, normalized to each other.

X_{chrg} is a generic system of charged decay products.

In a $\pi^0 \rightarrow \gamma_1 \gamma_2$ decay, where (E_i, \vec{p}_i) is the energy-momentum four-vector of the photon γ_i , the ΔE distribution of B candidates would peak at zero if the observed photon energy is an unbiased estimate of the true photon energy,

$$\Delta E = E_B^* - \sqrt{s}/2 \approx (E_1 + E_2 + E_{\text{chrg}}) - \sqrt{s}/2 \approx 0, \quad (4.2)$$

where $(E_{\text{chrg}}, \vec{p}_{\text{chrg}})$ is the energy-momentum four-vector associated with X_{chrg} and all quantities, in the above and following equations, are evaluated in the center-of-mass frame.

If the observed photon energy $E'_i = E_i - \delta^*$ is affected by a bias $-\delta^*$, the reconstructed photon energy in the center-of-mass frame is also affected by a bias $-\delta$, and the observed $\Delta E'$ distribution is biased toward negative values,

$$\Delta E' = E_B^{*'} - \sqrt{s}/2 \approx (E_1 - \delta + E_2 - \delta + E_{\text{chrg}}) - \sqrt{s}/2 \approx -2\delta. \quad (4.3)$$

Comparison of ΔE and $\Delta E'$ shows that the energy difference is a specially sensitive probe of photon-energy biases, which enter *linearly* in determining the position of the signal peak.

Let us now examine what would be the impact of the bias on the M_{bc} and $m(\gamma\gamma)$ distributions. For the $m(\gamma\gamma)$ distribution, the unbiased case leads to

$$\begin{aligned} m(\gamma_1 \gamma_2) &= \sqrt{(E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2} = \sqrt{E_1^2 + E_2^2 + 2E_1 E_2 - (E_1^2 + E_2^2 + 2E_1 E_2 \cos \theta)} \\ &= \sqrt{2E_1 E_2 (1 - \cos \theta)}, \end{aligned} \quad (4.4)$$

where θ is the opening angle between photon directions. If a negative bias affects the photon-energies, the diphoton mass becomes

$$\begin{aligned} m'(\gamma_1 \gamma_2) &= \sqrt{(E'_1 + E'_2)^2 - (\vec{p}'_1 + \vec{p}'_2)^2} = \sqrt{E_1'^2 + E_2'^2 + 2E'_1 E'_2 - (E_1'^2 + E_2'^2 + 2E'_1 E'_2 \cos \theta)} \\ &= \sqrt{2E'_1 E'_2 (1 - \cos \theta)} = \sqrt{2[E_1 E_2 - \delta(E_1 + E_2) + \mathcal{O}(\delta^2)](1 - \cos \theta)} \\ &\approx \sqrt{2E_1 E_2 (1 - \cos \theta)} \left(1 - \frac{\delta}{E}\right) \\ &= m(\gamma_1 \gamma_2) \left(1 - \frac{\delta}{E}\right). \end{aligned} \quad (4.5)$$

The impact on the $m(\gamma\gamma)$ distribution is suppressed proportionally to δ/E , which makes harder the detection of a miscalibration using this observable.

A similar result holds for the M_{bc} distribution. In the absence of a photon-energy bias one gets

$$\begin{aligned} M_{\text{bc}} &= \sqrt{\frac{s}{4} - (\vec{p}_{\text{chrg}} + \vec{p}_{\pi^0})^2} = \sqrt{\frac{s}{4} - (\vec{p}_{\text{chrg}} + \vec{p}_1 + \vec{p}_2)^2} \\ &= \sqrt{\frac{s}{4} - (p_{\text{chrg}}^2 + p_1^2 + p_2^2 + 2|\vec{p}_1 \vec{p}_2| \cos \theta + 2|\vec{p}_{\text{chrg}} \vec{p}_1| \cos \phi_1 + 2|\vec{p}_{\text{chrg}} \vec{p}_2| \cos \phi_2)} \\ &= \sqrt{\frac{s}{4} - (p_{\text{chrg}}^2 + E_1^2 + E_2^2 + 2E_1 E_2 \cos \theta + 2|\vec{p}_{\text{chrg}}| E_1 \cos \phi_1 + 2|\vec{p}_{\text{chrg}}| E_2 \cos \phi_2)}, \end{aligned} \quad (4.6)$$

where ϕ_1 and ϕ_2 are the opening angles between the directions of the system of charged

decay products and each of the photons. A negative photon-energy bias implies

$$\begin{aligned}
 M'_{bc} &= \left[\frac{s}{4} - (p_{\text{chrg}}^2 + E_1'^2 + E_2'^2 + 2E_1'E_2' \cos \theta + 2|\vec{p}_{\text{chrg}}|E_1' \cos \phi_1 + 2|\vec{p}_{\text{chrg}}|E_2' \cos \phi_2) \right]^{1/2} \\
 &= \left[\frac{s}{4} - (p_{\text{chrg}}^2 + (E_1 - \delta)^2 + (E_2 - \delta)^2 + 2(E_1 - \delta)(E_2 - \delta) \cos \theta \right. \\
 &\quad \left. + 2|\vec{p}_{\text{chrg}}|(E_1 - \delta) \cos \phi_1 + 2|\vec{p}_{\text{chrg}}|(E_2 - \delta) \cos \phi_2) \right]^{1/2} \\
 &= \left[\frac{s}{4} - (p_{\text{chrg}}^2 + E_1^2 + E_2^2 + 2E_1E_2 \cos \theta + 2|\vec{p}_{\text{chrg}}|E_1 \cos \phi_1 + 2|\vec{p}_{\text{chrg}}|E_2 \cos \phi_2) \right. \\
 &\quad \left. + \delta(2E_1 + 2E_2 + 2(E_1 + E_2) \cos \theta + 2|\vec{p}_{\text{chrg}}| \cos \phi_1 + 2|\vec{p}_{\text{chrg}}| \cos \phi_2) + \mathcal{O}(\delta^2) \right]^{1/2} \\
 &\approx \left[\frac{s}{4} - (p_{\text{chrg}}^2 + E_1^2 + E_2^2 + 2E_1E_2 \cos \theta + 2|\vec{p}_{\text{chrg}}|E_1 \cos \phi_1 + 2|\vec{p}_{\text{chrg}}|E_2 \cos \phi_2) \right. \\
 &\quad \left. + \delta(2E_1 + 2E_2 + 2(E_1 + E_2) \cos \theta + 2|\vec{p}_{\text{chrg}}| \cos \phi_1 + 2|\vec{p}_{\text{chrg}}| \cos \phi_2) \right]^{1/2} \\
 &= M_{bc} \left(1 + \frac{\delta}{2M_{bc}^2} (2E_1 + 2E_2 + 2(E_1 + E_2) \cos \theta + 2|\vec{p}_{\text{chrg}}| \cos \phi_1 + 2|\vec{p}_{\text{chrg}}| \cos \phi_2) \right) \\
 &= M_{bc} \left(1 + \frac{\delta}{M_{bc}^2} \left((E_1 + E_2)(1 + \cos \theta) + |\vec{p}_{\text{chrg}}|(\cos \phi_1 + \cos \phi_2) \right) \right). \tag{4.7}
 \end{aligned}$$

It is not immediately intuitive to visualize the qualitative features of M'_{bc} , since they depend on the masses of the charged particles contributing to X_{chrg} , on their momenta, and on the decay-angle configuration. I perform a simple simulation to calculate the value of

$$f = (E_1 + E_2)(1 + \cos \theta) + |\vec{p}_{\text{chrg}}|(\cos \phi_1 + \cos \phi_2), \tag{4.8}$$

for all kinematically allowed configurations. The resulting distribution of f is shown in Fig. 4.4 normalized by $5.279 \text{ GeV}/c^2$, the expected M_{bc} value for signal. The value of f/M_{bc} is nonnegative and peaked at about one. M'_{bc} is always larger than M_{bc} , and its distribution peaks at $\approx M_{bc}(1 + \delta/M_{bc})$.

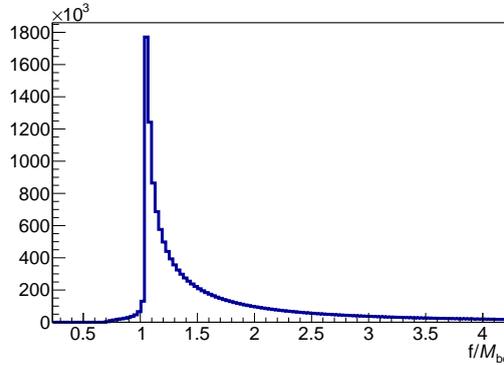


Figure 4.4: Distribution of the term f in Eq. (4.8) divided by $5.279 \text{ GeV}/c^2$, the expected M_{bc} for signal events for all possible physical configurations.

Even under our coarse approximations, the kinematic relationships show that findings on data are compatible with our initial hypothesis of a photon-energy-bias $-\delta$. The impact is linear in δ for the ΔE distribution, linear in the suppressed δ/E term for the diphoton mass, and linear in the suppressed term δ/M_{bc} for M_{bc} .

	$\delta = 0$ MeV	5 MeV	10 MeV	15 MeV	20 MeV
ΔE shift (MeV)	0	-9.7 ± 0.2	-19.3 ± 0.2	-29.0 ± 0.2	-38.8 ± 0.2
M_{bc} shift (MeV/ c^2)	0	2.4 ± 0.1	4.9 ± 0.1	7.4 ± 0.1	9.8 ± 0.1
$m(\gamma\gamma)$ shift (MeV/ c^2)	0	-0.8 ± 0.1	-1.5 ± 0.1	-2.1 ± 0.1	-2.7 ± 0.1

Table 4.3: Observed shifts in ΔE , diphoton mass, and M_{bc} as functions of the input bias δ in signal-only simulated $B^+ \rightarrow K^+\pi^0$ decays.

It is therefore not surprising that the bias went undetected prior to my work on $B^+ \rightarrow K^+\pi^0$ decays. On the one hand standard Belle II calibrations monitor calorimeter energy by using π^0 and η invariant masses, where biases are suppressed. In addition, no ΔE shift was observed in B decays with final-state neutral pions because they were $B \rightarrow D(\rightarrow \pi^0 X)h$ decays reconstructed by using a constraint on the D mass in the kinematic fit. The ‘mass constraint’ is a feature of the candidate-reconstruction algorithm used when an intermediate-resonance participates in the decay (D in this case). It corrects the four-momenta of the intermediate resonance decay products so as their invariant mass matches the known mass of the intermediate particle, thus achieving a more precise B decay reconstruction. However, these corrections reduce or remove the ΔE shift thus ‘hiding’ the effect.

4.4 Simulation

I further support the analytical findings using events simulated using standard Belle II simulation.

I reconstruct candidates in simulated signal-only $B^+ \rightarrow K^+\pi^0$ decays after subtracting from the reconstructed photon-energies in the center-of-mass frame 0–20 MeV biases, which are similar in size to the typical energy-shifts observed in data. This allows for checking independently from previous findings the impact of this bias on the ΔE , $m(\gamma\gamma)$, and M_{bc} distributions. The results, summarized in Table 4.3, are consistent with the analytical derivations. They show a ΔE shift linearly proportional to the size of the photon-energy bias, as expected. The table allows estimating the actual size of the bias in the ECL, which is approximately -5 MeV in the center-of-mass frame. The value in the lab frame is similar, due to the small center-of-mass boost produced in SuperKEKB collisions.

4.5 Summary

The first reconstruction of $B^+ \rightarrow K^+\pi^0$ decays, performed in this work, exposes a photon-energy miscalibration in the electromagnetic calorimeter. This effect went previously unidentified as it was either suppressed in the standard Belle II control variables or removed by mass constraints in B decay fits. Following my detailed investigation the issue has been acknowledged by the collaboration, which decided to include the ΔE distribution of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^+$ decays as an additional, sensitive control variable in the standard set of calorimeter monitoring informations.

Chapter 5

Studies of charged-hadron identification

This chapter reports my studies of charged-hadron identification aimed at optimizing the charmless B reconstruction performance.

5.1 Motivation and strategy

Charged-hadron identification is a crucial background-discrimination capability in most charmless physics analyses because final states involve kaons and pions, which are easily misidentified with each other when using kinematic information only. Data collected by dedicated and generic subdetectors are processed and combined by reconstruction algorithms to obtain information about the identity of the reconstructed charged particles. This information is typically encapsulated into a particle identification quantity, which is directly used in physics analysis as a variable subjected to a background-suppressing selection, or to a fit.

The goal of this part of my work is an in-depth study of the performance of the current Belle II PID information for B physics measurements in Belle II. I use pure kaon and pion samples from $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays reconstructed in Belle II data to compare the discriminating information provided by the individual subdetectors with the information from their combination.

5.2 Sample and tools

To calibrate K^\pm and π^\pm identification, samples of kaons and pions that are pure, and unbiased with respect to the PID information under study, are needed. I use an abundant sample of $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays reconstructed from Belle II 2019–2020 data and corresponding to an integrated luminosity of 34.6 fb^{-1} . No Belle II simulation is used.

The $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ channel has a sufficiently high branching fraction ($\approx 2.6\%$) 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 is 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 \bar{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 . On top of that, the large magnitude of V_{cs} (≈ 0.97) cause the D^0 meson to decay into a $K^-\pi^+$ pair much more

frequently than into a $K^+\pi^-$ pair, which is a doubly Cabibbo-suppressed transition with a branching fraction one hundred times smaller. Hence, identification of the charge of the pion coming directly from the $D^{*\pm}$ provides identification of the D flavor, which in turn allows identifying the pion and the kaon based on their charges, in a totally independent way from any dedicated PID information.

In addition, 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. This also provides effective background discrimination. The invariant mass of the D^0 allows for suppressing combinatorial two-track background. Most importantly, further suppression comes from the peak in the mass difference between the D^{*+} and the D^0 meson, which is narrow because common resolution effects cancel in the difference. 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 PID performance.

I reconstruct $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays by applying a loose selection on the quality of tracks and vertices and then combine the final-state candidates in kinematic fits consistent with the desired decay topology. I select only high-quality tracks reconstructed in the SVD and in the CDC by demanding a minimum number of measurement hits; the transverse (longitudinal) distance of the reconstructed tracks from the interaction point is $|dr| < 2$ cm ($|dz| < 4$ cm), to suppress tracks pointing outside the interaction region, which are mainly associated with beam-background particles that interact with the material.

I impose that the $K\pi$ mass is consistent with a D^0 decay, $1.84 < m(K^-\pi^+) < 1.90$ GeV/ c^2 , and the mass difference with a $D^{*+} \rightarrow D^0\pi^+$ decay, $0.144 < m(K^-\pi^+\pi^+) - m(K^-\pi^+) < 0.147$ GeV/ c^2 . These requirements suppress combinatorial background and ensure that D^0 mesons come from D^{*+} decays. The D^{*+} momentum in the center-of-mass frame is restricted to exceed 2.5 GeV/ c to suppress $B \rightarrow D^*X$ decays, where the available energy for the D meson is typically lower. Selecting only D^{*+} particles from $e^+e^- \rightarrow c\bar{c}$ continuum simplifies the analysis. These selections choices are the result of maximizing $S/\sqrt{S+B}$, where S is the signal yield and B is the background yield observed in data.

Figure 5.1 shows the effect of the selection on the $m(K^-\pi^+\pi^+) - m(K^-\pi^+)$ and $m(K^-\pi^+)$ distributions. The candidates shown in the $m(K^-\pi^+)$ distribution are restricted in $m(K^-\pi^+\pi^+) - m(K^-\pi^+)$ and viceversa. The selection achieves a signal efficiency of 70% and a background retention of less than 5% with respect to our baseline starting point, yielding high-purity kaon and pion samples for PID studies.

5.3 Particle identification at Belle II

Particle identification at Belle II is achieved by combining information from several subdetectors. The trajectories of charged particles reconstructed by the tracking detectors, PXD, SVD, and CDC, are extrapolated outward to the TOP, ARICH, ECL and KLM detectors, where geometric matching between the tracks and observed signal clusters is attempted. Offline reconstruction associates to each matching track PID-detector information sensitive to its identity. For example, the drift chamber output encodes information on the specific ionization energy loss associated with each track. The raw information is further processed to provide higher-level quantities that are more convenient for usage in analysis. These are typically ‘likelihood’ values associated to the track. For each of six possible mass hypotheses, kaon, pion, electron, muon, proton and deuteron, the likelihood expresses the probability to observe the reconstructed PID information if the mass hypothesis was true.

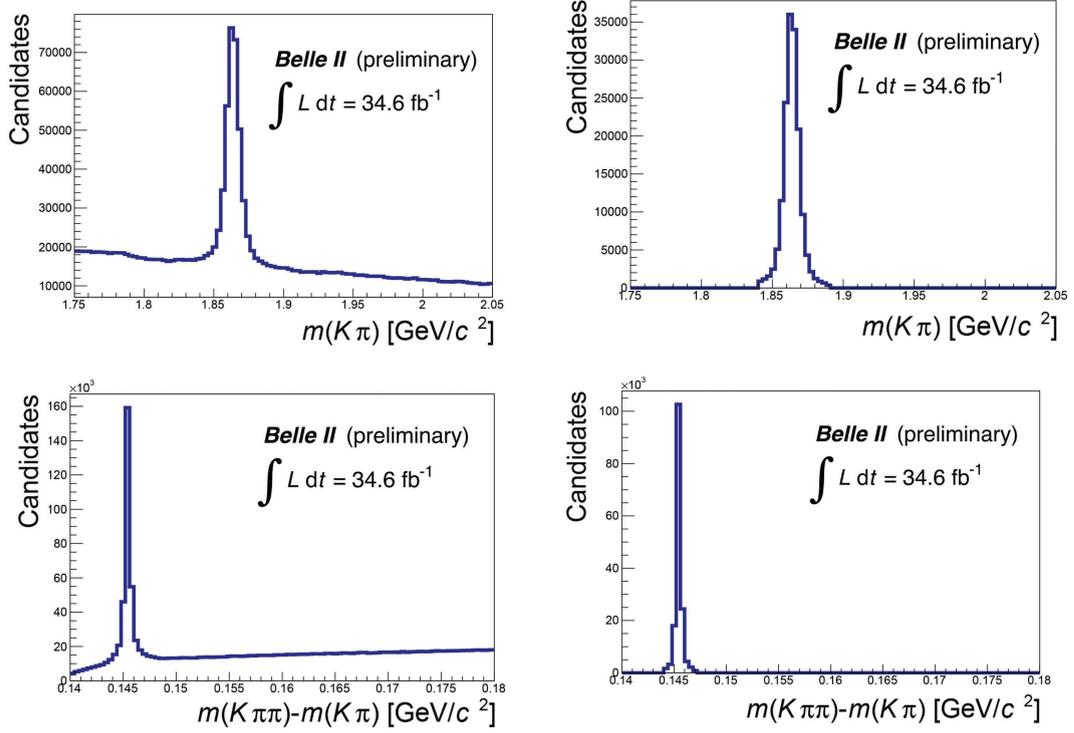


Figure 5.1: Distribution of (top) $K^- \pi^+$ mass and (bottom) difference between $K^- \pi^+ \pi^+$ and $K^- \pi^+$ masses (left) before and (right) after the selection for $D^{*+} \rightarrow D^0(\rightarrow K^- \pi^+) \pi^+$ candidates reconstructed in 34.6 fb^{-1} of 2019–2020 Belle II data.

5.3.1 Detector-specific likelihoods and PID

For each detector and particle-hypothesis, the likelihood is usually obtained by comparing the expected and the observed value of the raw information. For instance, in the CDC such information is $\frac{dE}{dx}_{\text{obs}}(h)$, the specific-ionization energy-loss observed for a charged particle h , averaged across the CDC wires. The (natural logarithm of) the likelihood is

$$\ln \mathcal{L}_{\text{hyp}}^{\text{CDC}}(h) = -\frac{1}{2} \left[\frac{\frac{dE}{dx}_{\text{obs}}(h) - \frac{dE}{dx}_{\text{exp-hyp}}(h)}{\sigma_{\text{obs}}(h)} \right]^2, \quad (5.1)$$

where ‘hyp’ represents the particle hypothesis and $\sigma_{\text{obs}}(h)$ is the observed uncertainty on $\frac{dE}{dx}_{\text{obs}}(h)$, which mainly depends on the number of CDC hits associated to h . The expected value $\frac{dE}{dx}_{\text{exp-hyp}}(h)$ is the average ionization-energy loss from a charged particle h that has the observed momentum, assuming the hypothesis ‘hyp’. This is calculated using the Bethe-Bloch relation [56, 57], which expresses the mean ionization energy-loss per unit path in gas of a charged particle as a function of velocity. Minor empirical adjustments are introduced to adapt the Bethe-Bloch prediction to the details of the CDC response. Figure 5.2 shows the $\frac{dE}{dx}_{\text{obs}}(h)$ distribution for various particle species in Belle II data and the expected energy loss for each of the six mass hypotheses considered.

In the TOP, the likelihood is calculated by comparing the observed number of detected photons associated to the charged particle with the expected photon yield, obtained from simulation [59, 60],

$$\ln \mathcal{L}_{\text{hyp}}^{\text{TOP}}(h) = \sum_{i=1}^N \ln \left[\frac{S_{\text{hyp}}(x_i, t_i, h) + B(x_i, t_i)}{N_e(h)} \right] + \ln P_N(N_e(h)), \quad (5.2)$$

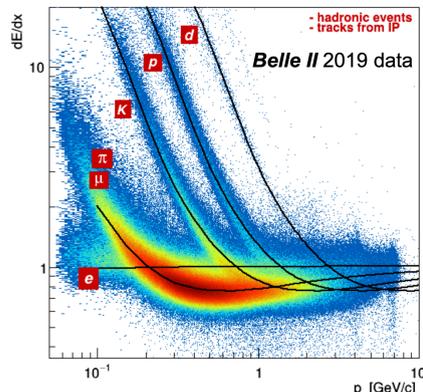


Figure 5.2: Distribution of (points) observed ionization-energy loss as a function of momentum for tracks from hadronic events reconstructed in Belle II data, along with the average expected values (solid lines) [58].

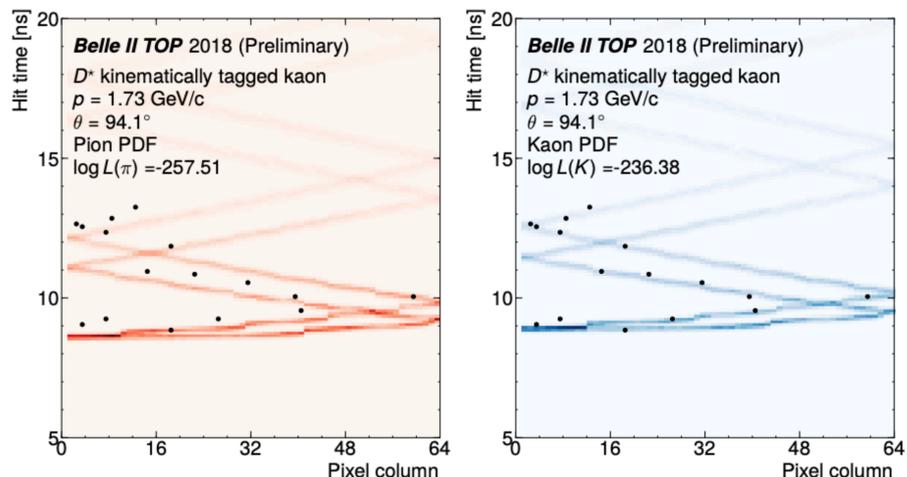


Figure 5.3: Example of kaon identification in the TOP detector. Arrival time of the Cherenkov photons as a function of position is compared with the expectations for (left) a pion and (right) a kaon passing in the TOP [58].

where x_i and t_i are, respectively, the positions and times of arrival of the N Cherenkov photons excited by the charged hadron h . The term $S_{\text{hyp}}(x, t, h)$ is the signal distribution for the hypothesis ‘hyp’; $B(x, t)$ is the distribution of background; and $N_e(h) = N_{\text{hyp}}(h) + N_B$ is the expected number of detected photons, which is the sum of the expected number of signal photons $N_{\text{hyp}}(h)$ for hypothesis ‘hyp’ and the expected number of background photons N_B . The second term in Eq. (5.2) is a probability for a Poisson with mean N_e to generate N photons [59]. Figure 5.3 shows an example of the identification of a kaon in the TOP detector: the positions and times of arrival of the Cherenkov photons are compared with the values expected for a pion or a kaon.

Using the likelihoods for the various mass hypotheses, Belle II algorithms construct a particle identification variable $\text{PID}_{\text{hyp}}^{\text{det}}$ for every detector, which is directly used in physics analyses and expresses the probability for observing the particle-identification information associated with the reconstructed track, assuming true a mass hypothesis ‘hyp’ out of six possibilities. As an example, the $\text{PID}_{\text{hyp}}^{\text{det}}$ expression associated with the pion mass

hypothesis is¹

$$\text{PID}_\pi^{\text{det}}(h) = \frac{\mathcal{L}_\pi^{\text{det}}(h)}{\mathcal{L}_\pi^{\text{det}}(h) + \mathcal{L}_K^{\text{det}}(h) + \mathcal{L}_e^{\text{det}}(h) + \mathcal{L}_\mu(h) + \mathcal{L}_p^{\text{det}}(h) + \mathcal{L}_d^{\text{det}}(h)}. \quad (5.4)$$

The $\text{PID}_{\text{hyp}}^{\text{det}}$ values for the other mass hypotheses are obtained by replacing the likelihood at the numerator $\mathcal{L}_\pi^{\text{det}}(h)$ with the corresponding hypothesis-specific value $\mathcal{L}_{\text{hyp}}^{\text{det}}(h)$. $\text{PID}_{\text{hyp}}^{\text{det}}(h)$ is defined similarly to a likelihood ratio $\mathcal{L}_0/\mathcal{L}_1$, which is the best-performing quantity to test two alternative simple statistical hypotheses [61]. However, the sum of likelihoods at the denominator is not a likelihood itself and therefore $\text{PID}_{\text{hyp}}^{\text{det}}(h)$ departs from a true likelihood ratio. The $\text{PID}_{\text{hyp}}^{\text{det}}$ variable assumes values from 0 to 1. The larger the $\text{PID}_{\text{hyp}}^{\text{det}}$, the higher the probability of observing the reconstructed track assuming true the chosen mass hypothesis.

Figure 5.4 shows an example of the TOP $\text{PID}_{\text{hyp}}^{\text{det}}$ distribution ($\text{PID}_{\text{hyp}}^{\text{TOP}}$) in the pion and kaon hypotheses, for samples of kaons and pions. Under the pion hypothesis, kaons cluster at small values, corresponding to low probabilities of the observation assuming the pion hypothesis, but a peak at ≈ 1 is also present, corresponding presumably to residual background pions and kaons incorrectly identified as pions. Pions are more uniformly distributed. The proportion between correctly and mistakenly identified pions is lower than the proportion between correctly and mistakenly identified kaons. For the kaon hypothesis, kaons peak at high values, corresponding to higher probabilities of observing the reconstructed TOP information if the track was a kaon, but a peak at ≈ 0 is also present, showing that a fairly large fraction of kaons is not identified correctly. Pions under the kaon hypothesis peak at zero, corresponding to low probabilities of observing the reconstructed TOP information if the track was a kaon, with a much smaller peak at $\text{PID}_{\text{hyp}}^{\text{TOP}} \approx 1$. The prominent peak at $\text{PID}_{\text{hyp}}^{\text{TOP}} = 0.33$ corresponds to candidates for which the particle identification algorithms cannot distinguish between the kaon, proton, and deuteron hypotheses and therefore associate the same dummy value to their likelihoods. These candidates correspond to charged particles with momentum lower than the threshold for exiting a signal in the TOP detector.

5.3.2 Detector combination

Information from individual detectors is combined to improve the identification performance. The detector-specific likelihoods are combined together as a product,

$$\mathcal{L}_{\text{hyp}}(h) = \mathcal{L}_{\text{hyp}}^{\text{TOP}}(h) \times \mathcal{L}_{\text{hyp}}^{\text{CDC}}(h) \times \mathcal{L}_{\text{hyp}}^{\text{SVD}}(h) \times \mathcal{L}_{\text{hyp}}^{\text{ARICH}}(h) \times \mathcal{L}_{\text{hyp}}^{\text{ECL}}(h) \times \mathcal{L}_{\text{hyp}}^{\text{KLM}}(h), \quad (5.5)$$

and the result is used in Eq.(5.4) to obtain the detector-combined PID. If a particle does not get reconstructed in a detector because, for instance, it escapes its acceptance, no PID information from that detector is available and the corresponding individual likelihood is set to one.

Equation (5.5) shows that Belle II *assumes that the single-detector likelihoods are independent from each other*, and therefore factorizes the combined likelihood as the product of single likelihoods. This is an important and critical assumption that is investigated later in this chapter (Sec. 5.4).

¹In practice, the Belle II software expresses PID_{hyp} using only the natural logarithm of the likelihood values,

$$\text{PID}_{\text{hyp}} = \frac{e^{\ln \mathcal{L}_{\text{hyp}} - \ln \mathcal{L}_{\text{MAX}}}}{\sum_i (e^{\ln \mathcal{L}_i - \ln \mathcal{L}_{\text{MAX}}})} = \frac{e^{\Delta \ln \mathcal{L}_{\text{hyp}}}}{\sum_i (e^{\Delta \ln \mathcal{L}_i})} \quad (5.3)$$

where \mathcal{L}_{MAX} is the largest of the likelihood values over the six hypotheses.

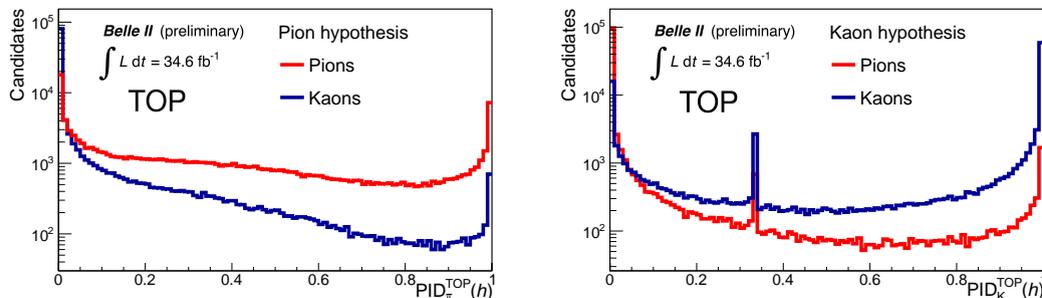


Figure 5.4: $\text{PID}_{\text{hyp}}(h)$ distribution (in logarithmic scale) from the TOP detector in the (left) pion hypothesis and (right) kaon hypothesis for a sample of (blue) kaons and (red) pions from $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays reconstructed in 2019–2020 Belle II data, corresponding to 34.6 fb^{-1} of integrated luminosity.

My first goal is to investigate whether the performance combined across detectors is unbiased and optimal. Since I am dealing with calibration samples of kaons and pions only, I restrict my study to the detectors that are most sensitive to the discrimination between kaons and pions, TOP, CDC, and ARICH. This is also expedient physics-wise, since kaons and pions are the main final-state particles in charmless B decays. From now on

- the label TC denotes the combined performance of TOP and CDC detectors;
- the label AC denotes the combined performance of ARICH and CDC detectors,

where, for example,

$$\ln \mathcal{L}_{\text{hyp}}^{\text{TC}} = \ln \mathcal{L}_{\text{hyp}}^{\text{TOP}} + \ln \mathcal{L}_{\text{hyp}}^{\text{CDC}}. \quad (5.6)$$

I restrict the sample to subsets of tracks having non-null information from all detectors considered in each combination, to achieve consistent results. For instance, to compare the individual performance of TOP and CDC with the combined-TC performance, I consider only tracks having non-null information from both TOP and CDC. Otherwise the resulting degradation in combined performance would be an artifact of the nonuniformity of the sample, since tracks that only have TOP or CDC information dilute the combined performance.

TOP and ARICH detectors have approximately exclusive geometrical acceptances. Hence, the fraction of tracks having information from both TOP and ARICH detectors is negligible and not studied further.

5.3.3 Performance metrics

I use two figures of merit to assess separation performance: the receiver operating characteristic curve (see Sec 3.3) and the separation power s [62]. The ROC curve is constructed as the true positive rate (e.g., the number of genuine kaons identified as kaons) as a function of the complement to the false positive rate (one minus the number of pions identified as kaons) at various threshold settings on the discriminating quantity. The larger the area under the ROC curve, the better the separating performance.

The separation power s is the observed separation between two distributions *relative to their ideally optimal* separation. The optimal separation corresponds to completely non-overlapping distributions, where classification performance is only affected by the statistical fluctuations due to the finite sizes of the samples. In a sample with two classes of events

(in our case, kaons and pions), each contributing a fraction f and $(1 - f)$ of the total, let be x a discriminating observable that distinguishes between those classes. We take the resolution σ on the observed value of f as a measure of the separation performance of x . Rather than determining σ through simulation, we estimate it using the minimum variance bound, which is the lower bound to the variance that can be achieved, according to likelihood theory [63]. For this simple two-component sample case, the minimum variance bound of the resolution on f is

$$\sigma^2(f) = \frac{1}{N} \left(\int \frac{(p_1(x) - p_2(x))^2}{fp_1(x) + (1-f)p_2(x)} dx \right)^{-1}, \quad (5.7)$$

where $p_1(x)$ is the *pdf* of the variable x for one class of events and $p_2(x)$ is the *pdf* for the other; N is the total number of events in the sample. If the two classes of events are totally separated in x , the uncertainty on the fraction f is just the statistical uncertainty due to the statistical fluctuations in the finite-size sample, i.e., the binomial uncertainty

$$\sigma_{\text{best}}^2(f) = \frac{f(1-f)}{N}. \quad (5.8)$$

The separation power s of the observable x is then the ratio between the resolution and the limiting resolution,

$$s = \frac{\sigma_{\text{best}}(f)}{\sigma(f)} = \sqrt{f(1-f) \int \frac{(p_1(x) - p_2(x))^2}{fp_1(x) + (1-f)p_2(x)} dx}. \quad (5.9)$$

No analytic expression is needed for $p_1(x)$ and $p_2(x)$. Observed distributions (histograms) are sufficient to determine s . Separation s is adimensional and independent from the sample size: it shows the power of the observable x in separating the samples, from 0 (no separation) to 1 (absolute maximum achievable with the given sample). The separation power s has the further advantage over the ROC curve that it generalizes straightforwardly to a multidimensional case \vec{x} .

5.4 Detector-specific and detector-combined performances

Figures 5.4, 5.5, and 5.6 show the single-detector distributions of the $\text{PID}_\pi(h)$ and $\text{PID}_K(h)$ variables. Figures 5.7 and 5.8 show the distributions based on the TC and AC combinations, respectively. All distributions but $\text{PID}_{\text{hyp}}^{\text{CDC}}$ show a peak at small values of PID, corresponding to low probabilities of observing the reconstructed PID information for the given hypothesis, and a peak at high values, corresponding to high probabilities.

Tables 5.1 and 5.2 show the corresponding separation powers, ranging from about 40% to 50% for the CDC alone, from about 60% to 75% for the TOP alone, and from about 62% to 75% for the ARICH alone. The uncertainty on the separation s is estimated using simulation to be 0.4% for all values and is not reported henceforth.

The separation power under the kaon hypothesis is consistently larger by 3%–16% than that under the pion hypothesis in all configurations. An explanation requires investigation of low-level information from the single detectors that is not available for this work.

In all cases, the combination of detectors (TC and AC) improves the separation power by 5%–10% over the largest individual detector performance. In the TC combination under the kaon hypothesis, PID_K separation improves from 73% to 79%. In the AC combination, PID_K separation improves from 75% to 82%.

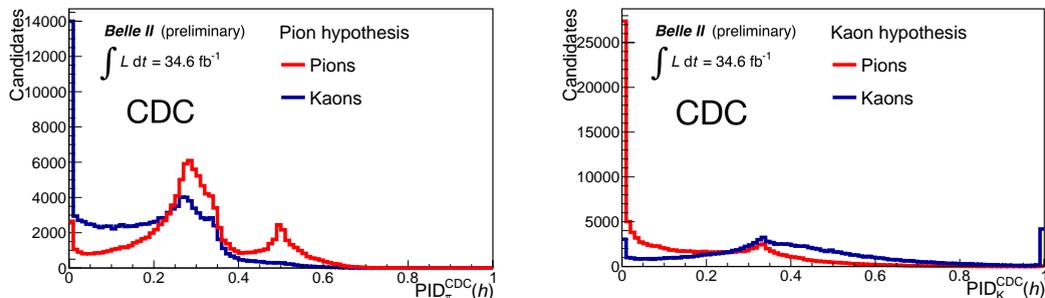


Figure 5.5: PID distribution from the CDC in the (left) pion hypothesis and (right) kaon hypothesis for a sample of (blue) kaons and (red) pions from $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays reconstructed in 2019–2020 Belle II data, corresponding to 34.6 fb^{-1} of integrated luminosity.

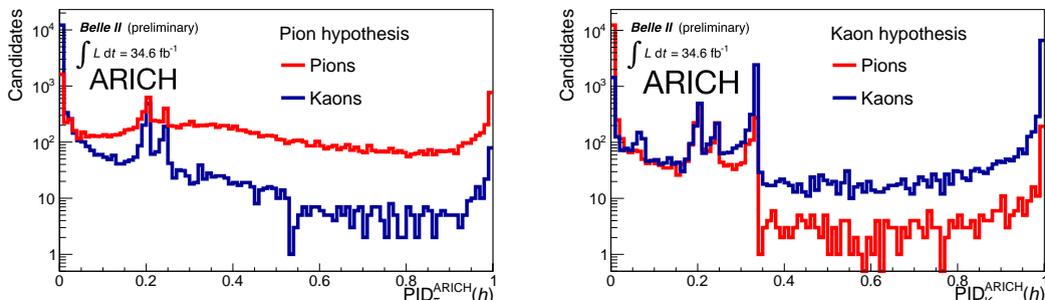


Figure 5.6: PID distribution (in logarithmic scale) from the ARICH in the (left) pion hypothesis and (right) kaon hypothesis for a sample of (blue) kaons and (red) pions from $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays reconstructed in 2019–2020 Belle II data, corresponding to 34.6 fb^{-1} of integrated luminosity.

Detector	KaonID	PionID
TOP	0.73	0.57
CDC	0.51	0.38
TC	0.79	0.63

Table 5.1: PID separation power s under the kaon and pion hypotheses. Only events having non-null information from both TOP and CDC detectors are considered. The uncertainty on the separation s is estimated by a Monte Carlo study and found to be 0.4% for all values.

Figures 5.9 and 5.10 show the associated ROC curves. Each data point corresponds, respectively, to the fraction of kaons correctly identified as kaons as a function of one minus the fraction of pions incorrectly identified as kaons for the kaon hypothesis and viceversa for the pion hypothesis. The ROC curves confirm our findings based on separation s . The combination of detectors outperforms any of single detectors. The ROCs also confirm a better performance under the kaon hypothesis with respect to the pion hypothesis, as indicated by a larger area under the curve.

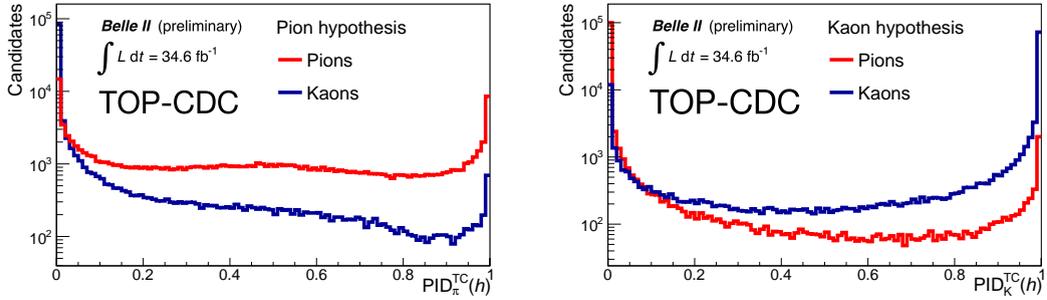


Figure 5.7: PID distribution (in logarithmic scale) from the TC combination in the (left) pion hypothesis and (right) kaon hypothesis for a sample of (blue) kaons and (red) pions from $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays reconstructed in 2019–2020 Belle II data, corresponding to 34.6 fb^{-1} of integrated luminosity. Only events having non-null information from both TOP and CDC detectors are considered.

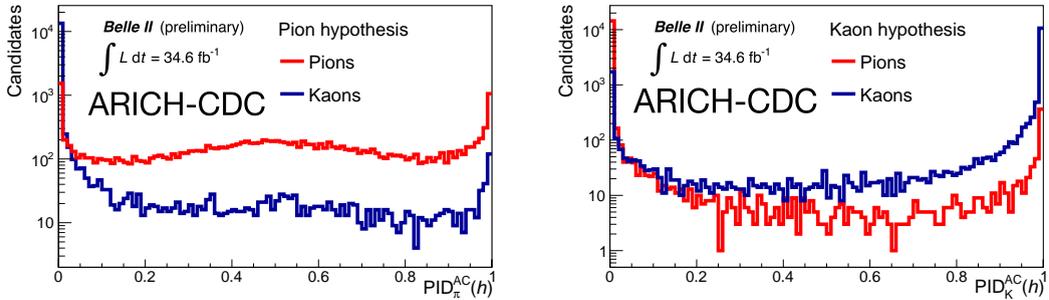


Figure 5.8: PID distribution (in logarithmic scale) from the AC combination in the (left) pion hypothesis and (right) kaon hypothesis for a sample of (blue) kaons and (red) pions from $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays reconstructed in 2019–2020 Belle II data, corresponding to 34.6 fb^{-1} of integrated luminosity. Only events having non-null information from both ARICH and CDC detectors are considered.

Detector	KaonID	PionID
ARICH	0.75	0.72
CDC	0.53	0.45
AC	0.82	0.78

Table 5.2: PID separation power s under the kaon and pion hypotheses. Only events having non-null information from both ARICH and CDC detectors are considered. The uncertainty on the separation s is estimated by a Monte Carlo study and found to be 0.4% for all values.

This study offers an important benchmark for the performance of the combination of multiple-detector PID-informations at Belle II, which is based on a fairly complicated software infrastructure whose full control is not obvious at the first stages of validation of detectors and algorithms.

The immediate follow-up is to verify if the size of the improvement brought by the

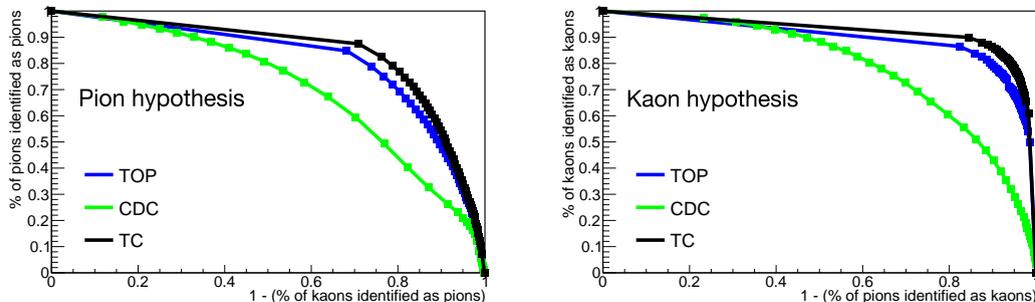


Figure 5.9: Signal efficiency as function of (1-background efficiency) for the PID variable in the (left) pion hypothesis and (right) kaon hypothesis using $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays reconstructed in 2019–2020 Belle II data, corresponding to 34.6 fb^{-1} of integrated luminosity. The green curve corresponds to the CDC detector, the blue curve corresponds to the TOP detector and the black curve corresponds to the TC combination.

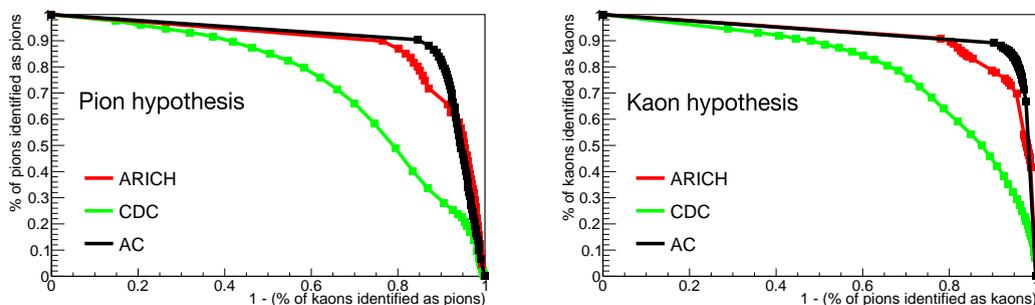


Figure 5.10: Signal efficiency as function of (1-background efficiency) for the PID variable in the (left) pion hypothesis and (right) kaon hypothesis using $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays reconstructed in 2019–2020 Belle II data, corresponding to 34.6 fb^{-1} of integrated luminosity. The green curve corresponds to the CDC detector, the red curve corresponds to the ARICH detector and the black curve corresponds to the AC combination.

combination is consistent with expectations. I use a simple simulation of the TOP-CDC combination under the kaon hypothesis to check if the observed combined performance is reproduced based on individual-detector values.

I generate two Gaussian distributions, one representing PID for kaons and one for pions. The distributions are simulated to have a separation of 0.74 (similar to the separation from the TOP detector only). In addition, I generate two Gaussian distributions separated by 0.45, close to the one observed for the CDC. The generated distributions are shown in Figure 5.11. I then combine the simulated information from the CDC and TOP detectors, assuming independence, to obtain the separation power of the detector-combined PID. The resulting value of 0.79, the same as found in data, confirms that no major technical or conceptual issues interfere in the combination workflow.

The final task is to check to which extent the approximation of independent detector informations is valid.

5.5 Exploring likelihood dependences

When dealing with information associated with multiple likelihoods, a joint multidimensional likelihood is generally needed, as it includes properly the possible dependences across variables. Only if no dependences exist, the joint likelihood factorizes as the product of the individual likelihoods, $pdf(x, y) = pdf(x) \times pdf(y)$. In Belle II, the likelihoods from the various subdetectors are treated as independent.

This is likely to be an approximation, as dependences are expected from the correlation between the kinematic properties of the track and the kinematics-dependent performances of most PID detectors. For instance, the identification by means of specific ionization depends on the momentum of the particle, as average ionizations from different particle species converge to, or depart from, each other depending on momentum. Similarly, simulations of the TOP and ARICH responses have shown that the identification performance improves for charged particles with momenta below 2 GeV/c [35]. Therefore, *dependent* PID information across detectors is expected due to kinematic correlations, even if such information is provided by physically *independent* PID detectors.

When two variables, say x and y , are independent, the conditional distribution of one variable conserves the same shape regardless of the conditioning value chosen for the other. I study possible dependences between the CDC and TOP detectors and between the CDC and ARICH detectors using this simple property.

Figures 5.12 and 5.13 show the distributions of $\ln \mathcal{L}$ for TOP and ARICH, respectively, for various ‘slices’ (restricted ranges) of $\ln \mathcal{L}^{\text{CDC}}$. Figure 5.12 shows no evidence of differences between the various conditional distributions, indicating that the assumption of TOP-CDC independence is fairly satisfied. In contrast, Fig. 5.13 shows significant differences in the shapes of the conditional distributions. The $\ln \mathcal{L}^{\text{ARICH}}$ shape is approximately bimodal, with a broad Gaussian component and an overlapping narrow peak at values near zero. The proportions between these components differ across $\ln \mathcal{L}^{\text{CDC}}$ ranges. For the pion hypothesis, the fraction of events in the narrow component changes from about 2% in the first $\ln \mathcal{L}^{\text{CDC}}$ slice, to 1.2% in the second, and 0.1% in the third. Such differences indicate a statistical dependence between $\ln \mathcal{L}^{\text{CDC}}$ and $\ln \mathcal{L}^{\text{ARICH}}$ that invalidates the approximation of independent likelihoods embedded in standard Belle II PID. Figure 5.13 indicates also that the size of the dependences differs across hypotheses. While the shape-difference between the slices is striking under the pion hypothesis, no large change is observed for the kaon hypothesis.

As interdetector dependences are expected from kinematics, I repeat the study after re-

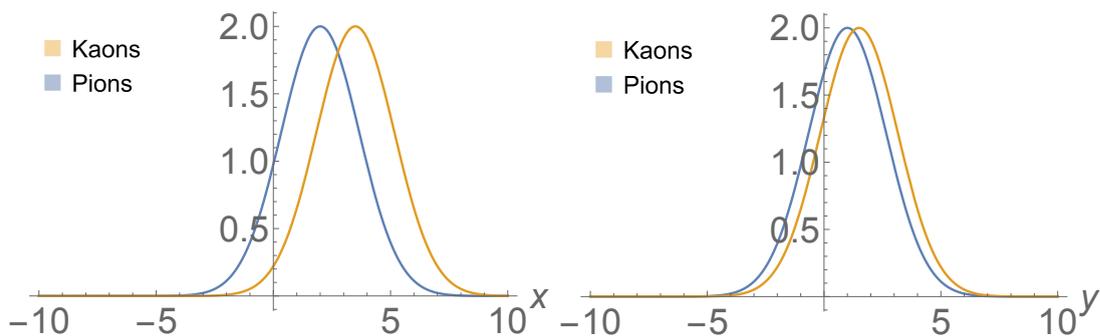


Figure 5.11: Simulated PID distributions for (blue) pions and (yellow) kaons with separation powers corresponding to (left) TOP information only and (right) CDC information only.

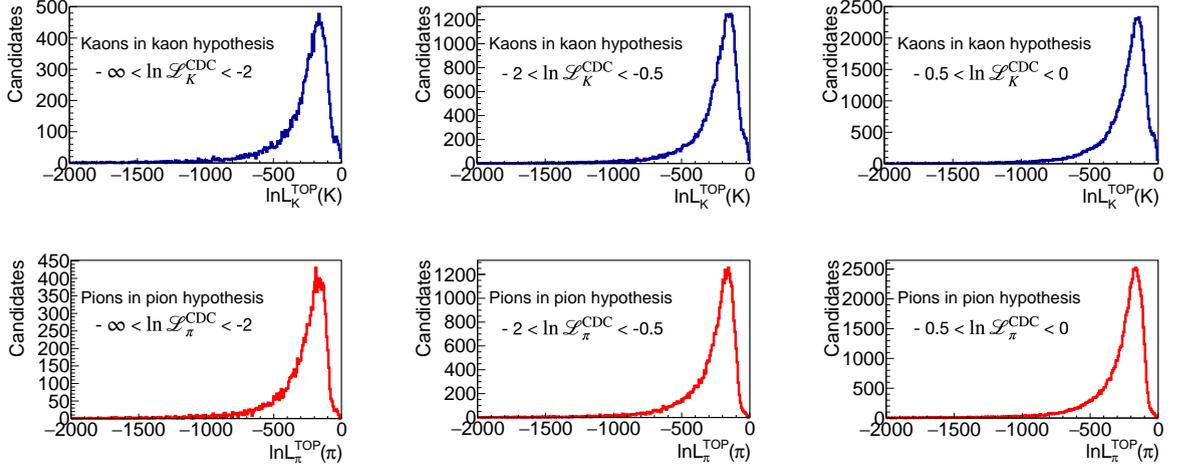


Figure 5.12: Distributions of (top) $\ln \mathcal{L}_K^{\text{TOP}}$ for kaons and (bottom) $\ln \mathcal{L}_\pi^{\text{TOP}}$ for pions for various slices of the $\ln \mathcal{L}^{\text{CDC}}$ distribution, from $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays reconstructed in 2019–2020 Belle II data and corresponding to 34.6 fb^{-1} of integrated luminosity.

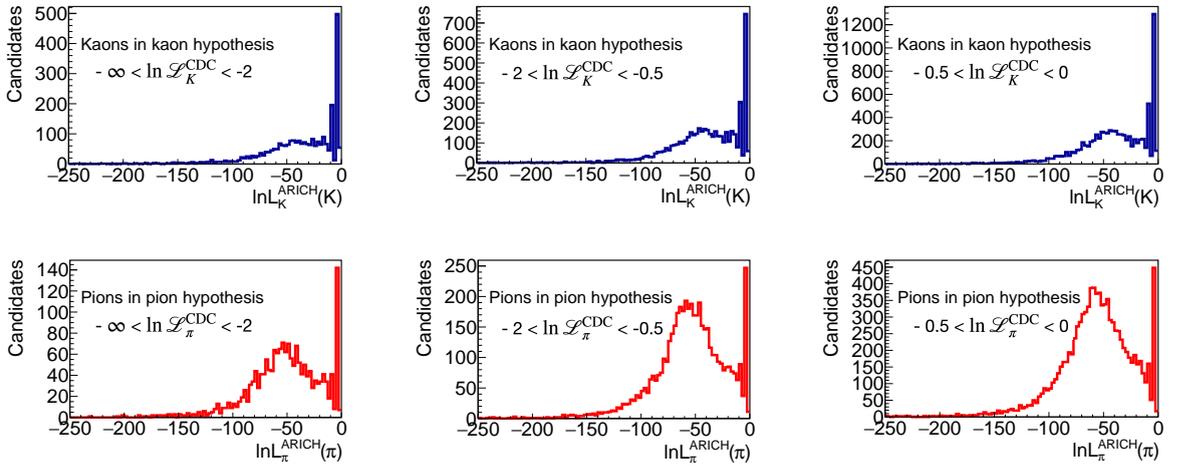


Figure 5.13: Distributions of (top) $\ln \mathcal{L}_K^{\text{ARICH}}$ for kaons and (bottom) $\ln \mathcal{L}_\pi^{\text{ARICH}}$ for pions for various slices of the $\ln \mathcal{L}^{\text{CDC}}$ distribution, from $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays reconstructed in 2019–2020 Belle II data and corresponding to 34.6 fb^{-1} of integrated luminosity.

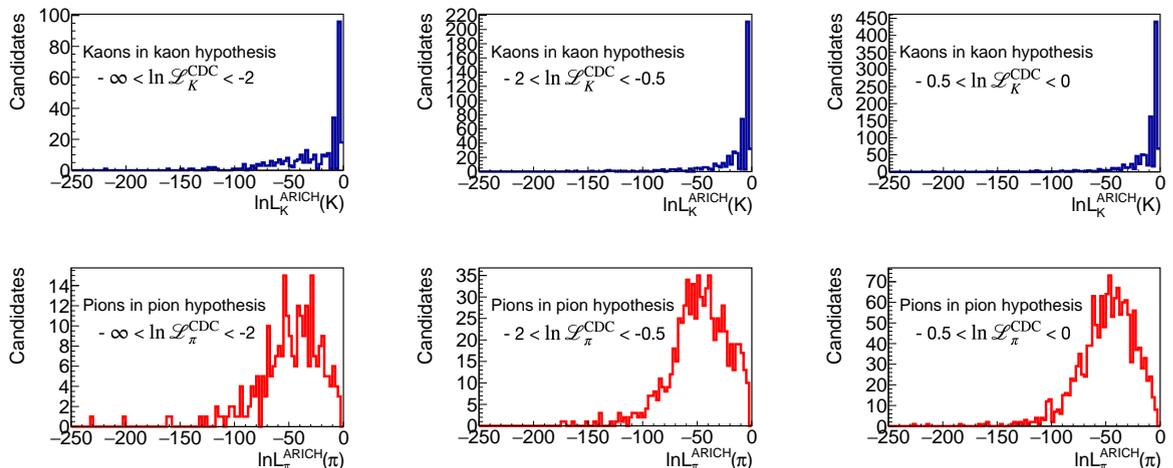


Figure 5.14: Distributions of (top) $\ln \mathcal{L}_K^{\text{ARICH}}$ for kaons and (bottom) $\ln \mathcal{L}_\pi^{\text{ARICH}}$ for pions for various slices of the $\ln \mathcal{L}^{\text{CDC}}$ distribution, from $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays reconstructed in 2019–2020 Belle II data and corresponding to 34.6 fb^{-1} of integrated luminosity. Particles are restricted to $0.5 < p < 1 \text{ GeV}/c$.

stricting the particles to a narrow region in momentum. I first restrict the sample to kaons and pions with momentum $0.5 < p < 1 \text{ GeV}/c$. Figure 5.14 shows the corresponding conditional ARICH distributions in ranges of CDC likelihood. The distributions show evidence of residual dependences, in particular for the pion hypothesis: for small negative values of $\ln \mathcal{L}^{\text{CDC}}$, the mean of the distribution shifts towards values nearer to zero, passing from -50 (at $-\infty < \ln \mathcal{L}^{\text{CDC}} < -2$) to -40 (at $-0.5 < \ln \mathcal{L}^{\text{CDC}} < 0$). A similar phenomenology occurs when restricting the sample to kaons and pions with momenta $1.5 < p < 2 \text{ GeV}/c$ (Figure 5.15). The likelihood dependences observed are reduced (indicatively by a order of magnitude) but not completely removed. This shows that not only the expected kinematic correlations are at play, but possibly other unknown effects that have to be investigated and studied.

The correct approach to study dependences and possibly remove or model them would be to use the joint interdetector likelihood of each track. As of today, this information is not available in mdst Belle II data. Approximate or simplified approaches can be attempted to investigate further. One possibility would be to form a pseudo-joint likelihood constructed by combining together the conditional likelihoods studied in Figs.5.12 and 5.13. However, the resulting multidimensional function would be inaccurate, as its features would depend on the granularity of the sampling.

The further simplified approach I choose consists in using simulation to have a first grasp of how likelihood dependences affect physics performances.

5.6 Impact of neglected likelihood dependences on performance

I simulate a sample of charged particles composed by only pions and kaons. I assume that two detectors x and y provide PID information about these particles, and I simulate these informations to be dependent. I then compare the performance of the combined PID obtained by properly accounting for dependences or assuming independent information

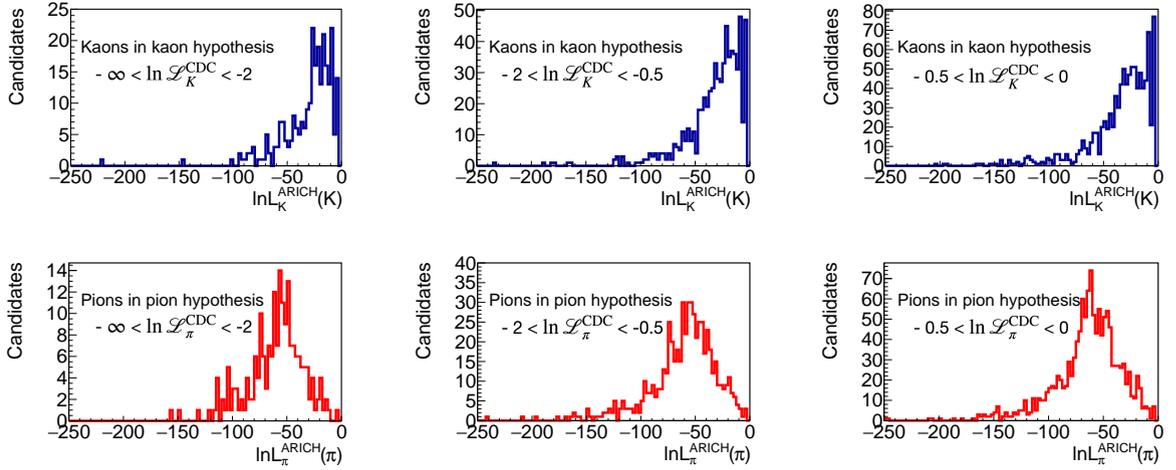


Figure 5.15: Distributions of (top) $\ln \mathcal{L}_K^{\text{ARICH}}$ for kaons and (bottom) $\ln \mathcal{L}_\pi^{\text{ARICH}}$ for pions for various slices of the $\ln \mathcal{L}^{\text{CDC}}$ distribution, from $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays reconstructed in 2019–2020 Belle II data and corresponding to 34.6 fb^{-1} of integrated luminosity. Particles are restricted to $1.5 < p < 2 \text{ GeV}/c$.

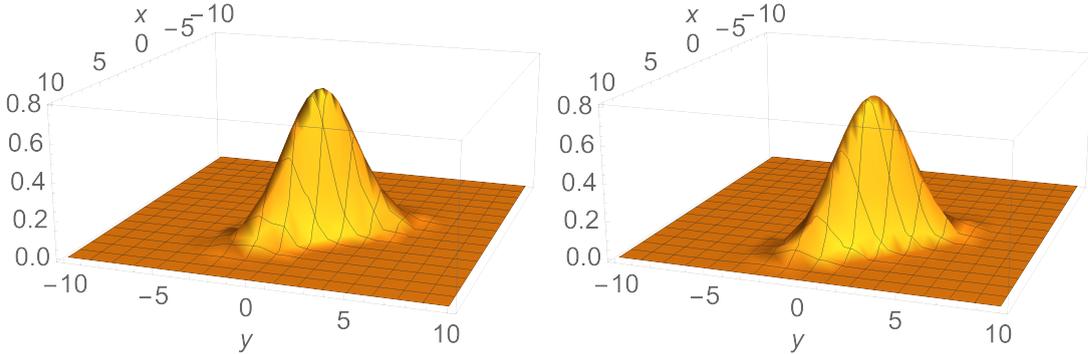


Figure 5.16: Simulated joint probability distribution for detectors x and y , (left) associated with pions, $f(x, y)$, and (right) associated with kaons, $g(x, y)$.

from the detectors, as done at Belle II. The chosen shape and size of the dependence is arbitrary, but this is irrelevant to our goal of illustrating qualitatively the possible impact.

5.6.1 Simplified simulation

I assume that the PID information on pions follows the two-dimensional *pdf* (left panel in Fig. 5.16)

$$f(x, y) = \frac{1}{0.5\sqrt{2\pi}} e^{-[\frac{1}{2}(x-2)^2 + \frac{4}{5}(x-2)(y-1) + \frac{1}{2}(y-1)^2]}.$$

The variables x and y are dependent and represent PID information from detectors x and y . The corresponding marginal distributions $f_y(x)$ and $f_x(y)$ are shown in Fig. 5.17. I assume that the PID information on kaons follows the *pdf* (right panel in Fig. 5.16)

$$g(x, y) = \frac{1}{0.5\sqrt{2\pi}} e^{-[\frac{1}{2}(x-3.5)^2 + \frac{4}{5}(x-3.5)(y-1.5) + \frac{1}{2}(y-1.5)^2]},$$

with marginal distributions $g_y(x)$ and $g_x(y)$ (Fig. 5.18).

I then compare the separation between kaons and pions under the kaon hypothesis using two approaches. I first account *properly* for the known dependence between x and y obtaining s as the bidimensional separation between the kaon and pion distributions shown in Fig. 5.19. I then compare the result with the one obtained *ignoring the dependence* and combining the separation powers obtained from the individual marginal distributions, as shown in Fig. 5.20 and done in Belle II.

5.6.2 Results and impact on physics

After generating 10^5 kaons (x_K, y_K) and 10^5 pions (x_π, y_π) that follow the f and g distributions, respectively, I calculate their likelihoods under kaon and pion hypotheses and compute the proper value of PID associated with the kaon hypothesis,

$$\text{PID}_K(\pi) = \frac{\mathcal{L}_K(\pi)}{\mathcal{L}_K(\pi) + \mathcal{L}_\pi(\pi)} = \frac{g(x_\pi, y_\pi)}{g(x_\pi, y_\pi) + f(x_\pi, y_\pi)}, \quad (5.10)$$

$$\text{PID}_K(K) = \frac{\mathcal{L}_K(K)}{\mathcal{L}_K(K) + \mathcal{L}_\pi(K)} = \frac{g(x_K, y_K)}{g(x_K, y_K) + f(x_K, y_K)}. \quad (5.11)$$

The resulting distributions are shown in the left panel of Fig. 5.21. The separation between the pion and kaon distributions obtained by *properly accounting for the dependence* is $s = 0.728$. The uncertainty on this value is estimated using simulation and found to be 0.004.

I now simulate the Belle II-like combination by calculating the separation between the PID distributions of pions and kaons *assuming x and y to be independent*. The total likelihoods are then the products of the marginal likelihoods, and the resulting PID expressions are

$$\text{PID}_K(\pi)^{\text{ind}} = \frac{\mathcal{L}_K(\pi)}{\mathcal{L}_K(\pi) + \mathcal{L}_\pi(\pi)} = \frac{g_y(x_\pi) \cdot g_x(y_\pi)}{g_y(x_\pi) \cdot g_x(y_\pi) + f_y(x_\pi) \cdot f_x(y_\pi)}, \quad (5.12)$$

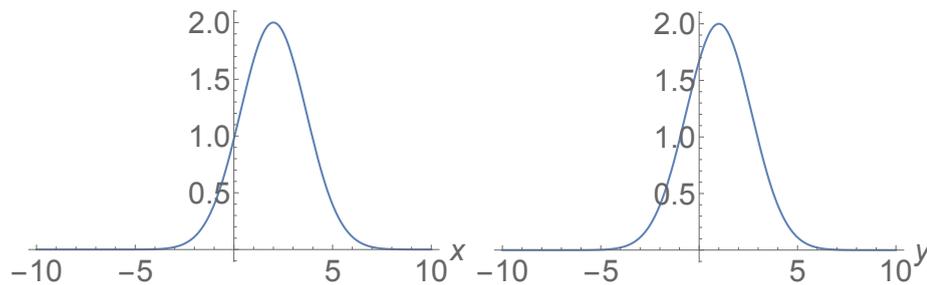


Figure 5.17: Simulated marginal probability distributions (left) $f_y(x)$ and (right) $f_x(y)$.

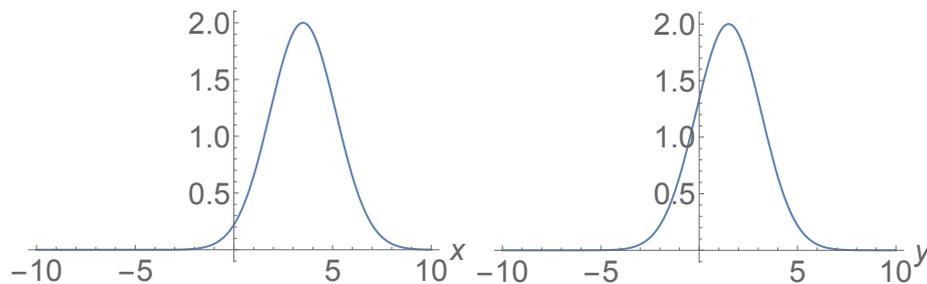


Figure 5.18: Simulated marginal probability distributions (left) $g_y(x)$ and (right) $g_x(y)$.

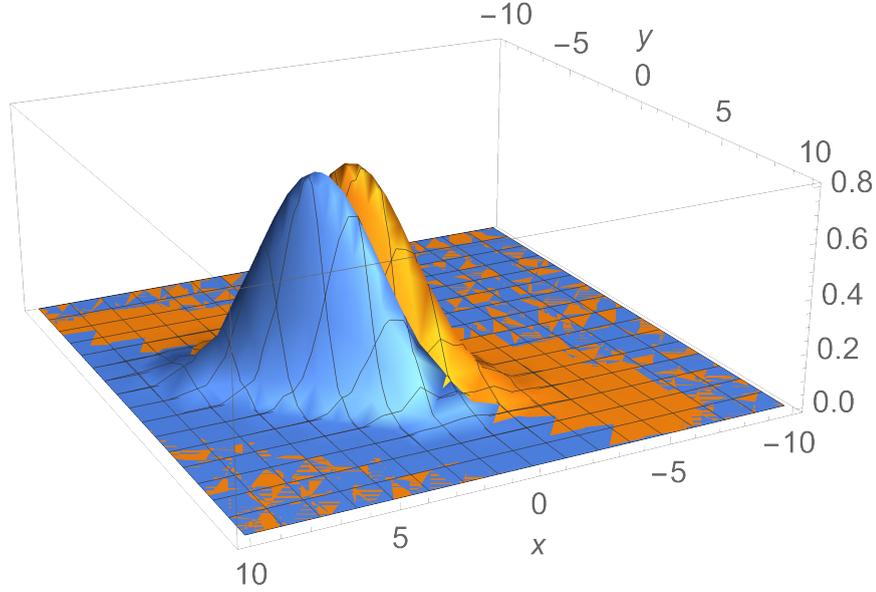


Figure 5.19: Simulated probability distributions $f(x, y)$ in blue and $g(x, y)$ in yellow.

$$\text{PID}_K(K)^{\text{ind}} = \frac{\mathcal{L}_K(K)}{\mathcal{L}_K(K) + \mathcal{L}_\pi(K)} = \frac{g_y(x_K) \cdot g_x(y_K)}{g_y(x_K) \cdot g_x(y_K) + f_y(x_K) \cdot f_x(y_K)}. \quad (5.13)$$

The right panel of Fig. 5.21 shows the resulting distributions for pions and kaons. The separation between pion and kaon distributions calculated by neglecting the dependence is $s = 0.788$, significantly larger than the proper value.

The apparent larger separation is misleading and due to incorrectly double-counting the same separating information from detectors x and y .

In this simplified simulated case, the assumption of independent variables produces a 6% overestimation in separation, which is a significant bias given the strongly nonlinear character of s . These findings are confirmed by the ROC curves (see Fig. 5.22), which indicate a false overperformance.

The severity of the mismodeling is best appreciated by comparing the PID distributions, obtained by accounting for the dependence between detectors x and y and the distributions obtained assuming independence (Fig. 5.21). The PID distributions have strongly different shapes.

As particle identification enters directly many Belle II physics results, an incorrect

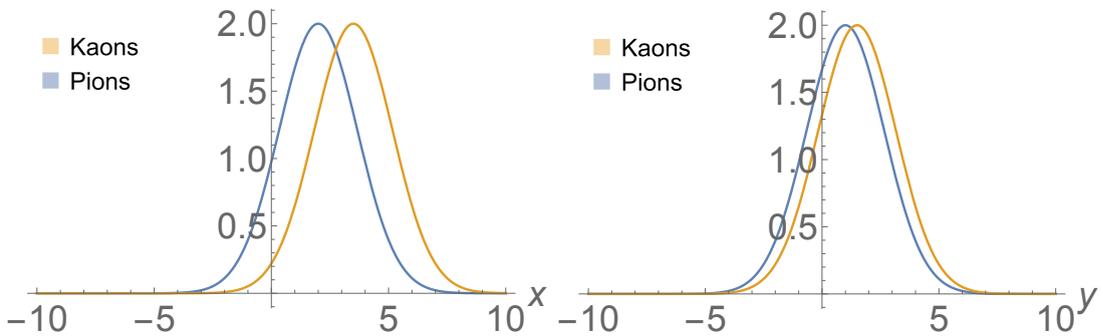


Figure 5.20: Simulated marginal probability distributions (left) $f_y(x)$ in blue and $g_y(x)$ in yellow and (right) $f_x(y)$ in blue and $g_x(y)$ in yellow.

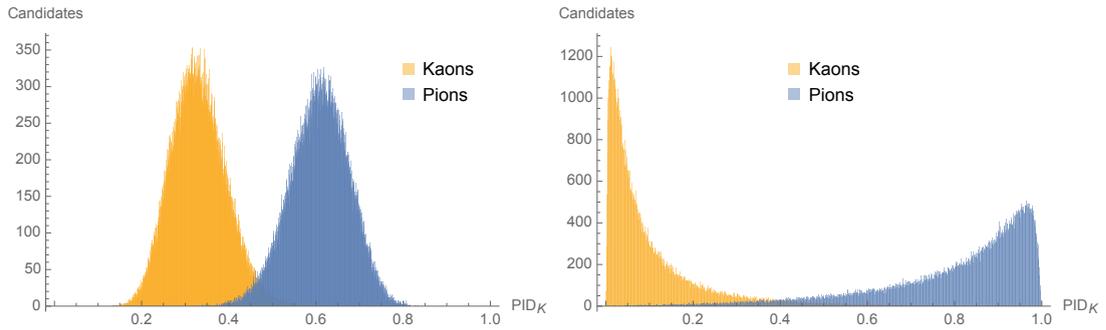


Figure 5.21: PID distributions under the kaon hypothesis for a simulated sample of (blue) pions and (yellow) kaons and calculated (left) by taking into account the dependence between detectors x and y and (right) assuming x and y to be independent.

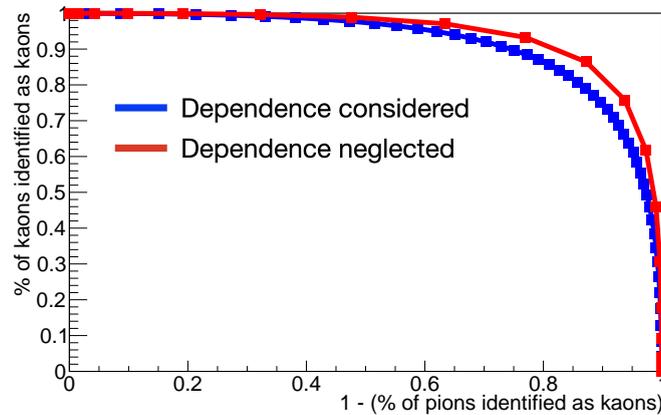


Figure 5.22: Signal efficiency as function of (1-background efficiency) for the simulated PID variable obtained (blue) taking into account the dependence between x and y and (red) assuming them to be independent in simulated data.

model is prone to cause biases in measurements. For instance, the PID selection efficiency, which is fundamental in branching fraction measurements where PID is used, is expected to be directly affected if evaluated based on the distributions in the left panel of Fig. 5.21 rather than with the correct distributions in the right panel. Even if the size and functional form of the dependence introduced in the simulation is arbitrary and might differ from the actual dependence present in Belle II data, this study exposes that approximating dependent variables as independent may lead to a strikingly incorrect model.

These considerations are not merely conceptual. Data-simulation discrepancies in the PID distributions are observed in Belle II data. Similar phenomena were observed at the predecessor experiment Belle, where per-track empirical correction factors were applied to the PID selection efficiencies thus adding to the systematic uncertainty of results. While no study exist to establish that those discrepancies are all due to neglected dependences, it is reasonable to assume that the latter may have at least a role.

Following this work, we are collaborating with the particle identification working group to explore possible approaches to improve the PID variable algorithms by making available the joint likelihood function to analyzers. This would take properly into account all the dependences between subdetectors, thus reducing or eliminating the need for final corrections.

5.7 Summary

In this part of my work I report in-depth studies of Belle II PID performance. I find dependences in PID quantities that Belle II assumes independent and studied the impact of this approximation on the PID performance using a simplified simulated model.

Chapter 6

First measurement of \mathcal{CP} -violating asymmetries in three-body charmless decays at Belle II

This chapter reports the last part of my work, which is devoted to the first measurement of charge-parity violating asymmetries in three-body $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$ decays reconstructed in Belle II data.

6.1 Motivation and strategy

An asymmetry in the rate of a decay ($X \rightarrow f$) and its \mathcal{CP} -conjugate decay ($\bar{X} \rightarrow \bar{f}$) due to differing decay amplitudes is called direct \mathcal{CP} violation. Direct \mathcal{CP} -violation measurements are important as they may uncover the presence of non-SM physics if the observed \mathcal{CP} -violating effects are inconsistent with expectations. Measurements of direct \mathcal{CP} violation are especially informative in multibody B -decays, where asymmetries localized in specific regions of the multibody kinematics can provide dynamic information on specific subsets of amplitudes.

The goal of this last part of my work is the first measurement of direct \mathcal{CP} -violating asymmetries in three-body charmless decays using Belle II data. Direct \mathcal{CP} -violation for the decay channels $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$ has been measured with high precision in the past decade by the Belle, BaBar, and LHCb experiments, which observed interestingly large asymmetries in specific regions of the kinematics [31]. Belle II does not expect a comparable precision in the 2019–2020 sample, which is still much smaller than those of previous experiments; however, this work represents a necessary preparation in view of more precise upcoming results. In addition, and perhaps more importantly, this work prompts the development of a data-driven in-depth investigation of charge-dependent detection and reconstruction asymmetries, which are an essential ingredient in many future measurements of \mathcal{CP} violation.

The strategy mirrors the one adopted for the $B^+ \rightarrow K^+\pi^0$ analysis, discussed in Chapter 3. The analysis is conducted in a blind fashion to avoid experimenter's biases in the results. I use simulated data to develop and finalize all the analysis details; once the analysis choices are finalized, I apply the analysis on the total sample of experimental data collected by Belle II for the summer 2020 conferences. Finally, I compare the obtained results with known references.

After an optimization of the baseline selection, continuum background suppression and

PID, I measure the asymmetries by fitting separately but simultaneously the ΔE distributions of the B^+ and B^- candidates using directly the raw partial-decay-rate asymmetry as a fit parameter,

$$\mathcal{A} = \frac{N(b) - N(\bar{b})}{N(b) + N(\bar{b})}, \quad (6.1)$$

where N are signal yields and b (\bar{b}) indicates the meson containing a bottom (antibottom) quark. I then correct the raw asymmetry \mathcal{A} for possible charge-dependent reconstruction asymmetries determined from control samples in data. An additional challenge with respect to the $B^+ \rightarrow K^+\pi^0$ analysis of Chap. 3 is the treatment of insidious peaking backgrounds, which, if not properly accounted for, bias the final results.

6.2 Raw charge-asymmetry determination

6.2.1 Tools, selection, and reconstruction

I reconstruct the three-body decays $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$, which constitute my signal, in the Belle II data set collected as of May 14, 2020, corresponding to an integrated luminosity of 34.6 fb^{-1} . In addition, I reconstruct the control channels $D^0 \rightarrow K^-\pi^+$ and $D^+ \rightarrow \bar{K}^0\pi^+$ to determine the instrumental asymmetries.

I use simulated generic data (see Sec. 3.3), corresponding to an integrated luminosity of 600 fb^{-1} , and 2×10^6 signal events for each channel to develop the analysis, optimize the signal selection, and study peaking backgrounds.

Three-body decays can proceed through multiple intermediate states whose amplitudes interfere with each other. An accurate decay model in simulation is important to ensure that reconstruction efficiencies are realistic. I use an approximated decay model that takes into account the major resonant components but neglects interferences. A convenient way to model straightforwardly the contributions of intermediate processes is the Dalitz plot, a phase-space representation based on the correlations between the squared invariant masses of two two-body combinations of final-state particles in a three-body decay. Figure 6.1 shows the Dalitz plot for simulated $B^+ \rightarrow K^+\pi^-\pi^+$ decays. A cluster of events centered at a specific two-body mass indicates the presence of an intermediate resonance. In this case, the darker horizontal band is due to $B^+ \rightarrow K^+f_0(980)(\rightarrow \pi^+\pi^-)$ decays, while the darker vertical band is due to $B^+ \rightarrow K^*(1430)^0(\rightarrow K^+\pi^-)\pi^+$ decays.

To reconstruct B -meson candidates, I select triplets of high-quality tracks reconstructed in the full CDC polar acceptance through a converged track fit. I also require their transverse (longitudinal) distance to the interaction point to meet $dr < 0.5 \text{ cm}$ ($|dz| < 3 \text{ cm}$), to remove tracks produced outside the interaction region, which mainly associate with particles from beam-background events that interact with the interaction region material.

The resulting K^\pm and π^\pm candidates are combined through kinematic fits consistent with $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$ decays. Only candidates with $|\Delta E| < 0.3 \text{ GeV}$ and $M_{bc} > 5.20 \text{ GeV}/c^2$ are kept in the analysis signal region.

6.2.2 Selection optimization

I use the same baseline track selections discussed in Chap. 3 for the $B^+ \rightarrow K^+\pi^0$ analysis. As for higher level quantities, I simultaneously optimize the selection on continuum suppression (CS) output and kaon identification (PID). The first variable aims at suppressing continuum by exploiting information on the topology and the kinematics of the event (see

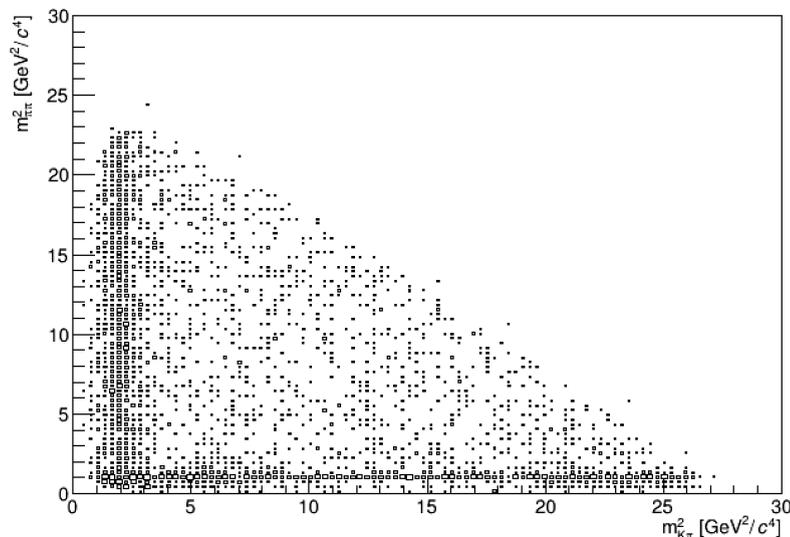


Figure 6.1: Dalitz plot for simulated $B^+ \rightarrow K^+ \pi^- \pi^+$ decays. The darker horizontal (vertical) band is due to $B^+ \rightarrow K^+ f_0(980) (\rightarrow \pi^+ \pi^-)$ [$B^+ \rightarrow K_0^*(1430)^0 (\rightarrow K^+ \pi^-) \pi^+$] decays.

Sec. 3.5). Kaon identification is useful for reducing backgrounds from hadron misidentification. As these quantities can be correlated, the optimization is simultaneous to achieve optimal performance.

The optimization consists in varying the selection criteria on both the CS output and the PID to maximize the figure of merit $\text{FoM} = S/\sqrt{S+B}$, where S is the signal yield, and B is the background yield. Both yields are obtained from extended maximum likelihood fits of the unbinned ΔE distributions of $B^+ \rightarrow K^+ K^- K^+$ and $B^+ \rightarrow K^+ \pi^- \pi^+$ decays reconstructed in simulated e^+e^- collisions and restricted to the signal region $|\Delta E| < 0.15 \text{ GeV}$ and $M_{bc} > 5.27 \text{ GeV}/c^2$. The resulting FoM values as functions of CS-PID selections are shown in Fig. 6.2. For $B^+ \rightarrow K^+ K^- K^+$ decays, a high-FoM plateau is observed for $\text{PID} > 0.2$ and for CS outputs from > 0.6 to > 0.9 . For $B^+ \rightarrow K^+ \pi^- \pi^+$ decays, the CS selection that optimizes the FoM consistently corresponds to $\text{CS} > 0.9$ for all possible PID selections. The optimized configurations are shown in Table 6.1.

Decay channel	CS output	Kaon PID
$B^+ \rightarrow K^+ K^- K^+$	> 0.75	> 0.2
$B^+ \rightarrow K^+ \pi^- \pi^+$	> 0.9	> 0.4

Table 6.1: Optimized requirements for the continuum suppression output and kaon PID.

6.2.3 Sample composition

An additional challenge in the analysis of multibody B decays with respect to the $B^+ \rightarrow K^+ \pi^0$ analysis is the treatment of peaking backgrounds that contaminate the $B^+ \rightarrow K^+ K^- K^+$ and $B^+ \rightarrow K^+ \pi^- \pi^+$ samples. These are non-charmless transitions that generate ΔE distributions peaked under, or in proximity, of the signals of interest. As the fit that determines the signal yields generally assumes the background to be smoothly distributed, peaking backgrounds may be mistakenly absorbed into signal and bias the re-

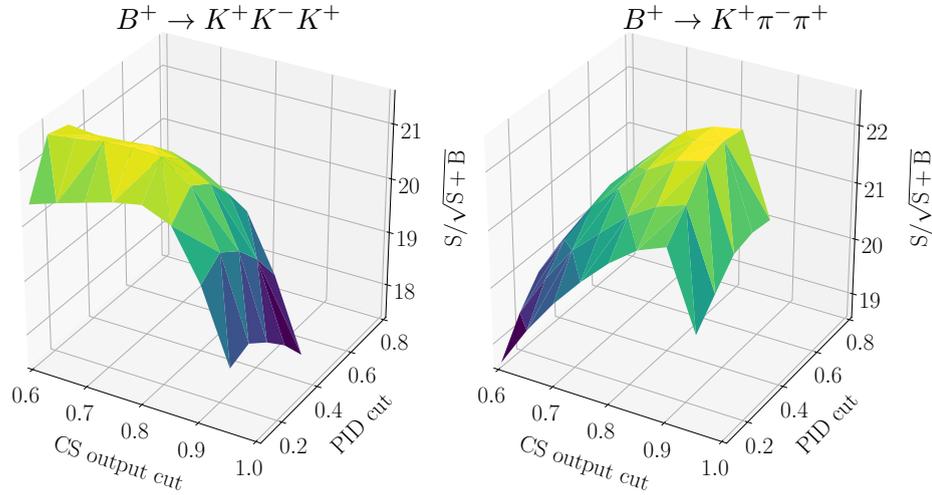


Figure 6.2: $S/\sqrt{S+B}$ as a function of the lower threshold on continuum suppression output and PID requirements for (left) $B^+ \rightarrow K^+K^-K^+$ decays and (right) $B^+ \rightarrow K^+\pi^-\pi^+$ decays.

sults if not taken into account. Figure 6.3 shows the ΔE distribution of $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$ decays reconstructed in generic simulation. The fraction of each component with respect to the signal in the samples is shown in Fig. 6.4. A contribution from $B \rightarrow cX \rightarrow K^+K^-K^+$ decays, amounting to about 1/4 of signal, contaminates the $B^+ \rightarrow K^+K^-K^+$ distribution. A contribution of $b \rightarrow c$ decays $B \rightarrow DX \rightarrow K^+\pi^-\pi^+$ five times larger than the signal spoils the $B^+ \rightarrow K^+\pi^-\pi^+$ distribution.

The strategy to identify and remove, or model, these contributions is to inspect the distributions of the two-body masses of intermediate states, where narrow peaks allow for straightforward vetoes, as shown in Fig. 6.5. I suppress the dominant $B^+ \rightarrow \bar{D}^0(\rightarrow K^+K^-)K^+$, $B^+ \rightarrow \eta_c(\rightarrow K^+K^-)K^+$, and $B^+ \rightarrow \chi_{c1}(\rightarrow K^+K^-)K^+$ contributions to the $B^+ \rightarrow K^+K^-K^+$ sample by excluding the two-body mass ranges $1.84 < m(K^+K^-) < 1.88$ GeV/c^2 , $2.94 < m(K^+K^-) < 3.05$ GeV/c^2 , and $3.50 < m(K^+K^-) < 3.54$ GeV/c^2 , respectively. In the $B^+ \rightarrow K^+\pi^-\pi^+$ channel, I suppress the $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$, $B^+ \rightarrow \eta_c(\rightarrow \pi^+\pi^-)K^+$, $B^+ \rightarrow \chi_{c1}(\rightarrow \pi^+\pi^-)K^+$, and $B^+ \rightarrow \eta_c(2S)(\rightarrow \pi^+\pi^-)K^+$ contributions by excluding the two-body mass ranges $1.8 < m(K^+\pi^-) < 1.92$ GeV/c^2 , $2.93 < m(\pi^+\pi^-) < 3.15$ GeV/c^2 , $3.45 < m(\pi^+\pi^-) < 3.525$ GeV/c^2 , and $3.62 < m(\pi^+\pi^-) < 3.665$ GeV/c^2 . In addition, the distribution of $B^+ \rightarrow K^+\pi^-\pi^+$ candidates receives an important contribution from B -to-charmonium decays, such as $B^+ \rightarrow J/\psi(\rightarrow \mu^+\mu^-)K^+$ and $B^+ \rightarrow \psi(2S)(\rightarrow \mu^+\mu^-)K^+$, in which the muons are incorrectly reconstructed as pions. I suppress these distributions by excluding the two-body mass ranges $2.93 < m(\pi^+\pi^-) < 3.15$ GeV/c^2 and $3.67 < m(\pi^+\pi^-) < 3.72$ GeV/c^2 . Finally I veto the genuine charmless $B^+ \rightarrow K^*(892)^0\pi^+$ subcomponent too by excluding candidates with $0.82 < m(K^+\pi^-) < 0.98$ GeV/c^2 to remain consistent with Ref. [17] where this component is not included.

Figure 6.6 shows the impact of the vetoes. The contribution of peaking backgrounds is greatly suppressed in both distributions. The remainder component (about 1% of the total) from decays with an intermediate D -meson in the $B^+ \rightarrow K^+\pi^-\pi^+$ distribution is accounted for while fitting the ΔE distribution.

More than one candidate per event populates the resulting ΔE distributions. Average candidate multiplicities are 1.01 in both simulated and Belle II data for $B^+ \rightarrow K^+K^-K^+$

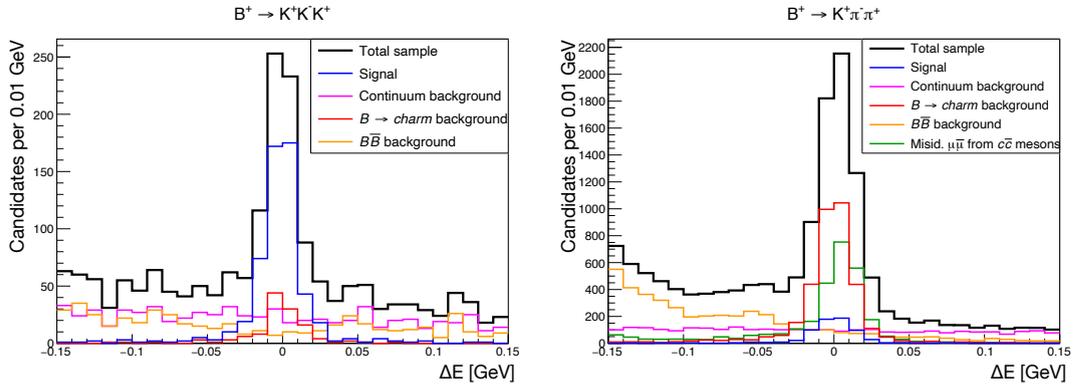


Figure 6.3: Distributions of ΔE for (left) $B^+ \rightarrow K^+K^-K^+$ and (right) $B^+ \rightarrow K^+\pi^-\pi^+$ candidates reconstructed in 50 fb^{-1} of simulated data. Each color indicate a leading contribution to the sample.

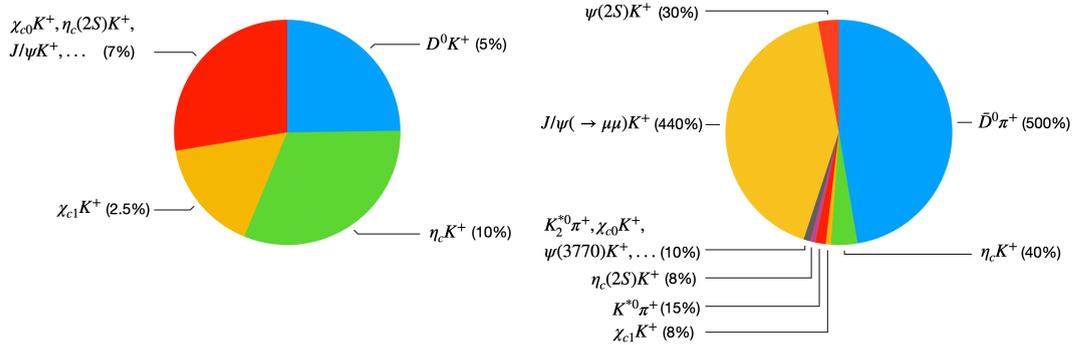


Figure 6.4: Fraction of the main peaking backgrounds with respect to the signal for (left) $B^+ \rightarrow K^+K^-K^+$ decays and (right) $B^+ \rightarrow K^+\pi^-\pi^+$ decays, obtained from simulation.

events and 1.17 in simulated data and 1.20 in experimental data for $B^+ \rightarrow K^+\pi^-\pi^+$ events. The average multiplicity is larger for $B^+ \rightarrow K^+\pi^-\pi^+$ events as two final-state pions allow for more combinatorial combinations with continuum which is dominated by pions. I select randomly a single B candidate from each event for further analysis. While this choice might be suboptimal, it is unbiased.

The resulting ΔE distributions of B candidates reconstructed in simulated data are shown in Figs. 6.7 and 6.8, separated according to flavor. Both distributions feature a signal peak at $\Delta E \approx 0$, overlapping a smooth, slowly-decreasing background, with a signal-to-background ratio at peak of about five for $B^+ \rightarrow K^+K^-K^+$ and of about one for $B^+ \rightarrow K^+\pi^-\pi^+$. The latter features a larger combinatorial background due to the presence of two pions for the reasons discussed above. In the ΔE distribution of $B^+ \rightarrow K^+\pi^-\pi^+$ candidates, I expect two Gaussian-like components on either side of the signal peak, due to $B^+ \rightarrow K^+K^-\pi^+$ and $B^+ \rightarrow \pi^+\pi^-\pi^+$ decays where a kaon is incorrectly reconstructed as a pion or viceversa, respectively. Only the peak of $B^+ \rightarrow K^+K^-\pi^+$ misidentified events is visible in simulated data at $\Delta E \approx -0.05$.

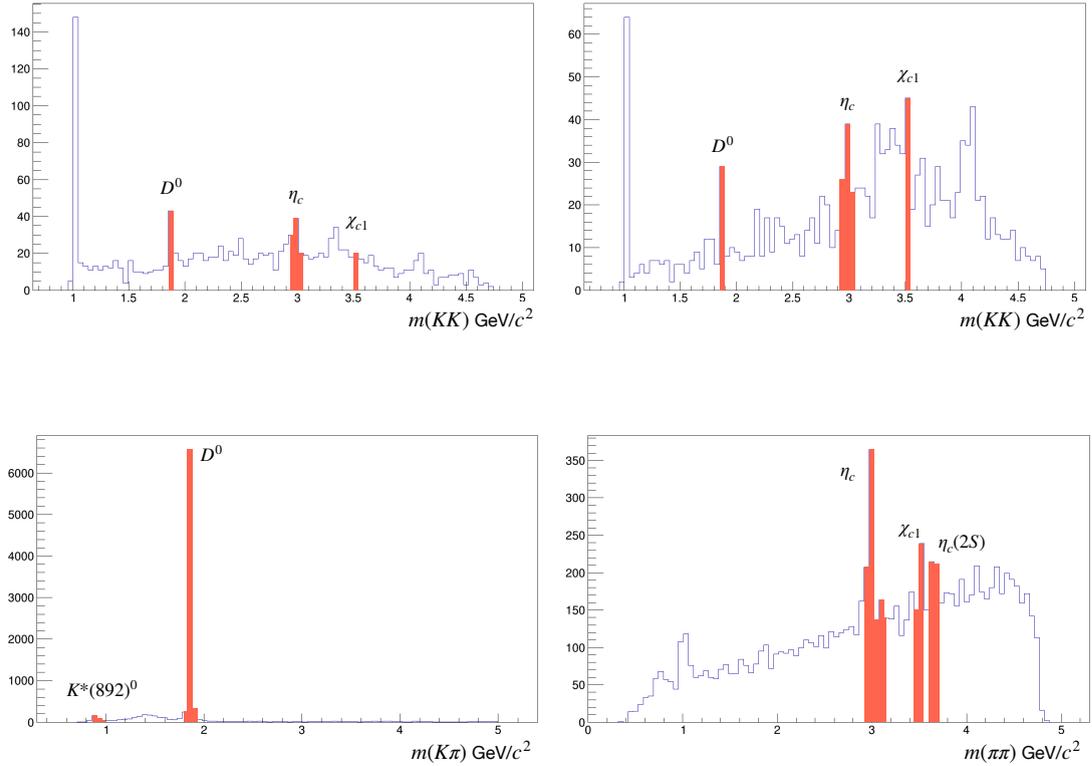


Figure 6.5: Invariant mass distributions for pairs of final-state particles in (top) $B^+ \rightarrow K^+K^-K^+$ decays and (bottom) $B^+ \rightarrow K^+\pi^-\pi^+$ decays. The vetoes are indicated by the shaded regions. Decays $B^+ \rightarrow J/\psi(\rightarrow \mu^+\mu^-)K^+$ and $B^+ \rightarrow \psi(2S)(\rightarrow \mu^+\mu^-)K^+$ misidentified as $B^+ \rightarrow K^+\pi^-\pi^+$ are not shown.

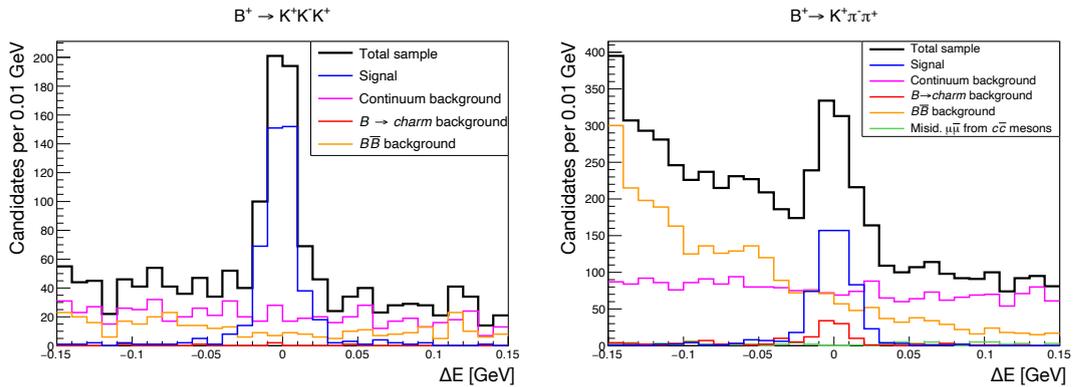


Figure 6.6: Distributions of ΔE for (left) $B^+ \rightarrow K^+K^-K^+$ and (right) $B^+ \rightarrow K^+\pi^-\pi^+$ candidates reconstructed in 50 fb^{-1} of simulated data after applying the vetoes to remove peaking backgrounds. Each color indicate a leading contribution to the sample.

6.2.4 Asymmetry determination

I simultaneously fit the unbinned ΔE distributions of bottom and antibottom candidates using the raw partial-decay-rate asymmetry \mathcal{A} defined in Eq. (6.1) as a fit parameter. The probability density functions (*pdf*) \mathcal{P} and $\bar{\mathcal{P}}$ describe the distributions of the B^+ and B^- candidates, respectively

$$\mathcal{P} = N \left(\frac{f_{\text{sig}}}{2} (1 - \mathcal{A}) \mathcal{P}_{\text{sig}} + \left(1 - \frac{f_{\text{sig}}}{2} (1 - \mathcal{A})\right) \mathcal{P}_{\text{bkg}} \right), \quad (6.2)$$

$$\bar{\mathcal{P}} = \bar{N} \left(\frac{f_{\text{sig}}}{2} (1 + \mathcal{A}) \mathcal{P}_{\text{sig}} + \left(1 - \frac{f_{\text{sig}}}{2} (1 + \mathcal{A})\right) \mathcal{P}_{\text{bkg}} \right), \quad (6.3)$$

where $N(\bar{N})$ is the total number of $B^+(B^-)$ candidates, f_{sig} is the fraction of signal events, \mathcal{P}_{sig} is the signal *pdf* and \mathcal{P}_{bkg} is the background *pdf*. The total yields are known and fixed in the fit, as their fluctuations are irrelevant in an asymmetry.

6.2.4.1 Fit model

I use the same signal model for the B^+ and B^- samples, as I expect the signal features to be independent of charge. Only the charge-specific signal and background yields are determined by the fit. I empirically describe signal shapes for both channels using the sum of a Gaussian and a Crystal Ball function (see Sec. 3.3.5 for details), with shape parameters fixed from simulation. The background is empirically described by an exponential function with floating slope common to the B^+ and B^- samples. The $B^+ \rightarrow K^+\pi^-\pi^+$ channel needs models for the two additional misidentified backgrounds, $B^+ \rightarrow K^+K^-\pi^+$ and $B^+ \rightarrow \pi^+\pi^-\pi^+$. I model their contributions with the same shape of the signal, fixing their displacement from the signal from simulation. In addition, I account for the residual peaking backgrounds surviving the vetoes, $B^+ \rightarrow K_2^{*0}(\rightarrow K^+\pi^-)\pi^+$, $B^+ \rightarrow \chi_{c0}(\rightarrow \pi^+\pi^-)K^+$, $B^+ \rightarrow D^0(\rightarrow \pi^+\pi^-)K^+$, and $B^+ \rightarrow \psi(3770)(\rightarrow \pi^+\pi^-)K^+$ decays (Fig. 6.6), with the sum of three Gaussians, as suggested by Monte Carlo. Charge-specific yields of all peaking backgrounds are fixed to the expected proportions from simulation.

6.2.4.2 Results using simulated data

I first apply the analysis on 600 fb^{-1} of simulated data. Figures 6.7–6.8 show the charge-specific ΔE distributions of $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$ candidates with fit projections overlaid. No evidence for obvious B^+ -vs- B^- asymmetries is visible in simulated data. In Fig. 6.8, a background mismodeling is visible at $\Delta E < -0.05 \text{ GeV}$, due presumably to additional unmodeled peaking backgrounds. The large size of the simulated sample exposes the limits of the model. However, we do not expect this to be a concern in data, where the expected mismodelings are smaller compared to statistical fluctuations due to the small sample size, and be accommodated by the flexibility of the model. In Section 6.5 further checks are reported to rule out biases induced by such background mismodelings.

The fit yields the following raw charge-asymmetries in simulated data

$$\mathcal{A}(B^+ \rightarrow K^+K^-K^+) = 0.020 \pm 0.015, \quad (6.4)$$

$$\mathcal{A}(B^+ \rightarrow K^+\pi^-\pi^+) = -0.020 \pm 0.023, \quad (6.5)$$

where only the statistical component of the uncertainty is reported.

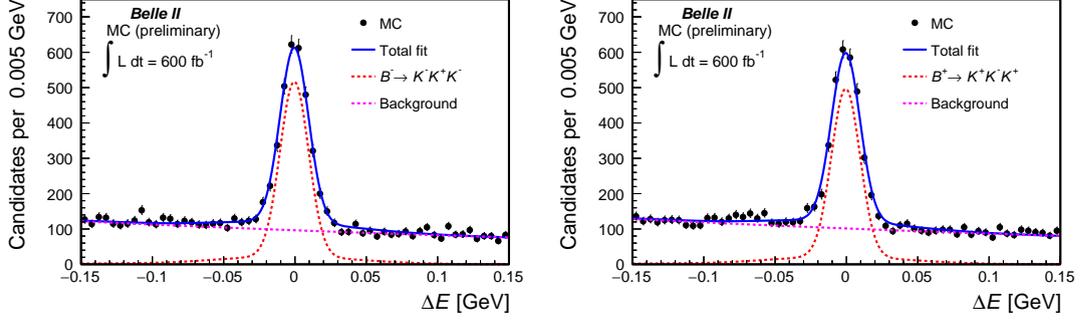


Figure 6.7: Distributions of ΔE for (left) $B^- \rightarrow K^- K^+ K^-$ and (right) $B^+ \rightarrow K^+ K^- K^+$ candidates reconstructed in 600 fb^{-1} of simulated data. The projection of an unbinned maximum likelihood fit to the charge asymmetry is overlaid.

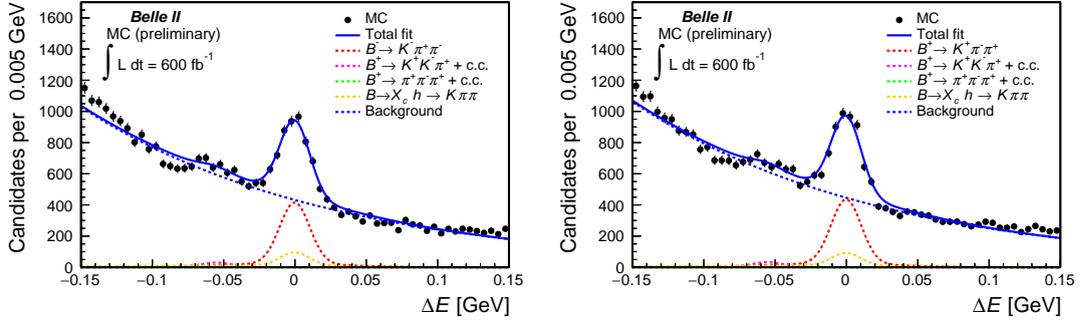


Figure 6.8: Distributions of ΔE for (left) $B^- \rightarrow K^- K^+ K^-$ and (right) $B^+ \rightarrow K^+ K^- K^+$ candidates reconstructed in 600 fb^{-1} of simulated data. The projection of an unbinned maximum likelihood fit to the charge asymmetry is overlaid.

6.2.4.3 Results using Belle II data

After having developed and performed the analysis on simulated data, I apply it on the summer-2020 Belle II data set, corresponding to an integrated luminosity of 34.6 fb^{-1} . I first test for the need of additional degrees of freedom to the signal model to allow sufficient flexibility to account for possible discrepancies between the simulated and experimental distributions. These degrees of freedom are a signal peak-shift that accounts for possible miscalibrations in track reconstruction measurements and a signal width scale-factor that accounts for a possible mismatch of momentum resolution in data and simulation.

By performing a likelihood-ratio-based study similar to the study of Sec. 3.10, I obtain the p -values associated with the various models, presented in Table 6.2, which show that data favor adding a peak-shift degree of freedom as free parameter.

The background is described by an exponential function as in simulation. In the $B^+ \rightarrow K^+ \pi^- \pi^+$ fit the yields of the $B^+ \rightarrow K^+ K^- \pi^+$ and $B^+ \rightarrow \pi^+ \pi^- \pi^+$ contributions are also floating, while the expected charge-specific residual background yields at $\Delta E \approx 0$ are fixed from simulation.

Figures 6.9–6.10 show the charge-specific ΔE distributions with fit projections overlaid for $B^+ \rightarrow K^+ K^- K^+$ and $B^+ \rightarrow K^+ \pi^- \pi^+$ candidates reconstructed in Belle II data. In Fig. 6.9 both distributions feature a narrow signal peak centered at $\Delta E \approx 0$ overlapping a uniform background, with a signal-to-background ratio of about five at peak. The shape is comparable with that obtained in simulation and no evidence for yield asymmetries is

Degree of freedom	p -value
$B^+ \rightarrow K^+ K^- K^+$:	
Common peak-shift	0.3171
Charge-specific peak-shift	0.6063
Common peak-shift, common scale-factor	0.6063
Charge-specific peak-shift, common scale-factor	0.801
$B^+ \rightarrow K^+ \pi^- \pi^+$:	
Common peak-shift	0.3170
Charge-specific peak-shift	0.6062
Common peak-shift, common scale-factor	0.6061
Charge-specific peak-shift, common scale-factor	0.8009

Table 6.2: p -values for the alternate signal models in the $B^+ \rightarrow K^+ K^- K^+$ and $B^+ \rightarrow K^+ \pi^- \pi^+$ analyses.

visible in data. Figure 6.10 shows a large signal peak over a more abundant background, with a signal-to-background ratio of about one at peak. In the distribution of B^+ candidates, peaks due to misidentified $B^+ \rightarrow K^+ K^- \pi^+$ and $B^+ \rightarrow \pi^+ \pi^- \pi^+$ decays are visible on the sides of the signal peak, with sizes larger than observed in simulation. These contributions are less prominent in the distribution of B^- candidates, suggesting a possible charge-dependent data-simulation mismatch in misidentification probability.

Table 6.3 summarizes the resulting charge-specific signal yields and raw charge-asymmetries in Belle II data, compared with simulation. The 0.25 ratio between the obtained statistical uncertainties in simulation and Belle II data is in good agreement with the 0.24 expectation from the comparison of the sample sizes. The results in data and simulation show good agreement, indicating that Belle II performances in these decays match already the expectations from simulation.

Decay	Yield (data)		Raw asymmetry	
	B^+	B^-	Data	Simulation
$B^+ \rightarrow K^+ K^- K^+$	191 ± 16	168 ± 16	-0.064 ± 0.063	0.020 ± 0.015
$B^+ \rightarrow K^+ \pi^- \pi^+$	241 ± 26	206 ± 26	-0.078 ± 0.081	-0.020 ± 0.023

Table 6.3: Summary of charge-specific signal yields for the measurement of \mathcal{CP} -violating asymmetries in Belle II data and comparison of raw charge-asymmetry obtained in data and simulation. Only the statistical contributions to the uncertainties are reported.

6.3 Determination of instrumental asymmetries

In measurements of \mathcal{CP} -violating asymmetries, the observed charge-specific raw asymmetries \mathcal{A} are in general due to the combination of genuine \mathcal{CP} -violating effects in the weak-interaction dynamics and instrumental asymmetries associated with the process of detecting and reconstructing the final-state particles. For instance, in measurements involving final-state charged kaons, we expect an asymmetry from the known difference in

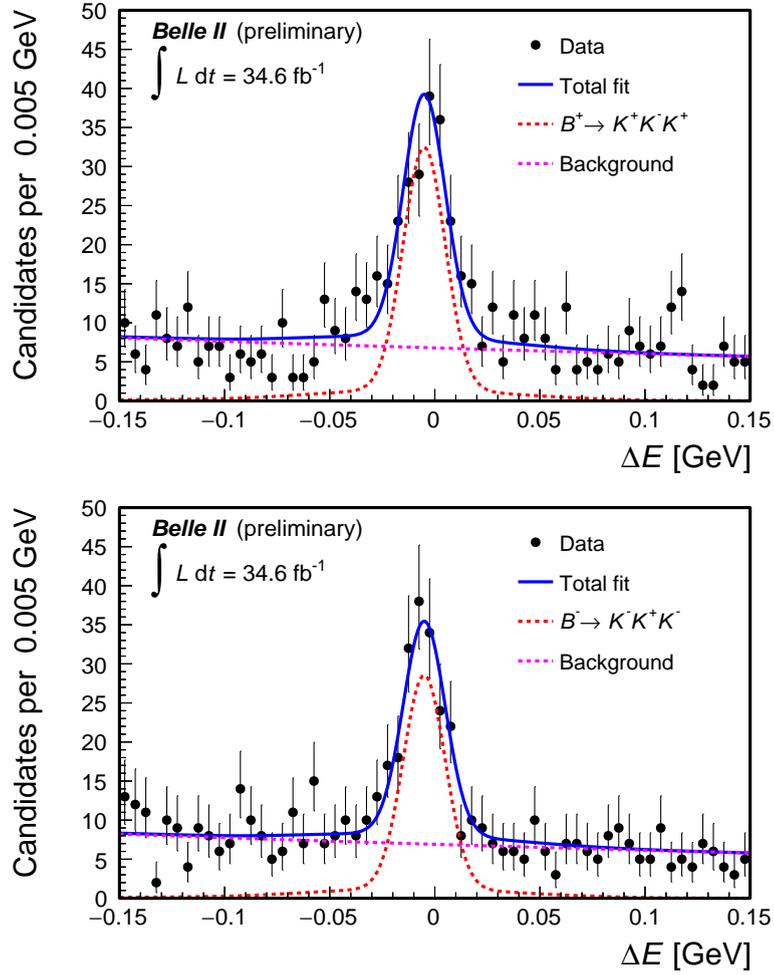


Figure 6.9: Distributions of ΔE for (top) $B^+ \rightarrow K^+K^-K^+$ and (bottom) $B^- \rightarrow K^-K^+K^-$ candidates reconstructed in 2019–2020 Belle II data, corresponding to 34.6 fb^{-1} of integrated luminosity. The projection of an unbinned maximum likelihood fit to the charge asymmetry is overlaid.

K^+/K^- interaction probabilities with matter due to differing strangeness content between K^+ and K^- mesons: the presence of the \bar{u} quark in the negative kaon allows a wider variety of interaction processes, as the production of strange baryons $\Sigma^*(1385)$, $\Lambda(1405)$, $\Sigma(1660)$, which is forbidden in the K^+ case. The asymmetry depends on the particle momentum and position inside the tracking volume, as interaction cross-sections depend on the incident-particle energy and particles with different trajectories traverse different amounts and species of materials. The combination of genuine weak effects ($\mathcal{A}_{\mathcal{CP}}$) and instrumental effects (\mathcal{A}_{det}) is additive for small asymmetries, $\mathcal{A} = \mathcal{A}_{\mathcal{CP}} + \mathcal{A}_{\text{det}}$, with

$$\mathcal{A}_{\text{det}}(X) = \frac{X - \bar{X}}{X + \bar{X}},$$

where X corresponds to a given final state and \bar{X} to its charge-conjugate. Hence, observed raw charge-specific decay yields need be corrected for instrumental effects to determine the genuine \mathcal{CP} -violating asymmetries.

The channels $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$ are only affected by the instrumental asymmetry due to the reconstruction of the charged kaon, $\mathcal{A}(K^-)$, as the net

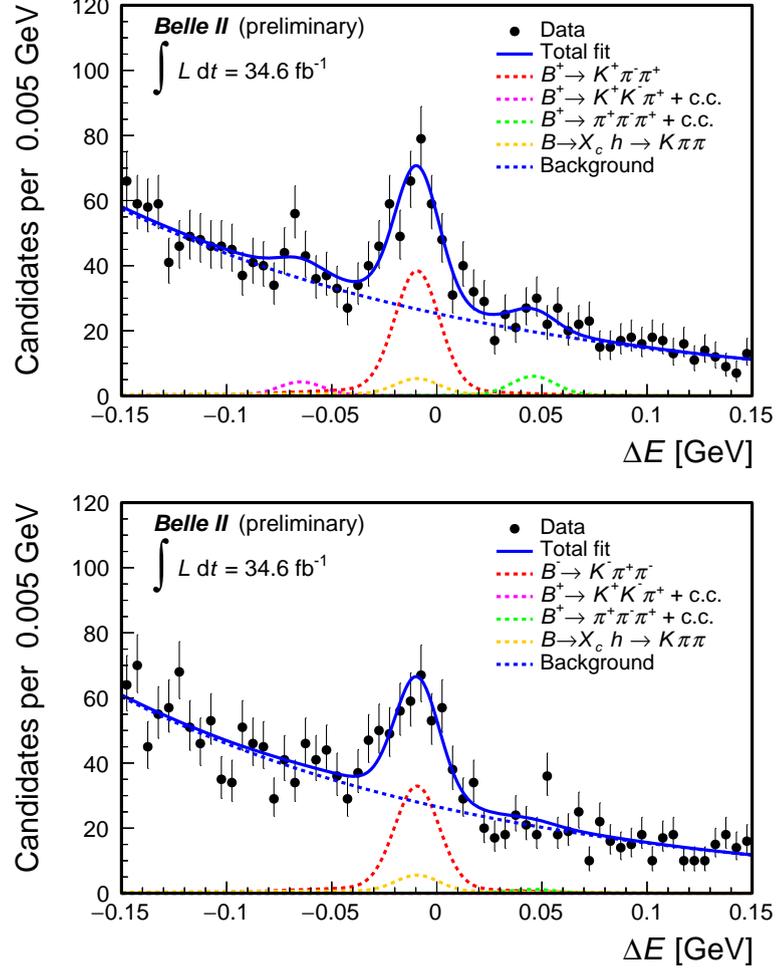


Figure 6.10: Distributions of ΔE for (top) $B^+ \rightarrow K^+\pi^-\pi^+$ and (bottom) $B^- \rightarrow K^-\pi^+\pi^-$ candidates reconstructed in 2019–2020 Belle II data, corresponding to 34.6fb^{-1} of integrated luminosity. The projection of an unbinned maximum likelihood fit to the charge asymmetry is overlaid.

instrumental asymmetry of two opposite-charge hadrons of the same species (K^+K^- or $\pi^+\pi^-$) vanishes.

I determine $\mathcal{A}(K^-)$ by first measuring the charge-asymmetry $\mathcal{A}(K^-\pi^+)$ in an abundant sample of $D^0 \rightarrow K^-\pi^+$ decays. For these decays, \mathcal{CP} -violation is expected to be smaller than 0.1%, if any [17]. I therefore attribute any nonzero observed asymmetry $\mathcal{A}(K^-\pi^+)$ in $D^0 \rightarrow K^-\pi^+$ decays to instrumental effects. However, the instrumental asymmetry $\mathcal{A}(K^-\pi^+)$ determined from $D^0 \rightarrow K^-\pi^+$ decays includes a contribution associated with possible π^+/π^- asymmetries, which need to be subtracted. I use an abundant sample of $D^+ \rightarrow \bar{K}^0\pi^+$ decays, where direct \mathcal{CP} -violation is expected to vanish, to determine $\mathcal{A}(\bar{K}^0\pi^+)$ which includes the π^+/π^- asymmetry I need to subtract, but also folds-in an additional asymmetry $\mathcal{A}(\bar{K}^0)$, associated with neutral kaons. I finally estimate the instrumental asymmetry related to the charged kaon reconstruction alone by combining all inputs in the relationship

$$\mathcal{A}(K^-) = \mathcal{A}(K^-\pi^+) - \mathcal{A}(\bar{K}^0\pi^+) + \mathcal{A}(\bar{K}^0). \quad (6.6)$$

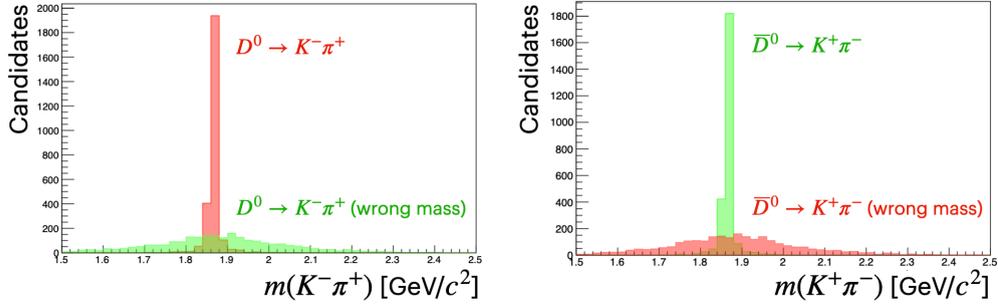


Figure 6.11: Distribution of $m(K\pi)$ for simulated (red) $D^0 \rightarrow K^-\pi^+$ decays (left) properly reconstructed and (right) misreconstructed with swapped mass assignment and (green) simulated $\bar{D}^0 \rightarrow K^+\pi^-$ decays (left) misreconstructed with swapped mass assignment and (right) properly reconstructed.

6.3.1 Samples, tools and selection

I use simulated generic data (see Sec. 3.3) corresponding to an integrated luminosity of 50 fb^{-1} to determine fit models and the details of sample composition. I use the full sample corresponding to 34.6 fb^{-1} for the $D^+ \rightarrow \bar{K}^0\pi^+$ analysis, and reduce the abundant control channel $D^0 \rightarrow K^-\pi^+$ by uniformly sampling it from the full data set to maintain the analysis sample size at a manageable level.

Because charge-dependent instrumental effects depend on the kinematic properties of the sample at hand, I apply to both samples selections similar to those used for the $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$ channels, so that the K^+ are kinematically similar. I select only high-quality tracks reconstructed in the full CDC polar acceptance and featuring a converged track fit. I also require their transverse (longitudinal) distance to the interaction point to meet $dr < 0.5 \text{ cm}$ ($|dz| < 3 \text{ cm}$), to remove tracks associated with particles from beam-background. I reconstruct K_S^0 candidates by selecting pairs of opposite-charge pions reconstructed from a converged vertex fit with dipion mass $450 < m(\pi^+\pi^-) < 550 \text{ MeV}/c^2$. I also require $1.78 < m(K^-\pi^+) < 2.00 \text{ GeV}/c^2$ and $1.78 < m(\bar{K}^0\pi^+) < 2.00 \text{ GeV}/c^2$ to suppress combinatorial background. Finally I require the charm momenta to exceed $2.5 \text{ GeV}/c$ in the center-of-mass frame to suppress the contribution of $B \rightarrow D\pi$ events, which could complicate the analysis.

The invariant mass distributions of genuine D^0 and \bar{D}^0 candidates are expected to show a narrow peaking component of properly reconstructed D -meson decays into $K^\mp\pi^\pm$ final states, along with a broad component of D -meson decays into $K^\pm\pi^\mp$ final states misreconstructed assuming swapped $K^\mp\pi^\pm$ masses as shown in Fig. 6.11. In each sample, I veto events that, reconstructed with swapped mass assignments, are consistent with a D -meson decay, to reduce the contamination from misreconstructed opposite-flavor charm decays that would dilute the resulting asymmetry. A candidate is excluded if the ‘swapped invariant mass’ m_{sw} is within twice the typical mass resolution from the known D -meson mass, $m_{\text{sw}}(K\pi) < 1.857 \text{ GeV}/c^2$ or $m_{\text{sw}}(K\pi) > 1.872 \text{ GeV}/c^2$.

The resulting invariant mass distributions in Belle II data are shown in Fig. 6.12 for $D^0 \rightarrow K^-\pi^+$ candidates and in Fig. 6.13 for $D^+ \rightarrow \bar{K}^0\pi^+$ candidates. Both distributions show a peak at the corresponding charmed-meson mass, overlapping a smooth background dominated by continuum. The $D^+ \rightarrow \bar{K}^0\pi^+$ distribution features more abundant background, due to the lower decay rate (about one third of the $D^0 \rightarrow K^-\pi^+$) and the larger combinatorial background due to an additional final-state track.

6.3.2 $K^-\pi^+$ instrumental asymmetry

I fit simultaneously the unbinned $K\pi$ mass distributions of D^0 and \bar{D}^0 candidates, setting the raw signal-yield $K^-\pi^+$ charge asymmetry

$$\mathcal{A}(K^-\pi^+) = \frac{N(K^+\pi^-) - N(K^-\pi^+)}{N(K^+\pi^-) + N(K^-\pi^+)}, \quad (6.7)$$

as a free parameter. The parameter N indicates the signal yield in each sample.

Data are shown in Fig. 6.12, with fit projections overlaid. In both distributions, a narrow signal peak centered at the D^0 mass overlaps a flat, smoothly decreasing background, with signal-to-background ratio at peak of about three.

The fit results are in Table 6.4. The charge asymmetries are about 1% with 0.3% uncertainties and consistent in simulation and data.

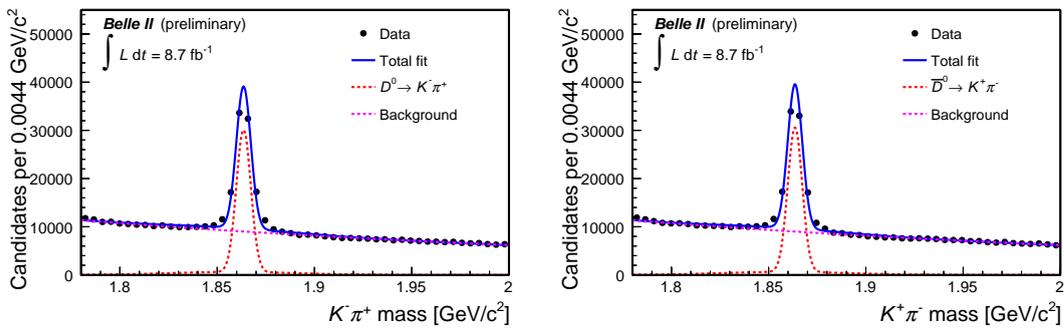


Figure 6.12: Invariant $K\pi$ -mass distributions for (left) $D^0 \rightarrow K^-\pi^+$ and (right) $\bar{D}^0 \rightarrow K^+\pi^-$ candidates simultaneously reconstructed in a sample of Belle II data. Fit projections are overlaid.

	Data	Simulation
$\mathcal{A}(K^-\pi^+)$	0.008 ± 0.003	0.012 ± 0.001

Table 6.4: Charge-asymmetry results for $D^0 \rightarrow K^-\pi^+$ decays selected through the $B^+ \rightarrow K^+\pi^-\pi^+$ selection and reconstructed in simulated and experimental data.

We expect that the instrumental asymmetries depend on kinematics. Interaction cross-sections with the detector materials depend on the particle momentum, and the thickness and composition of material encountered by the incident particle depends on its trajectory.

I therefore measure $\mathcal{A}(K^-\pi^+)$ as a function of D -meson momenta, polar angles, and azimuthal angles. Tables 6.5, 6.6, and 6.7 show the results. No evidence of kinematic dependencies of the charge asymmetry within 0.4–0.7% uncertainties is found.

D^0 momentum (GeV/c)	< 1.1	1.1 – 1.8	> 1.8
$\mathcal{A}_{\text{data}}(K^-\pi^+)$	0.002 ± 0.007	0.018 ± 0.005	0.009 ± 0.004
$\mathcal{A}_{\text{MC}}(K^-\pi^+)$	0.009 ± 0.002	0.010 ± 0.002	0.016 ± 0.002

Table 6.5: Charge-asymmetry as a function of D momentum in simulated and experimental data.

$ \cos\theta $	> 0.8	$0.4 - 0.8$	< 0.4
$\mathcal{A}_{\text{data}}(K^-\pi^+)$	0.015 ± 0.006	0.011 ± 0.005	0.004 ± 0.006
$\mathcal{A}_{\text{MC}}(K^-\pi^+)$	0.011 ± 0.001	0.012 ± 0.003	0.013 ± 0.004

Table 6.6: Charge-asymmetry as a function of D polar-angle in simulated and experimental data.

$ \phi $	> 0.6	$0.3 - 0.6$	< 0.3
$\mathcal{A}_{\text{data}}(K^-\pi^+)$	0.012 ± 0.005	0.017 ± 0.006	0.004 ± 0.005
$\mathcal{A}_{\text{MC}}(K^-\pi^+)$	0.011 ± 0.002	0.011 ± 0.003	0.013 ± 0.003

Table 6.7: Charge-asymmetry as a function of D azimuthal-angle in simulated and experimental data.

6.3.3 $\bar{K}^0\pi^+$ instrumental asymmetry

I use an abundant sample of $D^+ \rightarrow \bar{K}^0\pi^+$ decays, where only $\bar{K}^0 \rightarrow K_S^0(\rightarrow \pi^+\pi^-)$ decays are reconstructed, to subtract the π^+/π^- asymmetry component from the instrumental asymmetry $\mathcal{A}(K^-\pi^+)$ reported in the previous section.

A analysis similar to the previous determines the charge-asymmetry $\mathcal{A}(\bar{K}^0\pi^+)$ from a simultaneous fit on the charge-conjugate samples. Figure 6.13 shows the $K_S^0\pi^\pm$ mass distribution for the charge-conjugate samples with fit results overlaid. In both distributions, a peak centered at the D^+ mass overlaps an abundant, smoothly decreasing background, with a signal-to-background ratio at peak of about 0.25.

Table 6.8 summarizes the final charge-asymmetries for the channel $D^+ \rightarrow \bar{K}^0\pi^+$ in data and simulation. The resulting asymmetries are consistent with zero both in data and simulation.

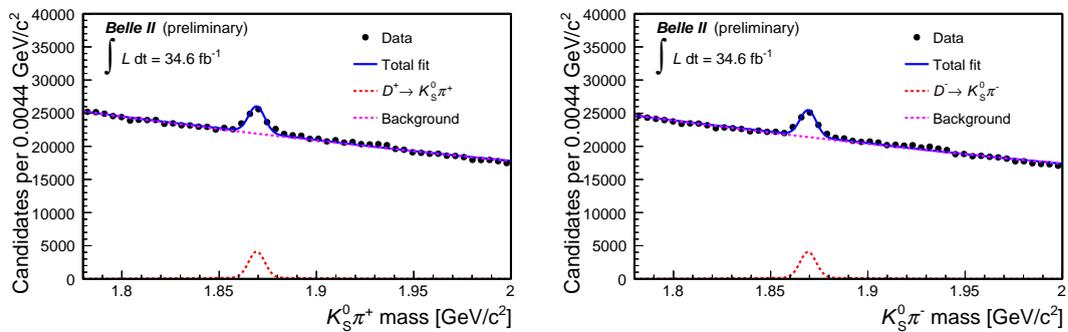


Figure 6.13: Invariant $\bar{K}^0\pi$ -mass distributions for (left) $D^+ \rightarrow \bar{K}^0\pi^+$ and (right) $D^- \rightarrow \bar{K}^0\pi^-$ candidates simultaneously reconstructed in Belle II data corresponding to 34.6 fb^{-1} . Fit projections are overlaid.

	Data	Simulation
$\mathcal{A}(\bar{K}^0 \pi^+)$	-0.007 ± 0.022	0.014 ± 0.022

Table 6.8: Charge-asymmetry results for $D^+ \rightarrow \bar{K}^0 \pi^+$ decays selected through the $B^0 \rightarrow K^+ \pi^- \pi^+$ selection and reconstructed in simulated and experimental data.

6.3.4 Determination of the final K^+/K^- instrumental asymmetry

I obtain the instrumental asymmetry associated with the reconstruction of a charged kaon from the relationship

$$\mathcal{A}(K^-) = \mathcal{A}(K^- \pi^+) - \mathcal{A}(\bar{K}^0 \pi^+) + \mathcal{A}(\bar{K}^0), \quad (6.8)$$

where $\mathcal{A}(K^- \pi^+)$ and $\mathcal{A}(\bar{K}^0 \pi^+)$ are determined as per previous discussion and $\mathcal{A}(\bar{K}^0)$ is the asymmetry associated with the neutral kaon \bar{K}^0 , which is detected as a K_S^0 . This term receives two contributions, an asymmetry due to \mathcal{CP} violation in neutral kaons, which is known and $\mathcal{O}(10^{-3})$, and an asymmetry from strangeness-dependent interactions with the detector material. The flavor eigenstates K^0 and \bar{K}^0 are distinct superpositions of the weak eigenstates K_S^0 and K_L^0 . In Belle II, equal numbers of D^+ and D^- mesons are produced. Hence, final-state K^0 and \bar{K}^0 too are in equal numbers at $t = 0$. In the absence of \mathcal{CP} violation, the probability of observing a K_S^0 after a time t if the initial state was a K^0 equals the probability of observing a K_S^0 if the initial state was a \bar{K}^0 . \mathcal{CP} violation in neutral kaons generates a difference between these probabilities: therefore, the number of K_S^0 reconstructed at time t associated with an initial K^0 differs from the number of reconstructed K_S^0 associated with an initial \bar{K}^0 , leading to an asymmetry in the yields of reconstructed $D^+ \rightarrow K_S^0 \pi^+$ and $D^- \rightarrow K_S^0 \pi^-$ decays.

This asymmetry is further modified by the presence of material. Due to its strangeness content, K^0 meson can only interact (quasi-)elastically with nucleons in the detector, while \bar{K}^0 can also excite a nucleon into a Λ or a Σ hyperon or their excited states. This results in a difference in interaction cross sections, which depends on the momentum of the kaon and on the density of nucleons, that modifies the amplitudes of the eigenstates K_S^0 and K_L^0 and therefore the probability to observe a K_S^0 from an initial K^0 or \bar{K}^0 state.

I estimate $\mathcal{A}(\bar{K}^0)$ by using the results of the LHCb collaboration, obtained by simulating neutral kaon mixing and the propagation through the detector [64]. Since the Belle II detector has less material than LHCb, the LHCb values for the asymmetry provide a (conservative) upper limit for the effect, even taking into account the differences in kinematics. As the resulting correction is smaller than 0.1%, we assume it to be zero with an associated uncertainty of 0.2%. This is consistent with previous studies performed at Belle [65].

By combining all inputs, I obtain $\mathcal{A}(K^-) = 0.015 \pm 0.022$, where the uncertainty is only statistical and dominated by the uncertainty on the measurement of $\mathcal{A}(\bar{K}^0 \pi^+)$.

I finally correct the raw charge-asymmetries in the decay channels $B^+ \rightarrow K^+ K^- K^+$ and $B^+ \rightarrow K^+ \pi^- \pi^+$ using the instrumental asymmetry $\mathcal{A}(K^+)$,

$$\mathcal{A}_{\mathcal{CP}} = \mathcal{A} - \mathcal{A}_{\text{det}}(K^+) = \mathcal{A} + 0.015. \quad (6.9)$$

6.4 Systematic uncertainties

I consider three sources of systematic effects and evaluate the associated uncertainties.

Signal modeling. In the asymmetry fits, the signal model is empirically fixed from simulated data, and the model is the same for both the B^+ and B^- samples. If the B^+ and B^- signal ΔE shapes differ due, for example, to differences in reconstruction of positively and negatively charged particles, the fit would misinterpret the difference in model as a difference in yield, and the raw-asymmetry measurement would be biased. I evaluate the impact of the assumption by comparing the results obtained by fitting data with charge-symmetric or charge-specific models and take the difference between the results as systematic uncertainty. Table 6.9 shows the results of an example of an alternative ΔE fit performed considering independent and floating peak-shifts for B^+ and B^- signal decays. The resulting charge-asymmetries show little difference (2% of the statistical uncertainty), confirming no evidence for asymmetric signal shapes.

	Default model	Alternative model
$B^+ \rightarrow K^+ K^- K^+$	-0.067 ± 0.063	-0.068 ± 0.063
$B^+ \rightarrow K^+ \pi^- \pi^+$	-0.079 ± 0.081	-0.077 ± 0.081

Table 6.9: Raw charge-asymmetries in experimental data for $B^+ \rightarrow K^+ K^- K^+$ and $B^+ \rightarrow K^+ \pi^- \pi^+$ decays using the default fit model and an alternative model where the charge-conjugate samples have independent peak-shifts.

Peaking background modeling. In the $B^+ \rightarrow K^+ \pi^- \pi^+$ channel, I estimate the size of the residual peaking backgrounds that survive the vetoes by using simulation. The final results can be biased if the simulation of peaking backgrounds does not describe properly experimental data, as events from peaking backgrounds could be mistakenly absorbed into the signal by the fit. I therefore compare the results obtained by fixing the charge-conjugate components to the expectations from simulation and by fixing them to half of the total peaking background yield. The observed difference between results using these alternative models is only 5% of the statistical uncertainty on \mathcal{A} .

Instrumental asymmetry correction. The statistical uncertainty on the determination of $\mathcal{A}(K^-)$ contributes a systematic uncertainty.

Table 6.10 shows a summary of systematic uncertainties. The total systematic uncertainty, obtained by summing in quadrature the individual components, amounts to about 25% of the statistical uncertainty. This is reassuring in view of future, larger samples, showing how the current limiting factor on the precision is the small size of the data sample, and not a systematic effect.

Source	$B^+ \rightarrow K^+ K^- K^+$	$B^+ \rightarrow K^+ \pi^- \pi^+$
Signal model	0.001	0.003
Pkg. background model	–	0.004
Instrumental asymmetry correction	0.022	0.022
Total	0.022	0.023

Table 6.10: Summary of (absolute) systematic uncertainties in the $\mathcal{A}_{\mathcal{CP}}$ measurements.

6.5 Consistency checks

A further effect that could bias the raw asymmetry is the possible presence of charge-asymmetries in background. Large background asymmetries under the signal could be leaking into the signal asymmetry if the fit does not separate properly signal from background. I therefore determine the background charge-asymmetry in various independent background-rich sideband regions of the ΔE distribution. I identify four different regions, chosen to feature approximately uniform background composition in each. In the $B^+ \rightarrow K^+\pi^-\pi^+$ sample, the *extreme* regions ($0.10 < |\Delta E| < 0.15$ GeV) only include continuum background, while the *intermediate* regions ($0.05 < |\Delta E| < 0.10$ GeV) receive contributions from misidentified $B^+ \rightarrow K^+K^-\pi^+$ and $B^+ \rightarrow \pi^+\pi^-\pi^+$ decays too. No major peaking backgrounds are expected in the $B^+ \rightarrow K^+K^-K^+$ sample, but the same ΔE subdivision is applied for consistency. The results are shown in Table 6.11 and in Fig. 6.14. All results but one (corresponding to the ΔE region $-0.10 < \Delta E < -0.05$ GeV in the $B^+ \rightarrow K^+K^-K^+$ case) show asymmetries compatible with zero. No evidence of background charge-asymmetry is found.

ΔE sideband region [GeV]	$[-0.15, -0.10]$	$[-0.10, -0.05]$	$[0.05, 0.10]$	$[0.10, 0.15]$
$B^+ \rightarrow K^+K^-K^+$	0.043 ± 0.079	0.224 ± 0.080	-0.008 ± 0.089	-0.103 ± 0.092
$B^+ \rightarrow K^+\pi^-\pi^+$	0.024 ± 0.031	-0.002 ± 0.035	-0.005 ± 0.051	0.030 ± 0.061

Table 6.11: Background charge-asymmetries obtained in various sideband ΔE regions for the $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$ decay channels.

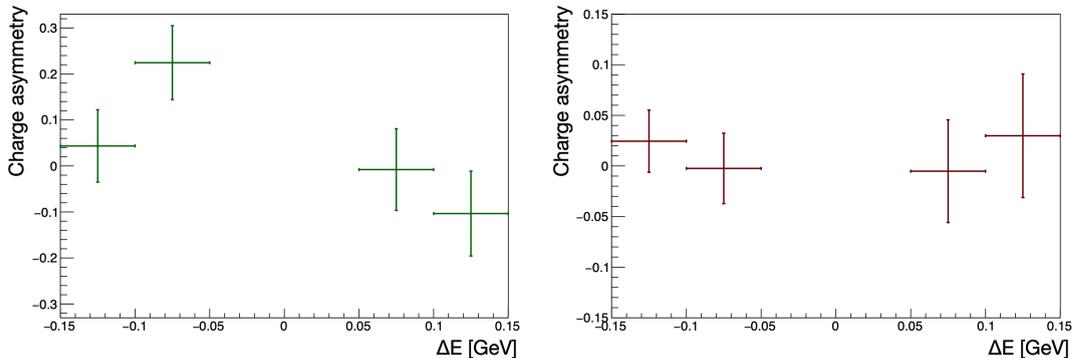


Figure 6.14: Background charge-asymmetries measured in sideband regions of the ΔE distribution of (left) $B^+ \rightarrow K^+K^-K^+$ decays and (right) $B^+ \rightarrow K^+\pi^-\pi^+$ decays.

6.6 Results and summary

By applying the instrumental correction to the raw charge-asymmetries and considering the systematic uncertainties discussed in the previous section, I obtain the following measurements of direct \mathcal{CP} -violating asymmetries:

$$\mathcal{A}_{\mathcal{CP}}(B^+ \rightarrow K^+K^-K^+) = -0.049 \pm 0.063(\text{stat}) \pm 0.022(\text{syst}), \quad (6.10)$$

$$\mathcal{A}_{\mathcal{CP}}(B^+ \rightarrow K^+\pi^-\pi^+) = -0.063 \pm 0.081(\text{stat}) \pm 0.023(\text{syst}) \quad (6.11)$$

These are the first such measurements in three-body charmless decays reported by Belle II. The statistical and systematic uncertainties are still large with respect to current best values, but the results are compatible with known determinations and represent the first step for future studies and analyses of charmless B decays at Belle II.

This study offers significant contributions to the characterization of the detector as it provides the first data-driven determination of instrumental asymmetries in Belle II data, which serve for many other \mathcal{CP} -violation measurements.

Summary

The Standard Model (SM) describes accurately thousands of measurements associated with three (out of the four) fundamental interactions spanning many orders of magnitude in energy. However, many open questions suggest that the SM ought to be completed by a more general theory that extends it at energies higher than the TeV scales explored so far. Determining the theory that completes the SM is the main goal of today's high-energy physics. Flavor physics offers a powerful probe into higher-energy completions of the SM. Contributions from virtual (off mass-shell) non-SM particles of arbitrary high mass to known lower-energy processes may be identified, if measurements and SM predictions are sufficiently precise to discern inconsistencies.

The Belle II experiment, with its target of nearly 1000 B -meson decays collected per second with a newly upgraded detector, will advance the progress in flavor physics for the next decade, guiding the next steps in indirect searches for the completion of the Standard Model at the intensity frontier. Charmless B decays play a key role in this endeavor, as they provide unique access to parameters of the weak interactions highly sensitive to non-SM dynamics.

Belle II started to collect data in 2019. Chief priority in this initial stage is studying known physics processes to understand, calibrate, and optimize the performance of the new detector and reconstruction algorithms. Because of the variety of the decay topologies and final states involved, and their rarity, charmless B decays are sensitive probes of many essential aspects of the detection and reconstruction performance.

This thesis reports three synergic contributions. I firstly target the first reconstruction in Belle II data of the decay channel $B^+ \rightarrow K^+ \pi^0$ using the sample collected in 2019 and corresponding to 8.7 fb^{-1} of integrated luminosity. The task is challenging because $B^+ \rightarrow K^+ \pi^0$ decays are rare, contaminated by abundant backgrounds, and exposed to nontrivial instrumental challenges associated with π^0 reconstruction. The key is to achieve an efficient discrimination between signal and background while maintaining signal efficiency high. I meet the challenge by optimizing in detail the baseline and advanced sample selection using simulation and control samples. In addition, anomalies in the ΔE spectrum of signal events prompt an in-depth investigation of the calorimeter-energy scale using a control sample, analytical kinematic derivations, and simulation.

In the central part of my work, I study thoroughly the charged-hadron identification performance, which is essential to suppress misidentification backgrounds in B physics analyses. Belle II reconstruction algorithms combine identification informations from various detectors assuming their independence. I perform a detailed investigation of the assumption and its impact on physics, using abundant, low-background samples of kaons and pions from $D^{*+} \rightarrow D^0(\rightarrow K^- \pi^+) \pi^+$ decays and simulation.

Finally, I target the first measurement of \mathcal{CP} -violating asymmetries in $B^+ \rightarrow K^+ K^- K^+$ and $B^+ \rightarrow K^+ \pi^- \pi^+$ decays using the data sample available for summer 2020, corresponding to 34.6 fb^{-1} of integrated luminosity. These channels pose additional challenges with

respect to the $B^+ \rightarrow K^+\pi^0$ study due to significant peaking backgrounds and the need to identify possible instrumental charge-asymmetries that could bias the results. I use simulation to identify and suppress the principal peaking backgrounds and $D^0 \rightarrow K^-\pi^+$ and $D^+ \rightarrow \bar{K}^0\pi^+$ decays to determine the instrumental effects directly on data. The main results of my work follow.

- I achieve the first reconstruction of $B^+ \rightarrow K^+\pi^0$ decays in Belle II data, observing 27 ± 8 decays corresponding to a statistical significance of 3.95σ . This channel probes key aspects of the detection and reconstruction in Belle II, such as the background suppression and the reconstruction of high-energy photons;
- prompted by anomalies observed in the $B^+ \rightarrow K^+\pi^0$ signal, I identify and investigate in detail a previously undetected energy-miscalibration in photon reconstruction, which biases the physics measurements on channels with final-state photons. This motivates the collaboration to include novel probes of calorimeter-calibration performance that are more sensitive than previous Belle II methods;
- I identify unmodeled dependences in the information used for charged-hadron identification, exposing an important limitation in the current approach to charged-hadron identification in Belle II that may introduce biases in measurement using that capability;
- I achieve the first determination of charge-parity violation in $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$ decays using Belle II data. This includes a first thorough investigation, based entirely on data, of instrumental charge-asymmetries, which is a fundamental stepping stone for any future measurement of \mathcal{CP} -violation at Belle II.

All physics results are consistent with expectations and show (luminosity-independent) performance on par with the best performance of the predecessor experiment. This work provides solid ground for upcoming analyses of charmless B decays at Belle II and promises significant impact on the quality of the physics produced by Belle II in the next few years, when competitive precisions with current world-best results will be available.

This work is documented in two published Belle II preprints [1, 2] and in five internal Belle II documents [3–7].

Bibliography

- [1] F. Abudinén *et al.*, (Belle II Collaboration), *Charmless B decay reconstruction in 2019 Belle II data*, 2020 . BELLE2-CONF-PH-2020-001.
- [2] F. Abudinén *et al.*, (Belle II Collaboration), *Measurements of branching fractions and CP-violating charge asymmetries in charmless B decays reconstructed in 2019–2020 Belle II data*, 2020 . BELLE2-CONF-PH-2020-012.
- [3] F. Abudinén, E. Ganiev, R. Manfredi, S. Raiz, and D. Tonelli, *Measurement of instrumental $K^- \pi^+$, $K_S^0 \pi^+$, K^- and π^+ asymmetries and π^0 reconstruction efficiency*, 2020. BELLE2-NOTE-TE-2020-024.
- [4] F. Abudinén, E. Ganiev, R. Manfredi, S. Raiz, and D. Tonelli, *Impact of photon-energy biases on candidate-reconstruction variables*, 2020. BELLE2-NOTE-TE-2020-011.
- [5] F. Abudinén *et al.*, *Skimming selections for summer 2020 charmless analyses*, 2020. BELLE2-NOTE-PH-2020-041.
- [6] F. Abudinén, E. Ganiev, R. Manfredi, S. Raiz, and D. Tonelli, *Charmless B decay reconstruction in 2019 data*, 2020. BELLE2-NOTE-PH-2020-007.
- [7] F. Abudinén *et al.*, *Charmless B decay reconstruction and first measurements in 2019 and 2020 data*, 2020. BELLE2-NOTE-PH-2020-041.
- [8] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.
- [9] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961) 579.
- [10] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519** (1968) 367.
- [11] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.
- [12] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.
- [13] G. Guralnik, C. Hagen, and T. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585.
- [14] 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 separated and free quarks.

BIBLIOGRAPHY

- [15] G. Lüders, *Proof of the TCP theorem*, *Annals of Physics* **2** (1957) 1.
- [16] W. Pauli, *Exclusion Principle, Lorentz Group, and reversal of space-time and charge*, Niels Bohr and the Development of Physics (1955) 1.
- [17] P. Zyla *et al.*, Particle Data Group, *Review of Particle Physics*, *Prog. Theor. Exp. Phys.* **2020** (2020) 083C01.
- [18] M. Gell-Mann and M. Lévy, *The axial vector current in beta decay*, *Il Nuovo Cimento* **16** (1960) 705.
- [19] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531.
- [20] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, *Phys. Rev. D* **2** (1970) 1285.
- [21] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652.
- [22] 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.
- [23] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, *Phys. Rev. Lett.* **51** (1983) 1945.
- [24] 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.
- [25] C. Jarlskog, *A Basis Independent Formulation of the Connection Between Quark Mass Matrices, CP Violation and Experiment*, *Z. Phys. C* **29** (1985) 491.
- [26] I. Dunietz, O. W. Greenberg, and D.-D. Wu, *A priori definition of maximal CP nonconservation*, *Phys. Rev. Lett.* **55** (1985) 2935.
- [27] M. Gell-Mann and A. Pais, *Behavior of Neutral Particles under Charge Conjugation*, *Phys. Rev. Lett.* **97** (1955) 1387.
- [28] M. Battle *et al.*, (CLEO Collaboration), *Observation of B^0 decay to two charmless mesons*, *Phys. Rev. Lett.* **71** (1993) 3922.
- [29] S. Herb *et al.*, *Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions*, *Phys. Rev. Lett.* **39** (1977) 252.
- [30] M. Gronau, *A Precise sum rule among four $B \rightarrow K\pi$ CP asymmetries*, *Phys. Lett. B* **627** (2005) 82.
- [31] R. Aaij *et al.*, (LHCb Collaboration), *Measurements of CP violation in the three-body phase space of charmless B^\pm decays*, *Phys. Rev. D* **90** (2014) 112004.
- [32] Y. Ahmis *et al.*, (Belle II Collaboration), *Averages of b-hadron, c-hadron, and τ -lepton properties as of summer 2016*, *Eur. Phys. J. C* **77** (2017) 895.

BIBLIOGRAPHY

- [33] J. Charles *et al.*, (CKMfitter Group), *Unitarity Triangle plot*, Eur. Phys. J. C41, 1 (2005) [hep-ph/0406184], updated plots available at <http://www.ni.com/white-paper/4752/en/> .
- [34] R. Aaij, (LHCb Collaboration), *Observation of CP Violation in Charm Decays*, Phys. Rev. Lett. **122** (2019) 211803.
- [35] Belle II Collaboration and B2TiP theory Community, *The Belle II Physics Book*. Prog. Theor. Exp. Phys., 1st ed., 2018.
- [36] Y. Ohnishi *et al.*, *Accelerator design at SuperKEKB*, Prog. Theo. Exp. Phys. (2013) 03A011.
- [37] 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.
- [38] P. Raimondi *et al.*, *Beam-beam issues for colliding schemes with large Piwinski angle and crabbed waist*, arXiv:physics/0702033.
- [39] C. Marinas, DEPFET, *The Belle II pixel detector: High precision with low material*, Nucl. Instrum. Meth. **A731** (2013) 31.
- [40] K. Adamczyk *et al.*, (Belle-II SVD Collaboration), *The Belle II silicon vertex detector assembly and mechanics*, Nucl. Instrum. Meth. **A845** (2017) 38.
- [41] N. Taniguchi, (Belle II Collaboration), *Central Drift Chamber for Belle-II*, JINST **12** (2017) C06014.
- [42] A. Abashian *et al.*, (Belle II Collaboration), *The K_L/μ detector subsystem for the BELLE experiment at the KEK B factory*, Nucl. Instrum. Meth. **A449** (2000) 112.
- [43] T. Abe *et al.*, (Belle II Collaboration), *Belle II Technical Design Report*, arXiv:1011.0352.
- [44] 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.
- [45] A. Ryd *et al.*, *EvtGen: A Monte Carlo Generator for B-Physics*, 5, 2005. <https://evtgen.hepforge.org/doc/EvtGenGuide.pdf>.
- [46] T. Sjöstrand *et al.*, *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159.
- [47] T. Kuhr *et al.*, (Belle-II Framework Software Group), *The Belle II Core Software*, Comput. Softw. Big Sci. **3** (2019) 1.
- [48] J. Gaiser, *Appendix-F Charmonium Spectroscopy from Radiative Decays of the J/ψ and ψ'* , SLAC-R-255 (1982) 178.
- [49] T. Keck *et al.*, *The Full Event Interpretation*, Computing and Software for Big Science **3** (2019) 1.
- [50] D. Asner *et al.*, (CLEO Collaboration), *Search for exclusive charmless hadronic B decays*, Phys. Rev. D **53** (1996) 1039.

BIBLIOGRAPHY

- [51] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Ann. Math. Statist. **9** (1938) 60.
- [52] Q.-D. Zhou, (Belle II Collaboration), *Momentum bias correction and systematics*, 2020 . BELLE2-NOTE-PH-2020-030.
- [53] J. Libby, N. Rout, and K. Trabelsi, (Belle II Collaboration), *Beam Energy Calibration*, 2020 . BELLE2-NOTE-PH-2020-002.
- [54] S. Longo, *Neutrals bias*, 2020, Internal Belle II talk .
https://indico.belle2.org/event/2419/contributions/11928/attachments/5986/9410/B2GM_220620_NeutralsBias_SL.pdf.
- [55] H. Ikeda *et al.*, (Belle Collaboration), *NIM A441*, 41 (2000) 1.
- [56] H. Bethe, *Theory of the Passage of Fast Corpuscular Rays Through Matter*, Annalen Phys. **5** (1930) 325.
- [57] F. Bethe, *Stopping power of atoms with several electrons*, Z. Phys. **81** (1933) 363.
- [58] M. Milesi, *Lepton identification in Belle II using observables from the electromagnetic calorimeter and precision trackers*, Talk at CHEP 2019 ,
https://indico.cern.ch/event/773049/contributions/3476156/attachments/1932835/3217425/mmilesi_mva_CHEP_07_11_19.pdf.
- [59] M. Starič *et al.*, *Likelihood analysis of patterns in a time-of-propagation (TOP) counter*, Nucl. Instrum. Meth. A **595** (2008) 252.
- [60] M. Starič, *Pattern recognition for the time-of-propagation counter*, Nuclear Instruments and Methods in Physics Research Section A **639** (2011) 252.
- [61] J. Neyman and E. S. Pearson, *IX. On the problem of the most efficient tests of statistical hypotheses*, Phil. Trans. R. Soc. Lond. A. 231 (1933) 289.
- [62] G. Punzi, *Notes on statistical separation of classes of events*, 2003.
<https://arxiv.org/abs/physics/0611219v1>.
- [63] Although the inequality was first (implicitly) derived by Aitken and Silverstone, it is generally referred to as Cramer-Rao (or Frechet) inequality, following the papers below: C. R. Rao, *Information and accuracy attainable in the estimation of statistical parameters*, Bull. Calcutta Math. Soc., **37**, 81 (1945); H. Cramer, *Mathematical Methods of Statistics*, Princeton University Press (1946).
- [64] A. Davis *et al.*, (LHCb Collaboration), *Measurement of the instrumental asymmetry for $K^- \pi^+$ -pairs at LHCb in Run 2*, LHCb Public Notes 2018.
- [65] B. R. Ko, E. Won, B. Golob, and P. Pakhlov, *Effect of nuclear interactions of neutral kaons on CP asymmetry measurements*, Physical Review D **84** (2011) 111501.