UNIVERSITÀ
DEGLI STUDI DI TRIESTE

MASTER DEGREE IN NUCLEAR AND SUB-NUCLEAR PHYSICS

PHYSICS DEPARTMENT, UNIVERSITY OF TRIESTE

Study of the π^0 reconstruction efficiency for the $B^0 \to \pi^0\pi^0$ decay at the Belle II experiment

---

Author
Francesco ROSSI

Supervisor
Doctor Mirco DORIGO
Co-supervisor
Doctor Sebastiano RAIZ

October 21, 2022

UNIVERSITÀ
DEGLI STUDI DI TRIESTE

LAUREA MAGISTRALE IN FISICA NUCLEARE E SUB-NUCLEARE

DIPARTIMENTO DI FISICA, UNIVERSITÀ DI TRIESTE

Studio dell'efficienza di ricostruzione del π^0 per il decadimento $B^0 \to \pi^0\pi^0$ all'esperimento Belle II

Autore

Francesco ROSSI

Relatore

Dott. Mirco DORIGO

Correlatore

Dott. Sabastiano RAIZ

21 Ottobre 2022

Abstract

The aim of this Thesis is to study the efficiency of the reconstruction of the π^0 meson in $\pi^0 \to \gamma\gamma$ decays at the Belle II experiment. Using D^0-meson decays reconstructed in data and simulation, I calculate a correction factor for an unbiased determination of the π^0-reconstruction efficiency from simulation, and I investigate possible dependencies of this factor on the π^0 kinematics and on the location of the photon clusters in the Belle II electromagnetic calorimeter. The π^0-reconstruction efficiency is studied by comparing the measured signal yields of the decays $D^{*+} \to D^0(\to K^-\pi^+\pi^0)\pi_s^+$ and $D^{*+} \to D^0(\to K^-\pi^+)\pi_s^+$, reconstructed in the same data set. I use a sample which corresponds to an integrated luminosity of 190 fb^{-1} recorded by Belle II experiment in energy-asymmetric e^+e^- collisions delivered by the SuperKEKB collider at the center-of-mass energy of $\sqrt{s} = 10.58$ GeV. By comparing the signal-yield ratio in data and simulation, I obtain the correction factor which is determined for the photon- and π^0-selection requirements used in the analysis of $B^0 \to \pi^0\pi^0$ decays. This factor is found to be $1.037 \pm 0.007\,(\text{stat}) \pm 0.037\,(\text{syst})$, where the first uncertainty is statistical and the second is systematic, and it is dominated by the uncertainty on the branching fractions of D^0 decays. The method detailed in this Thesis has been used to assess the systematic uncertainty due to the π^0-reconstruction efficiency on the Belle II measurement of the $B^0 \to \pi^0\pi^0$ branching fraction, presented at the XLI International Conference of High Energy Physics held in July 2022 in Bologna.

Sunto

L'obiettivo di questa Tesi è lo studio dell'efficienza di ricostruzione del mesone π^0 nei decadimenti $\pi^0 \to \gamma\gamma$ all'esperimento Belle II. Usando due decadimenti del mesone D^0 ricostruiti nei dati e nella simulazione, ho calcolato un fattore di correzione per determinare senza bias l'efficienza di ricostruzione del π^0 dalla simulazione. Ho investigato le possibili dipendenze del fattore di correzione dalla cinematica del π^0 e dalla posizione dei cluster dei fotoni nel calorimetro elettromagnetico di Belle II. L'efficienza di ricostruzione del π^0 è studiata comparando il numero di eventi di segnale misurato nei decadimenti $D^{*+} \to D^0(\to K^-\pi^+\pi^0)\pi_s^+$ e $D^{*+} \to D^0(\to K^-\pi^+)\pi_s^+$, ricostruiti nello stesso set di dati. Il campione usato corrisponde ad una luminosità integrata di 190 fb^{-1}, acquisito tramite il detector Belle II usando collisioni e^+e^- ad energie asimmetriche, fornite dal collisionatore SuperKEKB ad un'energia nel centro di massa di $\sqrt{s} = 10.58$ GeV. Comparando il rapporto degli eventi di segnale misurati nei dati e nella simulazione, ho ottenuto il fattore di correzione che è determinato per le selezioni dei fotoni e dei π^0 usate nell'analisi dei decadimenti $B^0 \to \pi^0\pi^0$. Il fattore trovato è $1.037 \pm 0.007\,(\text{stat}) \pm 0.037\,(\text{syst})$, dove la prima incertezza è statistica e la seconda è sistematica. Quest'ultima è dominata dall'incertezza sul rapporto dei branching ratio dei decadimenti del mesone D^0. Il metodo descritto in questa Tesi è stato usato per valutare l'incertezza sistematica dovuta all'efficienza di ricostruzione del π^0 nella misura del branching fraction del decadimento $B^0 \to \pi^0\pi^0$ eseguita dalla collaborazione Belle II, presentata alla conferenza "XLI International Conference of High Energy Physics" tenutasi a luglio del 2022 a Bologna.

Contents

Introduction

The Belle II experiment analyses energy-asymmetric e^+e^- collisions delivered by the SuperKEKB collider at centre-of-mass energy $\sqrt{s} = 10.58$ GeV, to study the heavy-quark sector of the Standard Model (SM) of particle physics. Key targets of its physics program are high-precision measurements of the elements of the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix. The unitarity can be shown as a triangle in the complex plane: measuring the angles and the sides of this triangle, the SM weak interactions of quarks are tested. Any inconsistency between the redundant measurements of the properties of this triangle would signal physics beyond the SM.
The work presented here is tuned for the measurement of the branching fraction of the $B^0 \to \pi^0\pi^0$ decays, which is instrumental for the measurement of α, currently the least known angle of the unitarity triangle. The parameter α is measured through an analysis of several charmless decays of B mesons, such as two-body decays into pions ($B \to \pi\pi$). The rationale behind the analysis is to combine measurements of branching fractions and $\mathcal{CP}$ asymmetries of isospin-related decays to cancel hadronic unknowns, that spoil the determination of α. Among the $B \to \pi\pi$ decays, the $B^0 \to \pi^0\pi^0$ features the largest uncertainties on the branching fraction and $\mathcal{CP}$ asymmetry. This is limiting the precision on α determined from $B \to \pi\pi$ decays. Leveraging efficient reconstruction of low-energy π^0, improved measurements in $B^0 \to \pi^0\pi^0$ decays will be unique to Belle II. Although precision is still limited by the sample size, preliminary results based on early data collected in 2019 and 2020 [1], showed that the π^0-reconstruction efficiency contributes a sizeable part to the uncertainty on the measurement of the branching fraction. My work targets at reducing this systematic uncertainty.
I measure the efficiency of reconstructing $\pi^0 \to \gamma\gamma$ decays in Monte Carlo (MC) simulation and in data using the signal yields of two D^0-meson decays, $D^{*+} \to D^0(\to K^-\pi^+\pi^0)\pi_s^+$ and $D^{*+} \to D^0(\to K^-\pi^+)\pi_s^+$, which final states differ only for the presence of a single π^0. These favoured decays guarantee high statistics, high-purity samples, and a wide range of momentum of the π^0 meson. Taking the ratio of the efficiencies measured in data and simulation, I obtain a correction factor for an unbiased determination of the π^0 reconstruction efficiency from simulation. The precision of this method is limited by the uncertainty on the charm-decays branching fractions, which contributes an irreducible 3.5% uncertainty. However, this is sufficient precision for the current measurement of the $B^0 \to \pi^0\pi^0$ branching fraction at Belle II. I also investigate possible sample-dependence of the correction factor, by studying its variation for photons hitting different

regions of the electromagnetic calorimeter, and its dependence on the momentum and direction of the neutral pion.
This document is structured as follows:

Chapter 1: an introduction to the SM, the weak interactions of quarks and the Cabibbo-Kobayashi-Maskawa matrix is given; the measurement of α through the isospin analysis is outlined.

Chapter 2: the SuperKEKB collider and the Belle II detector are presented. Each sub-detector is described, with a particular focus on the electromagnetic calorimeter which is the relevant detector for this work. A brief description of the Belle II Monte Carlo simulations is also reported.

Chapter 3: general description of the reconstruction techniques for B-meson decays introduces an overview of the $B^0 \to \pi^0\pi^0$ analysis. The photons and π^0 selections are presented in details.

Chapter 4: a detailed description of measurement method for the π^0 reconstruction efficiency and the correction factor is reported. Here, the studies of selections and background compositions of the D^0-meson decays, along with the achieved results using Monte Carlo simulations and experimental data, are discussed. The correction factor for the π^0 reconstruction efficiency is obtained.

Chapter 5: the correction-factor dependencies on different region of the electromagnetic calorimeter, on the direction of the neutral pion, and on its momentum, are investigated.

Chapter 6: the systematic uncertainties on the correction factor and additional checks are presented.

Chapter 7: all results are summarized and possible extension of this work are considered.

Chapter 1

Quark-flavour physics and the CKM angle α

This Chapter sets the motivation of the work presented in this Thesis, which stems from the analysis of $B^0 \to \pi^0\pi^0$ decays for the measurement of the angle α, a combination of elements of the Cabibbo-Kobayashi-Maskawa quark-mixing matrix. After a brief introduction to the Standard Model of particle physics, I focus on the quark sector to outline the main theoretical concepts subtending the theory of weak interactions of quarks and the CKM matrix. I briefly discuss the experimental determination of the elements of this matrix, and I expand on the method to measure the angle α from a combined analysis of two-body charmless decays of beauty mesons.

1.1 The Standard Model of particle physics

The Standard Model (SM) is a quantum field theory that describes the elementary constituents of matter and their fundamental interactions (electromagnetic, weak and strong interactions, while gravity in not included) [2, 3, 4, 5, 6, 7]. Particles are represented as excitation of quantum fields that interact with each other: the SM describes their dynamics at energy scales relevant for the subatomic world. Particles and their interactions are described in a Lagrangian formalism. Local gauge symmetry, *i.e.* the invariance of the Lagrangian under space-time-dependent transformations applied to the phases of fields, is a key concept. Interaction terms appear in the free-field Lagrangian after requiring it to be invariant under local gauge symmetries. The SM is based on the symmetry group

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

where $SU_C(3)$ is the standard unitary group that describes the strong interactions (quantum chromodynamics, QCD), and C stands for the color charge; $SU_L(2) \otimes U_Y(1)$ is the product of groups that describe the combination of the weak and electromagnetic interactions (electroweak, EW),

with L standing for *left* and Y standing for *hypercharge*.[1] Spin-1 particles, called *gauge bosons* as they correspond to the generators of the symmetry group, mediate the interactions. Strong interactions are mediated by eight massless particles called *gluons*: they carry a charge called *color*, that can be of three kinds (red, green or blue). 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:

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

where 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. The Higgs potential has a typical mexican-hat shape, where the potential minimum does not preserve the global symmetry of the Lagrangian causing a spontaneous symmetry breaking. The vacuum state is found to be charged under the generators of the SM Lagrangian symmetry group. The interaction with this charged ground state force the particles to acquire a mass [5, 6].

Matter particles correspond to excitations of spin-$\frac{1}{2}$ fields and are called *fermions*. Their masses, or, equivalently, their couplings with the Higgs field, are free parameters. Each fermion is also associated with an anti-particle of the same mass but opposite internal quantum numbers. Fermions are further classified into two classes, quarks 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 it is distinguished by its *flavour*, a quantum number which comes in six varieties and which is conserved in the electromagnetic and strong interactions but not in the weak interactions. Due to color confinement free quarks are not observable. They are only observed in their colorless bound states, which include mesons, composed of a quark and an anti-quark, and baryons, composed of three quarks (or three antiquarks). 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 particle, called neutrino, and a massive particle with electric charge $-e$;

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

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

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 found to not be conserved only for neutrinos that can oscillate between states of different lepton flavour. Gauge bosons, quarks, and leptons constitutes all the fundamental interactions and particles of the SM, which are sketched in Figure 1.1.

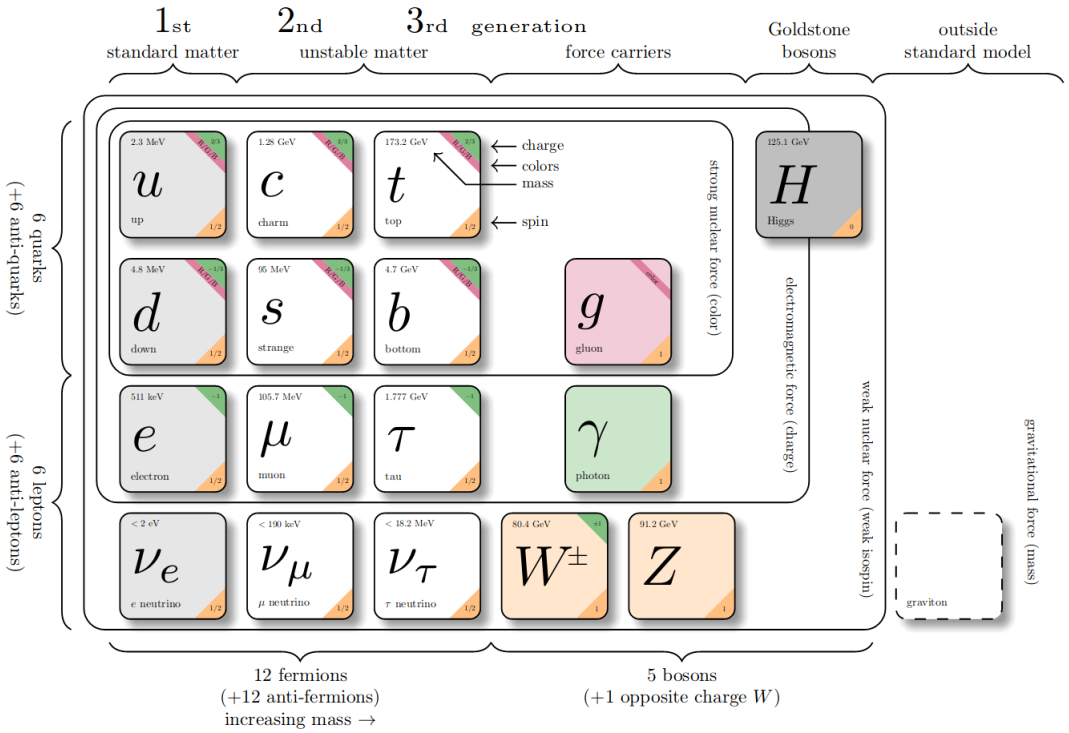


Figure 1.1: Scheme of particles and interactions in the Standard Model.

1.1.1 Searching beyond

The SM was completed in the 1970's and has been successfully tested since, in thousands of measurements whose fractional precisions reach one part per trillion [8]. However, observations and theoretical considerations suggest that the SM 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 the origin for the observed asymmetry between matter and antimatter in the universe, the strikingly large differences observed between fermion masses, 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 SM can broadly be classified into two synergical approaches, the energy and intensity frontiers.

The energy-frontier is a *direct* approach that aims at using high-energy collisions to produce on-shell particles 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. 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.
The intensity-frontier exploits an *indirect* approach which consists in searching for significant differences between precise measurement and equally precise SM predictions in processes sensitive to non-SM contributions. The idea is that exchanges of new virtual (off-mass-shell) particles of arbitrary high mass can occur in the amplitude of lower energy processes, thus altering the SM-expected dynamics in an observable manner. Experimental evidence of new particles is typically harder to establish. However, the reach is not bounded by the maximum energy reachable by experiments. A large portion of the effort in this approach is centered on the so-called *flavour physics*. Currently, the only hints of possible physics beyond the SM are emerging from a series of intriguing anomalies in B decays observed mainly in pp collisions at the LHCb experiment [9]. Since plans for a higher-energy collider in the near future are not set yet, flavour physics seems a very promising asset to search for non-SM in the next decade.
The term *flavour physics* usually refers to the study of the dynamics of both quarks and leptons. However, hereafter I will restrict the scope by referring solely to the quark sector, and with particular focus on the weak interactions of quarks which offer a rich avenue for testing the SM and search for further extensions.

1.2 Weak interactions of quarks

A crucial role in the development of the theory of weak interactions has been played by discrete symmetries. Those are parity, charge conjugation, and time reversal. 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 direction. The product of the three discrete symmetry transformations is found to be conserved in all interactions, but they are not conserved individually [10, 11]. Parity and charge conjugation symmetries are 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 violation of the $\mathcal{CP}$ symmetry in weak interactions has been observed for the first time in kaons decays in 1964 by Cronin and Fitch [12]. However, only in the early 1970's, when only three quarks were known, Kobayashi and Maskawa have been able to accommodate it in the SM, by generalizing Cabibbo's theory [13] from a four-quark model to a six-quark model [14]. They introduced a matrix to describe the relations between mass and weak interaction eigenstates of

[2] On-shell particles satisfy the energy-momentum conservation in the production process. *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.

quarks as seen by $W^{\pm}$ bosons. This is known as the Cabibbo-Kobayashi-Maskawa quark-mixing matrix or $V_{\rm CKM}$, a complex matrix that is unitary because the sum of all the transition probabilities is equal to one. The $N \times N$ CKM matrix has $(N-1)^2$ free parameters, where N is the number of quarks families.[3] 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. The rotation of the quarks eigenstates through the CKM 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* [15], an expansion in the small parameter $\lambda = \sin\theta_C \approx 0.23$ that makes explicit the observed hierarchy between its elements,

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

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 the imaginary part that allows for $\mathcal{CP}$ violation. The unitarity condition $V_{\rm CKM} V_{\rm CKM}^{\dagger} = \mathbb{1}$ yields nine relations,

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

which are sums of three complex numbers each, 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 $J = \Im(V_{us} V_{cb} V_{ub}^* V_{cs}^*)$ [16, 17, 18]. 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' (UT), shown in Figure 1.2. Conventionally, side sizes are normalized to the length of the base, and the

[3] 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

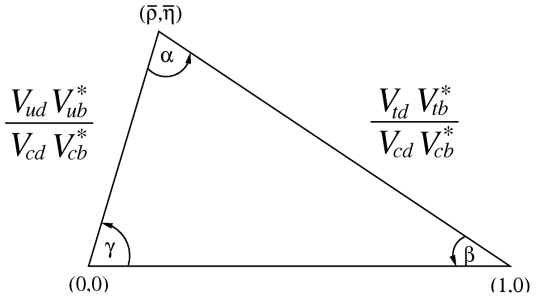


Figure 1.2: Graphical representation of the Unitarity Triangle. The base is normalized to 1*, the length of the other side is shown in the figure [19].*

three angles are labelled α (or ϕ_2), β (or ϕ_1), and γ (or ϕ_3). The UT angles are defined as follows,

$$\alpha = arg(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{V_{ub}^*}}), \quad \beta = arg(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}), \quad \gamma = arg(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}) \tag{1.1}$$

1.2.1 Experimental status

Measurements of parameters associated with quark-flavour physics have been performed in many dedicated, or general-purpose, experiments in the last three decades, including CLEO, CPLEAR, NA32, NA48, KTeV, SLD, OPAL, L3, ALEPH, DELPHI, BaBar, Belle, CDF, CDFII, LHCb, Atlas, and CMS [20]. The resulting global picture is that the CKM interpretation of quark-flavour 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 Figure 1.3 [21].

The sides of the UT are determined through the measurements of the magnitude of the CKM matrix elements $|V_{cb}|$, $|V_{ub}|$, $|V_{ts}|$ and $|V_{td}|$. Semileptonic B decays, mainly studied in e^+e^- collisions, are used to determine $|V_{cb}|$ and $|V_{ub}|$, which are known with precision of about 1.5% and 3%, respectively. However, their determination is spoiled by a 20-year-long mild discrepancy between different measurement methods. The elements $|V_{ts}|$ and $|V_{td}|$ are accessed with the measurement of the frequency of neutral $B_{(s)}$-mesons flavour oscillations. This frequency is known in both systems with better than 1% precision from pp and $p\bar{p}$ collisions. Angles of the UT are determined by measurement of $\mathcal{CP}$ violation in B decays. Measurements of $\sin 2\beta$ reached a precision of 2.5%, mainly due to the availability of large samples of $B^0 \to J/\psi K_S^0$ decays in e^+e^- collisions. The angle γ is measured with 5% precision using combinations of several measurements involving $B \to DK$ decays (B, D, K charged or neutral) reconstructed in e^+e^- and pp collisions. The angle α is known down to a 6% precision from $B \to hh$ decays (where B is charged or neutral, and h represents a charged or neutral π or ρ) in e^+e^- and pp collisions. More details about the determination of the angle α are presented in Section 1.3.1. In addition, many other measurements in charm and kaon physics contribute that are not straightforwardly represented in the UT.

Despite the first-order consistency of the experimental flavour 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

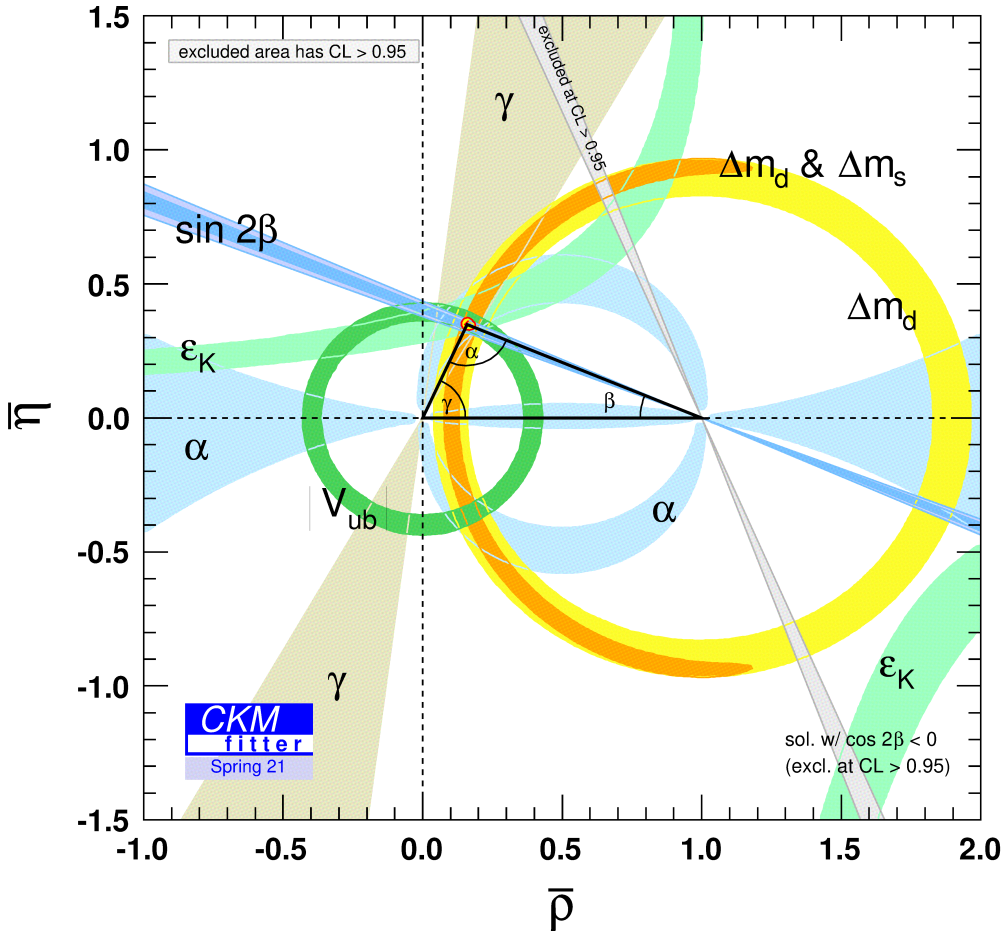


Figure 1.3: Current constrains on sides and angles of the Unitarity Triangle in the $(\bar{\rho}, \bar{\eta})$ plane. The red hashed region of the global combination corresponds to 68% confidence level (CL) [22].

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 pp collisions at energies of up to 14 TeV, where incoherent QCD production yields large samples of $b\bar{b}$ pairs ($\approx$ 45 kHz) and $c\bar{c}$ pairs ($\approx$ 1 MHz). The high production rate enables low statistical uncertainties, and TeV collision energy allows the production of all species of *beauty* 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 $B\overline{B}$ pairs from $\Upsilon(4S)$ decays at $\approx$ 1 kHz rate. Belle II aims at operating until

2033 at the intensity frontier of the flavour sector to explore indirectly of non-SM physics, in competition and synergy with LHCb [23]. Compared to LHCb, Belle II has a lower production rate thus, thanks to a much lower background and additional kinematic constraints provided by the point-like nature of colliding particles, it features a better sensitivity final states that include neutral particles. Owing to comparably strong reconstruction efficiencies and resolutions of neutral and charged particles, Belle II has the unique capability of studying jointly, and within the same experimental environment, all relevant final states of the decays required for the determination of the angle α, as explained in the next Section. The current uncertainty on this parameter is among the major limitations on the global-consistency check of the UT. Belle II is in a unique position for measuring α at an unprecedented precision.

1.3 Charmless B decays and the angle α

Decays governed by transitions of b quarks into quarks other than c, so-called charmless B decays, proceed through the quark-level processes $b \to u$, $b \to d$, and $b \to s$. These processes yield final-state hadrons without net charm, such as kaons, pions, ρ mesons, and ϕ mesons. Typical examples of charmless decays include $B^0 \to \pi^0\pi^0$, $B^0 \to \pi^+\pi^-$, $B^+ \to \rho^+\rho^0$ and many others.

Because the magnitude of the CKM element V_{ub} ($\approx 4\times10^{-3}$) is much smaller than that of V_{cb} ($\approx 4\times10^{-2}$), the tree-level amplitudes are suppressed.[4] The resulting branching fractions are typically small (less than 2×10^{-5}), which makes charmless decays difficult to observe. 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. Nonetheless, charmless decays play a key role in flavour physics. Transitions $b \to du\overline{u}$ provide, for example, the only reliable determination of the α angle.

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 $\approx 1\,\text{GeV}$). 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.

Various approaches have been proposed to overcome the issue. Currently the most promising strategy is to exploit exactly (or approximately) valid flavour symmetries existing between quark-level processes to reduce, with appropriate approximations, the QCD unknowns. An approximation often used is based on $SU(2)$ symmetry, where $m_u = m_d$ is assumed. The comparison be-

[4] Tree-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.

tween precise measurements and theoretical predictions is therefore performed on combinations of experimental results, based on channels related by $SU(2)$. This is at the base of the so-called *isospin analysis*.

1.3.1 The determination of α

The B-mesons, considered in the isospin analysis, are charged and neutral and their decay into a pair of light unflavoured mesons is described by the weak transition $\overline{b} \to \overline{u}u\overline{d}$. The successive hadronisation includes the spectator quark of the isodoublet. The leading-order amplitudes are tree and penguin transitions $\overline{b} \to \overline{u}(u\overline{d})$ and $\overline{b} \to \overline{d}(u\overline{u})$, respectively.[5]
Tree-level amplitudes enjoy accurate predictions, whilst the penguin amplitudes, due to the presence of an internal quark line, are sensible for non-SM contributions, but makes the calculations difficult. The process can be written as:

$$B_q \to (\overline{u}u\overline{d}, q) \to h_1^i h_2^j \tag{1.2}$$

where h_1 and h_2 are two mesons (ρ or π), the possible value of $j = -, 0, +$ identify the charge of the barions and q is the spectator quark of the B-meson. The Figure 1.4 shows tree, QCD and EW penguin diagrams. Penguin contributions are not negligible, especially for decays $B^0 \to h^0 h^0$, where the leading-order tree contributions is color suppressed. In fact, in the right Figure 1.4b, the quarks resulting from the hadronic decay of the W boson, must match the color of the B-meson quarks in order to create two colorless mesons. The fine-structure constant α is smaller than the strong coupling coefficient α_s, for this reason the EW penguin amplitudes are expected to be of $\mathcal{O}(0.1)$ of QCD penguin amplitudes.
The corresponding amplitude, considering three different up-type quark occuring in the W loop, is:

$$A^{ij} = \langle h_1^i h_2^j | \mathcal{H}_{eff} | B_d \rangle = V_{ud}V_{ub}^*(\mathcal{T}_u^{ij} + \mathcal{P}_u^{ij}) + V_{cd}V_{cb}^*\mathcal{P}_c^{ij} + V_{td}V_{tb}^*\mathcal{P}_t^{ij} \tag{1.3}$$

where $\mathcal{H}_{eff}$ is the effective Hamiltonian of the transition, $\mathcal{T}_u^{ij}$ is the tree-level amplitude and $\mathcal{P}_{u,c,t}^{ij}$ are the penguin-level amplitude. Using the unitarity of the CKM matrix it is possible to write:

$$A^{ij} = V_{ud}V_{ub}^*(\mathcal{T}_u^{ij} + \mathcal{P}_u^{ij} - \mathcal{P}_c^{ij}) + V_{td}V_{tb}^*(\mathcal{P}_t^{ij} - \mathcal{P}_c^{ij}) \tag{1.4}$$

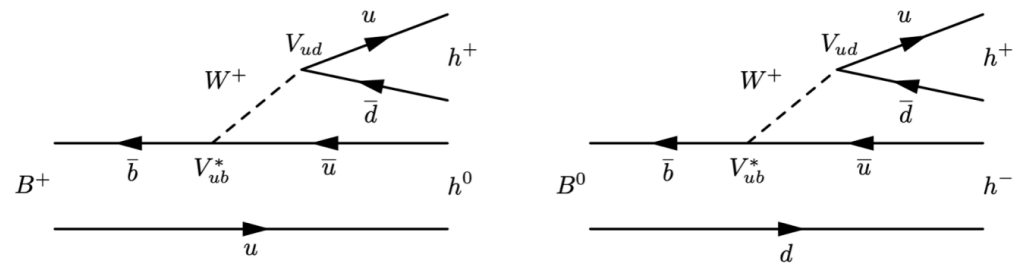
Remembering the definitions of the CKM angles and defining two new variables:

$$T^{ij} = |V_{ud}V_{ub}^*|(\mathcal{T}_u^{ij} + \mathcal{P}_u^{ij} - \mathcal{P}_c^{ij}), \quad P^{ij} = -|V_{td}V_{tb}^*|(\mathcal{P}_t^{ij} - \mathcal{P}_c^{ij}). \tag{1.5}$$

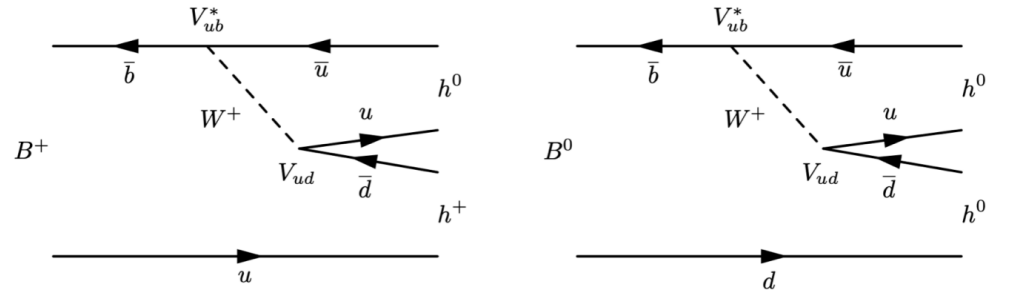
It is possible to rewrite the amplitudes as:

$$A^{ij} = -e^{i\gamma}T^{ij} + e^{-i\beta}P^{ij}, \quad \frac{p}{q}\overline{A}^{ij} = -e^{-i\gamma}T^{ij} + e^{i\beta}P^{ij} \tag{1.6}$$

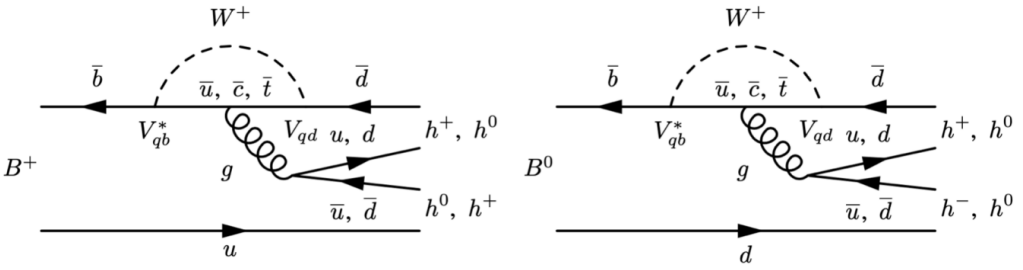
[5] A penguin amplitude involves the emission and reabsorption of a W boson and typically the presence of a gluon.



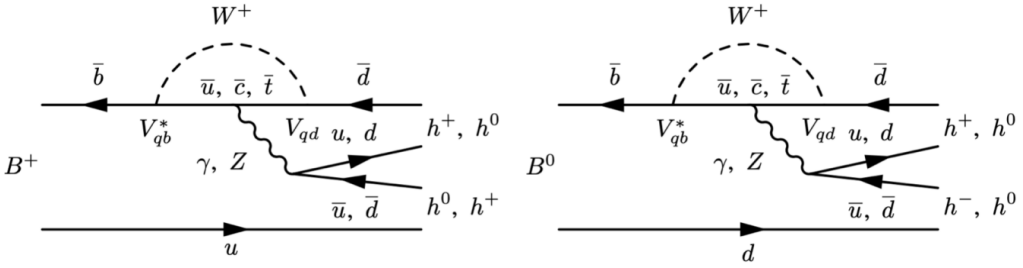
(a) Color-favoured tree diagrams.



(b) Color-suppressed tree-diagrams.



(c) QCD-penguin diagrams.



(d) Color-suppresed EW penguin diagrams.

Figure 1.4: Dominant tree-level and QCD-penguin Feynman diagrams together with the color-suppressed EW-penguin diagrams contributing to $B^+ \to h^- h^0$ (left) and to $B^0 \to h^+ h^-,\ h^0 h^0$ (right). The generic shorthand h indicates a pion or ρ meson [24].

where $\frac{p}{q} \sim e^{i2\beta}$ takes into account the $B^0 - \overline{B^0}$ mixing phase and $\overline{A}^{ij}$ is amplitude of the $\mathcal{CP}$-conjugate isodoublet. In this case the hadronic amplitudes T^{ij} and P^{ij} are the same of above, because the strong interactions are invariant under $\mathcal{CP}$ transformations, but the weak phases change sign. Rotating all the amplitude by the weak phase β and using $\alpha = \pi - \beta - \gamma$, a single relation for α is obtained:

$$e^{i2\alpha} = \frac{e^{i\alpha}T^{ij}}{e^{-i\alpha}T^{ij}} = \frac{\overline{A}^{ij} - P^{ij}}{A^{ij} - P^{ij}}. \tag{1.7}$$

If penguin contributions where negligible ($|P^{ij}| \sim 0$), it would be possible to obtain $sin(2\alpha)$ using the time dependent analysis of $B^0/\overline{B}^0 \to h_1^+ h_2^-$ decay yields the $\mathcal{CP}$ asymmetry:

$$a_{\mathcal{CP}}(t) = \frac{\Gamma(\overline{B}^0(t) \to h_1^+ h_2^-) - \Gamma(B^0(t) \to h_1^+ h_2^-)}{\Gamma(\overline{B}^0(t) \to h_1^+ h_2^-) + \Gamma(B^0(t) \to h_1^+ h_2^-)} = \mathcal{S}^{+-} sin(\Delta m_d t) - \mathcal{C}^{+-} cos(\Delta m_d t) \tag{1.8}$$

where Δm_d is the $B^0 - \overline{B}^0$ oscillation frequency, t is the time difference between the $\mathcal{CP}$ and tag-side decays and, defining $\lambda = \frac{\overline{A}^{+-}}{A^{+-}}$, the coefficients are:

$$\mathcal{S}^{+-} = \frac{2Im\lambda}{1 + |\lambda|^2} = sin(2\alpha)\ , \quad c^{+-} = \frac{1 - |\lambda|^2}{1 + |\lambda|^2}. \tag{1.9}$$

As mentioned before, penguin contributions have to be considered; so including P^{ij} a direct determination of α is not possible and a new "penguin pollutted" angle is introduced as α_{eff}. The relations become:

$$\lambda = |\lambda| e^{i2\alpha_{eff}}\ , \mathcal{S}^{+-} = \sqrt{1 - (C^{+-})^2} sin(2\alpha_{eff}). \tag{1.10}$$

Currently, the precision in the determination of α is dominated by the $B \to \rho\rho$ decays, to a lesser extent by the $B \to \pi\pi$ system, whereas the $B \to \rho\pi$ play only a limited role. This is clearly visible in Figure 1.5, which presents the confidence intervals of α determined from the average of measurements for each channel. The overall combination of all experimental information give a value of

$$\alpha = (85.2^{+4.8}_{-4.4})^\circ. \tag{1.11}$$

An equivalent solution shifted by 180° also exists. The solution in Equation. 1.11 is consistent with the its indirect determination from all other CKM elements and assuming unitarity, which reads $\alpha_{\text{ind}} = 91.9^{+1.6}_{-1.2}$ [22].

This work is centered on the $B^0 \to \pi^0\pi^0$ decay, so the next section focuses on the isospin analysis of $B \to \pi\pi$ decays.

1.3.2 Isospin analysis of $B \to \pi\pi$ decays

The π mesons are the lightest pseudoscalar (spin-0) and they are an isospin triplets ($I = 1$):

$$\pi(u\overline{d}), \quad \pi^0(\frac{u\overline{u} - d\overline{d}}{\sqrt{2}}), \quad \pi^-(\overline{u}d) \tag{1.12}$$

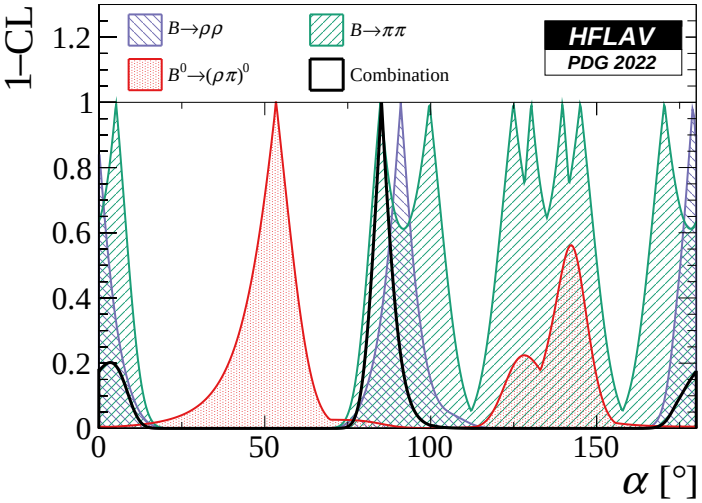


Figure 1.5: World average of α in terms of $1-\mathrm{CL}$ *split by decay mode [25].*

The isospin analysis are able to overcome the pollution due to penguin diagrams [26]. Hadronisation preserves the isospin, using this property it is possible to get relations between decay amplitudes, and use them to separate the effect of tree and penguin diagrams. The mesons in the final state are identical bosons so the amplitudes shown in Equation 1.3 simplify to:

$$A^{ij} = \langle \pi^i \pi^j | \mathcal{H}_{eff} | B^{i+j} \rangle, \quad A^{ij} = A^{+0},\, A^{+-},\, A^{00}. \tag{1.13}$$

The total final state wave function must be symmetric under particle exchange, for $i \neq j$ the state is

$$|\pi^i \pi^j\rangle = \sqrt{\frac{1}{2}} \left(|\pi_1^i \pi_2^j\rangle + |\pi_1^j \pi_2^i\rangle \right). \tag{1.14}$$

B mesons are spin-0 particles, so the total angular momentum of the final state is $J = 0$. Spin sum rules dictate that the total final-state isospin for a pion pair can be $0\,, 1$ or 2. The Bose statistics allows only the $I_f = 0$ or $I_f = 2$ final states. The final states can be expressed as linear combination of state with $I_f = 0$ and $I_f = 2$, resulting in

$$|\pi^+ \pi^0\rangle = |2, 1\rangle \tag{1.15}$$

$$|\pi^+ \pi^-\rangle = \sqrt{\frac{1}{3}} |2, 0\rangle + \sqrt{\frac{2}{3}} |0, 0\rangle \tag{1.16}$$

$$|\pi^0 \pi^0\rangle = \sqrt{\frac{2}{3}} |2, 0\rangle - \sqrt{\frac{1}{3}} |0, 0\rangle. \tag{1.17}$$

The decay amplitude in Equation 1.13 can be factorized in the weak decay $b \to u\bar{u}d$ and the hadronisation into two light mesons. Using the Wigner-Eckhart theorem, it is possible to introduce amplitudes expressed in terms of reduced matrix elements $A_{\Delta I,\, I_f}$, where ΔI is the isospin shift and I_f is the final-state isospin. These reduced matrix elements take into account hadronisa-

tion and strong rescattering, yielding:

$$A^{+0} = \langle \pi^+ \pi^0 | \mathcal{H}_{eff} | B^+ \rangle = \sqrt{\frac{3}{4}} A_{\frac{3}{2},2}, \tag{1.18}$$

$$A^{+-} = \langle \pi^+ \pi^- | \mathcal{H}_{eff} | B^0 \rangle = \sqrt{\frac{1}{6}} A_{\frac{3}{2},2} + \sqrt{\frac{1}{3}} A_{\frac{1}{2},0} \tag{1.19}$$

$$A^{00} = \langle \pi^0 \pi^0 | \mathcal{H}_{eff} | B^0 \rangle = \frac{1}{3} A_{\frac{3}{2},2} - \sqrt{\frac{1}{6}} A_{\frac{1}{2},0}. \tag{1.20}$$

Combining Equations 1.18, 1.19, 1.20 and considering analogous relations for the $\mathcal{CP}$-conjugated amplitudes $\overline{A}_{\Delta I, I_f}$, two relations can be obtained:

$$A^{+0} - A^{00} = \sqrt{\frac{1}{2}} A^{+-}, \quad \overline{A}^{+0} - \overline{A}^{00} = \sqrt{\frac{1}{2}} \overline{A}^{+-}. \tag{1.21}$$

These are referred to as isospin triangles, since they can be represented as triangles in the complex space. Strong penguin amplitudes can lead only to $\Delta I = \frac{1}{2}$ transitions. Since the amplitude in Equation 1.18 has no $\Delta I = \frac{1}{2}$ components the decays $B^\pm \to \pi^\pm \pi^0$ occurs purely as tree amplitude. Using Equation 1.7 one obtains

$$e^{-2i\alpha} = \frac{\overline{A}^{+0}}{A^{+0}}. \tag{1.22}$$

Using Equation 1.7, where α must be replace by α_{eff}, the CKM triangles (Equation 1.21) and the relation of Equation 1.22, yields all the ingredients to determine α through the isospin symmetry. The branching fractions $\mathcal{B}^{+0}$, $\mathcal{B}^{+-}$ and $\mathcal{B}^{00}$, so as the $\mathcal{CP}$ parameters $\mathcal{A}^{+-}_{\mathcal{CP}}$, $\mathcal{A}^{00}_{\mathcal{CP}}$ and $\mathcal{S}^{+-}_{\mathcal{CP}}$ can be expressed as linear combinations of six real positive variables: $|A^{+-}|$, $|A^{+-}|$, $|\overline{A}^{+-}|$, δ, α_{eff} and α. Where δ is the phase difference between T^{ij} and P^{ij} of Equation 1.5.

Using the measurements of these quantity, one obtains a system of six linear independent equations with six real positive variables. The value of α is determined up to an eight-fold ambiguity in the range $[0, \pi]$ because each isospin triangle has two possible orientations and exist an additional symmetry that makes $\mathcal{S}^{+-}_{\mathcal{CP}}$ invariant under the reflection:

$$(\alpha_{eff}, \alpha, \delta) \leftrightarrow (\frac{\pi}{2} - \alpha_{eff}, \frac{\pi}{2} - \alpha, \frac{\pi}{2} - \delta). \tag{1.23}$$

The Figure 1.6 shows an illustration of the eight-fold α ambiguity. An auxiliary measurement of the decay-time-dependent $\mathcal{CP}$-violating asymmetry in $B^0 \to \pi^0\pi^0$ decays, would bring additional insight into α by suppressing the mirror solutions in the isospin analysis. While restricting to diphoton decays of the π^0 makes a time-dependent analysis impossible, the full Belle II data set of 50 ab^{-1} will enable exploiting photon conversions and π^0 Dalitz decays (further details can be found in [27]).

This work focuses on the $B^0 \to \pi^0\pi^0$ decay, in particular on the improvement of one of the largest systematic uncertainty affecting the measurement of its branching fraction: the efficiency of the reconstruction of π^0 mesons in the Belle II experiment. The description of the Belle II experiment

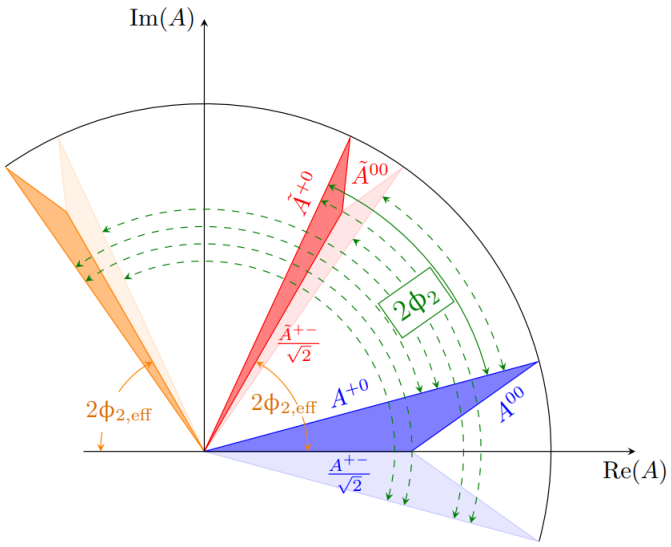


Figure 1.6: Geometrical representation of the isospin triangular relations (Equation. 1.21) in the complex plane of $B^{ij} \to h^i h^j$ amplitudes. The blue and the red shaded areas correspond to the isospin triangles. The angle between the $\mathcal{CP}$ conjugate charged amplitudes A^{+-} and $\tilde{A}^{+0}$ (indicates the same $\overline{A}^{+0}$ amplitude as referred in the text) corresponds to twice the weak phase α_{eff} (orange solid lines). The angle between the $\mathcal{CP}$ conjugate charged amplitudes A^{+0} and $\tilde{A}^{+0}$ corresponds to twice the CKM angle α (green solid line). The other triangles with lighter shade represent the mirror solutions allowed by the discrete ambiguities of the branching ratios and $\mathcal{CP}$ parameters, with the corresponding values for α represented by the green dashed lines.

is the topic of the next Chapter. Further details on the $B^0 \to \pi^0\pi^0$ analysis are presented in Chapter 3.

Chapter 2

The Belle II detector at the SuperKEKB collider

The data used in this work have been collected with the Belle II experiment. In this Chapter, I briefly introduce the collider SuperKEKB, which delivers e^+e^- collisions to Belle II, and I describe the detectors that constitute the experiment. I expand more on the electromagnetic calorimeter which is the detector used to reconstruct the π^0 mesons. I finally give also an overview of the simulation used in Belle II.

2.1 The SuperKEKB collider at KEK laboratory

SuperKEKB is a high-luminosity electron-positron (e^+e^-) energy-asymmetric collider, designed to produce nearly 1000 $B\overline{B}$ pairs ($B^0\overline{B^0}$ and B^+B^- in approximately equal proportions) per second via decays of $\Upsilon(4S)$ mesons produced at threshold [28]. 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 $B\overline{B}$ pairs at high rates.
Intense beams of electrons and positrons are brought to collision at the energy corresponding to the $\Upsilon(4S)$ meson mass, 10.58 GeV, which is just above the $B\overline{B}$ production kinematic threshold. Such finely tuned collision energy is key. The ensuing production of $\Upsilon(4S)$ mesons, which decay in $B\overline{B}$ pairs 96% of the times with little available energy to produce additional particles, suppresses backgrounds, 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, flavour is conserved, and the null net bottom content of the initial state implies production of a flavourless $B\overline{B}$ pair; even though B^0 and $\overline{B^0}$ undergo flavour oscillations before decaying, their time-evolution is quantum-correlated in such a way that no B^0B^0 or $\overline{B^0B^0}$ pairs are present at any time. In fact, angular-momentum conservation implies that the decay of a spin-1 particle

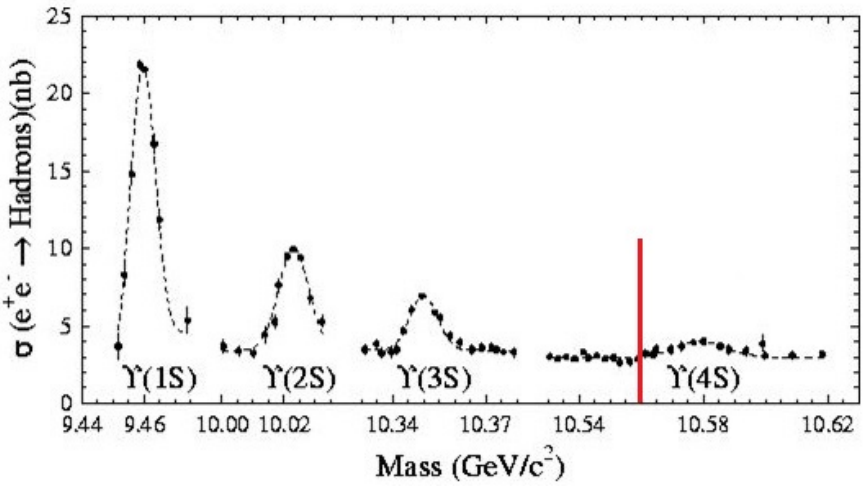


Figure 2.1: Hadron production cross section from e^+e^- collisions as a function of the centre-of-mass energy. The vertical red line indicates the $B\overline{B}$ production threshold.

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 $B^0\overline{B^0}$ particle-antiparticle pair until either one decays. This allows identification of the bottom (or antibottom) content of one meson at the time of decay of the other, if the latter decays in a final state accessible only by either bottom or antibottom states. This important capability is called 'flavour tagging' and allows measurements of flavour-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 ($e^+e^- \to q\bar{q}$, 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 $e^+e^- \to e^+e^-(\gamma)$, $e^+e^- \to e^+e^-e^+e^-$ and $e^+e^- \to \mu^+\mu^-(\gamma)$. These amount to about 94% of the total cross-section, but can straightforwardly be discriminated online owing to their distinctive final states.

Because the $\Upsilon(4S)$ mesons are produced at threshold, they would be nearly at rest in the laboratory frame in an energy-symmetric collider. The resulting B mesons too would be produced with low momentum ($\approx 10\,\mathrm{MeV}/c$) in the laboratory, because of the $21\,\mathrm{MeV}/c^2$ difference between the $\Upsilon(4S)$ mass and the mass of a $B\overline{B}$ pair. With such low momenta they would only travel approximately 1 µm before decaying. The 10 µm 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 [29]. SuperKEKB implements a 7-4 GeV energy-asymmetric double-ring design, which achieves a vertex displacement of about 130 µm.

SuperKEKB is designed to deliver collisions corresponding to $50\,\mathrm{ab}^{-1}$ of integrated luminosity by 2033, corresponding $\approx 5.3 \times 10^{10}\ B\overline{B}$ pairs, about 50 times the total amount collected by B Factories to date.

To achieve high luminosities, a nano-beam, large crossing-angle collision scheme is implemented [30]. This is an innovative configuration based on keeping small horizontal and vertical emittance and large crossing angle, as shown in Figure 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.

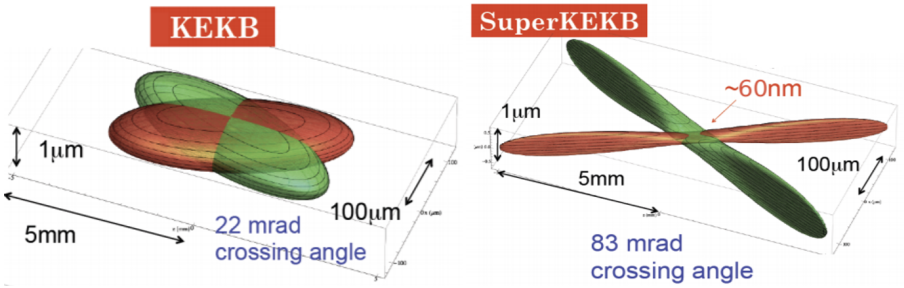


Figure 2.2: Two-dimensional sketch of the nano-beam mechanism implemented in SuperKEKB (right) compared with the previous KEKB collision scheme (left).

2.2 The Belle II detector

The Belle II detector 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. It is located in the same experimental hall and has a similar design to its predecessor, the Belle detector. The main subsystems are summarized in Table 2.1 and shown in Figure 2.3. 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.
- **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.

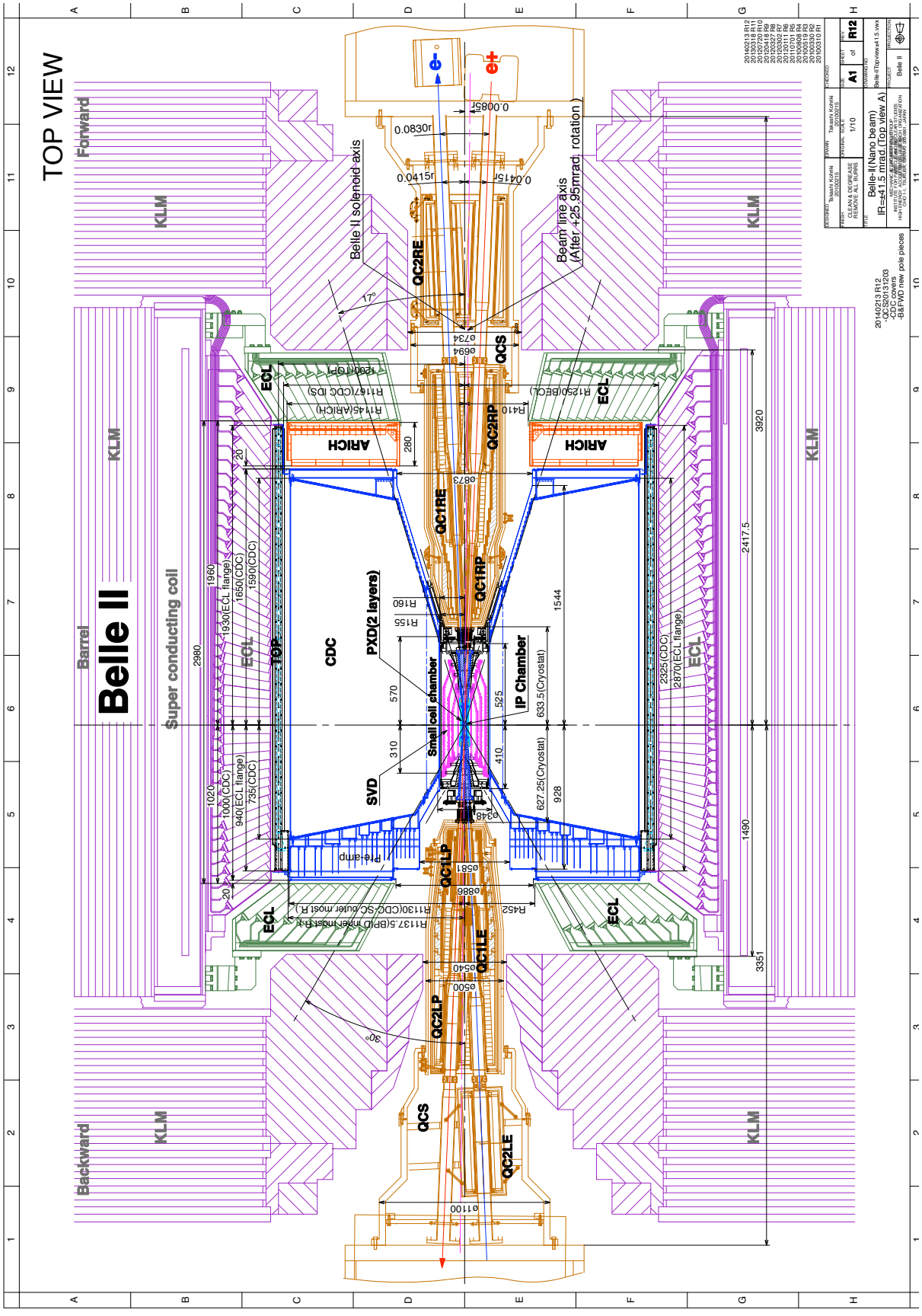


Figure 2.3: Top view of the Belle II detector [31].

Purpose	Name	Component	Channels	θ coverage
Tracking	PXD	Silicon Pixel (DEPFET)	10M	$[17°; 150°]$
	SVD	Silicon Strip	245k	$[17°; 150°]$
	CDC	Drift Chamber (He-C_2H_6)	14k	$[17° : 150°]$
Particle ID	TOP	RICH with quartz radiator	8k	$[31°; 128°]$
	ARICH	RICH with areogel radiator	78k	$[14°; 30°]$
Calorimetry	ECL	CsI(Tl)	6624 (Barrel)	$[12.4°; 31.4°]$
			1152 (FWD)	$[32.2°, 128.7°]$
			960 (BWD)	$[130.7°; 155.1°]$
Muon ID	KLM	barrel: RPCs and scintillator strips	θ, ϕ 16k	$[40°; 129°]$
	KLM	end-cap: scintillator strips	17k	$[25°; 40°]$
				$[129°; 155°]$

Table 2.1: Summary of the detector components

2.2.1 Tracking detectors

The innermost detectors are used for charged-particle reconstruction. Effective track reconstruction is of great importance since flavour-physics analysis rely strongly on precise momenta measurements and on precise determination of the decay positions. Precisely measured momenta and vertices allow separation of signal from backgrounds, thanks to invariant-mass narrower signal peak and therefore more distinctive from smoothly distributed backgrounds. Moreover, precise vertices measurements are key to determine the decay time, a fundamental quantity for $\mathcal{CP}$-violating asymmetries measurements 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.

Silicon-pixel vertexing detector

The innermost detector, shown in Figure 2.4, 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 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 [32]. They are made of p-channel MOSFET integrated on a silicon substrate, which is fully depleted by applying an appropriate voltage. Incident particles generate electron-hole pairs in the depleted region, and thus induce a current passing through the MOSFET. Sensors are 75 µm thick, which allows on-pixel integration of most of the electronics.
The PXD 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, $50 \times (50 - 55)$ µm^2 (inner layer) and $50 \times (70 - 85)$ µm^2 (outer layer) each. The polar acceptance ranges from 17° to 150°. Currently, only one octant of the outer layer is mounted at the experiment. The remaining part will be installed in 2023. The

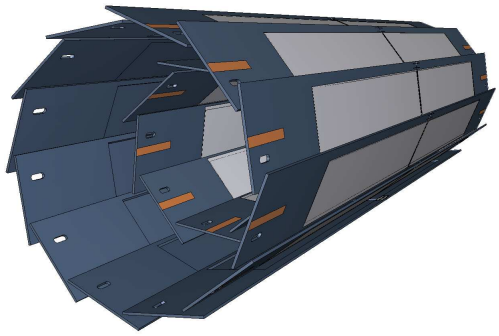


Figure 2.4: Scheme of the PXD detector geometry.

transverse impact-parameter resolution is 12 µm, achieved by weighting the charge deposited in neighbouring pixels.

Silicon-microstrip vertexing detector

Around the PXD is the silicon-microstrip vertexing (SVD), a silicon detector aimed at reconstructing decay vertices and charged-particle tracks at high spatial resolution [33], an exploded view is reported in Figure 2.5.

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.

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

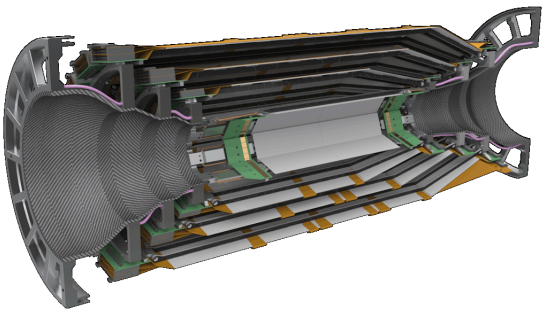


Figure 2.5: Exploded view of a SVD detector half.

Central drift chamber

The central drift chamber (CDC) [34] samples charged-particle trajectories at large radii. It provides trigger signals for events containing charged particles, and 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. Figure 2.6 shows the configuration of the wires inside the CDC. The chamber is composed of 14′336 sense wires, with a diameter of 30 µm, divided in 56 layers, immersed in a gaseous mixture of 50% He and 50% C_2H_6, while 42′240 aluminum wires, with a diameter of 126 µm, shape the electric field. The azimuthal acceptance ranges from 17° to 180°.
The single hit spatial resolution is about 100 µm and the dE/dx resolution is 11.9% for an incident angle of 90°. The typical transverse momentum resolution is $\sigma(p_T)/p_T^2 \approx 0.5\%/[\mathrm{GeV}/c]$.

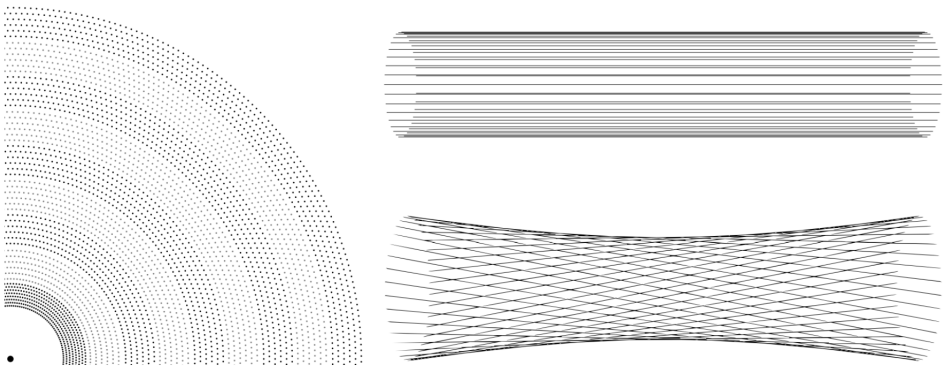


Figure 2.6: The left side shows a quadrant of a slice of the r-ϕ projection of the drift chamber [35]. The innermost superlayer contains eight layers, all the others contain six. The right side shows a visualization of axial wires (top) relative to stereo wires (bottom). The skew is exaggerated.

2.2.2 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). The ECL is presented in more details in Section 2.2.3, since is the detector more relevant for the study presented in this Thesis.

Time of propagation counter

The time of propagation (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, as shown in figure2.7: 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 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.

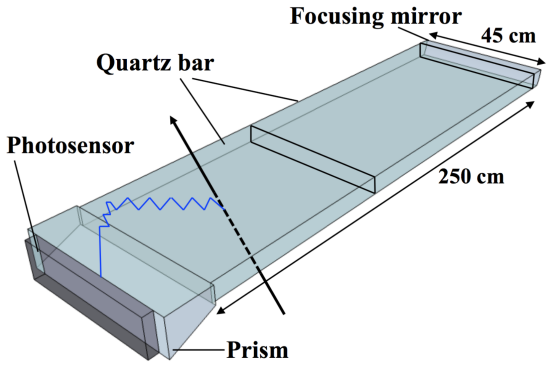


Figure 2.7: Sketch of a TOP quartz bar. A charged particle is shown passing through the radiator and emitting a Cherenkov photon.

Aerogel ring-imaging Cherenkov

Charged-particle identification in the forward end-cap is provided by the aerogel ring-imaging Cherenkov (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 GeV/c. When charged particles pass through the aerogel radiator, Cherenkov photons are produced; they propagate in a expansion volume where they form a ring on a photon sensitive surface made by position-sensitive photo-diodes. Photocatodes are used to convert photons in photoelectrons and generate electrical signals. As shown in Figure 2.8, two aerogel radiators with different refraction indexes are used to increase the number of generated photons without degrading the Cherenkov-angle resolution [23].

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 allows for a $5\,\sigma$ separation between kaons and pions of 0.4-4 GeV/c momenta, and a $4\,\sigma$ separation between pions, muons and electrons with momenta smaller than 1 GeV/c.

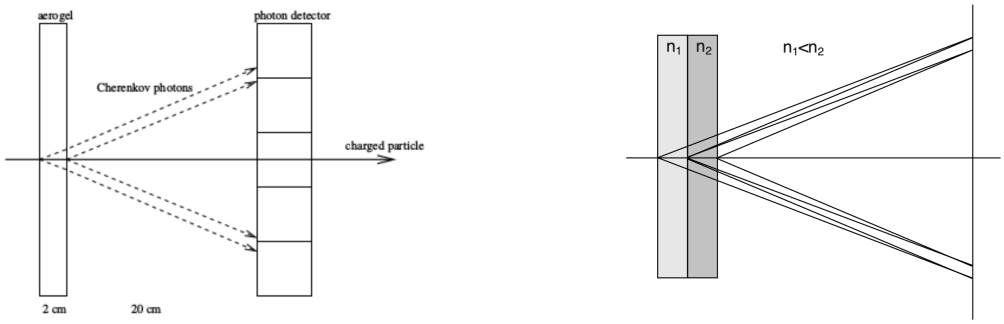


Figure 2.8: Scheme of the ARICH counter.

K_L^0 and muon detection system

The K_L^0 and muon detection system (KLM) detects muons and neutral particles that do not get absorbed in the inner detectors, such as K_L^0 [36]. 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.

2.2.3 Electromagnetic calorimeter

The electromagnetic calorimeter (ECL) [37] is a very important part of the Belle II detector, since one third of B-decay products are π^0 or neutral particles that provide photons. These photons cover a wide energy range, from 20 MeV to 4 GeV and a high electromagnetic resolution is fundamental. The ECL reconstruct energy and position for photons and neutral hadrons. It is useful, also for electron or charged hadron reconstruction in regions with limited tracking coverage. ECL permits to identify electrons and separate them from hadrons by extrapolation of CDC tracks. If the incident particle stops in the calorimeter, the total energy is measured. 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. In addition, the sum of all the reconstructed showers constrains the missing energy in

decays involving neutrinos. The ECL also allows for determining luminosity by measuring the Bhabha [1] scattering rate and using its precisely known cross-section.

The ECL is a highly-segmented array of 8′736 thallium-doped, cesium iodide crystals (CsI(Tl)) assembled in a projective geometry pointing to the interaction region (Figure 2.3). 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 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.

The ECL consists of a 3 m-long *barrel* section with an inner radius of 1.25 m and annular endcaps at $z = 1.96\,\text{m}$ (*forward*) and $z = -1.02\,\text{m}$ (*backward*) from the interaction point. The polar angle coverage ranges from 12.4° to 155.1°. Between the barrel and the endcaps exists two gaps of $\sim 1°$. The barrel has a tower structure that projects to a region near the interaction point of the beams. It contains 6′624 CsI(Tl) crystals with 29 different shapes. Each crystal is a truncated pyramid of an average size of $6 \times 6\,\text{cm}^2$ in cross section and 30 cm in length (corresponding to $16.1 X_0$ radiations lengths). Each crystal is enveloped in a 200 µm thick Gore-Tex teflon layer, coated by a 25 µm laminated aluminium and mylar sheets. The readout of the scintillation light is performed with two $10 \times 20\,\text{mm}^2$ Hamamatsu Photonics S2744-08 photodiodes, glued to the rear surface of each crystal. The average output signal of the crystals is about 5′000 photoelectrons per MeV. Figure 2.9 shows the material budget in front of the calorimeter [38]. It is possible to see that the thickness (X/X_0) as a function of $cos(\theta)$ is higher in the endcap, especially in the backward region where all the readout electronics is placed.

The intrinsic energy resolution of the calorimeter can be expressed as [37]:

$$\frac{\sigma_E}{E} = \sqrt{\left(\frac{0.066\%}{E}\right)^2 + \left(\frac{0.81\%}{\sqrt[4]{E}}\right)^2 + (1.34\%)^2} \tag{2.1}$$

where E is the energy in GeV and the first term is the contribute of the electronic noise. 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 Belle II ECL is the same used for the Belle detector. The increment of luminosity for Belle II data acquisition, has a big impact on four main aspects: photodiodes dark current, pile up noise, high-energy photon backgrounds and radiation damages. The photodiode dark current increases due to higher neutron bombardment. The pile up noise is caused by soft background photons with average energy around 1 MeV and the fluctuations of their number. Energetic background photons produce random cluster candidates in the ECL, resulting in combinatorial background

[1] Bhabha scattering is the electron-position scattering process $e^+e^- \to e^+e^-$

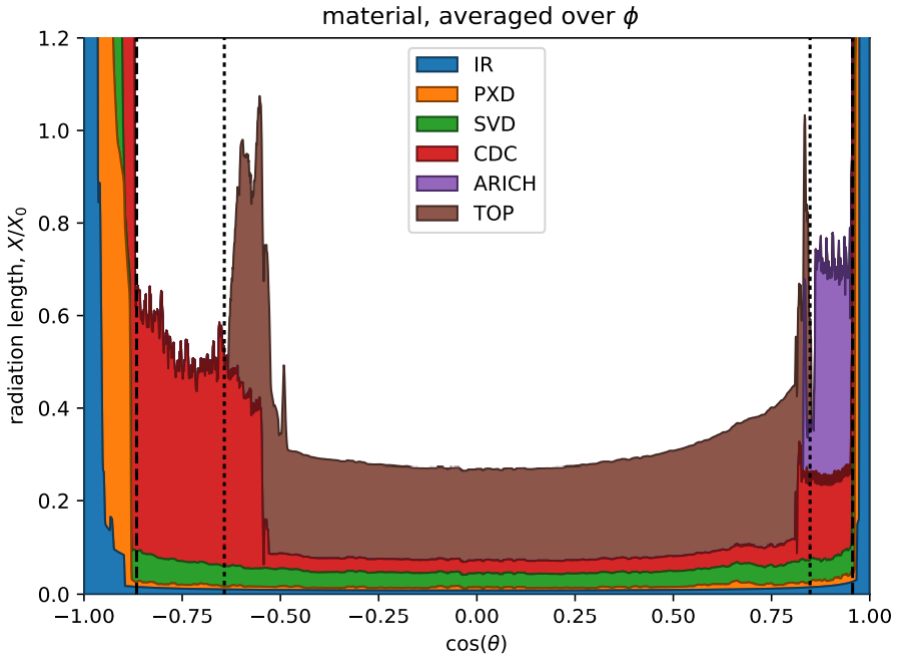


Figure 2.9: Number of electromagnetic radiation lengths X/X_0 in front of the calorimeter as a function of $\cos(\theta)$, averaged on ϕ [38].

for the event reconstruction. These first three problems have been resolved using an improved redout electronics and incrementing the photon energy threshold from 20 to 30 MeV. The crystals have not degraded their performances due to radiation damages. During the operations of the Belle program, the absorbed dose collected by the crystals has been measured. The results, for an integrated luminosity of 900 fb^{-1}, are shown in Figure 2.10. The light output loss is shown in Figure 2.11. The integrated dose is around 100 rad for the barrel crystals, and the highest result is about four times more in the endcap. The corresponding light loss is $\sim$ 7% in the barrel and up to 13% in the endcap region closest to the beam pipe. These results are in good agreement with previous measurements of the crystal radiation hardness. The same studies showed the loss of light output to be less than 30% at 3.6 krad. For all these reasons, the absorbed dose due to Belle II operations, one order of magnitude greater than the Belle one, is not a serious problem.

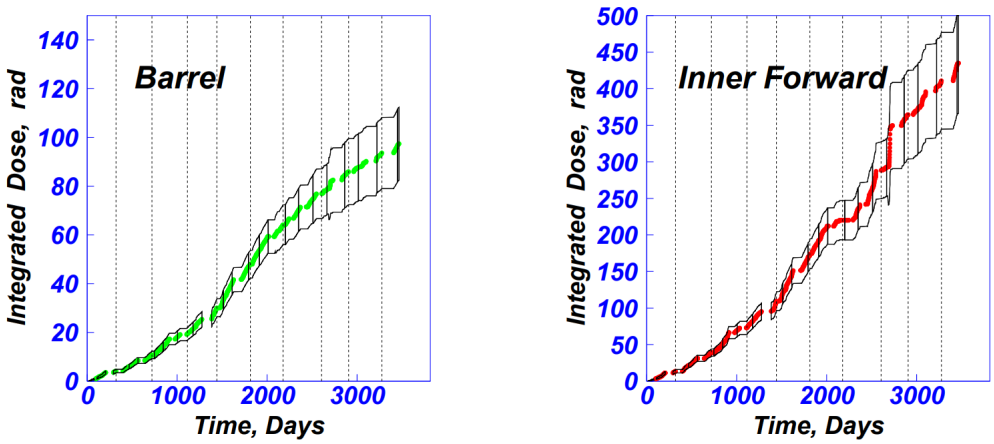


Figure 2.10: Absorbed dose received by the CSI(Tl) crystals during the Belle experiment [37].

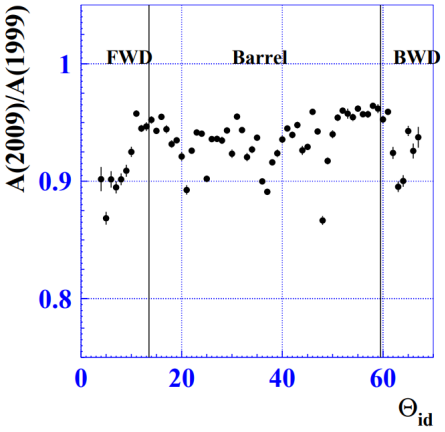


Figure 2.11: Decrease in crystal light output after ten years of Belle operation [37].

2.2.4 Online event selection

Various processes occur in 10 GeV e^+e^- collisions (Figure 2.12). 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 $\Upsilon(4S) \rightarrow B\overline{B}$ and for the continuum to a manageable level, up to a maximum accept rate of about 30 kHz, due to data-acquisition restrictions. $B\overline{B}$ 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) [37] followed by a software high-level stage (HLT) [37]. The L1 decision is mainly based on information from CDC and ECL, but also TOP and KLM information are available. The L1 decision feeds the global decision logic [37], that sends out the final trigger based on the information 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 $e^+e^- \rightarrow hadron$ processes.

2.3 Status of Belle II and future perspectives

SuperKEKB produced its first electron-positron collisions in April 2018, while Belle II started taking physics data in January 2019, after a commissioning period. The SuperKEKB collider has achieved the world's highest instantaneous luminosity for a colliding-beam accelerator, setting a record of $2.40 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, in June 2020.[2] During 2022, constantly higher instantaneous luminosity have been recorded, with a maximum value of $\sim 4.4 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. In June 2022,

[2]Previous world's highest instantaneous luminosity has been set by LHC proton-proton collider at the European Organization for Nuclear Research (CERN) with a value of $2.14 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

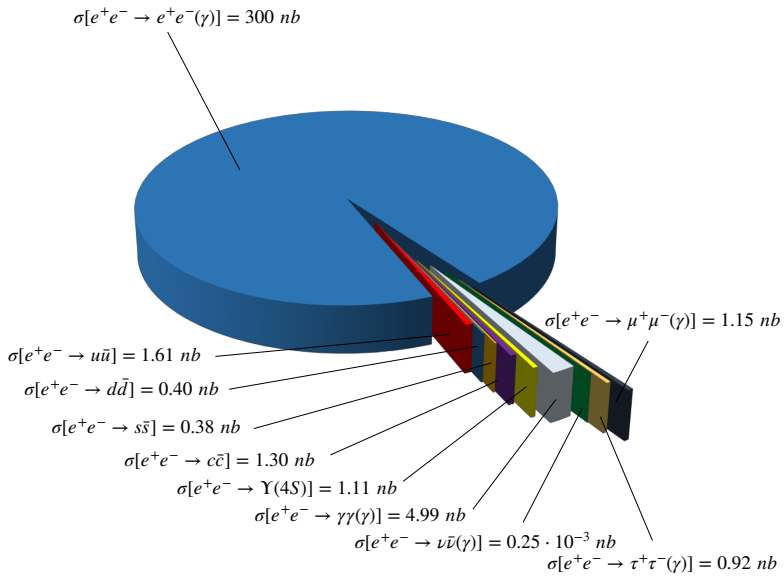


Figure 2.12: Cross sections of the main final states produced in e^+ e^- collision at the $\Upsilon(4S)$ center-of-mass energy.

both SuperKEKB and Belle II ended operations to enter a long shutdown period (LS1) that will last for about 15 months. During LS1, the PXD detector will be extracted and replaced with a new PXD that will have a completed outer layer of pixels. At the same time the TOP photomultipliers will be changed and other minors updates, regarding both the detector and the collider, will be done. Physics program restart is scheduled in early 2024 and it will go on until a second long shutdown (LS2), expected in 2027, will take place.

The complete data sample available today corresponds to an integrated luminosity of $428\,\text{fb}^{-1}$. This roughly corresponds to 80% of the data set collected by the BaBar experiment in its lifetime (9 years) and about to 40% of the Belle full sample (collected in about 11 years). Of the total Belle II sample, $367\,\text{fb}^{-1}$ have been collected at the $\Upsilon(4S)$ resonance, while $61\,\text{fb}^{-1}$ is the size of the off-resonance data sample ($42\,\text{fb}^{-1}$ with collision energy below the $\Upsilon(4S)$ and $19\,\text{fb}^{-1}$ above it). The total integrated luminosity as function of time from January 2019 to June 2022 is shown in Figure 2.13.

To achieve the goal of collecting $50\,\text{ab}^{-1}$, SuperKEKB needs to reach a peak luminosity of $\sim 6.5 \times 10^{35}\,\text{cm}^{-2}\text{s}^{-1}$ and to be upgraded during LS2. An international task force has been formed to provide advice to SuperKEKB on the possible upgrade options, which include a redesign of the interaction region and of the final focusing system. LS2 provides the possibility to upgrade parts of the Belle II detector as well. A new vertex detector might be required to accommodate the new interaction-region design, and other sub-detectors might require improved robustness against increasing machine background [39].

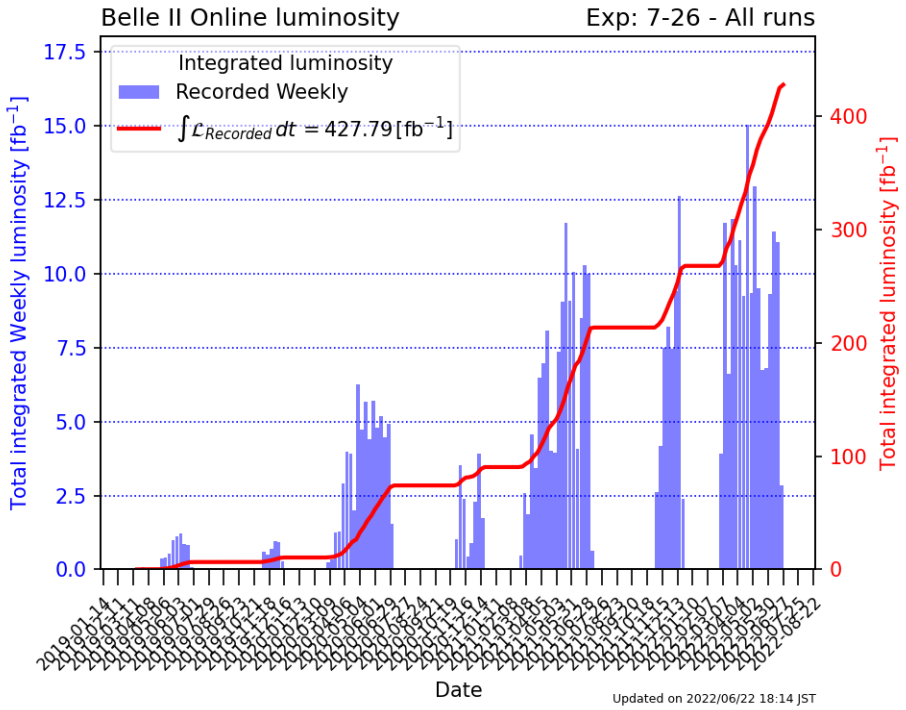


Figure 2.13: Weekly (blue histogram) and total (red line) integrated luminosity recorded using the Belle II detector during 2019-2022 operations [40]. This is the full sample collected before long shutdown.

2.4 The Belle II software and simulation

The Belle II analysis software framework (basf2) [41] contains open source tools and algorithms for event simulation, reconstruction and analysis. In addition to basf2, other packages are used for event generation in Belle II: EvtGen [42] simulates the B-meson decays, PYTHIA [43] simulates hadronisation, KKMC [44, 45] generates $e^+e^- \rightarrow \tau^+\tau^-$ events, TAUOLA [46, 47] generates τ decays and GEANT4 [48] is used for the detector response.

MC samples are used for three main aspects: to study the expected number of events, to study the background composition of the samples and the selection efficiency of the applied cuts.

There are two different types of MC simulation used in Belle II: run-indipendent and run-dependent. The biggest difference between the two is in the beam background description. Run dependent MC uses random triggers (delayed Bhabha) to record samples of non-signal detector occupancy. These are overlaid with the generated MC signal. For run-independent MC, the background overlays are simulated, a combination of various single beam backgrounds (*e.g.* beam gas) and low-angle high-cross section physics processes, such as radiative Bhabha or two photon fusion production of e^+e^-. A lot of work has gone into making these simulated backgrounds as realistic as possible, but they do not necessarily reflect reality. The run-dependent MC is also intended to reflect the varying status of the detector, such as dead channels, and the variations in the trigger. Run independent does not. Run-dependent simulation has been implemented only recently in the last months. In this work, I study the π^0 reconstruction efficiency using MC run-dependent for the first time.

Chapter 3

Overview of the $B^0 \to \pi^0\pi^0$ analysis

In this Chapter, I outline the general principles of the reconstruction of B and D decays at Belle II, which introduce the reconstruction and selection of the $B^0 \to \pi^0\pi^0$ decay analysed for the measurement of the CKM angle α. I give and overview of the most recent analysis of this decay in Belle II, which motivated the study of the π^0 reconstruction efficiency using D^0 decays presented in this Thesis. Therefore, I expand on the photon and π^0 selection of the $B^0 \to \pi^0\pi^0$ analysis.

3.1 B and D decays reconstruction at Belle II

The work presented in this Thesis targets a measurement of the reconstruction efficiency of the π^0 meson for the analysis of $B^0 \to \pi^0\pi^0$ decay. The study of this efficiency exploits D^0 meson decays as control channels. I give here a general overview of the reconstruction of B and D decays.

The collisions between electrons and positrons, produce various final states, as shown in Figure 2.12. In 0.4% of collisions, a $\Upsilon(4S)$ particle is produced, and 96% of them decay into a pair of B mesons, referred to as signal B-meson, B_{rec}, and partner B-meson, B_{tag}, in Figure 3.1. Charm mesons instead are produced by two mechanisms: from the primary process $e^+e^- \to c\bar{c}$, which yields an incoherent production of two charmed particles, or from decays of B mesons into charm final states. The first process has a cross section 20% larger than that of $e^+e^- \to \Upsilon(4S)$.

Both types of events, leading to either bottom or charm mesons, produce about ten tracks in acceptance on average, which are used by the trigger to recognize hadronic events against the majority of low-multiplicity processes. Figure 3.1 shows a sketch of an example of a B-meson decay, $B^+ \to K^+\pi^0$.

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 130 µm. Charm mesons also features significant lifetime (*e.g.* 0.4 ps for D^0), and have typically larger boost than B mesons, yielding displacements around 200 µm on average. This allows reconstructing the decay position for measurements of quantities depending on decay time (*e.g.* flavour oscillations, lifetime measurements) and, to some extend, for background discrimination. The trajectory of the charged particle is recon-

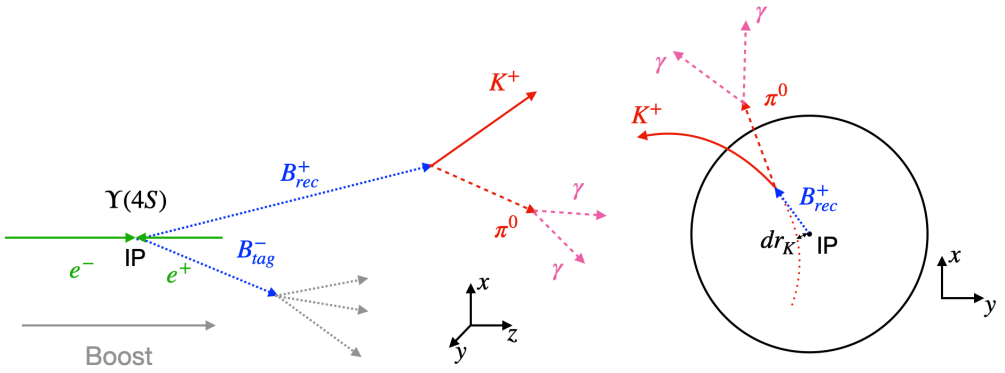


Figure 3.1: Sketch of the decay $B^+ \to K^+\pi^0(\to \gamma\gamma)$ in (left) longitudinal and (right) transverse views. The particle labeled as B^+_{rec} represents the signal B meson, while B^-_{tag} is the partner B-meson. Nothing is drawn to scale.

structed by a geometric fit using the hits on the active layers of the tracking detectors. Knowledge of trajectories and magnetic field 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. These photons are identified and reconstructed by using information from the electromagnetic calorimeter where their energy is measured. 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. 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: other decays (different from signal) and continuum. The first occur when one or more final-state particles are misidentified or when a different decay yields the same final state of that of the decay of interest. The second happen when 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. For B decays, charm mesons produced via $e^+e^- \to c\bar{c}$ are a source of continuum background too. To separate signal from background, Belle II exploits the available event information to construct discriminating variables at particle-, candidate-, and event-level.

3.1.1 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 resid-

ual 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 interaction point (IP) are effective to suppress such backgrounds, because beam-background tracks usually do not point back to the IP. The quantities typically used at Belle II are the transverse (dr) and longitudinal (dz) distances of a track from the IP. Examples of the distributions of these quantities for tracks coming from $D^0 \to K^+\pi^-$ decays are shown in Figure 3.2.

Hit multiplicity. The number of hits improves the resolution on momentum, and also false tracks made of accidental combinations of unrelated hits are suppressed by high hit-multiplicity requirements. This is particularly relevant for low-momenta tracks, such as those associated with the *slow* pion from D^* decays; an example is given by the plot of the CDC of this pion in $D^* \to D^0(\to K^+\pi^-)\pi$ decay in Figure 3.2.

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. A typical example is given in Figure 3.2, where the distribution of the kaon identification variable is shown for kaon candidates in $D^0 \to K^+\pi^-$ decays.

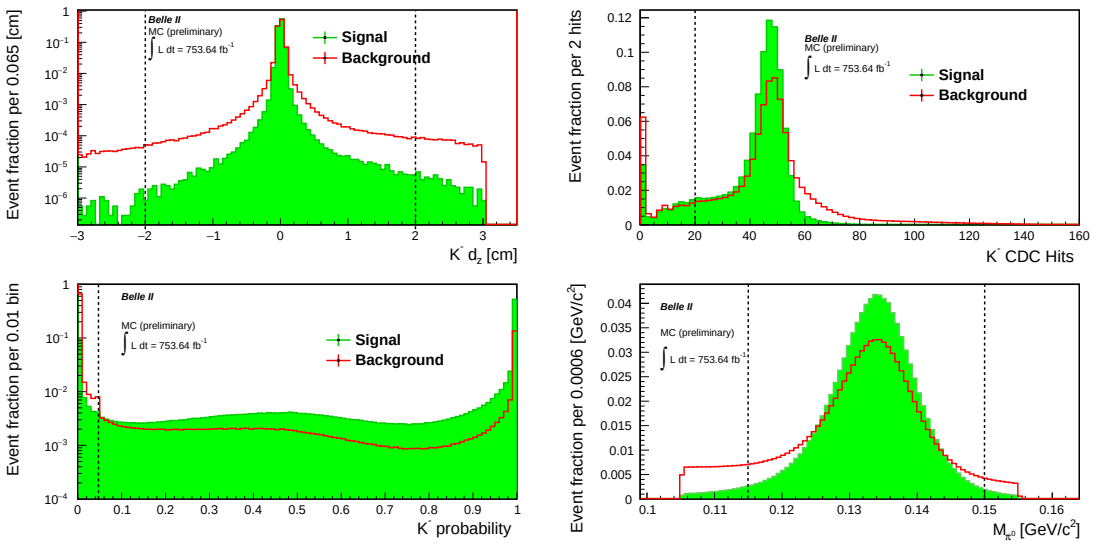


Figure 3.2: Longitudinal displacement from the interaction point for K^- (top-left) with a reconstruction-level cut set to ± 3 [cm]. Number of hits in the CDC for K^- (top-right). Distribution for the kaon ID output (bottom-right) with a cut value of 0.04717. Reconstructed invariant mass of π^0 candidates trough the decay $\pi^0 \to \gamma\gamma$ (bottom-left), the distribution has two cuts at 105 and 155 [MeV/c^2]. The histograms filled in green are the signal, the background is represented as a red line and the vertical dashed lines corresponds to the chosen cut values.

3.1.2 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 true signal.

Beam-constrained quantities. Using distinctive kinematic information about the signal is a common approach to suppress background. A widely used discriminator is the candidate's invariant mass, since fully reconstructed genuine signal events cluster at a specific mass and background shows typically broader distributions. An example is given by the diphoton invariant mass used to select π^0 candidates from the dominant $\pi^0 \to \gamma\gamma$ decay (see Figure 3.2). The peculiar kinematic environment of B-Factory colliders provides additional constraints that further background separation for B decays. The $\Upsilon(4S)$ is produced at rest and decays in two same-mass particles, B and $\overline{B}$. If the B meson is correctly reconstructed, the energy of its decay products equals half of the collision energy in the center-of-mass frame. This is optimally exploited by two variables specific of B-Factories,

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

$$M_{\mathrm{bc}} = \sqrt{s/4 - |\vec{p}_B^{\,*}|^2}, \tag{3.1}$$

  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 collision energy. The M_{bc} variable is similar to the B invariant mass ($5.28\,\mathrm{MeV/c^2}$), but it 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. The signal from B decays is expected to peak at the B meson mass, while background 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 $q\overline{q}$ continuum events.

- energy difference defined as:

$$\Delta E = E_B^* - \sqrt{s}/2 \tag{3.2}$$

  is the difference between the reconstructed B-candidate energy in the center-of-mass frame and 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.

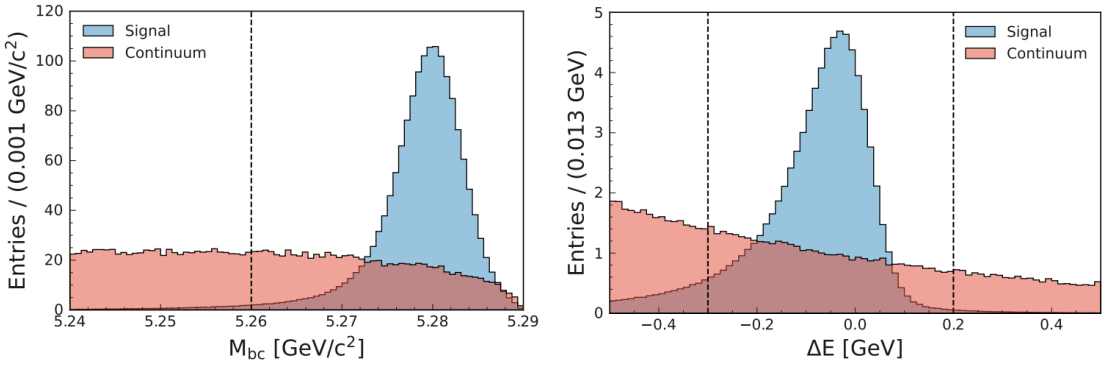


Figure 3.3: Comparison between signal (red) and background (blue) for M_{bc} (left) and ΔE (right) [49].

An example of $M_{\rm bc}$ and ΔE distribution for signal and background for $B^0 \to \pi^0\pi^0$ decays is reported in Figure 3.3.

3.1.3 Event-level variables for B decays

At a further level of reconstruction, global-event information are used too in signal-from-background discrimination. Relevant examples include event-shape variables. Hadronic e^+e^- cross-sections are dominated by continuum background, consisting in production of light $q\overline{q}$ pairs (where q is u, d, c, s, see Figure 2.12) that mostly yield pions and kaons. Because of the kinematic features associated with at-threshold $B\overline{B}$ production, variables capable to capture the spatial and phase-space distributions of final-state particles, such as the normalized Fox-Wolfram moments [50] or the Kakuno-Super-Fox-Wolfram moments (variants of the Fox-Wolfram moments with improved performances in case of neutral particles), offer powerful discrimination of $B\overline{B}$ events from continuum.

In a $B\overline{B}$ event, both B mesons are nearly at rest in the $\Upsilon(4S)$ frame. The B decay products are therefore emitted isotropically in that frame. 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 $B\overline{B}$ decay products are approximately spherical, compared to jet-like shapes for continuum.

Topological variables are usually given in input to a multivariate-analysis algorithm that combines them in a single output that features enhanced discrimination between signal and background. Those algorithms exploit machine-learning techniques such as boosted-decision-trees or neural networks, trained on simulation or control samples of data. As an example, I report in Figure 3.4 the distribution of the continuum-suppression fast-boosted-decision-tree (FBDT) output used in the $B^0 \to \pi^0\pi^0$ analysis. A requirement on the output is generally placed to select the signal sample; this requirement is based usually on the optimisation of a figure of merit which maximises signal significance (or other parameters of interest of the analysis).

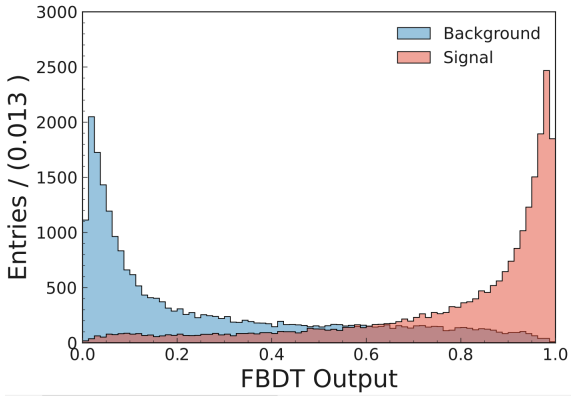


Figure 3.4: Plot of the continuum-suppression FBDT output from simulation for $B^0 \to \pi^0\pi^0$ decays for signal (red) and background (blue) [49].

3.2 The $B^0 \to \pi^0\pi^0$ analysis

The $B^0 \to \pi^0\pi^0$ is analysed to measure the branching fraction and the direct $\mathcal{CP}$ asymmetry $\mathcal{A}_{\mathcal{CP}}$, which is defined as:

$$\mathcal{A}_{\mathcal{CP}} = \frac{N(\overline{B}^0 \to \pi^0\pi^0) - N(B^0 \to \pi^0\pi^0))}{N(\overline{B}^0 \to \pi^0\pi^0)) + N(B^0 \to \pi^0\pi^0))}. \tag{3.3}$$

These two quantities, along with precise measurements of the branching fraction and CP asymmetries for each $B \to \pi\pi$ decay, are required to determine the CKM angles α through the isospin analysis described in Section 1.3.2.

The $B^0 \to \pi^0\pi^0$ decay is suppressed due to the smallness of the CKM matrix element $|V_{ub}|$ and its is also color suppressed;[1] its branching ratio is expected to be $\mathcal{O}(10^{-6})$. The decay is also difficult to reconstruct and prone to large background since there are large numbers of neutral pions in continuum background events that can be combined to mimic the signal. Moreover, $\pi^0 \to \gamma\gamma$ has a branching fraction of $98.823 \pm 0.034\%$ [51], meaning that the final state of the $B^0 \to \pi^0\pi^0$ decay is only four photons. This makes the reconstruction susceptible to photons from synchrotron radiation, and beam interactions with the beam pipe and residual gas.

The absence of charged particle in the final state reduces the variables with signal-background discrimination power to mainly M_{bc}, ΔE and the event-level variables. Other variables previously described as: track displacement, hit multiplicity and χ^2 probability of the vertex fit are not defined. The impossibility to determine precisely the vertex makes a decay-time-independent analysis impossible. Currently, the only active experiment that can study this decay and improve the previous measurements is Belle II. The latest results from Belle II has been presented as preliminary at the International Conference of High Energy Physics in Bologna in July 2022 [52]. The analysis is based on a data sample that corresponds to an integrated luminosity of $190\,\text{fb}^{-1}$, which is roughly half of the total sample currently available.

The Belle II analysis exploits three variables to determine the signal yields of $B^0 \to \pi^0\pi^0$ decays: M_{bc}, ΔE (Equations 3.1 and 3.2) and a log-transform of a FBDT output for the continuum sup-

[1] A decay is color suppressed when the quark produced from the decay of a W boson or a gluon, must match the color of the spectator quark, in order to produce colorless final states.

Particle	Variable	Requirement	Signal loss (%)
B^0	M_{bc} [GeV/c^2]	$[5.26, 5.29]$	0.5
	ΔE [GeV]	$[-0.3, 0.2]$	3.2
π^0	3D angle [rad]	< 0.4	1.6
	$\Delta\phi$ [rad]	$[-0.4, 0.4]$	1.6
	cosHelicityAngleMomentum	$[-0.99, 0.99]$	0.0
	p [GeV/c]	> 1.5	0.1
	M [MeV/c^2]	$[115, 150]$	6.8
γ	E [MeV]	> 30	1.2
	clusterTiming [ns]	$[-200, 200]$	0.0
	clusterNHits	> 1.5	0.2
	clusterTheta [rad]	$[0.2967, 2.6180]$	0.6
	photonMVA	$[0.2, 1.0]$	4.0
	CSMVA	$[0.74, 1.0]$	28.0

Table 3.1: Selections used in the $B^0 \to \pi^0\pi^0$ analysis, divided for the B meson, the π^0 and the photons. For each cut is reported the percentage of signal lost [49]. The different variables are explained in the text.

pression (see Figure 3.4). The unbinned distribution of these variables are fit in seven bins of r, a variable that accounts for the probability of tagging an event as $\bar{B}^0$ or B^0 to measure the $\mathcal{A}_{\mathcal{CP}}$ asymmetry. Flavour tagging is provided by the algorithm described in Ref. [53].

The fit follow a selection of the candidates that is summarized in Table 3.1. All combinations of two photons that have the requested requirements create a π^0 candidate. The combination of two π^0 that pass the selection are combined and create a B^0 candidate. The signal should have a ΔE distribution peaking at zero. The actual peak is shifted from zero due to energy loss in the π^0 reconstruction via energy leakage from the ECL cluster. The distribution of the beam-constrained mass and ΔE, with the corresponding cut values have been already presented in figure 3.3. In events with multiple candidates, the event with minimum deviation from the known value of each π^0 invariant mass is the selected. An requirement on the FBDT output for the continuum-suppression (see Figure 3.4) is determined using simplified simulation to study different cut values from 0.02 to 0.98 and selecting the requirement that minimises the relative signal-yield uncertainty. The FBDT output is then transformed to

$$T_c = \log\left(\frac{x - x_{cut}}{x_{max} - x}\right), \tag{3.4}$$

where x is the continuum suppression variable, x_{cut} is the cut value and x_{max} is the maximum value of the CSMVA output. This variable is used in the three-dimensional unbinned fit to determine the signal yields.

The selections of the neutral pions and the photons are detailed in the next Section.

3.2.1 Photon and π^0 selection

The photon selection aims to reduce the false photons contributions, coming mainly from beam background and interaction of other particles in the ECL that are misidentified with photon clus-

ters (in Section 4.1 a more detailed description is given).
The photon energy must exceeds 30 MeV to suppress low-energy contribution due to beam backgrounds. The variable clusterNHits is used to reduce the number of soft-photons and is the sum of weights $\omega_i(\omega_i < 1)$ of all crystals in an ECL cluster. For non-overlapping clusters this is equal to the number of crystals in the cluster. In case of energy splitting among nearby clusters, this can be a non-integer value.
The time of a cluster in the calorimeter is evaluated as a difference between the photon timing and the event time t_0. The event time is the trigger time and can be measured by several subdetectors including the CDC, ECL, and TOP. The t_0 variable is the final combined value of all the event time measurements. Photon timing is given by the fitted time of the recorded waveform of the highest energy crystal in the cluster. After all calibrations and corrections (including Time-Of-Flight), photons from the interaction point (IP) should have a photon timing that corresponds to t_0. For an ECL cluster produced at the interaction point in time with the event, the cluster time should be consistent with zero within the uncertainties. Requiring $|\text{clusterTiming}| < 200\text{ns}$ the contributions due to Bhabha scattering is suppressed.
The ECL cluster's polar angle, generally different from the photon polar angle, is the clusterTheta variable. The direction of a cluster is given by connecting the center of the laboratory reference frame $(0, 0, 0)$ with the cluster centroid position in the ECL. The centroid position is a logarithmically weighted average of the energy deposition evaluated at the crystal centers, calculated using up to 21 crystals. Cluster centroids are generally biased towards the centers of the highest energetic crystal. This effect is larger for low energetic photons. The selected region is equivalent to the CDC geometric acceptance.
The distributions of all these variables are shown in Figure A.2.

Photon MultiVariate Analysis

Electron-positron scattering events can deposit a large amount of energy in the crystals of the calorimeter. This energy can still be present when an another hadronic event occurs. The combination of a photon with this residual energy can form a π^0. Successively, the misreconstructed particle together with a true π^0 can create a B^0 candidate. To suppress these events, a fast boosted decision-tree (FBDT) has been trained to distinguish between true and false photon using ECL variables. The chosen variables must posses high discriminating power and good data-MC agreement, quantified as a reduced chi-squared of less than 5. The full list of the variables used and the corresponding reduced χ^2 is shown in Table 3.2.
ECL variables can vary significantly according to the region of the calorimeter, thus three different photon MultiVariate Analysis (photonMVA) have been trained (forward, barrel and backward) using photons from simulated $B^0 \to \pi^0\pi^0$ decays and, applying the same selections used in the analysis, the hits on ECL not caused by photons are classified as miresconstructed while those originating from photons are classified as genuine. Two independent datasets, with the same number of pure and misreconstructed γ, are used for training (80%) and validation (20%).

ECL variable	Reduced χ^2
clusterAbsZernikeMoment40	2.50
clusterAbsZernikeMoment51	1.31
clusterE1E9	3.64
clusterE9E21	2.79
clusterHighestE	1.32
clusterSecondMoment	3.63
clusterZenikeMVA	1.09
minC2TDist	1.95
p_t	1.05
E	0.93

Table 3.2: Variables used in the photonMVA and the reduced χ^2 for 24 degrees of freedom [49]. An explanation of the single variables is reported in appendix B.

The Receiving Operating Charateristic (ROC) curve helps to understand the discrimination power between signal and background for a given variable. On a plane with signal efficiency against background rejection, these two parameters are calculated for a given cut value of a variable. Changing cut value and plotting the results, the bigger the area under the curve, the better the discrimination power of the variable.

The photonMVA output, the ROC curve, the signal and background efficiency for the barrel region of the calorimeter are shown in Figure 3.5. The decay $D^{*+} \to \bar{D}^0(\to K^0_s(\to \pi^+\pi^-)\pi^0)\pi^+$ was used to validate the photonMVA, with a cut value of 0.2 in each region obtained maximizing the corresponding figure of merit (FOM) $\frac{S}{\sqrt{S+B}}$, where S is the number of signal events and B is the number of background events. A summary of true photons retention and misreconstructed photons rejections is shown in Table 3.3.

The selected photons are combined to form π^0 candidates. The variables of the π^0 candidate used to discriminate signal and background are the angles between the photons produced in a π^0 decay, the momentum and the mass of the candidate. More precisely, the 3D angle and $|\Delta|\phi$ are respectively, the angle and the azimuthal angle between the photon pair produced by the π^0 decay. The cosine of the helicity angle momentum is the cosine of the angle between the line defined by the momentum difference of the two daughters, in the frame of the mother particle, and the momentum of the given particle in the lab frame.

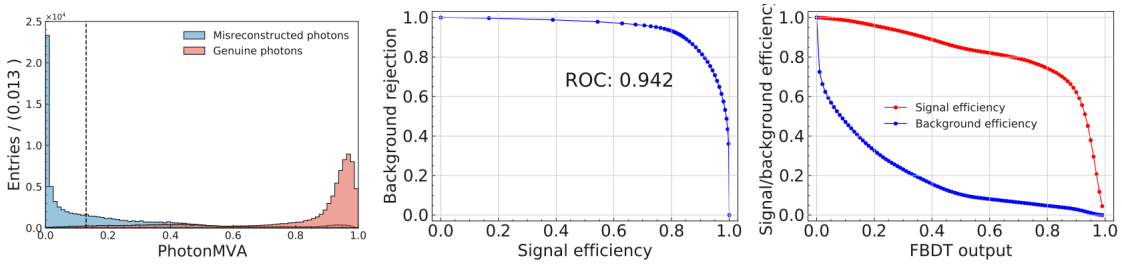


Figure 3.5: Plot of photonMVA output on testing data consisting of photons that originated from the signal-mode (left) for genuine (red) and misreconstructed (blue) photons where the black line is the photonMVA selection. Signal/background efficiency (middle) and the ROC curve (right) for the photonMVA in the barrel region [49]. These results are obtained on the validation sample ($D^{*+} \to D^0(\to K^0_s(\to \pi^+\pi^-)\pi^0)\pi^+$).

ECL Region	ROC curve	photonMVA cut	Signal retention (%)	Background rejection (%)
Backward	0.922	0.20	95.1	38.0
Barrel	0.942	0.20	95.9	32.5
Forward	0.904	0.20	95.0	45.9

Table 3.3: Areas of the ROC curves in each ECL region, photonMVA cut values with genuine photon retention and misrecontructed photon rejections calculated in the validation sample ($D^{*+} \to \bar{D}^0(\to K^0_s(\to \pi^+\pi^-)\pi^0)\pi^+$) [49].

The selection for the invariant mass of the π^0 corresponds to a $\pm 2.5\,\sigma$ about the nominal π^0 mass, where σ denotes the experimental resolution of approximately $8\,\mathrm{MeV/c^2}$. The momentum cut is set to $1.5\,\mathrm{GeV}/c$, where the spectrum of the signal starts, as shown in Figure A.1 (top-middle) along with the distributions of the other variables.

3.3 $B^0 \to \pi^0\pi^0$ analysis results

The measurement of branching fraction is obtained as:

$$\mathcal{B} = \frac{N}{\varepsilon \cdot 2 \cdot N_{B\bar{B}}} \tag{3.5}$$

where N is the signal yield obtained from the fits, ε is the signal efficiency and $N_{B\bar{B}}$ is the number of produced $B^0\bar{B}^0$ pairs. The latter is obtained, assuming that the $\Upsilon(4S)$ decays exclusively to $B\bar{B}$ pairs, using the integrated luminosity ($189.9\,\mathrm{fb}^{-1}$), the $e^+e^- \to \Upsilon(4S)$ cross section, and the $\Upsilon(4S) \to \mathrm{B^0\bar{B}^0}$ branching fraction. The signal yield and $\mathcal{A}_{CP}$ are determined via a simultaneous fit in 7 bins of $q \cdot r$ in three dimensions: M_{bc}, ΔE and T_c. The only free parameters are the signal, the background yields and the $\mathcal{A}_{\mathcal{CP}}$.

To validate the $B^0 \to \pi^0\pi^0$ analysis the full procedure has been repeated on the control mode $B^0 \to D^0(\to K^-\pi^+\pi^0)\pi^0$. For the control channel the overall branching fraction is higher, thus the expected number of events is 10-20 times higher than the signal, and $\mathcal{A}_{\mathcal{CP}} = 0$. The selections are the same (with addition of requirements for the tracks to form the $D^0 \to K^-\pi^+\pi^0$ candidates), only the cut on π^0 momentum is removed since one of the neutral pions in the control mode is soft [2].

The same procedure as explained before, has been repeated to perform the multidimensional unbinned maximum likelihood fit on the MC and then on the control mode data. The expected signal yield from simulation is 397 ± 27 and that found in data is 415 ± 29. The measured branching fraction is $3.90 \pm 0.29 \times 10^{-5}$, in agreement with the expected branching fraction $3.76 \pm 0.24 \times 10^{-5}$. The measured $\mathcal{A}_{\mathcal{CP}}$ is -0.053 ± 0.148 that is compatible with 0 within the uncertainties.

The analysis of the $B^0 \to \pi^0\pi^0$ has been carried out in a *blinded* way: the full analysis has been developed without using any experimental data in the signal region order to avoid possible bias in the final results. Only after full validation and interal review in the Collaboration, the unblinding of the dataset has been performed. The ΔE, M_{bc} and T_c distributions for the $189.9\,\mathrm{fb}^{-1}$ dataset

[2] The π^0 originating from the D^0 decay has low momentum, for this reason is named soft

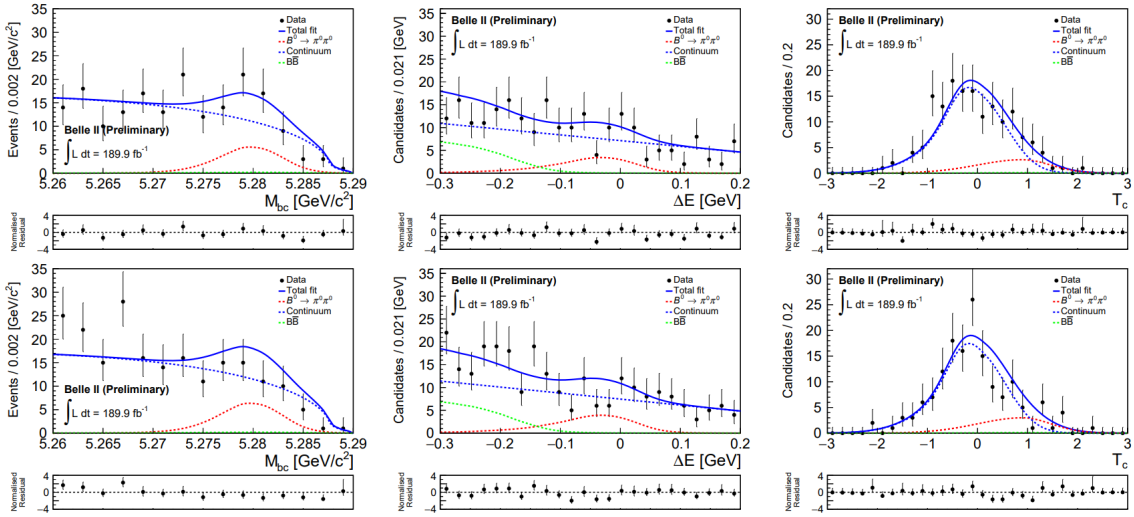


Figure 3.6: Distributions of M_{bc} (left), ΔE (middle) and T_c (right) for the $B^0 \to \pi^0\pi^0$ candidates for events with positive (top) and negative (bottom) q tags. The result of a fit to the sample is shown as a solid line. The fit components are shown as red dashed line (signal), blue dashed line (continuum background), and green dashed line ($B\overline{B}$ background). The plots are signal-enhanced, which correspond to restricting to candidates with $5.275 < M_{bc}\,[\mathrm{GeV}/c^2]$, $-0.10 < \Delta E\,[\mathrm{GeV}] < 0.05$ and $0 < T_c < 3$ when the respective variable is not displayed. This reject $\sim 96\%$ of background improving signal visualization. The pulls are shown below each distribution [49].

with fit projections overlaid are shown in Figure 3.6. With a signal efficiency of 35.5% and using the PDG [54] branching fraction, the expected number of signal events is 116 ± 16 in $189.9\,\mathrm{fb}^{-1}$. The yield observed in data is 93 ± 18. Using Equation 3.5 the branching fraction is determined to be [49]:

$$\mathcal{B}(B^0 \to \pi^0\pi^0) = (1.32 \pm 0.25 \pm 0.13) \times 10^{-6}\,. \tag{3.6}$$

The first uncertainty is statistical the second is systematic and will be explained later. The result agrees with the previous measurements, which averaged to $1.59 \pm 0.26 \times 10^{-6}$ [54]. The $\mathcal{A}_{\mathcal{CP}}$ asymmetry is found to be [49]:

$$\mathcal{A}_{\mathcal{CP}} = 0.14 \pm 0.46 \pm 0.07\,. \tag{3.7}$$

The first uncertainty is statistical the second is systematic and will be explained later. These results have been presented at the ICHEP 2022 conference of Bologna [52]. Currently the work is under final internal review in Belle II for the submission of a paper to a peer-review journal.

A comparison between the Belle, BaBar and Belle II measurements of $\mathcal{B}(B^0 \to \pi^0\pi^0)$ and $\mathcal{A}_{CP}(B^0 \to \pi^0\pi^0)$ are shown in Table 3.4. In all cases the statistical uncertainty is the dominant contribution. The size of the data sample used in this analysis ($189.9\,\mathrm{fb}^{-1}$) is roughly 3.5 times smaller than that of Belle ($693\,\mathrm{fb}^{-1}$). However, thanks to the refinement of the analysis techniques that led to a better background suppression, the two measurements have a comparable sensitivity.

The Belle II precision is limited only by the sample size: there are good prospect of improvement once new data will be accumulated. The greatest systematic uncertainty on the branching fraction measurement is that associated the π^0 reconstruction efficiency, which contributes a 7.6% relative uncertainty (compared to 13.1% statistical). The second largest systematic uncertainties

	Belle	BaBar	Belle II
$\mathcal{B}\ [\times 10^{-6}]$	$1.31 \pm 0.19 \pm 0.19$	$1.83 \pm 0.21 \pm 0.13$	$1.32 \pm 0.25 \pm 0.13$
$\mathcal{A}_{CP}$	$-0.14 \pm 0.36 \pm 0.10$	$-0.43 \pm 0.26 \pm 0.05$	$0.14 \pm 0.46 \pm 0.07$

Table 3.4: The branching fraction $\mathcal{B}$ and $\mathcal{A}_{CP}$ of $B^0 \to \pi^0\pi^0$ as measured at Belle [55], BaBar [56] and Belle II [49]. For each measurement, the first uncertainty is statistical and the second is systematic.

are related to the continuum modelling in the fit (7.4%) and on the uncertainty of the continuum-suppression selection efficiency (6.5%). Part of the work presented in this Thesis has been used to calculate the systematic uncertainty in the π^0 reconstruction efficiency (equal to 3.8% per π^0). The result achieved is summarized in an internal note of the Belle II experiment [57]. It represents an improvement over the previous method used in a preliminary results of $\mathcal{B}(B^0 \to \pi^0\pi^0)$ obtained on 63 fb^{-1} of data, which featured a systematic uncertainty equal to 10% per π^0 [58].
In addition to that, in this Thesis I carried out a study that uses an improved simulation of the experiment (featuring better description of the beam background as described in Section 2.4) and a new investigation of the sample-dependence of the π^0 reconstruction efficiency. This work is presented in the next Chapters.

Chapter 4

Study of the π^0 reconstruction efficiency

A major challenge in the measurement of the branching fraction of $B^0 \to \pi^0\pi^0$ decays is given by the determination of the π^0 reconstruction efficiency. In this Chapter, I present the method to study this efficiency using control decays of the D^0 meson. Comparing the value obtained in data and in simulation in the control channels, I derive a correction factor to obtain an unbiased measurement of the efficiency from simulation for the $B^0 \to \pi^0\pi^0$ decay, with an associated systematic uncertainty of about 3.5%. I outline the measurement method using control channels, their selections and study of sample composition, and finally conclude with the determination of the correction factor.

4.1 Reconstruction efficiency and simulation pitfalls

The efficiency of reconstructing a particle, such as the neutral pion in this case, quantifies the capabilities of detecting and correctly identifying that particle. It accounts for the geometrical acceptance of the detector, its intrinsic efficiency of detecting the particle (in this case, the photons that are produced from the π^0 decay), and the efficiency of the selection requirements used to discriminate that particle from other background processes. To measure the efficiency, an unbiased sample of neutral pions of known size should be used. A means to measure efficiency is the use of a full, realistic simulation: knowing the number of generated particles and setting the decay model, one can count the number of surviving particles after simulating the full chain of detection, reconstruction and selections. However, simulation could fail in reproducing the correct value of the reconstruction efficiency for approximation and simplification used to model the detection processes. In the case of the reconstruction of $\pi^0 \to \gamma\gamma$, we might have difference between data and simulation due to:

- Imperfect modeling of the material distribution in the detector. A photon can undergo pair production in the material of the detector before reaching the calorimeter. If the produced

tracks are reconstructed in the tracking detectors, the corresponding clusters in the calorimeter, if any, are tagged as being produced by a charged track and the photon candidate is lost. Even if the reconstruction algorithms still find a photon candidate, the energy resolution might be degraded, leading to a π^0 candidate with an incorrectly reconstructed energy or mass.

- Imperfect modeling of photon-shower shape. In order to discriminate electromagnetic from hadronic showers, shower shape variables, such as the number of crystals in a shower are used. Showers shapes are difficult to simulate as they require a detailed description of the particles interaction in the crystals. Approximation used in modelling this process can creates a small efficiency difference between data and MC.

- Split-offs. The particle showers created by hadrons interacting with the material in the calorimeter contain a fraction of neutral hadrons. Such secondary hadrons can travel a sizable distance in the calorimeter before interacting with the material and depositing (a part of) their energy. These so-called split-offs leave the signature of a calorimeter cluster without an associated track pointing to it, which resembles very closely a real photon. As for the photon-shower shape, a detailed modeling of hadronic showers is difficult: split-offs present a further potential source of difference between data and simulation.

- Additional background in data. Real data events typically contain more (soft) photon candidates, most of which originate from beam-related background. This background consists primarily of electrons and positrons from radiative Bhabha scattering which hit elements of the detector or the beam pipe, producing neutrons with energies in the MeV range, which then can produce low energy showers in the calorimeter. These additional photon candidates increase the number of photons combinations in data, giving rise to more π^0 candidates, especially at low momentum.

All these effects could bias the determination of the efficiency from simulation. It is therefore useful to define control sample of data where the determination of the efficiency from simulation can be checked, and derive corrections to be applied to simulation, when necessary. Several methods have been used in the past at B Factories experiments, involving decays of τ or η particles in clean $e^+e^- \to \tau^+\tau^-$ processes [59]. These control samples have the advantage of exploiting particle decays with very well known branching fractions; however, such class of events tend to produce substantially more activity in the detector than B decays (the number of tracks and ECL cluster are smaller: they are called *low-multiplicity* events). Thus, an additional step is the validation of the efficiency correction to make sure the correction is applicable to B decays. Alternatively, one could exploit control samples of charm-meson decays, that feature same detector condition of B decays, at the cost of a larger uncertainty associated with D-decays branching fractions. This is the method used in this work.

4.2 Measurement method with D^0-meson decays

This Thesis aims to measure the reconstruction efficiency of the π^0 mesons for the $B^0 \to \pi^0\pi^0$ decay. We used a method to validate the reconstruction efficiency determined in simulation using control samples of D^0 mesons. Charmed mesons, *i.e.* mesons with at least one charm quark, are produced in the $e^+e^- \to c\bar{c}$ processes with a cross section of 1.30 nb (higher than the 1.11 nb of $e^+e^- \to \Upsilon(4S)$). This provides abundant samples for reconstructing D^0-meson decays. The selected decays are:

$$D^{*+} \to D^0(\to K^-\pi^+\pi^0)\pi^+_{\text{soft}}\,, \quad D^{*+} \to D^0(\to K^-\pi^+ +)\pi^+_{\text{soft}} \tag{4.1}$$

because they guarantee samples with large statistics (branching fractions of $D^0 \to K^-\pi^+\pi^0$ and $D^0 \to K^-\pi^+\pi^0$ are 4% and 14.4%, respectively, and the D^{*+} goes into $D^0\pi^+_{\text{soft}}$ about 68% of the times) and are very clean channels thanks to the distinctive signature given by the D^{*+} reconstruction with the *soft* pion.[1] The $D^0 \to K^-\pi^+\pi^0$ decay features π^0 that spans a large range of momentum, from less than 0.2 GeV/c to 5 GeV/c, providing good overlaps with the typical momentum range of π^0 in $B^0 \to \pi^0\pi^0$ decays (see Figure 4.1). The key idea is to use these D^0 decays that differ only for the π^0 in the final state, such that a comparison of their reconstructed signal yields in the same data set provides access to the π^0 reconstruction efficiency.
The signal yield of *e.g.* $D^{*+} \to D^0 \to (K^-\pi^+\pi^0)\pi^+_{\text{soft}}$ decays is defined as:

$$Y_s = \left[\sigma(e^+e^- \to D^{*+}D^{*-}) \int \mathcal{L}(t)\,dt\right] \cdot \left[\mathcal{B}(D^{*+} \to D^0\pi^+_{\text{soft}})\mathcal{B}(D^0 \to K^-\pi^+\pi^0)\mathcal{B}(\pi^0 \to \gamma\gamma)\right] \cdot \left[\varepsilon(D^{*+})\varepsilon(D^0)\varepsilon(\pi^+_{\text{soft}})\varepsilon(K^-)\varepsilon(\pi^+)\varepsilon(\pi^0)\right] \tag{4.2}$$

where σ is the cross section, $\int \mathcal{L}(t)\,dt$ is the integrated luminosity of the data set considered, $\mathcal{B}$ are the branching fractions of the decays, and ε are the reconstruction efficiencies of the single particles assuming that the joint efficiency factorises.

[1] In the $D^{*+} \to D^0\pi^+_{\text{soft}}$ decay the pion is defined as soft, because due to the two-body decay dynamics the momentum spectra is distributed at low energy range (below 700 MeV/c).

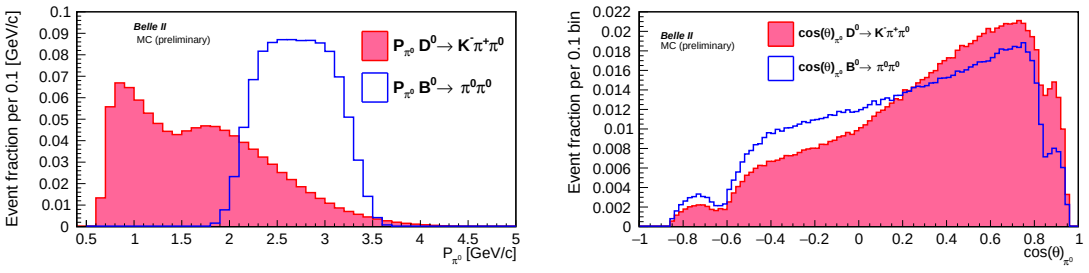


Figure 4.1: Comparison between the signal distributions of π^0 momentum (left) and $\cos(\theta)$, between the $B^0 \to \pi^0\pi^0$ and the $D^0 \to K^-\pi^+\pi^0$ decay. The π^0 momentum distribution for the $K^-\pi^+\pi^0$ decay has a cut at 0.6 GeV/c, applied at reconstruction level. The histograms are normalized to the same area.

The second decay features the same charged particles (tracks) in the final state, and the the signal yield of $D^{*+} \to D^0(\to K^-\pi^+)\pi^+_{\text{soft}}$ is:

$$Y_s = \left[\sigma(e^+e^- \to D^{*+}D^{*-}) \int \mathcal{L}(t)\, dt\right] \cdot \left[\mathcal{B}(D^{*+} \to D^0\pi^+_{\text{soft}})\mathcal{B}(D^0 \to K^-\pi^+)\right] \cdot \left[\varepsilon(D^{*+})\varepsilon(D^0)\varepsilon(\pi^+_{\text{soft}})\varepsilon(K^-)\varepsilon(\pi^+)\right] . \tag{4.3}$$

Using Equations 4.2, 4.3, and solving for $\varepsilon(\pi^0)$ the equation for π^0 reconstruction efficiency is:

$$\varepsilon(\pi^0) = \frac{Y_s(K^-\pi^+\pi^0)}{Y_s(K^-\pi^+)} \left(\frac{\mathcal{B}(D^0 \to K^-\pi^+)}{\mathcal{B}(D^0 \to K^-\pi^+\pi^0) \cdot \mathcal{B}(\pi^0 \to \gamma\gamma)}\right) . \tag{4.4}$$

The reconstructed D^0 mass distributions are fitted. The simulation and data samples used in the analysis corresponds to an integrated luminosity of $753.64\,\text{fb}^{-1}$ and $189.9\,\text{fb}^{-1}$, respectively. In the next Sections, I present the selection of the control samples and the study of the background composition. A limitation of the this method is due to the uncertainty associated with the D^0-meson branching fractions [60]:

$$\frac{\mathcal{B}(D^0 \to K^-\pi^+\pi^0) \cdot \mathcal{B}(\pi^0 \to \gamma\gamma)}{\mathcal{B}(D^0 \to K^-\pi^+)} = 3.65 \pm 0.13\,. \tag{4.5}$$

The relative uncertainty is 3.56%: this sets a boundary to the minimum uncertainty achievable. The correction factor for the simulation is derived by comparing the value of $\varepsilon(\pi^0)$ obtained with the same method in data and in simulation,

$$\frac{\varepsilon_{\text{data}}(\pi^0)}{\varepsilon_{\text{MC}}(\pi^0)}\,. \tag{4.6}$$

In this ratio, I assume that all difference between data and simulation are due to the π^0 reconstruction efficiency. Systematic uncertainties related to this assumption are computed in Chapter 6. A known difference between data and simulation is for PID variables: these variables are defined as ratios between particle-specific likelihood functions and are used to reduce mis-identification. An example is reported in Section 4.3 (in the charged-particle paragraph). A dedicated working-group in Belle II studies the efficiencies of the PID selections and proved corrections for the differences between data and simulation. I used the provided corrections through all the work. A systematic uncertainty related to this data-simulation difference is computed in Section 6.1.2.

4.3 Selection of the $D^0 \to K^-\pi^+\pi^0$ decay

The selection requirements applied for the $D^{*+} \to D^0 \to (K^-\pi^+\pi^0)\pi^+_{\text{soft}}$ decay are reported in Table 4.1. They are divided into three parts: the requirements applied to charged particles; those to form charm-mesons candidates; finally, requirements on π^0 and photons, which mirror exactly those used in the $B^0 \to \pi^0\pi^0$ analysis.

Particles	Variable	Requirement
K^-, π^+, π^+_{soft}	d_z [cm]	$[-2.0, 2.0]$
K^-, π^+, π^+_{soft}	SVD hits	> 0
K^-, π^+	CDC hits	> 20
π^+_{soft}	CDC hits	> 0
K^-	kaonID	$[0.1, 1.0]$
D^{*+}, D^0	Δm [MeV/c^2]	$[144.8, 146.1]$
D^0	m_{D^0} [GeV/c^2]	$[1.75, 1.97]$
D^{*+}	$p_{CMS}(D^{*+})$[GeV/c]	> 2.5
π^0	$< m_{\pi^0}$[MeV/c]	$[115, 150]$
	p_{π^0}[GeV/c]	> 1.5
	cosHelicityAngleMomentum	$[-0.99, 0.99]$
	3D angle (γ from π^0) [rad]	< 0.4
	$\Delta\phi$ (γ from π^0) [rad]	$[-0.4, 0.4]$
γ	photonMVA (forward)	$[0.2, 1.0]$
	photonMVA (barrel)	$[0.2, 1.0]$
	photonMVA (backward)	$[0.2, 1.0]$
	E_γ [MeV]	> 30
	γ cluster hits	> 1.5
	γ cluster θ [rad]	$[0.2967, 2.6180]$
	γ cluster timing ns	$[-200, 200]$

Table 4.1: Selection used on the charged particles (K^-, π^+, π^+_{soft}), charm-mesons candidate (D^{*+} D^0), π^0 and photons.

Charged particles

The requirements are imposed on the following variables: the longitudinal distance of the tracks from the IP ($|d_z|$); the number of hits on the SVD and CDC detector used to fit the track associated to a charged particles; particle-identification variable (kaonID) that discriminates true kaons and reduces particles mis-identification (mainly kaons misidentified as pions, in this case).
The requests on longitudinal displacement and number of hits are very loose to guarantee only negligible loss for the signal efficiency; they are used to suppress beam background and false tracks made of accidental combinations of unrelated hits. The kaonID variable range from 0 to 1 and quantifies the probability of an observed particle to be a true kaon, see the distribution for signal and background in the bottom-right panel in Figure 4.2. It is defined as $\mathcal{L}_K/(\mathcal{L}_e + \mathcal{L}_\mu + \mathcal{L}_K + \mathcal{L}_p + \mathcal{L}_d)$; were $\mathcal{L}_x$ identifies the likelihood function for a specific particle. False kaons constitute the biggest contribution to the total background. The request on kaonID reduce this component, resulting in a total background rejection of 77% with a loss of 12% of the signal. More details on the background composition of the sample are reported in Section 4.3.1.
The distributions of the longitudinal displacement, the number of hits in the CDC and the number of hits in the first SVD layer are respectively shown in Figure C.1 and C.4.

Charmed mesons

The selections used for the charmed mesons exploits the difference between the reconstructed mass of the D^{*+} and D^0 meson (Δm); the reconstructed mass of the D^0 candidates (m_{D^0}) and the

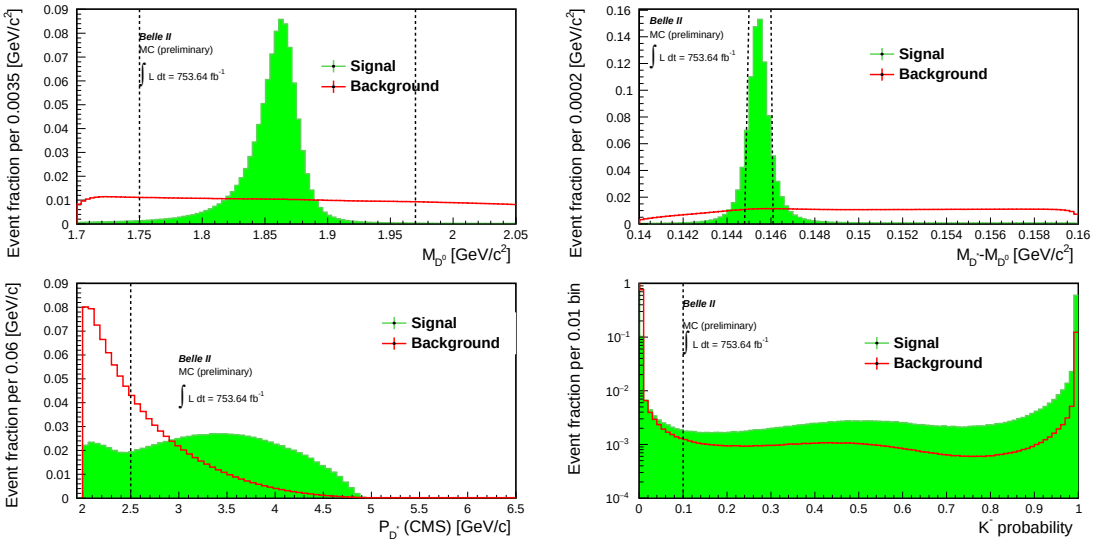


Figure 4.2: The distribution of the invariant mass of the D^0 candidates (top-left), the mass-difference between D^{*+} and the D^0 candidates (top-right), the momentum of the D^{*+} candidates in the center of mass reference frame (bottom-left), the kaon ID output (bottom-right). The distribution of the D^{*+} momentum in the center of mass has a cut at $2\,\mathrm{GeV}/c$ applied at the reconstruction level in order to reduce the size of the sample. The histogram filled in green is the signal, the background is represented as a red line and the vertical dashed lines corresponds to the chosen cut values.

momentum of the D^{*+} candidates in the center-of-mass frame ($p_{CMS}(D^{*+})$). The distributions of these variables are presented in Figure 4.2.

For correctly reconstructed signal candidates the Δm distribution should peak around $145\,\mathrm{MeV}/c^2$, while for the background this distribution is flat. The advantage of using Δm is that most of the experimental resolution in the measurement of the mass (driven by the π^0 energy resolution) cancels in the difference between the charm-candidate masses. A narrow window centered on the peak position rejects a lot of background events increasing significantly the signal purity. The range in D^0 mass is set to avoid partially reconstructed four-body decays of the D^0 meson. The requirement on $2.5\,\mathrm{GeV}/c < p_{CMS}(D^{*+})$ selects only prompt D^{*+}, meaning that only D^{*+} directly created in the e^+e^- collision are considered. This requirements removes a large fraction of background from continuum, but also charm mesons produced in B decays (secondary charm decays). While this latter component could be considered as signal, the benefit in background suppression from this requirement outdoes the gain in signal efficiency from considering secondary charm decays.

Neutral pion and photons

The selections applied for the neutral pion and those applied for the photons are the same requirements used in the $B^0 \to \pi^0\pi^0$ analysis, described in Section 3.2.1.

The $m_{\gamma\gamma}$ range spans $2.5\,\sigma$ around the nominal π^0 mass value, where σ is the experimental resolution which is about $8\,\mathrm{MeV}/c^2$. The π^0 momentum spectrum in the $D^0 \to K^-\pi^+\pi^0$ decay is wide and goes from $\sim 0.2\,\mathrm{GeV}/c$ to $\sim 5\,\mathrm{GeV}/c$. The requirement at $1.5\,\mathrm{GeV}/c$ is imposed to select pions with similar kinematics of those in the $B^0 \to \pi^0\pi^0$ decay (see the left panel of Figure 4.1).

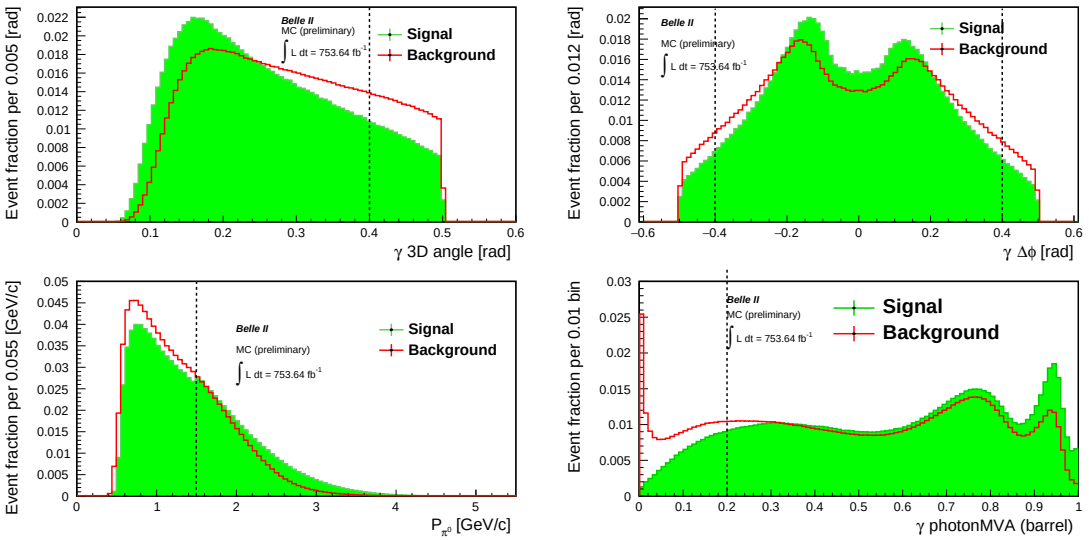


Figure 4.3: Distributions of: the the 3D angle (top-left), the difference in azimuthal angle (top-right) between photons coming from the decay of the π^0 candidates, the π^0 momentum (bottom-left) and the photonMVA (bottom-right) output for the photons of the barrel of the ECL. The 3D angle, $|\Delta\phi|$ and the momentum of the π^0 distributions show cuts applied at the reconstruction level. The histogram filled in green is the signal, the background is represented as a red line and the vertical dashed lines corresponds to the chosen cut values.

Figure 4.3 presents the distributions of 3D angle, $|\Delta\phi|$ and momentum of the π^0 candidates along with the photonMVA output in the barrel of the ECL. The distribution of the π^0 mass is reported in the bottom-right panel in Figure 3.2. The definition of the cosine of the helicity angle, 3D angle and $|\Delta\phi|$ are reported in Section 3.2.1. In each region of the ECL the photonMVA requirements are set to 0.2, analogously with what is explained in Section 3.2.1. The distribution of the photonMVA output in the remaining regions of the electromagnetic calorimeter (forward and backward) are presented in Figure C.2. The request on photon energy suppresses the beam background contributions due to low energy photons. A cut on the γ cluster θ (see Section 3.2.1) keeps only the photons in the CDC acceptance region $[17°, 150°]$. This choice ensure that the nearest area to the beam pipe are excluded, avoiding to consider the majority of the beam-background photons. Moreover, this request permits to recognise the photons produced by the passage of a charged particle. A requirement on the cluster timing (see Section 3.2.1) ensures that no beam background is included in the analysis. These interactions are not expected to be in time with the event that starts the trigger. The distribution of the photon energies, the number of hits, θ, and timing for each cluster are shown in Figure C.3 in Appendix C.1).

4.3.1 Background composition

I use simulation to study the background composition that survive the selection. In simulation we can access truth information about the particles nature and their decays. This are used to categorise candidates as genuine signal or background, and to identify the types of background so to tune the selection to suppress them efficiently. Figure 4.4 shows in orange the background distribution after a first pre-selection used to reconstruct the candidates. Almost all the selection

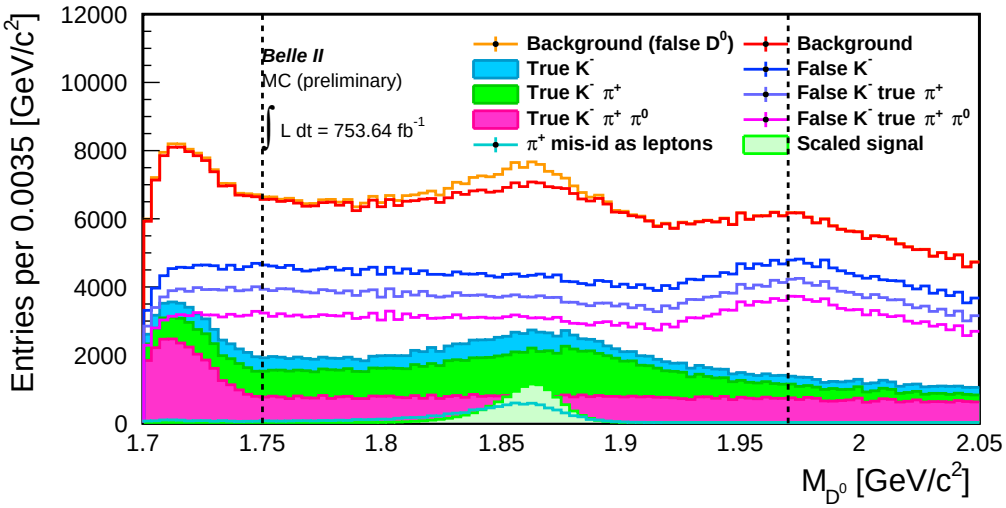


Figure 4.4: Study of the background composition. The orange distribution is the sum of the red histogram and the light blue line, that identifies π^+ mis-identified as leptons due to their decays. The red histogram is the background distribution without the contribution of candidates with decay-in-flight of pions from genuine signal decay. True K^-, $K^-\pi^+$ and $K^-\pi^+\pi^0$ are presented respectively as filled light-blue, green and pink histograms. False kaons, the combination false K^- - true π^+ and the combination false K^- - true π^+ - true π^0 contributions are respectively shown as the blue, violet and magenta histograms. The vertical dashed lines indicate the selected region used throughout the analysis. No cuts on kaonID and photonMVA are applied.

showed in Table 4.1 has been applied, the only exceptions are the p_{π^0}, kaonID and the photonMVA cuts.

We observe a small peak in the signal region, although genuine signal candidates have been removed by truth-level information. Studying the composition of this component, I identified a small contribution of charged pions coming from the D^0 mis-identified as leptons. The mis-identification may be due to the decay-in-flight of charged pions.[2] These D^0 events are true signal events where the π^+ can not be reconstructed as a pion, but it is detected as a lepton originated from the decay of the pion. These type of candidates can be considered as signal for the purpose of the measurement of the π^0 reconstruction efficiency. This contribution is shown in Figure 4.4 with a light-blue histogram. As a further check, I compared the shape between the mis-identified pion and the signal and found them very similar. The small difference between the two shapes is due to presence of a neutrino in the decay $\pi^+ \to \mu^+\nu_\mu$ that yields a worse mass resolution, causing a tail to lower mass value due to the missing energy of the neutrino.

The background composition without the contribution of pions mis-identified as leptons is presented in orange in Figure 4.4. The distribution has three different structures that can be separated considering true and false kaons. The contribution of true kaons is shown as a light-blue filled histogram, and two different peaks are visible between $1.7 - 1.75\,\mathrm{GeV/c^2}$ and $1.8 - 1.97\,\mathrm{GeV/c^2}$. The first one is composed by true K^-, π^+, π^0, and is due to a four-body decay of the D^0 where the fourth particle is not reconstructed. The second structure lays in the same region of the signal, and has a large contribution from fake π^0 candidates. The third bump is located between

[2] The π^+ has a mean lifetime of $2.6 \cdot 10^{-8}$ s and can decay in $\mu^+\nu_\mu$ with a branching ratio of $\sim 99.99\%$.

$1.92 - 2.05\,\mathrm{GeV/c^2}$ and is almost entirely made up by false kaons (presented as a blue histogram). These false kaons are pions that are mis-identified and a wrong-mass hypothesis is assigned, which shifts the reconstructed invariant mass of the D^0 candidates to higher values.

False kaons give one of the largest contribution to the background distribution. To reduce it, the requirement on the kaonID output is applied. Figure 4.5 show the effect of applying this requirement. On top of this, I add the requirement on the photonMVA: the peaking background due to fake π^0 candidates is clearly reduced. The purple distribution is the surviving background distribution after applying the full selection, that will be considered in the rest of the analysis. The structure laying between $1.7 - 1.75\,\mathrm{GeV/c^2}$ is removed by cutting the distribution at $1.75\,\mathrm{GeV/c^2}$.

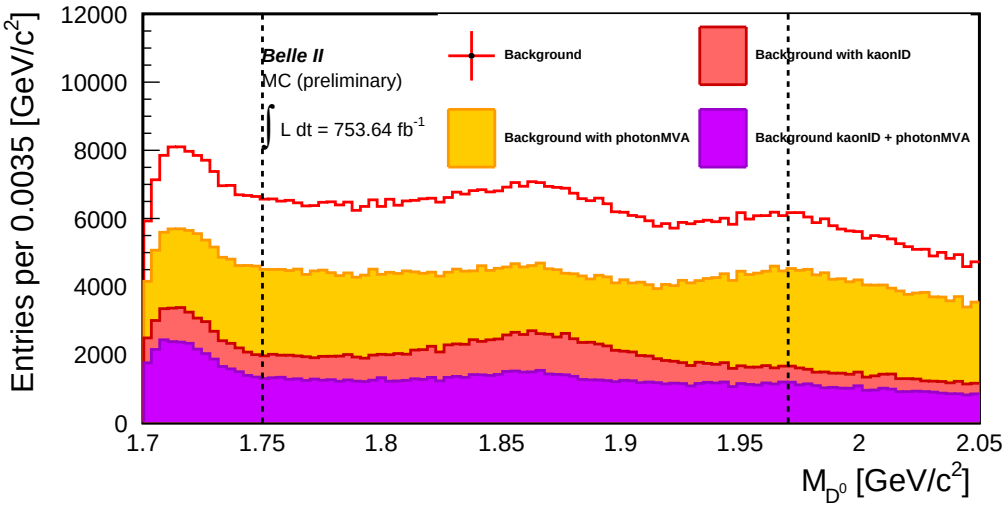


Figure 4.5: Effects of the cuts on kaonID and photonMVA on the total background distribution. The red histogram is the background distribution, the orange-filled histogram is the resulting distribution applying only the photonMVA cuts. Analogously, the histogram filled in red is obtained applying exclusively the cut on kaonID. The combination of the kaonID and the photonMVA cuts results in the histogram filled in purple. The vertical dashed lines indicate the selected region used throughout the analysis.

4.4 Selection for the $D^0 \to K^-\pi^+$ decay

The selection applied for the $D^{*+} \to D^0 \to (K^-\pi^+)\pi^+_{\mathrm{soft}}$ decay is divided in two parts: the requirement for the charged particles and those on the charm-meson candidates. The rationale to choose the selection requirement is to have the same signal efficiency of the $D^0 \to K^-\pi^+\pi^0$ decay such that the ratio of signal yields between the two channels through Equation 4.4 gives an unbiased determination of the π^0 reconstruction efficiency. Systematic uncertainties will be computed in Chapter 6 for the assumption that all efficiencies common to the two channels cancel in the ratio.

4.4.1 Charged particles

The selections applied on charged particles are summarised in Table 4.2. Apart from the kaonID cut value, the selections are the same used for the $D^0 \to K^-\pi^+\pi^0$ decay. The requirement on

Particles	Variable	Requirement
K^-, π^+, π^+_{soft}	d_z[cm]	$[-2.0, 2.0]$
K^-, π^+, π^+_{soft}	SVD hits	> 0
K^-, π^+	CDC hits	> 20
π^+_{soft}	CDC hits	> 0
K^-	kaonID	$[0.04717, 1.0]$
D^{*+}, D^0	Δm [MeV/c^2]	$[144.725, 146.710]$
D^0	m_{D^0} [GeV/c^2]	$[1.814, 1.912]$
D^{*+}	$p_{CMS}(D^{*+})$ [GeV]	> 2.3

Table 4.2: Selection used on the charged particles (K^-, π^+, π^+_{soft}) and charm-meson candidates (D^{*+}, D^0).

kaonID is chosen such to have the same signal efficiency of the $D^0 \to K^-\pi^+\pi^0$ decay.
I present a comparison of signal and background distribution of the variable used in the selection, normalising the histograms to the same unit area. The distributions of the longitudinal displacement from the interaction point of the π^+, the slow pion and the K^- are presented in Figure C.5 and 3.2 (top-left panel). The number of hits in the CDC and the first layer of the SVD are reported in Figure C.6. The kaonID distribution for signal, corresponding to the histogram filled in green, and for the background, red histogram, is shown in the bottom-right panel in Figure 4.6.

4.4.2 Charmed mesons

The selections used for the charmed mesons are reported in Table 4.2. Each selection is tuned to have the same signal efficiency of the corresponding requirements used for the $D^0 \to K^-\pi^+\pi^0$ decay. Figure 4.6 shows the signal and background distributions for each variable used in the selection. The histograms are normalized to the same area to compare the different shapes.

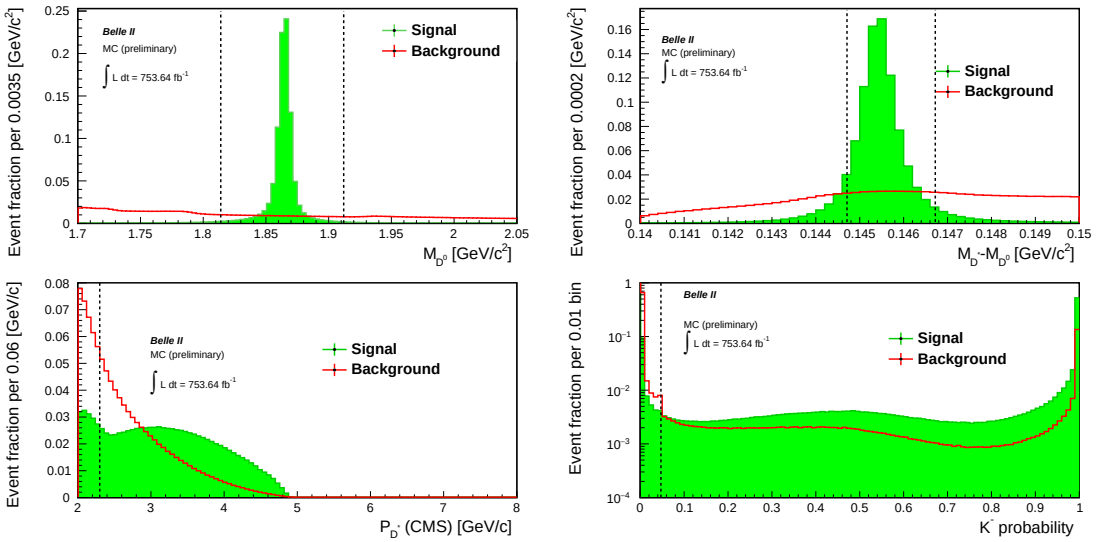


Figure 4.6: The distribution of the invariant mass of the D^0 candidates (top-left), the mass difference between D^{*+} and D^0 candidates (top-right) and the momentum distribution of the D^{*+} candidates in the center of mass reference frame (bottom-left) and the the kaon ID output (bottom-right). The histogram filled in green is the signal, the background is represented as a red line and the vertical dashed lines corresponds to the chosen cut values.

4.4.3 Background composition

The background composition for the $D^0 \to K^-\pi^+$ decay, with the corresponding cut values on the D^0 reconstructed invariant mass is shown in Figure 4.7.

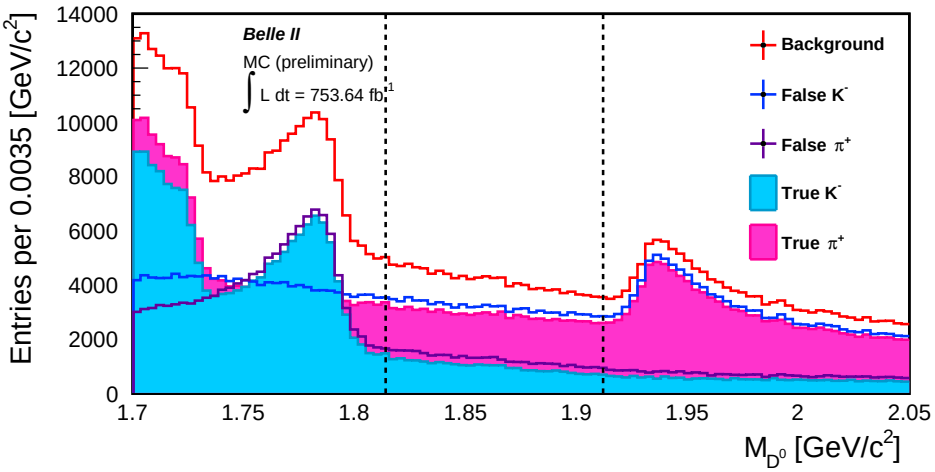


Figure 4.7: Background composition without the cut on kaonID. The contributions due to true kaons, true pions, false kaons and false pions are represented respectively as the blue-filled, the pink-filled, the blue and violet histograms. The vertical dashed line represents the cut values on the D^0 mass, used through all the analysis.

The distribution present three different structures. The first lays between $1.7 - 1.75\,\mathrm{GeV/c^2}$ and is made up of true kaons and pions. The mass of the D^0 candidates is shifted to lower value, because one of the decays products is missing. Thus, the three body decays of the D^0 where one the final state particle escapes the detector are responsible for this contribution. The second peak is between $1.75 - 1.80\,\mathrm{GeV/c^2}$ and is constituted by true kaons and false pions. Actually, the pions are misidentified as kaons and this contribution is due to the decay $D^0 \to K^-K^+$.
The third peak is situated above $1.9\,\mathrm{GeV/c^2}$ and the contributions are true pions and false kaons. In this case the kaons are mis-identified as pions and the decay $D^0 \to \pi^+\pi^-$ is responsible for the peak.

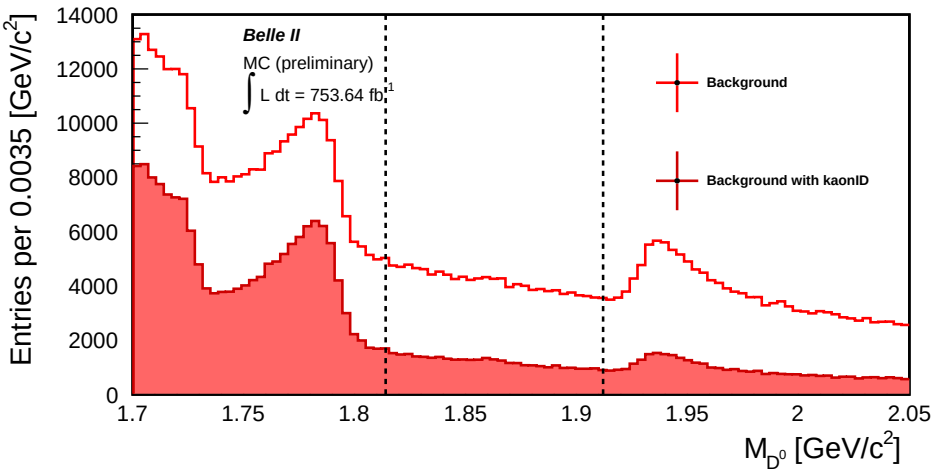


Figure 4.8: Comparison between two background distributions with and without the cut on kaonID.

The vertical dashed line represent the cut values on the reconstructed invariant mass of the D^0. The selected region has a flat background composed mainly by false kaons. Cutting on the kaonID variable cause a sensible reduction of the background distribution, as visible in Figure 4.8. The cut value is set to equal the signal efficiency of the cut on the same variable in the $D^0 \to K^-\pi^+\pi^0$ decay. The final result shows a smoother peak above $1.92\,\mathrm{GeV/c^2}$ and the selected mass region has still a flat background.

4.5 Measurement of the data-MC correction factor

To determine the signal yields of the control channel and compute the reconstruction efficiency with Equation 4.4, I carry out a binned maximum-likelihood fit to the D^0 mass distributions. I studied the probability density functions (PDFs) to use in the fit for the signal and background components of both decays, using MC simulations. The models are reported in Table 4.3.

Decay	Signal	Background
$D^0 \to K^-\pi^+\pi^0$	Sum of two Johnson functions	exponential + crystal-ball function
$D^0 \to K^-\pi^+$	Sum of three Johnson functions	1st order polynomial + exponential

Table 4.3: The PDF models used to fit the signals and background of both the selected decays.

Further details on the Johnson and crystal-ball functions are reported in Appendix C.3, in Equation C.1 and C.2. The fit for signal-only events are shown in Figure 4.9, left for $D^0 \to K^-\pi^+\pi^0$ and right for $D^0 \to K^-\pi^+$. Each Johnson function has four parameters: δ, γ, λ and μ. δ sets the asymmetry, γ and λ adjust the width and μ is responsible for the mode value of the distribution. Each one of these parameters is free to float during the fit. Only in the $D^0 \to K^-\pi^+$ case, the μ parameters are linked together; two distributions have the same parameter and the third is obtained using a shift factor. In addition to the Johnson parameters, the relative fractions of each distribution are free to vary. The fit on the signal distributions presented in the first row of Figure 4.9 have respectively 9 and 13 total free parameters and a reduced χ^2 of 1.28 (top-left) and 0.98 (top-right).

The fit for background-only events are shown in Figure 4.9, bottom-left for $D^0 \to K^-\pi^+\pi^0$ and bottom-right for $D^0 \to K^-\pi^+$. The fit in the left panel has a reduced χ^2 of 0.85 and 6 free parameters, corresponding to: the constant value of the exponential, the 4 parameters of the crystal-ball and the relative fraction between the two components. The crystal-ball is used to describe the small peak of the background in the signal region. The fit in the right panel has a reduced χ^2 of 1.07 and 4 free parameters, corresponding to: the two parameters of the straight line, the slope of the exponential and the relative fraction between the two components. The exponential describes the rising of the background at low mass values, as shown in the bottom-right panel of Figure 4.9. The fit to the full simulated sample and to the data of the $D^0 \to K^-\pi^+\pi^0$ and $D^0 \to K^-\pi^+$ samples are shown in Figures 4.10. In each plot, the lower panel presents the pull, which are

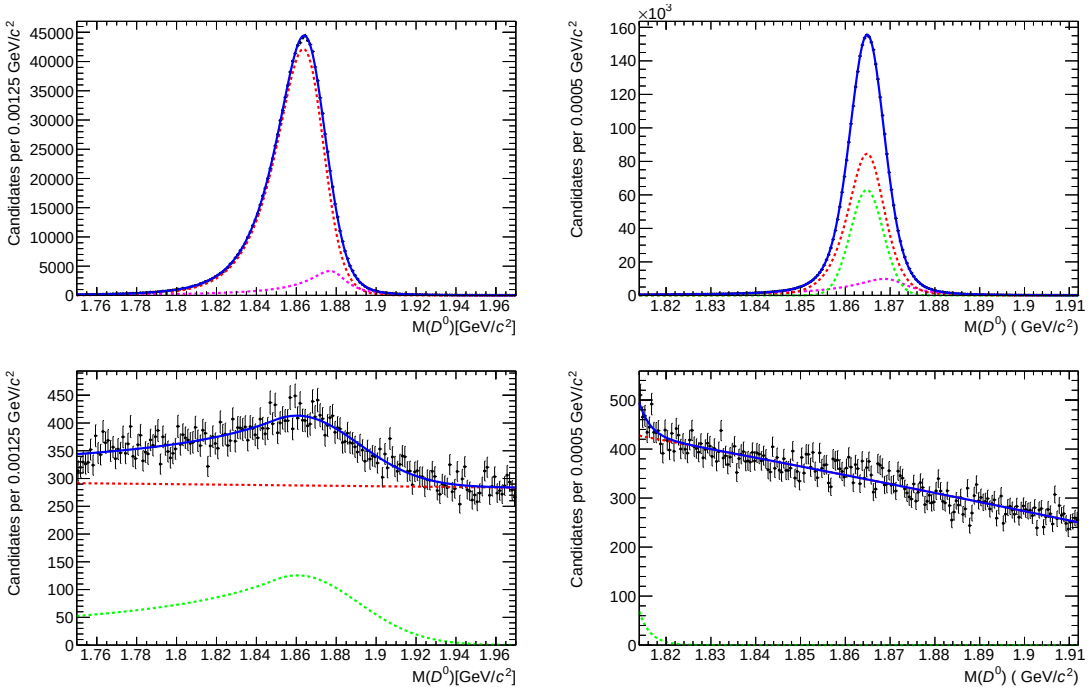


Figure 4.9: PDFs for the signal components, defined as correctly reconstructed D^0 candidate in MC simulations, for the $D^0 \to K^-\pi^+\pi^0$ decay (top-left), and the $D^0 \to K^-\pi^+$ decay (top-right). PDFs for the background components in MC simulations for the $D^0 \to K^-\pi^+\pi^0$ decay (bottom-left), and the $D^0 \to K^-\pi^+$ decay (bottom-right).

evaluated as:

$$\frac{x_i - f_{av}(x_i)}{\sigma(x_i)} \tag{4.7}$$

where x_i are the entries in the i-th bin, $f_{av}(x_i)$ is the the average between two value of the fit function evaluated at the lower and higher edges of the i-th bin, and $\sigma(x_i)$ is the the uncertainty of bin entries. The χ^2 value of the fit is obtained summing the square values of each pull, which distribution helps to understand the contribute of each bin to the total. These fits have two additional free parameters: the signal and background yields. The fit to experimental data is performed by fixing some shape parameters of the PDF from the value extracted in the simulation. The free parameters, in the $D^0 \to K^-\pi^+\pi^0$ case, are eight, corresponding to: the constant of the exponential (red line in the bottom-right panel of Figure 4.9), γ, λ, μ of the red Johnson in left panel of Figure 4.9, λ, μ of the pink Johnson in left panel of Figure 4.9, signal and background yield. For the $D^0 \to K^-\pi^+$ decay the free parameters are: γ, λ, μ (the same of the green Johnson in the right panel of Figure 4.9) of the red Johnson, λ of the green Johnson λ, μ of the pink Johnson both in the top-right panel of Figure 4.9, signal and background yield. The yields determined from the fit are in agreement with the true numbers of the MC, showing that the information extracted from the fit is unbiased. The results are summarized in Table 4.4.

Using Equation 4.4 and the signal yields obtained from the fit, the reconstruction efficiency in simulation and data is found to be:

$$\varepsilon_{\pi^0}^{MC} = 0.0984 \pm 0.0001, \quad \varepsilon_{\pi^0}^{data} = 0.1006 \pm 0.0003\,(\text{stat}) \pm 0.0356\,(\text{BR})\,. \tag{4.8}$$

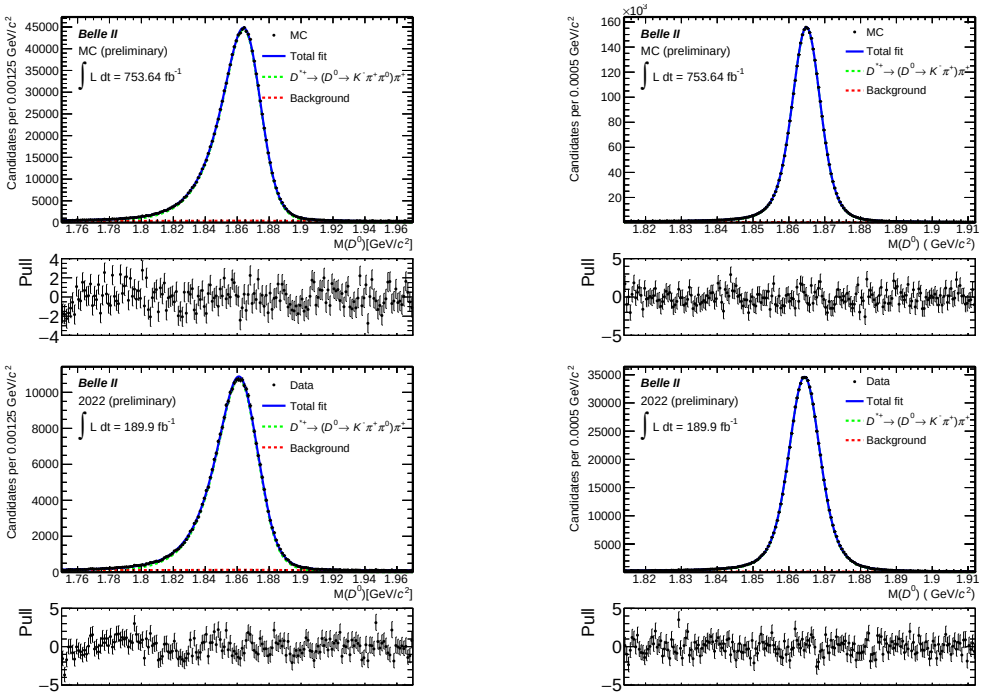


Figure 4.10: The top row shows the total fit to MC simulation (signal+background), while the bottom row presents the fit to experimental data. The pull distributions are showed in the lower panels. The $D^0 \to K^-\pi^+\pi^0$ decays is shown on the left side, the $D^0 \to K^-\pi^+$ decay is shown on the right side.

	Size	$B/(B+S)$	Signal yield	Background yield	Reduced χ^2
MC fit	$754\,\text{fb}^{-1}$	0.0479 ± 0.0006	$1'226'000 \pm 1'300$	$61'723 \pm 740$	1.31
MC truth-info		0.0479	$1'226'064$	$61'645$	–
Data	$190\,\text{fb}^{-1}$	0.0603 ± 0.0011	$311'130 \pm 640$	$19'958 \pm 346$	1.17
MC fit	$754\,\text{fb}^{-1}$	0.0193 ± 0.0003	$3'457'500 \pm 2'080$	$68'189 \pm 1'010$	0.96
MC truth-info		0.0190	$3'458'568$	$67'089$	–
Data	$190\,\text{fb}^{-1}$	0.0237 ± 0.0009	$858'120 \pm 1'220$	$20'801 \pm 800$	0.94

Table 4.4: Results of the fit for the $D^0 \to K^-\pi^+\pi^0$ (first three lines) and $D^0 \to K^-\pi^+$ (last three lines) decays reported in Figure 4.10. The MC truth-info is the real number of signal and background events in the simulated sample.

The term $0.0356\,(\text{BR})$ takes into account the relative uncertainty on the ratio of the branching fractions (see Equation 4.5), that is not considered in the MC efficiency, since the value in the simulation is exactly known. Thus, only the data efficiency is affected both from the uncertainties on the signal yields and the errors on the branching fraction. The correction factor is determined from the ratio of these two values:

$$\frac{\varepsilon_{\pi^0}^{data}}{\varepsilon_{\pi^0}^{MC}} = 1.023 \pm 0.003\,(\text{stat}) \pm 0.036(\text{BR})\,. \tag{4.9}$$

The ratio is compatible with 1 meaning that simulation reproduce fairly well the data at this level of precision. The associated uncertainty considers only the statistical component and that from the uncertainty on the charm-decay branching fractions, which is the dominant, irreducible contribution. The systematic uncertainties will be described in Chapter 6.

Chapter 5

Sample-dependence of the efficiency correction

The factor to correct the π^0-reconstruction efficiency determined from simulation has been calculated for π^0 mesons with momentum larger than 1.5 GeV/c. In this Chapter, I extend the study to investigate possible dependence of the correction factor on the region of the electromagnetic calorimeter and on the kinematics of the π^0 mesons.

5.1 Motivation to study the sample dependence

The study of the correction factor dependencies is interesting even if none is identified. In this case the method explained in Section 4.2 in the previous Chapter, is applicable without further adjustments to all those decays with a π^0 in the final, with kinematics characteristics replicable by the control sample $D^0 \to K^-\pi^+\pi^0$. This scenario is the simplest one and enhance the generality of D^0-meson method. Supposing that the efficiencies ratio has at least one dependency, *i.e.* from the momentum of the π^0, corrections in the measurement of the efficiencies should be applied. Indeed, assuming that the π^0 fractions in momentum bin in the control sample are different with respect to the signal decay, the mean correction factor is also different. This discrepancy bias the reconstruction efficiency measurement, unless the dependence from the sample is taken into account. The aim of this analysis is to develop the study presented in Chapter 4, delving into possible dependencies of the efficiencies ratio reported in Equation 4.9.
The possible dependencies of the correction factor studied in this chapter are: from the region of the electromagnetic calorimeter, from the momentum and direction [1] of the neutral pion. These three variables have been chosen because, as shown in Figure 2.9, the material budget varies among barrel and the endcaps; moreover the shape of the electromagnetic shower change with both momentum and direction of the incoming particle. This features may not be well reproduced in the simulation.

[1] the direction is studied in $\cos(\theta)$ bin, with θ the polar angle

5.2 Correction for different ECL regions

Using the same selections presented in the previous Chapter, the reconstruction efficiency in data and simulation and their ratio are calculated in each region of the electromagnetic calorimeter. The sample is divided in five different subsets according to the region of the ECL that detected the photon daughter from the π^0 decay: forward, forward-barrel, barrel, barrel-backward and backward. The fraction of events in each region in MC and data is reported in Table 5.1.

ECL region	MC event fraction	Data event fraction
Forward	4.5%	4.5%
Forward-Barrel	3.6%	3.4%
Barrel	89.3%	90.2%
Barrel-Backward	1.0%	0.8%
Backward	1.6%	1.2%

Table 5.1: The event fraction in MC and data in each region of the electromagnetic calorimeter.

Each region could feature a different energy-resolution, leading to different shape on the signal D^0 peak, and different background composition. The signal and background shapes are compared in simulation to check if the models used to fit the mass distribution are still able to describe them. Figure 5.1 shows six histograms filled with signal only events normalized to the same unit area. The red line shows the overall signal shape compared with the signals in the different ECL regions. The distributions have three different shapes. The barrel and the overall signal distributions are in good agreement, so as the forward and forward-barrel, barrel-backward and backward distributions. To describe the signal shapes, I used the same functions reported in Section 4.5, but the shapes and the value of the free parameters are fixed performing a fit on three different distributions. For the barrel region, I use the signal shape with parameters fixed from the fit of the total signal sample; for the forward and forward-barrel regions, I use the same PDF with parameters determined from the fit to simulation of the forward region (which feature higher statistic than forward-barrel); similarly for the backward and barrel-backward region, I use the simulated candidates in the backward region to determine the PDF parameters.
A comparison between background distributions (normalized to the same unit area) of the six regions from simulation is shown in Figure 5.1. The shapes look compatible within statistical uncertainties, thus a single model, determined from the total sample in simulation, is used.

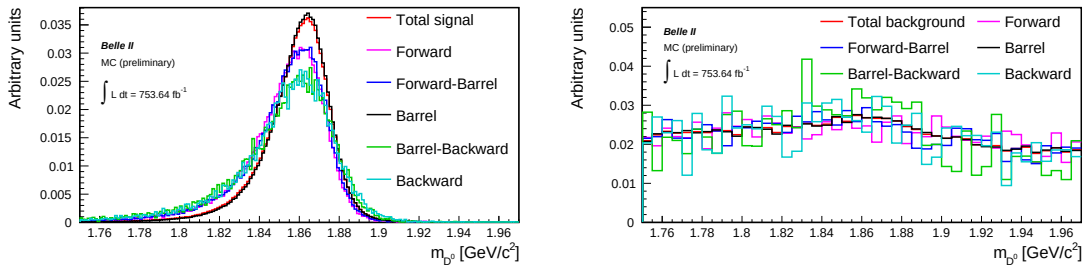


Figure 5.1: Normalized histograms referring to signal (left) and background (right) only events for each ECL region compared with the total distributions.

After determining the PDF model to each in each region, I fit the each sample of data and full simulation. The fit are presented in Figure 5.2. In each lower panel are shown the corresponding pull distributions. The extracted signal yields, the background fractions obtained from the fit and the normalized χ^2 values for both simulation and data are reported in Table 5.2.

ECL region	Signal yields		$\frac{B}{B+S}$		Fit χ^2	
	MC	Data	MC	Data	MC	Data
For.	$54'545 \pm 247$	$13'471 \pm 126$	0.0580 ± 0.0018	0.0892 ± 0.0043	1.10	1.24
For.-Bar.	$43'308 \pm 222$	$10'166 \pm 111$	0.0718 ± 0.0022	0.0981 ± 0.0053	1.43	1.35
Bar.	$1'098'600 \pm 1'220$	$282'840 \pm 620$	0.0447 ± 0.0006	0.0527 ± 0.0012	1.27	1.24
Bar.-Back.	$12'498 \pm 121$	$2'283 \pm 56$	0.0554 ± 0.0042	0.1175 ± 0.0138	1.39	0.95
Back.	$18'933 \pm 149$	$3'466 \pm 67$	0.0519 ± 0.0033	0.1014 ± 0.0103	1.20	0.87

Table 5.2: The signal yields (Y_s), the background fraction, and the normalized χ^2 of the fit, obtained for MC simulation and data in each ECL region.

The background fractions have some discrepancies between MC and data, especially in the forward, barrel-backward and backward region. An aspect to be considered here are the beam backgrounds that interest mainly the endcap regions, *i.e.* low-energy photons from processes difficult to be fully simulated.

The π^0 reconstruction efficiency is determined using the signal yields extracted from the fit as per Equation 4.4. The results for the simulation, the experimental data and their ratios that determines the correction factors are summarized in Table 5.3. The correction factors of the reconstructed efficiencies are also presented in Figure 5.3.

ECL region	ε_{MC}	ε_{data}	$\frac{\varepsilon_{data}}{\varepsilon_{MC}}$
Forward	0.00438 ± 0.00002	0.00436 ± 0.00004	0.995 ± 0.010
Forward-Barrel	0.00348 ± 0.00002	0.00329 ± 0.00004	0.946 ± 0.011
Barrel	0.0882 ± 0.0001	0.0915 ± 0.0002	1.037 ± 0.003
Barrel-Backward	0.00100 ± 0.00001	0.00074 ± 0.00002	0.736 ± 0.019
Backward	0.00152 ± 0.00001	0.00112 ± 0.00002	0.738 ± 0.016

Table 5.3: Reconstruction efficiencies for MC simulation, for data, and their ratio in each of the ECL regions.

The uncertainty reported on the correction factor accounts only for the statistical uncertainty of the $K^-\pi^+\pi^0$ sample, because the statistical uncertainty of the $K^-\pi^+$ sample and that on the charm-decay branching fractions (see Equation 4.4) are global uncertainty common (*i.e.* fully correlated) for each region. They are not relevant for a relative comparison between the regions. The correction factor of the forward, forward-barrel and barrel regions fluctuates between 5% around unity. Considering the common global uncertainties from the charm-decay branching fractions ratio (and that on the the $K^-\pi^+$ sample, although sub-leading), all three values are compatible with 1 within 2σ. The last two regions, barrel-backward and backward, present instead big discrepancies between MC and data. The ratios are significantly smaller then 1, thus the simulation overestimates the reconstruction efficiency.

To further investigate these differences, I compare the shapes of photon-energy spectra (normalised to same unit area) in data and simulation, so to see if the data-simulation discrepancies

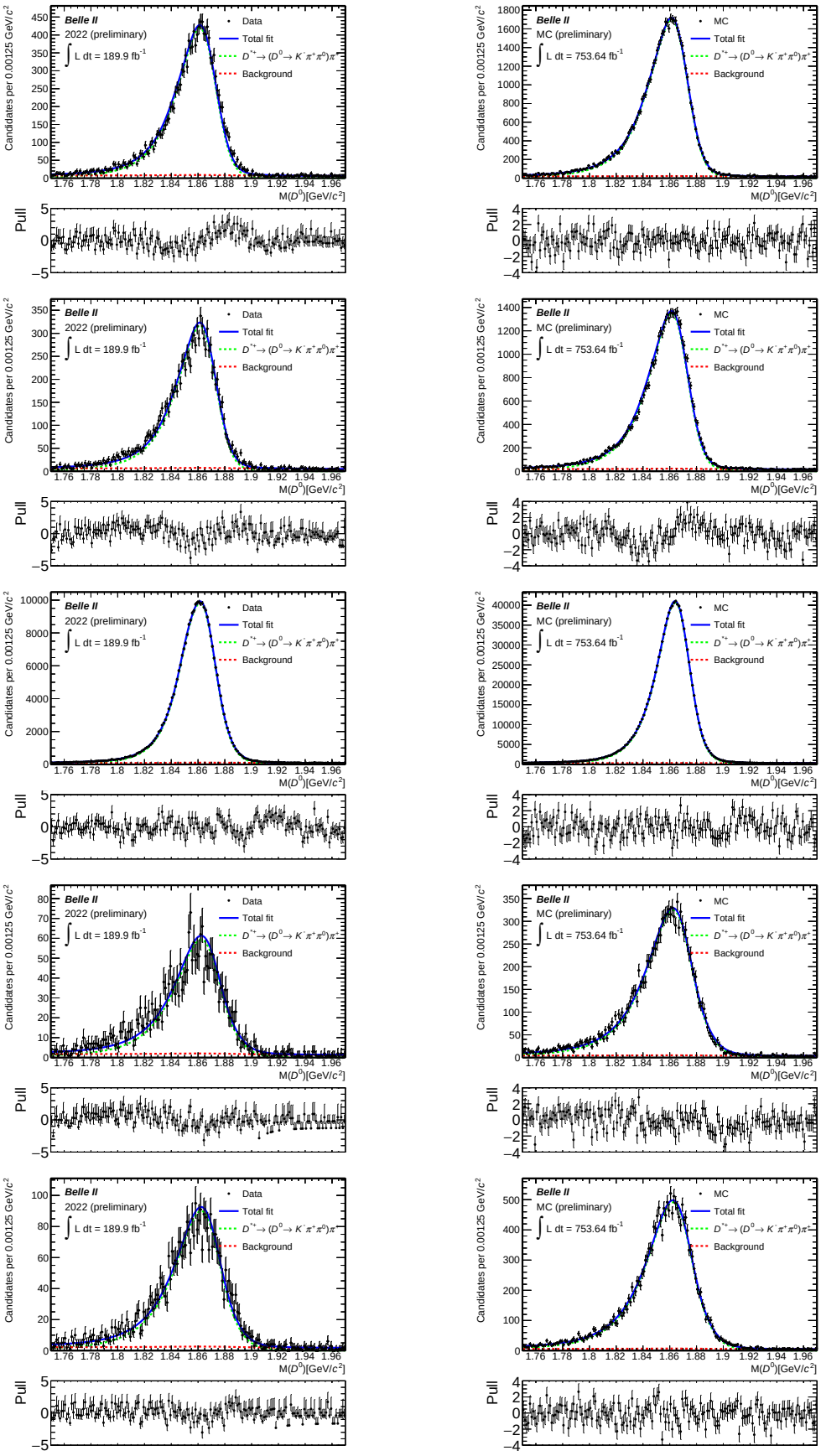


Figure 5.2: Fit of the D^0 mass distributions on data (left) and MC (right) for candidates of the different ECL regions: (first row) forward, (second row) forward-barrel, (third row) barrel, (fourth row) barrel-backward, and (fifth row) backward. The lower panels show the pull distributions.

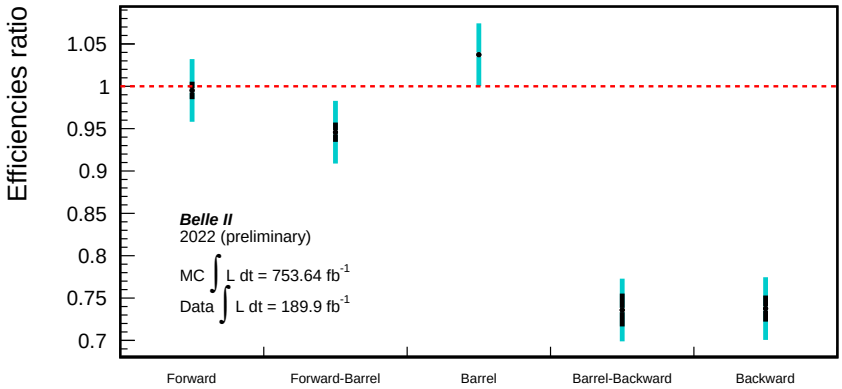


Figure 5.3: Ratio of the π^0 reconstruction efficiencies of data and MC simulation in five different ECL regions. The unit value is shown as a red line. The dark bars are the statistical uncertainties associated with the $D^0 \to K^-\pi^+\pi^0$ sample; the light blue lines are the total (statistical and systematic) uncertainty. The systematic uncertainties are discussed in Chapter 6.

are due to wrongly simulated energy distributions. In simulation, I also divide the signal (green-filled histogram) and the background (red-filled histogram) components, which are scaled by the total to keep their relative proportions. The plot relative to the forward, forward-barrel, barrel, barrel-backward and backward region are presented respectively in Figure 5.4. In the lower panel

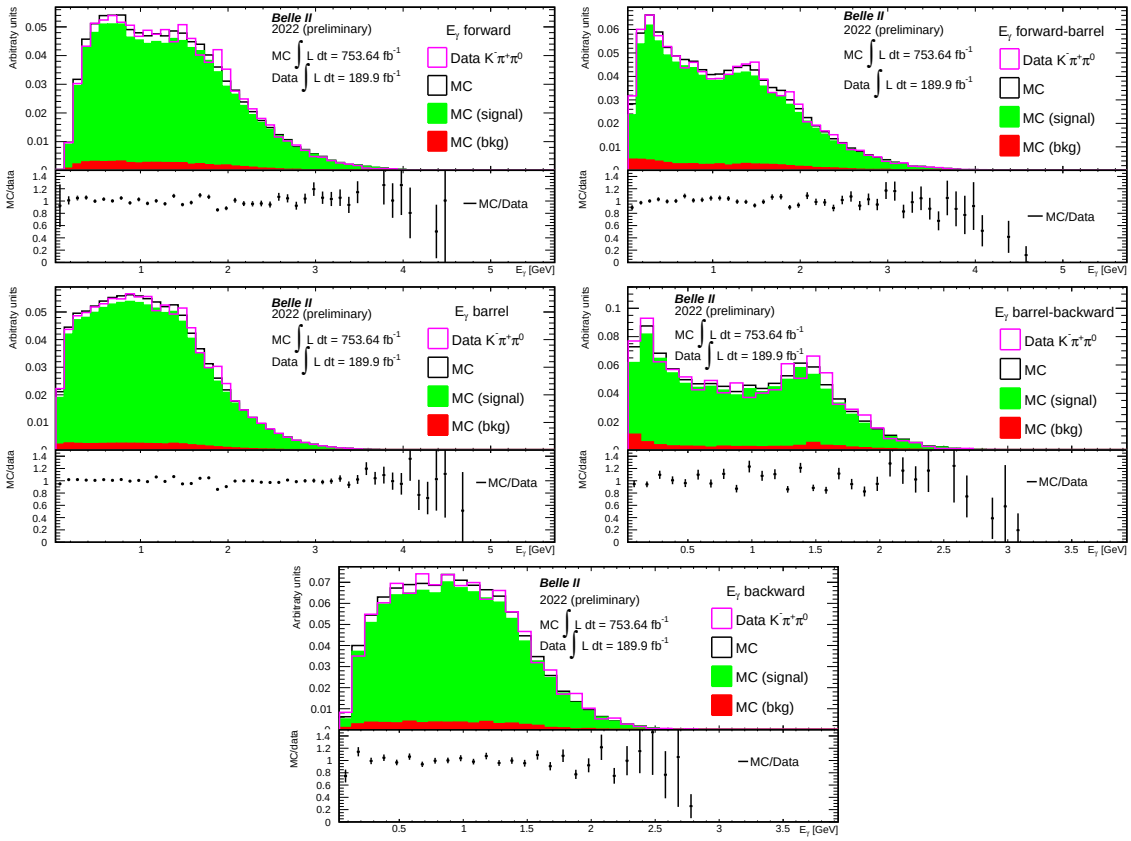


Figure 5.4: Photons energy spectra in each region of the ECL. The data (pink) and MC (dark) distributions are normalized to the same area. The signal (green) and background (red) are normalized by the integral of the MC histogram. In the lower panel the ratio of the MC distribution above the data is presented.

of each plot, the ratio of the MC over the data distribution is shown. If the distribution shape is well reproduced in the simulation, the ratio should be flat and around unity.
The overall agreement is quite good, although large fluctuations in the barrel-backward region, that corresponds to the sample with the smallest statistics. This indicates that simulation reproduces pretty well the energy spectrum observed in data, but fail to model the overall normalisation in the backward and backward-barrel regions, where the correction factors are significantly different from unity. This could be explained by a simplified description of the material before the calorimeter (see Figure 2.9) in simulation. Indeed, the forward and backward region of the ECL have the largest thickness of material composed mostly by the readout electronics.

5.3 Dependence on the π^0 direction

To study the dependence of the reconstruction-efficiency ratio on the π^0 direction, I split the sample in four subsets considering the cosine of the polar angle of the π^0 candidate, $\cos(\theta)$. To ensure enough statistics for this study, only photons hitting the calorimeter in the barrel region have been considered. The requirement on $\cos(\theta)$ are chosen in order to have roughly the same number of signal events in each subsets. Figure 5.5 report the $cos(\theta)$ distribution for data and simulation.

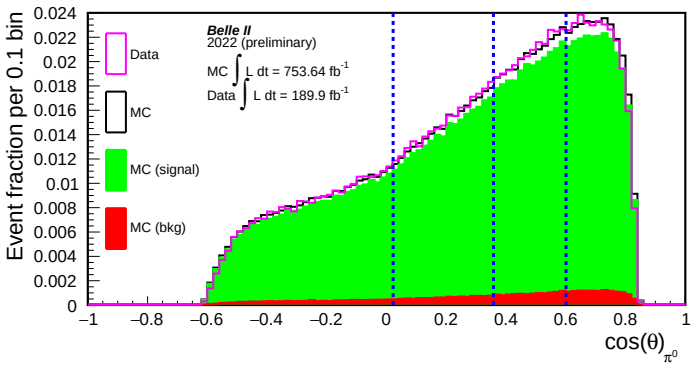


Figure 5.5: Normalized distribution of $\cos(\theta)_{\pi^0}$*, divided in four different intervals (blue lines), for data (pink) and MC (dark). The signal (green) and background (red) are scaled by the integral of MC histogram.*

Similarly to what I have done for the study in the ECL regions, I first investigate in simulation any possible variations of the shape of the signal and background distributions of the D^0-mass between the different $\cos\theta$ intervals, so to adjust the model to determine the signal yields. The normalized distributions of the signal events in each bin is plotted in Figure 5.6. No differences in the signal distributions are visible. To describe the signal shape, the same model previously described for the barrel region in Section 5.2 is used. The same comparison is done with the background distributions in each $cos(\theta)$ interval, as shown in Figure 5.6. Shape differences are still not present, so I use a the same PDF for all subsets.
Once signal and background models are determined, I fit the data and full simulation to determine the signal yields. The results are reported in Figure 5.7. In the lower panel of each plot

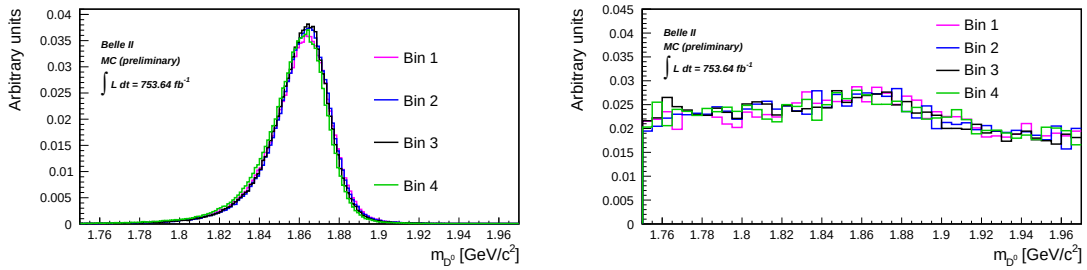


Figure 5.6: Normalized histograms referring to signal (left) and background (right) events for each $\cos(\theta)$ *interval:* $[-1.000, 0.022]$, $[0.022, 0.358]$, $[0.358, 0.601]$, $[0.601, 1.000]$. *Only photons hitting the barrel are considered.*

the pull distributions are presented. A summary of the information extracted through the fit, regarding the signal yields, the normalized χ^2 values and the background fractions is reported in Table 5.4.

$\cos(\theta)$ bin	Signal yields		$\frac{B}{B+S}$		Fit χ^2	
	MC	Data	MC	Data	MC	Data
$[-1.00, 0.022]$	$275'150 \pm 550$	$70'160 \pm 280$	0.0381 ± 0.0007	0.0481 ± 0.0015	1.32	0.89
$[0.022, 0.358]$	$274'830 \pm 550$	$71'510 \pm 282$	0.0418 ± 0.0007	0.0499 ± 0.0015	1.21	0.98
$[0.358, 0.601]$	$273'840 \pm 548$	$71'202 \pm 282$	0.0462 ± 0.0007	0.0543 ± 0.0015	1.41	1.09
$[0.601, 1.000]$	$274'770 \pm 553$	$69'735 \pm 282$	0.0525 ± 0.0008	0.0618 ± 0.0017	1.25	1.06

Table 5.4: The extracted signal yields, the background fractions, and the normalized χ^2 *values for both MC and data in each* $\cos(\theta_{\pi^0})$ *bin. Only photons detected by the barrel are considered.*

The π^0 reconstruction efficiencies in each $\cos(\theta)$ interval are measured using the signal yields from the fit. The results regarding both MC and data and their ratio are presented in Table 5.5. The uncertainty on the ratio accounts only for the statistical uncertainty of the $K^-\pi^+\pi^0$ signal yields, being the other two contributions (uncertainty on the charm-decay branching fractions and on the $D^0 \to K^+\pi^-$ signal yields) common to the measurement in each interval. We observe a dependence of the correction factor, as the agreement between data and simulation is better in the first and in the fourth bin, while the central region features a larger data-simulation discrepancy (around 5%). There is no clear explanation of this trend.
Including all uncertainties, the correction factor in the central region are compatible with unity within 2σ. The measured ratio in each $\cos(\theta)$ bin are shown in Figure 5.8.

$\cos(\theta)$	ε_{MC}	ε_{data}	$\frac{\varepsilon_{data}}{\varepsilon_{MC}}$
$[-1.0, 0.022]$	0.02206 ± 0.00004	0.02269 ± 0.00009	1.027 ± 0.005
$[0.022, 0.358]$	0.02206 ± 0.00004	0.02313 ± 0.00009	1.048 ± 0.005
$[0.358, 0.601]$	0.02198 ± 0.00004	0.02303 ± 0.00009	1.048 ± 0.005
$[0.601, 1.000]$	0.02206 ± 0.00004	0.02256 ± 0.00009	1.023 ± 0.005

Table 5.5: The π^0 *reconstruction efficiencies obtained from MC, data and the data over MC ratio in the four* $\cos(\theta)$ *bin.*

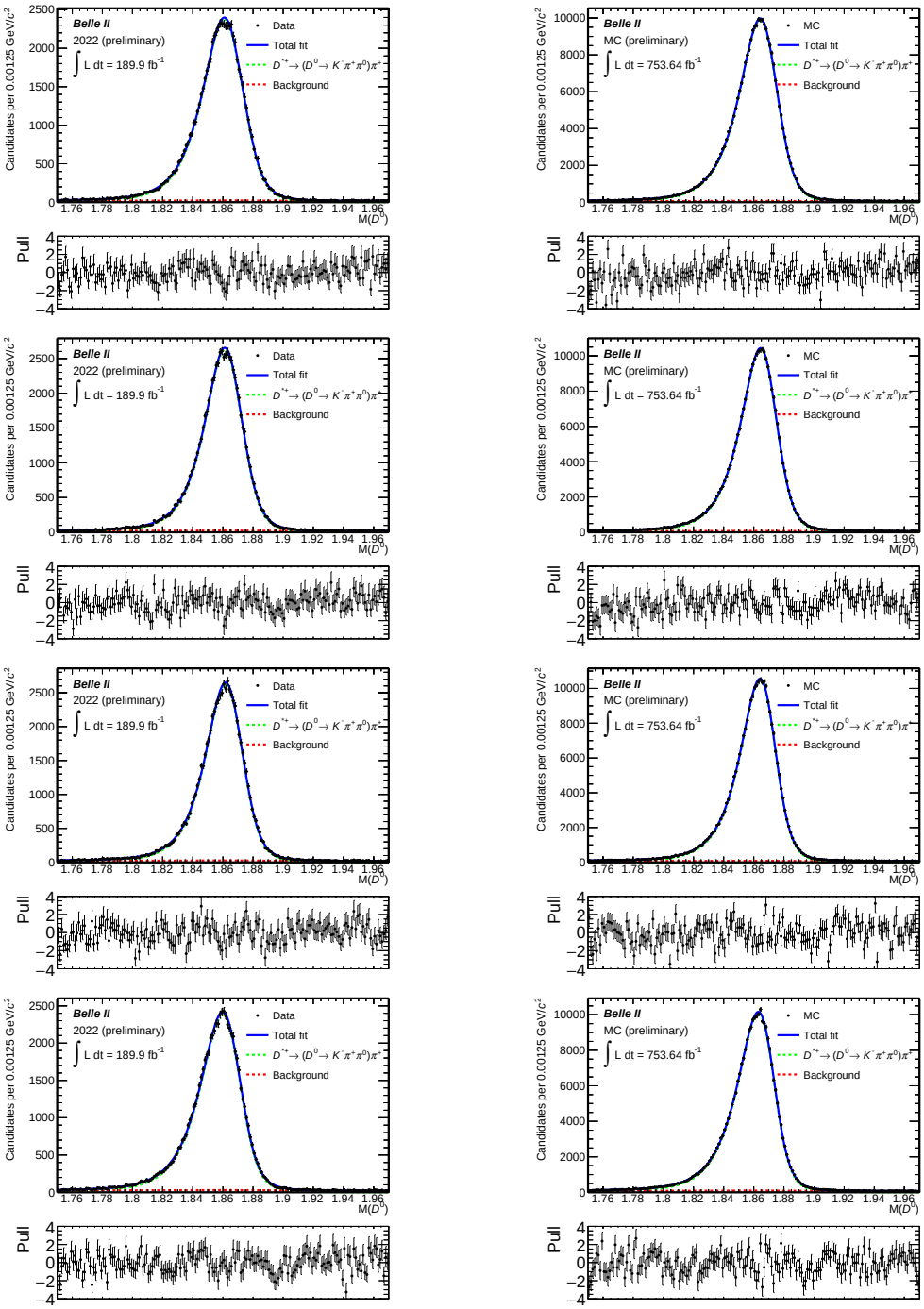


Figure 5.7: Fit on data (left) and MC (right) for each $\cos(\theta)$ bin. First row $[-1, 0.022]$, second row $[0.022, 0.358]$, third row $[0.358, 0.601]$ and fourth row $[0.601, 1]$. The corresponding pull distributions are shown in the lower panels. Only photons detected by the barrel are considered.

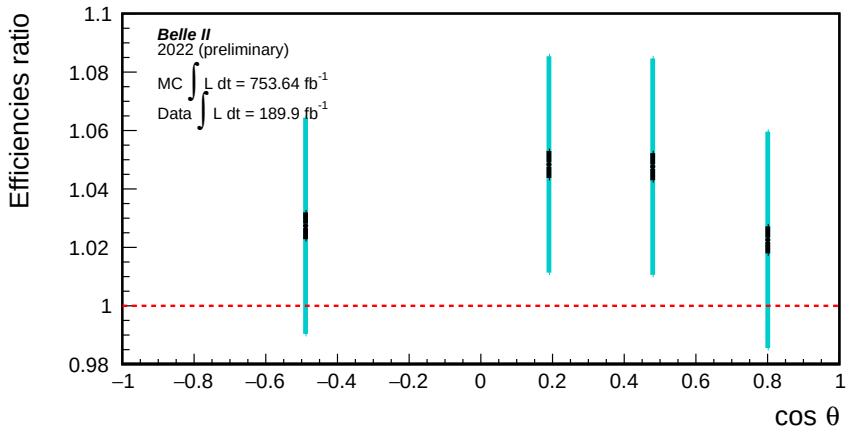


Figure 5.8: Ratios of the reconstruction efficiencies (data over MC) in four bin of $\cos(\theta)_{\pi^0}$*:* $[-1.000, 0.022]$*,* $[0.022, 0.358]$*,* $[0.358, 0.601]$ *and* $[0.601, 1.000]$*. The ideal value is shown as a red line, the dark lines are the statistical error associated with the* $D^0 \to K^-\pi^+\pi^0$ *sample, the light blue lines are the total (statistical and systematic) uncertainty. The systematic uncertainties are discussed in Chapter 6.*

5.4 Dependence on the π^0 momentum

The dependence of the reconstruction-efficiencies ratios on the π^0 momentum is investigated by splitting the sample in six subsets: I considered three intervals in the region above the requirement $p(\pi^0) > 1.5\,\mathrm{GeV/c^2}$ used in the $B^0 \to \pi^0\pi^0$ analysis, and I extend the study also in the region below this cut, considering three intervals in momentum. The distribution of the π^0 momentum in data and simulation are presented in Figure 5.9, where the momentum interval considered to split the samples are also reported with vertical lines. The width of the intervals are chosen such to consider approximately the same number of events in each subsets of the three bins above and below $1.5\,\mathrm{GeV/c^2}$.

The expected signal shapes in simulation for each momentum interval are shown in Figure 5.10. I observe that the signal gets narrower as the π^0 momentum decreases. This is expected from the ECL energy-resolution, and must be taken into account in the PDF of the signal with different width parameters in each momentum range.

By extending the region to lower momentum value, I need to perform an additional study of

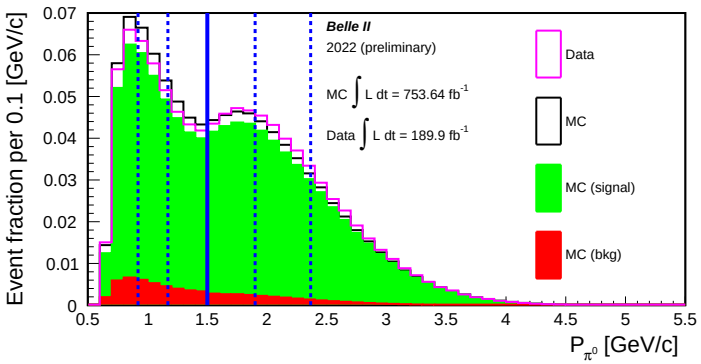


Figure 5.9: Normalized distribution of p_{π^0}*, divided in six different bin (blue lines), for data (pink) and MC (dark). The signal (green) and background (red) are scaled by the integral of MC histogram.*

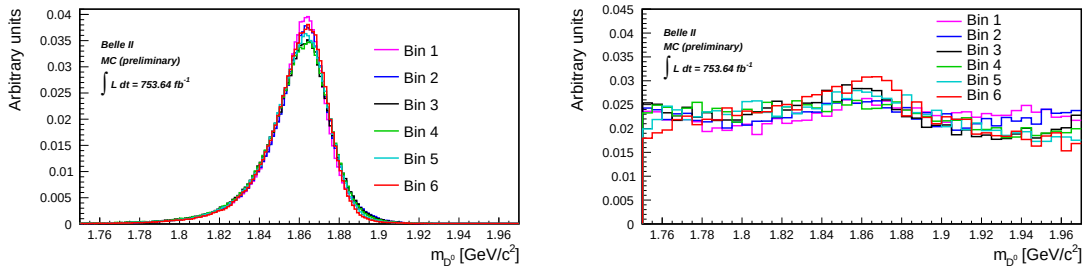


Figure 5.10: Normalized histograms referring to signal (left) and background only events for each π^0 momentum bin: $[0.60, 0.92]$, $[0.92, 1.17]$, $[1.17, 1.50]$, $[1.50, 1.90]$, $[1.900, 2.365]$, $[2.365, 5.000]$. *The signal and background shapes are obtained fitting the total distributions without the cut on π^0 momentum.*

the background distribution. By studying the background in simulation, I find that, in this low-momentum region, the small component of decay-in-flight of the pion candidates in $D^0 \to K^+\pi^-$ decays, causing the mis-identification of the pion with a muon, is negligible. In addition, the relative fractions of the background components (see study in Section 4.3.1) change. This is reflected in different background shapes for the six subset, as shown in Figure 5.10.

However, I use a single model for each bin, because the width of the distribution is kept as free parameter whilst performing the fit both on MC and data. This choice is possible thanks to high statistic of the sample and the cleanliness of the peaks. The shape of the background is obtained through a fit on the total background distribution without the cut on π^0's momentum. I will consider a systematic uncertainty for the fit modelling in Section 6.2.

The fit to the data and full simulation are shown in Figures 5.12 and 5.13. A summary of the signal yields, the normalized χ^2 values and the background fractions is presented in Table 5.6. The agreement between the background fraction expected from the MC and the fraction measured in data is quite good, with differences always smaller than 1%.

The π^0 reconstruction efficiencies in each $\cos(\theta)$ interval are measured using the signal yields from the fit. The results regarding both MC and data and their ratio are presented in Table 5.5. The uncertainty on the ratio accounts only for the statistical uncertainty of the $K^-\pi^+\pi^0$ signal yields, being the other two contributions (uncertainty on the charm-decay branching fractions and on the $D^0 \to K^+\pi^-$ signal yields) common to the measurement in each interval. We observe a dependence of the correction factor, as the agreement between data and simulation is better in the first and in the fourth bin, while the central region features a larger data-simulation discrepancy

p_{π^0} bin [GeV/c]	Signal yields		$\frac{B}{B+S}$		Fit χ^2	
	MC	Data	MC	Data	MC	Data
$[0.600, 0.920]$	$322'180 \pm 620$	$77'835 \pm 311$	0.1206 ± 0.0010	0.1299 ± 0.0021	2.01	1.21
$[0.920, 1.170]$	$324'220 \pm 612$	$77'624 \pm 310$	0.0944 ± 0.0009	0.0957 ± 0.0020	2.44	0.75
$[1.170, 1.500]$	$333'350 \pm 618$	$80'429 \pm 306$	0.0775 ± 0.0008	0.0747 ± 0.0017	1.37	1.46
$[1.500, 1.900]$	$408'980 \pm 725$	$104'810 \pm 357$	0.0561 ± 0.0009	0.0647 ± 0.0017	1.30	0.91
$[1.900, 2.365]$	$409'190 \pm 702$	$101'690 \pm 357$	0.0457 ± 0.0008	0.0604 ± 0.0016	1.37	1.20
$[2.365, 5.000]$	$411'260 \pm 670$	$106'120 \pm 346$	0.0339 ± 0.0006	0.0424 ± 0.0012	1.68	1.29

Table 5.6: The extracted signal yields, the background fractions, and the normalized χ^2 values for both MC and data in each p_{π^0} bin.

p_{π^0} [GeV/c]	ε_{MC}	ε_{Data}	$\frac{\varepsilon_{Data}}{\varepsilon_{MC}}$
$[0.600, 0.920]$	0.02587 ± 0.00005	0.0252 ± 0.0001	0.973 ± 0.004
$[0.920, 1.170]$	0.02603 ± 0.00005	0.0251 ± 0.0001	0.965 ± 0.004
$[1.170, 1.500]$	0.02676 ± 0.00005	0.0260 ± 0.0001	0.972 ± 0.004
$[1.500, 1.900]$	0.03283 ± 0.00006	0.0328 ± 0.0001	1.002 ± 0.004
$[1.900, 2.365]$	0.03285 ± 0.00006	0.0339 ± 0.0001	1.032 ± 0.004
$[2.365, 5.000]$	0.03302 ± 0.00005	0.0343 ± 0.0001	1.040 ± 0.004

Table 5.7: The π^0 reconstruction efficiencies obtained from MC, data and the data over MC ratio in the six π^0 momentum bin.

(around 5%). There is no clear explanation of this trend.

The π^0 reconstruction efficiencies for simulation and data are obtained using the signal yields extracted from the fit. The results along with the correction factors are shown in Table 5.7. The uncertainty considers only the statistical contribution from the $K^-\pi^+\pi^0$ signal yield, being the uncertainties on the charm-decay branching fractions and on the $D^0 \to K^+\pi^-$ signal yields common to the measurement in each interval. The overall agreement between MC simulation and experimental data is good, with discrepancies always smaller than 5%. Anyway, I observe a dependence of the efficiency ratio as a function of the π^0 momentum.

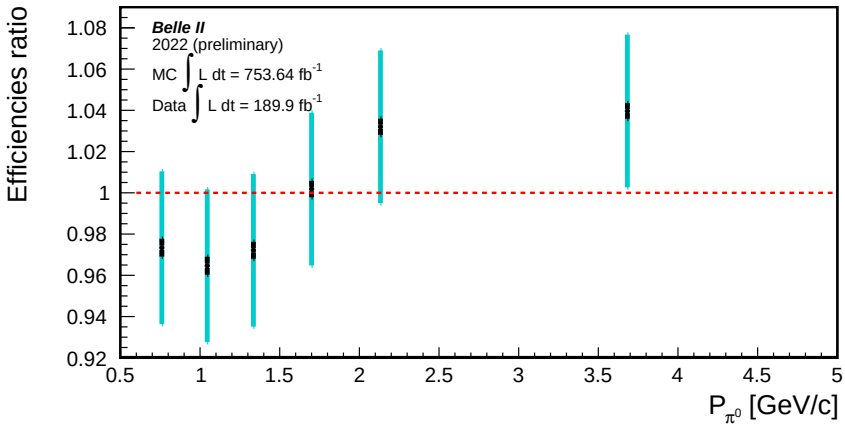


Figure 5.11: Ratios of the π^0-reconstruction efficiencies (data over MC) in six bin of π^0 momentum: $[0.60, 0.92]$, $[0.92, 1.17]$, $[1.17, 1.50]$, $[1.50, 1.90]$, $[1.900, 2.365]$ and $[2.365, 5.000]$. The unit value is shown as a red line. The dark bars are the statistical error associated with the $D^0 \to K^-\pi^+\pi^0$ sample, the light blue bars are the total (statistical and systematic) uncertainty. The systematic uncertainties are discussed in Chapter 6.

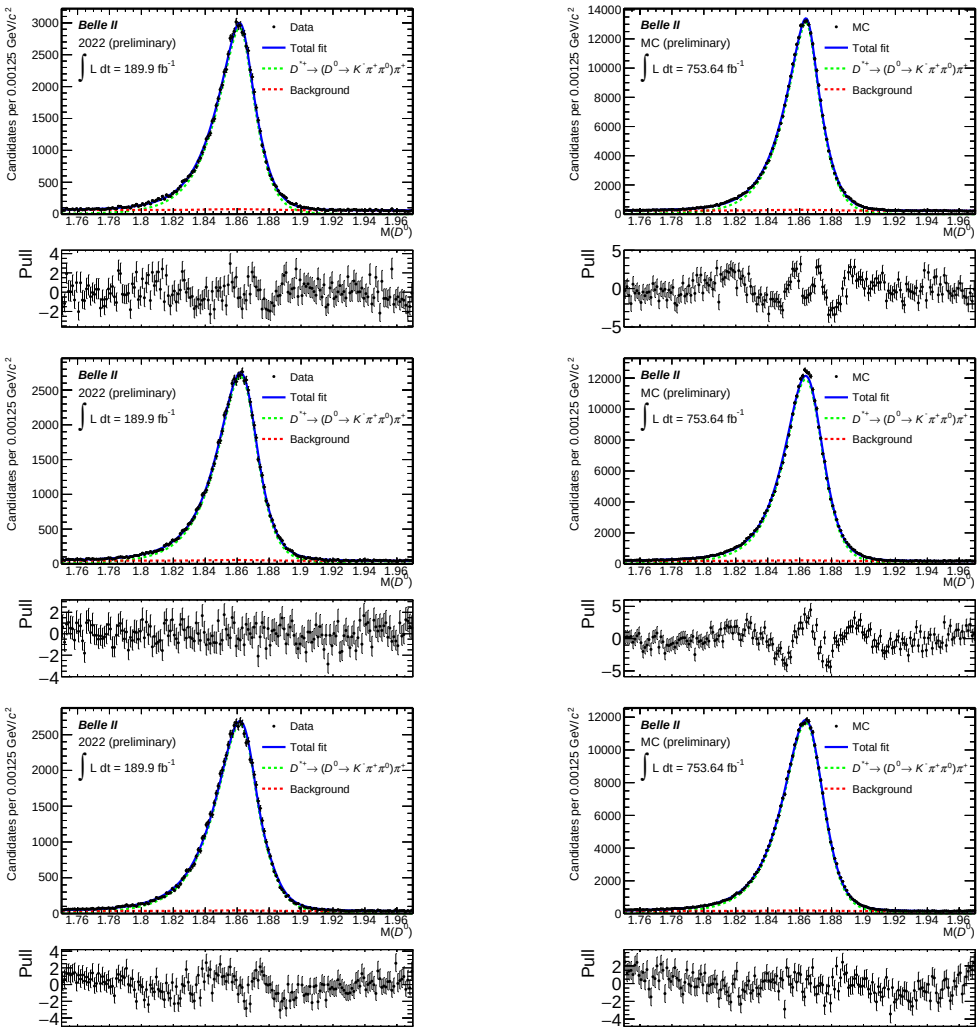


Figure 5.12: Fit on data (left) and MC (right) in the first $[0.60, 0.92]\,\mathrm{GeV/c}$ *(top), second* $[0.92, 1.17]\,\mathrm{GeV/c}$ *(middle) and third momentum bin* $[1.17, 1.50]\,\mathrm{GeV/c}$ *(bottom). The corresponding pull distributions are shown in the lower panels.*

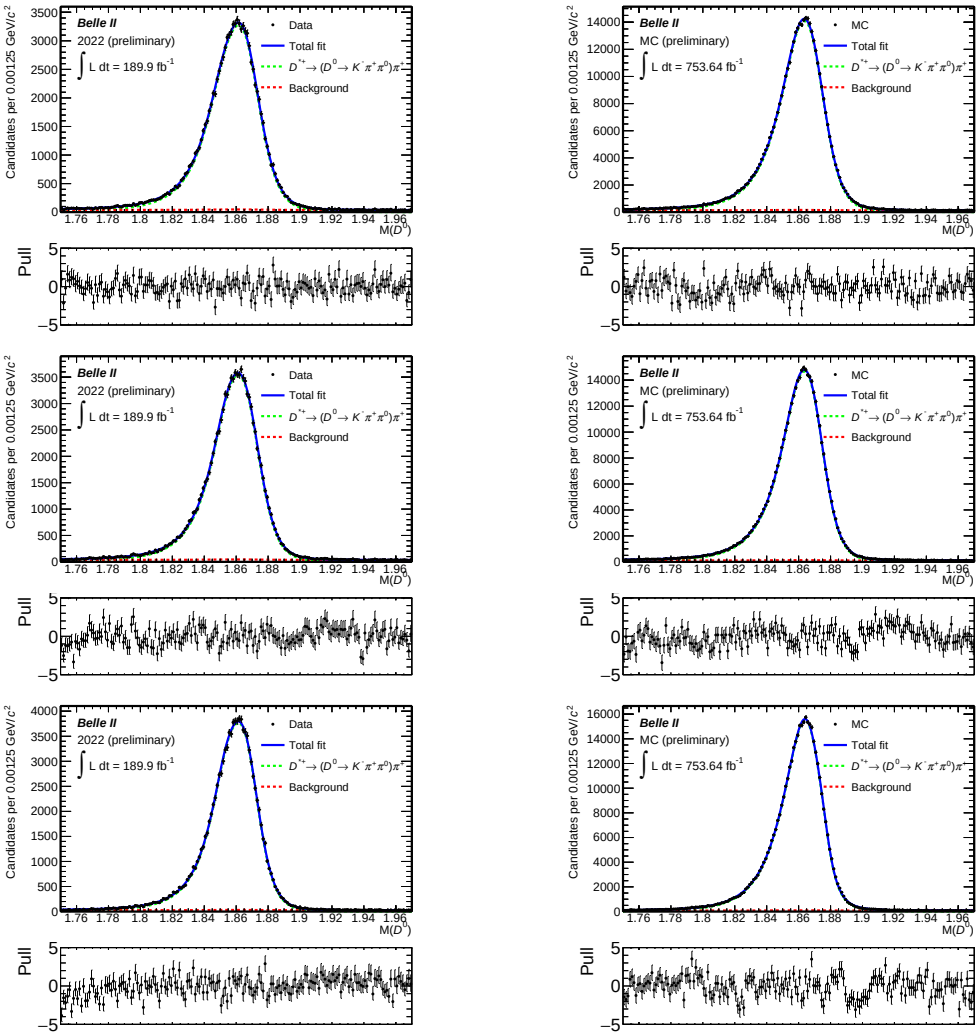


Figure 5.13: Fit on data (left) and MC (right) in the fourth $[1.50, 1.90]\,\mathrm{GeV/c}$ *(top), fifth* $[1.900, 2.365]\,\mathrm{GeV/c}$ *(middle) and sixth momentum bin* $[2.365, 5.000]\,\mathrm{GeV/c}$ *(bottom). The corresponding pull distributions are shown in the lower panels.*

Chapter 6

Systematic uncertainties

In this Chapter, I present the systematic uncertainties associated to the measurement of the data-to-simulation ratio of the π^0 reconstruction efficiency.

6.1 Systematic uncertainties related to common selections

The factor to correct the efficiency in simulation is obtained through the comparison in data and simulation of the π^0 reconstruction efficiency obtained from Equation 4.8. This method is based on the assumption that the efficiency factorises for each particle, and that the terms common to numerator ($D^0 \to K^+\pi^-\pi^0$) and denominator ($D^0 \to K^+\pi^-$), *i.e.* ε_K, ε_π, $\varepsilon_{\pi^+_{slow}}$, ε_{D^*} and ε_{D^0}, cancel in the ratio. The selection of the charged particles and the charm-mesons candidates has been tuned in order to guarantee such a cancellation. This has been done using simulation. In the following, we test any residual discrepancy between data and simulation that could bias the assumption of perfect cancellation between common terms.

6.1.1 Tracking efficiencies

Tracking efficiencies of different particles have been studied by the Belle II tracking working-group [61] and have been parameterised as a function of the kinematics of the particles, namely its momentum and the cosine of the polar angle, $\cos(\theta)$. Correction factors for simulation have been provided and are applied in the reconstruction of the simulated samples. Those corrections are expressed as a function of the momentum and polar angle of each track. So, in the case considered here, those have been applied to K^-, π^+ and π^+_{slow} for both decay channels. To search for possible residual differences between data and simulation, the momentum and the $\cos(\theta)$ distributions of each track are inspected, which are reported in Figures D.1, D.2, and D.3 of Appendix D. For each variable the ratio between the $D^0 \to K^-\pi^+\pi^0$ decay and the $D^0 \to K^-\pi^+$ decay in simulation and data have been considered. Residual data-simulation difference in each single channel can get suppressed in the ratio. Therefore, the ratio between MC and data is checked to control for significant discrepancies from 1. From now on, this procedure will be called "double ratio" for

sake of brevity. If this double ratio departs from unity, a systematic uncertainty is assigned due to data-simulation discrepancy in the tracking efficiency.

The plot of the double ratios as a function of momentum and cosine of the polar angle are shown in Figure 6.1, no cut on the π^0 momentum is applied.

For the kaon track, the agreement between simulation and data is good for the $\cos(\theta)$ distribution, while for the momentum it presents some differences. We quantify the deviation from a flat double-ratio by computing a χ^2 test for a constant-line hypothesis. The χ^2 values normalised to the degree-of-freedom (χ^2/dof) are 1.10 in $\cos\theta$ and 2.8 in momentum. The kaon double ratios show discrepancies from unity of $\mathcal{O}(10\% - 20\%)$, thus a systematic uncertainty is considered. Following recommendation from the Belle II tracking working-group, the systematic uncertainty is set to 0.3% for each kaon track.[62]. We conservatively assume that the uncertainty are not correlated, and consider independent uncertainty for the kaon of the $D^0 \to K^-\pi^+\pi^0$ decay and that of $D^0 \to K^-\pi^+$.

For the pion track, the double-ratio are flat both as a function of momentum and $\cos(\theta)$; the χ^2/dof are 0.95 and 0.9, respectively. No systematic uncertainty is considered in this case.

For the slow-pion, the double-ratio distribution as a function of $\cos(\theta)$ shows large departure from 1 especially for values below -0.4. The χ^2/dof is 1.7. The double-ratio distribution as a function of momentum features some departures too, with $\chi^2/\text{dof} = 1.9$.

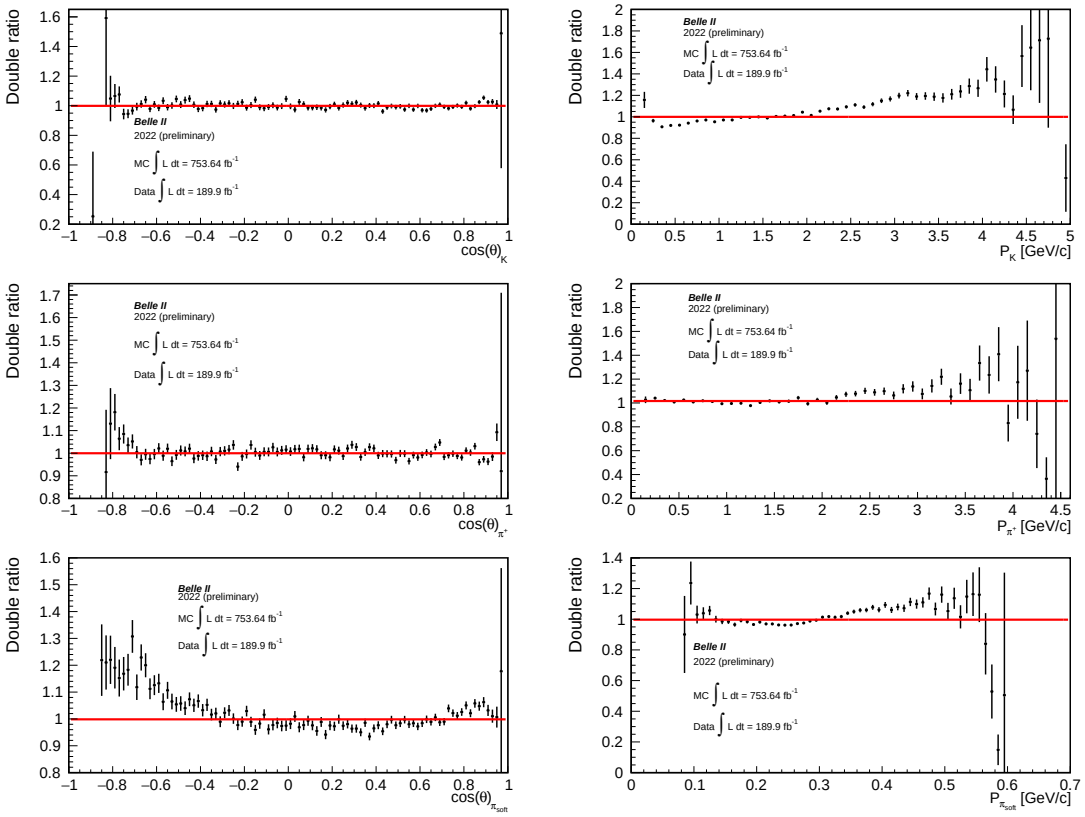


Figure 6.1: Double ratios (MC over data) for the $\cos(\theta)$ (left) and momentum (right) distributions for K^- (top), π^+ (middle) and π^+_{soft} (bottom). The double ratio is obtained taking the ratio between the $D^0 \to K^-\pi^+\pi^0$ decay and the $D^0 \to K^-\pi^+$ decay in simulation and data, and then taking the ratio between MC and data. Each histogram is overlaid with the result of a fit to a constant value.

We consider also in this case a systematic uncertainty which is 0.3% [62] for each π^+_{slow}, following the same procedure used for the kaons.
Assuming that the track uncertainties are independent, all the contributions are summed in quadrature:

$$\sqrt{(2 \cdot 0.3\%)^2_{K^-} + (2 \cdot 0.3\%)^2_{\pi^+_{slow}}} = 0.85\%. \tag{6.1}$$

The total uncertainty due to the tracking efficiency is found to be 0.85%.

6.1.2 Particle-identification efficiency for kaons

The simulation pitfall regarding PID variable has been described in Section 4.2. Previous studies in Belle II created a software framework that returns, as output for each candidate, one main PID correction, and one hundred variations around the mean value calculated from statistical and systematic uncertainties. These corrections can be used as weights to scale MC distributions.
Throughout all the analysis presented in this Thesis, for each candidate only the mean PID correction has been used as correction weight while filling the D^0 reconstructed invariant mass histograms.
To evaluate the PID systematic, a variation of $\pm 1\sigma$ around the mean weight has been considered, where σ is the standard deviation of the weights distribution for one single candidate. Applying synchronously the variation on $D^0 \to K^-\pi^+\pi^0$ and $D^0 \to K^-\pi^+$ decay and extracting the signal yields from the fit, three different MC efficiencies have been calculated: $\varepsilon_{MC}(+1\sigma)$, $\varepsilon_{MC}(\text{mean})$ and $\varepsilon_{MC}(-1\sigma)$. The fit on the MC distributions are presented in figure 6.2, whilst the fit results are summarized in Table 6.1. The $K^-\pi^+\pi^0$ sample considered does not have a cut on π^0 momentum, further details on the fit models are discussed in Section 6.2.
The fit on data is presented in the bottom-left panel of Figure 6.5 for $D^0 \to K^-\pi^+\pi^0$ and Figure 6.7 for $D^0 \to K^-\pi^+$. The results are presented in the first row of Table 6.4, the resulting reconstruction efficiency is $\varepsilon_{data} = 0.1787 \pm 0.0004$.
Taking the ratio between MC and data efficiencies, three different correction factors have been obtained. The systematic uncertainty due to PID selection has been evaluated as the largest difference between the correction factor presented in Table 6.2 and the value is found to be 0.4%.

PID weights	Signal yields	$\frac{B}{B+S}$	Fit χ^2
$+1\sigma$	$2'239'300 \pm 1'710$	0.0689 ± 0.0004	1.20
mean	$2'210'600 \pm 1'700$	0.686 ± 0.0004	1.19
-1σ	$2'182'100 \pm 1'680$	0.0682 ± 0.0004	1.18
$+1\sigma$	$3'488'600 \pm 1'040$	0.0193 ± 0.0003	0.96
mean	$3'457'500 \pm 2'080$	0.0193 ± 0.0003	0.96
-1σ	$3'426'500 \pm 2'080$	0.0193 ± 0.0003	0.96

Table 6.1: Signal yields, background fractions and the χ^2 value of the fit performed on MC distributions of $D^0 \to K^-\pi^+\pi^0$ (first and second line) and $D^0 \to K^-\pi^+$ (third and fourth line).

PID weights	ε_{MC}	$\frac{\varepsilon_{data}}{\varepsilon_{MC}}$
$+1\sigma$	0.1782 ± 0.0002	1.003 ± 0.003
mean	0.1775 ± 0.0002	1.007 ± 0.003
-1σ	0.1768 ± 0.0002	1.011 ± 0.003

Table 6.2: Reconstruction efficiency obtained using MC simulation. The efficiency measured in data is $\varepsilon_{data} = 0.1787 \pm 0.0004$, the corresponding correction factors are presented as function of the correction weights for PID selection.

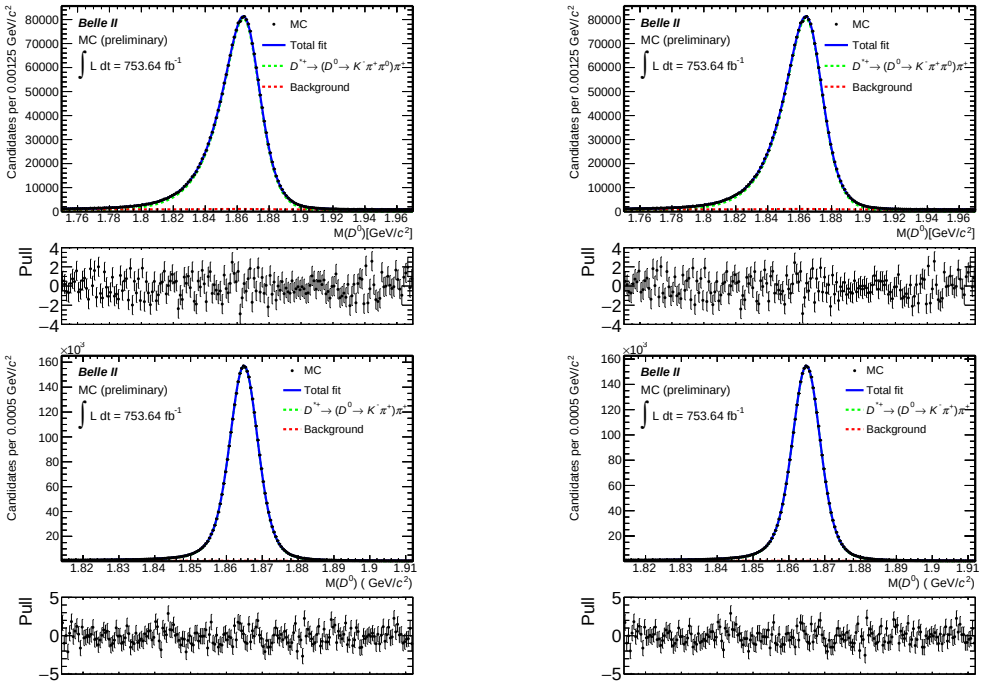


Figure 6.2: Fit on MC distributions for the $D^0 \to K^-\pi^+\pi^0$ (top) and $D^0 \to K^-\pi^+$ (bottom) decays. The correction weights applied are one standard deviation above (left) and below (right) the mean value used in the previous analysis. The sample considered does not have a cut on the π^0 momentum.

6.1.3 Δm and D^* selections

The selection applied to D^* candidates concerns D^* momentum in the CMS frame and the ΔM cut; that on D^0 candidate concerns only the D^0 mass range used in the fit to determine the signal yields.

We check the double ratio of the D^* momentum in the CMS frame, which is presented in Figure 6.3, to search for possible data-to-simulation discrepancies (see Figure D.4 in Appendix D). The distribution do not show any large departure from unity, with the fluctuation around 1 always below 5%. The χ^2/dof is equal to 0.98%, therefore a systematic uncertainty is not assigned.

We see no significant difference in the Δm distribution in data and simulation, as the peak position and the width of the signal are the same within $\mathcal{O}(1\%)$, and this has no impact in the selection. We decide not to add a systematic uncertainty related to ΔM selection. Similar considerations apply to the mass selection of D^0 candidates.

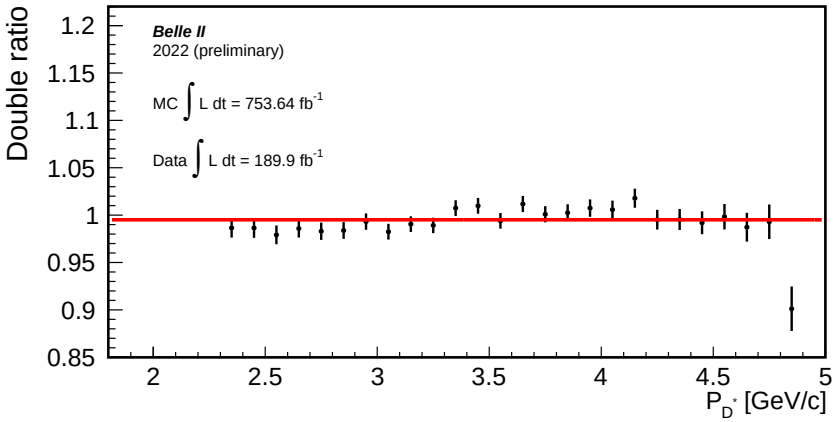


Figure 6.3: Double ratios (MC over data) of the D^* momentum in the center of mass frame. The double ratio is obtained taking the ratio between the $D^0 \to K^-\pi^+\pi^0$ decay and the $D^0 \to K^-\pi^+$ decay in simulation and data, and then taking the ratio between MC and data.

6.2 Uncertainties from the fit models

Another source of systematic uncertainty is related to the determination of the signal yields in Equation 4.4. This is achieved with the fit to the D^0 mass distributions, and we need to account for a systematic uncertainty related to the choice of the fit model for signal and background to describe the mass distributions of the two decays.

The quantify this uncertainty, four alternative models, one for each signal and background distributions of the two decays are considered. For the $D^0 \to K^+\pi^-\pi^0$ decay, we consider the full sample, without any requirement on the π^0 momentum, $\cos\theta$ or ECL region, to allow for the largest possible statistic and enhance possible relative differences between the different models used. We conservatively assume that the systematic uncertainty derived from this case, can be used also for all other case where a selection is applied. Therefore, we consider the systematic uncertainty on the fit model fully correlated in each interval of π^0 momentum, $cos\theta$, and in each ECL region. The comparison between the models used in the analysis and the alternative models is presented in Table 6.3 "Models 1" are those used for the nominal values already presented. "Models 2" are the alternative models used to determine the systematic uncertainty.

Model 1		Model 2	
$K^-\pi^+\pi^0$	$K^-\pi^+$	$K^-\pi^+\pi^0$	$K^-\pi^+$
3 johnson exp. + crystal ball	3 johnson exp. + 1st deg. pol.	3 crystal ball johnson + 2nd deg. pol.	2 crystal ball + johnson 1st deg. pol.

Table 6.3: Comparison between different signal (above) and background (below) fit models.

Figures 6.4 compares the models, while Figures 6.5 shows the fit of the total samples for the $D^0 \to K^-\pi^+\pi^0$ sample. Analogous plots are shown in Figures 6.6 and 6.7 for the $D^0 \to K^-\pi^+$ sample. All the fit results are summarised in Table 6.4. For each fit the signal yields, along with the normalized χ^2 values and the background fractions are reported. The signal yields do not change sensibly with the model.

$K^-\pi^+\pi^0$	Signal yields		$\frac{B}{B+S}$		Fit χ^2	
	MC	Data	MC	Data	MC	Data
Model 1	$2'210'600 \pm 1'700$	$552'350 \pm 1'120$	0.0686 ± 0.0004	0.0690 ± 0.0015	1.19	1.07
Model 2	$2'212'300 \pm 1'600$	$549'200 \pm 276$	0.0679 ± 0.0003	0.0743 ± 0.0021	1.45	1.32
$K^-\pi^+$						
Model 1	$3'457'500 \pm 2'080$	$858'120 \pm 1'220$	0.0193 ± 0.0003	0.0237 ± 0.0009	0.96	0.94
Model 2	$3'457'400 \pm 1'990$	$851'760 \pm 1'400$	0.0194 ± 0.0002	0.0309 ± 0.0012	1.16	1.91

Table 6.4: For each decay ($D^0 \to K^-\pi^+\pi^0$ and $D^0 \to K^-\pi^+$) and each model the results obtained from the fit to the MC and data are divided in: signal yield, background fractions and normalized χ^2 values.

Using the signal yields from the fit, the π^0 reconstruction efficiencies are measured, considering all possible combinations between models. The results are summarized in Table 6.5.

The voice "model 1", means that this model is applied on both the decays; similarly for "model 2". The combination "model 1 $K^-\pi^+\pi^0$+ model 2 $K^-\pi^+$" (and *vice versa*) states which model is applied for each decay.

The uncertainties of the efficiencies ratios are only statistical. The largest discrepancy from unity is less than 1.5%. The ratio obtained using the model 1 on both the decays is set as benchmark value, as this corresponds to the case of the nominal models used in the analysis. The systematic uncertainty is evaluated as a root-mean-square (RMS) from the benchmark value. The total systematic uncertainty due to the fit models is set to 0.58%.

	ε_{MC}	ε_{data}	$\frac{\varepsilon_{data}}{\varepsilon_{MC}}$
Model 1	0.17747 ± 0.00017	0.17867 ± 0.00044	1.0067 ± 0.0027
Model 2	0.17761 ± 0.00016	0.17897 ± 0.00031	1.0077 ± 0.0020
Model 1 $K^-\pi^+\pi^0$+ Model 2 $K^-\pi^+$	0.17748 ± 0.00017	0.18000 ± 0.00047	1.0142 ± 0.0028
Model 2 $K^-\pi^+\pi^0$+ Model 1 $K^-\pi^+$	0.17761 ± 0.00017	0.17765 ± 0.00027	1.0002 ± 0.0018

Table 6.5: The π^0 reconstruction efficiencies from MC, data and the data over MC ratio for the four possible combinations of the fit models.

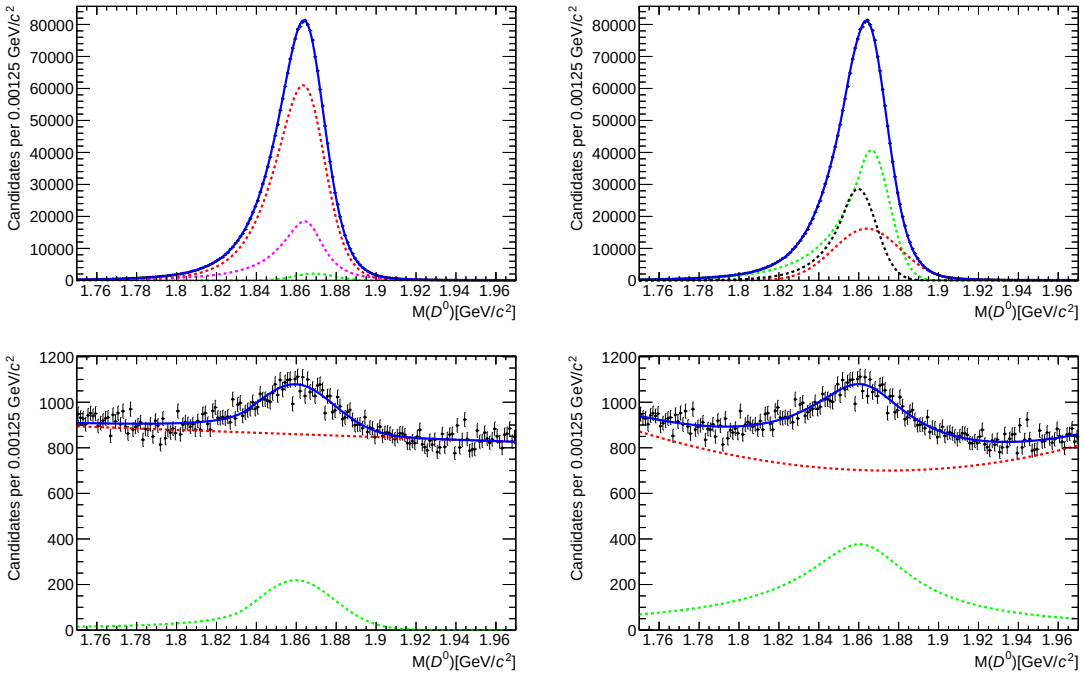


Figure 6.4: Signal (top) and background (bottom) descriptions for the $D^0 \to K^-\pi^+\pi^0$ decay without the cut on π^0 momentum. The model 1 *is shown on the left side, while model* 2 *is shown on the right.*

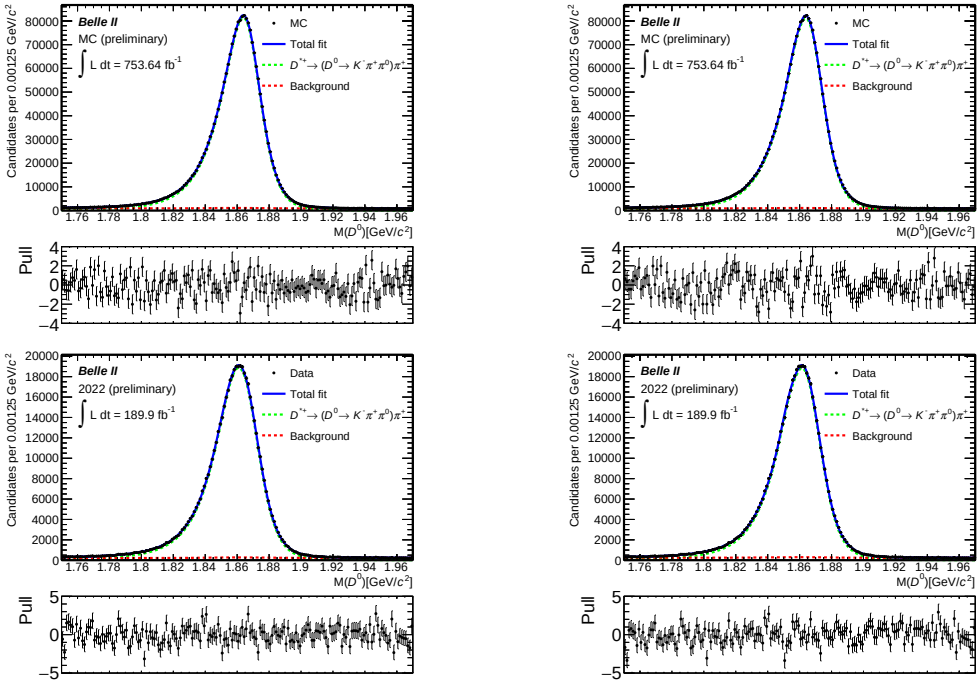


Figure 6.5: Total fit on MC (top) and data (bottom) distributions for the $D^0 \to K^-\pi^+\pi^0$ decay without the cut on π^0 momentum. On the left plots the model 1 *is applied, on the right plots the model* 2 *is applied.*

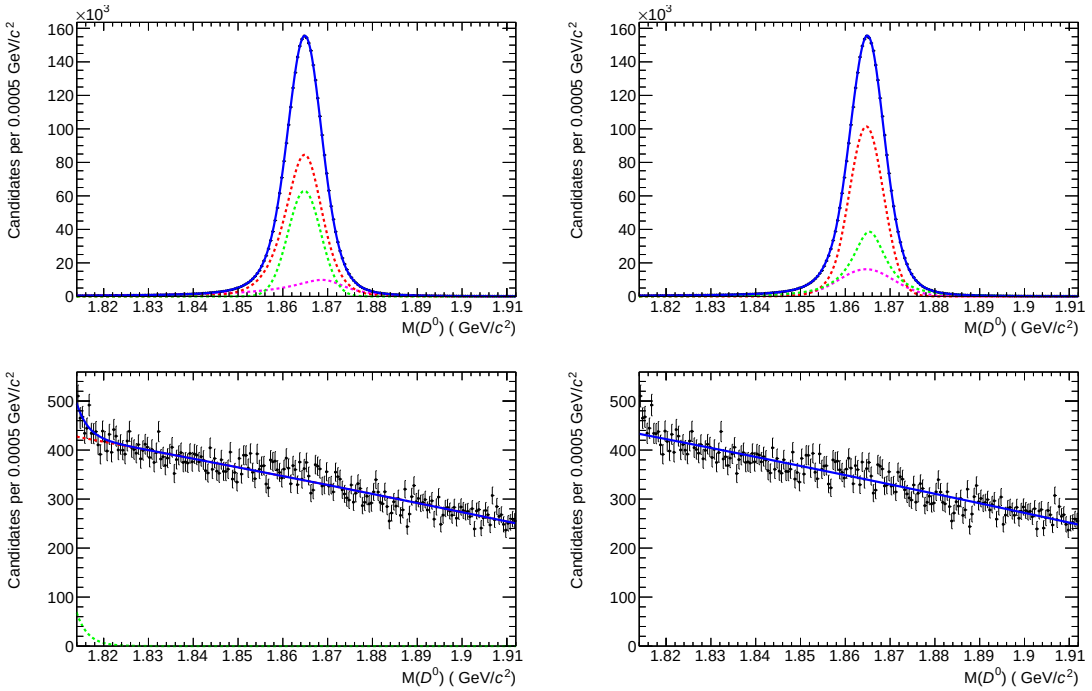


Figure 6.6: Signal (top) and background (bottom) descriptions for the $D^0 \to K^-\pi^+$ decay. The model 1 *is shown on the left side, while model* 2 *is shown on the right.*

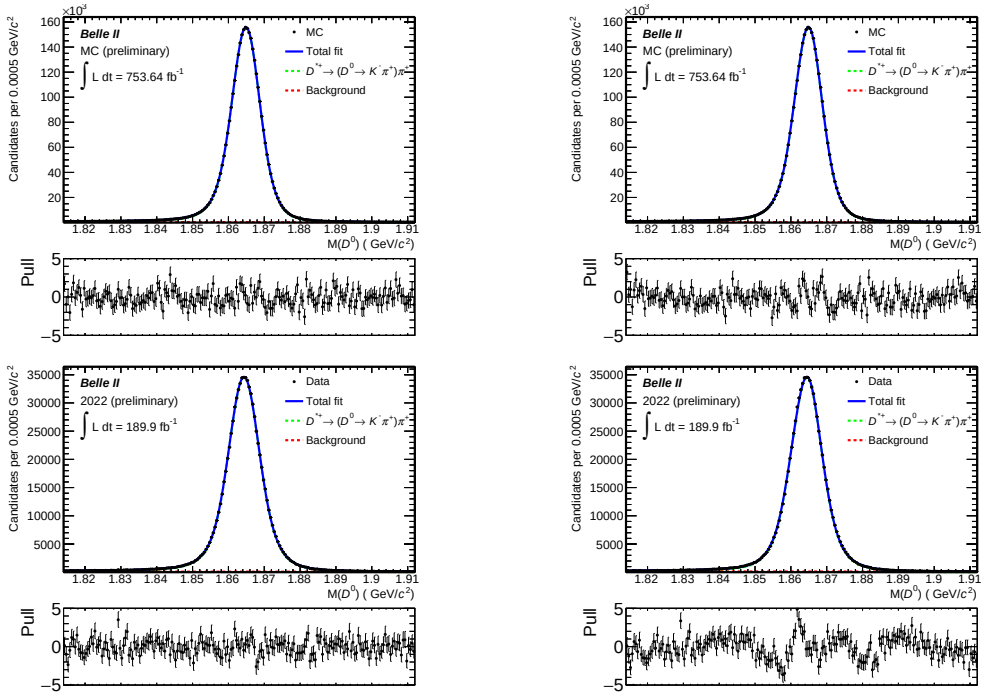


Figure 6.7: Total fit to MC (top) and data (bottom) distributions for the $D^0 \to K^-\pi^+$ decay. On the left plot the model 1 *is applied, on the right plot the model* 2 *is applied.*

6.3 Additional checks

The decay considered to measure the π^0 reconstruction efficiency is a tree-body decay which feature a rich Dalitz structure for the particle in the final state [63]. A poor description of this structure in simulation could bias the measurement of the correction factor, as it introduce another source of data-simulation difference. Possible discrepancies between simulation and data should appear as differences in the momentum distributions of the particles. The charged particles momentum distributions have already been investigate, the last check concerns the neutral pion. The ratio of MC over data for the $\cos(\theta)$ and the momentum distributions of the π^0 are reported in Figure 6.8.

The π^0 momentum and $\cos(\theta)$ distribution have normalized χ^2 values of 2.4 and 1.5. Residual data-MC discrepancy in this distribution can be either ascribed to differences in simulation or simulated kinematic (stemming from Dalitz structure), and the two effect cannot be disentangled. This is an unavoidable limitation of using a 3-body decay for the determination of the π^0 efficiency. However, since no selections are applied on the Dalitz plane, we do not consider a systematic uncertainty and ascribe any observed discrepancy to the data-simulation difference of the reconstruction efficiency.

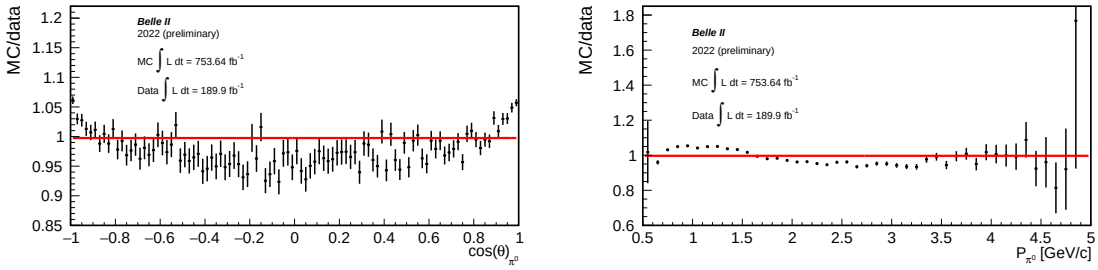


Figure 6.8: Ratio of MC over data of the π^0 $\cos(\theta)$ (left) and momentum (right) distributions.

6.4 Total systematic uncertainty

I found three different contributions to the systematic uncertainty: 0.85% due to uncertainty on tracking efficiencies, 0.40% due to the uncertainty on the kaonID selection efficiency, and 0.58% due to the choice of fit model. Finally we consider that on the external input of the charm-decay branching fraction, which is 3.56%. All systematic uncertainties are summarised in Table 6.6. Assuming that the single contributions are independent, the final systematic uncertainty is obtained summing each single contribution in quadrature:

$$\sqrt{(0.85\%)^2 + (0.58\%)^2 + (0.40\%)^2 + (3.56\%)^2} = 3.73\% \tag{6.2}$$

This systematic uncertainty is considered as a global common uncertainty for all the correction factors obtained in this work (*i.e.* in different ECL regions, range of of $\cos(\theta)_{\pi^0}$ and π^0 momentum).

Source	Contribution (%)
K^- tracking efficiency	0.60
π^+_{slow} tracking efficiency	0.60
PID efficiency	0.40
Fit model	0.58
Ratio of branching fractions	3.56
Total	3.70

Table 6.6: The systematic uncertainties for the measurement of the correction factor. The total is calculated by adding all the systematic uncertainties in quadrature.

Chapter 7

Conclusions

This Thesis targeted a study of the reconstruction efficiency of the π^0 meson in $\pi^0 \to \gamma\gamma$ decays at the Belle II experiment. The reconstruction efficiency and its systematic uncertainty are provided for the measurement of the branching fraction of the $B^0 \to \pi^0\pi^0$ decay carried out at Belle II. This measurement is instrumental for the determination of the angle α of the Unitarity Triangle from the isospin analysis of $B \to \pi\pi$ decays.
I used two D^*-tagged decays of the D^0 meson, $D^0 \to K^-\pi^+\pi^0$ and $D^0 \to K^-\pi^+$, to measure the π^0-reconstruction efficiency in MC simulation and experimental data, and I derived a correction factor for an unbiased determination of the efficiency from the simulation of $B^0 \to \pi^0\pi^0$ decays. The size of the data sample used in this work corresponds to an integrated luminosity of $190\,\text{fb}^{-1}$ collected by the Belle II experiment in e^+e^- collisions provided by the SuperKEKB collider at the centre-of-mass energy of the Y(4S) resonance.
The main results achieved are summarized as follows.

- The π^0-reconstruction efficiency correction factor is found to be:

$$\frac{\varepsilon^{\text{data}}_{\pi^0}}{\varepsilon^{\text{MC}}_{\pi^0}} = 1.023 \pm 0.003\,(\text{stat}) \pm 0.037(\text{syst}) \tag{7.1}$$

  The value is compatible with unity within 1σ. The systematic uncertainty for the $B^0 \to \pi^0\pi^0$ analysis is set to 3.7% per π^0. It represents an improvement over the previous method used in a preliminary results of $\mathcal{B}(B^0 \to \pi^0\pi^0)$ obtained on $63\,\text{fb}^{-1}$ of data, which featured a systematic uncertainty equal to 10% per π^0 [58].

- Three different dependencies of the correction factor have been identified: *i.e.* on the region of the electromagnetic calorimeter hit by the photons (see Figure 5.3); on the cosine of the polar angle of the π^0 candidate (see Figure 5.8); and on its momentum (see Figure 5.11). The last two dependencies can be further investigated taking into account the correlation between the momentum and the direction of the π^0. Indeed, the analysis presented in this Thesis could be extended considering a two dimensional map as a function of momentum and $\cos(\theta)$.

- The dependence on momentum can be used to have a better assessment of the correction factor needed for the $B^0 \to \pi^0\pi^0$ analysis. Considering the three correction factors determined above the requirement $p(\pi^0) > 1.5\,\mathrm{GeV}/c$, and using the fraction of events in different momentum intervals from the momentum distribution of $B^0 \to \pi^0\pi^0$ decays in simulation, a weighted average can be obtained. By doing so, the resulting correction is

$$\frac{\varepsilon^{data}_{\pi^0}}{\varepsilon^{MC}_{\pi^0}} = 1.037 \pm 0.007\,(\mathrm{stat}) \pm 0.037\,(\mathrm{syst}) \tag{7.2}$$

  The value is compatible with the previous result in equation 7.1 (obtained in Chapter 4) where no dependence has been considered. The efficiencies ratio between MC and data remains compatible with 1 within $1\,\sigma$.

  I used the method presented in this Thesis to determine the correction factor and to assess the systematic uncertainty on the π^0-reconstruction efficiency for the preliminary result on the branching fraction of the $B^0 \to \pi^0\pi^0$ decay presented by the Belle II Collaboration at the XLI International Conference of High Energy Physics held in July 2022 in Bologna [49, 52, 57]. The analysis discussed in this Thesis, extended that work using an improved MC simulation featuring a better description of beam backgrounds and the data-taking conditions. This is the first time that the π^0-reconstruction efficiency has been checked for this new simulation, which is used for the ongoing update of the $B^0 \to \pi^0\pi^0$ analysis with the full sample collected before the long shutdown of the experiment. The current method to determine the π^0 reconstruction efficiency is affected by an irreducible 3.5% uncertainty due to the uncertainty on charm-decays branching fractions. However, this is sufficient precision for the ongoing measurement, which is expected to have a statistical uncertainty of 15% on the $B^0 \to \pi^0\pi^0$ branching fraction.

Appendix A

$B^0 \to \pi^0\pi^0$

A.1 π^0 selection variables

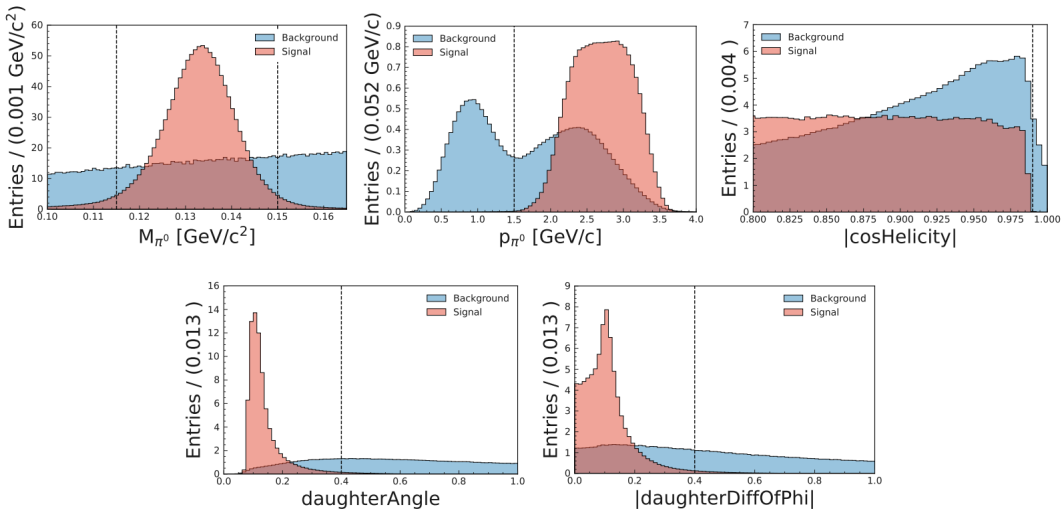


Figure A.1: Comparison between signal (red) and background (blue) for π^0 for mass (top left), momentum (top right), |cosHelicity| (middle left), daughterAngle (middle right) and |daughterDiffOfPhi| (bottom) [49].

A.2 Photons selection variables

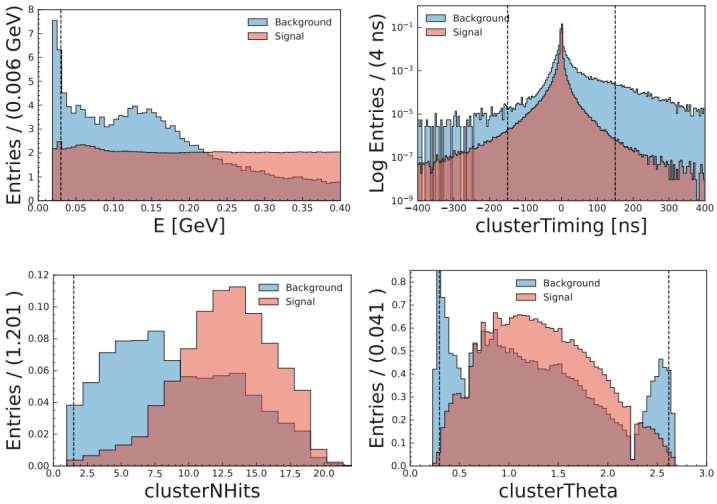


Figure A.2: Comparison between signal (red) and background (blue) for γ for energy (top-left), cluster timing (top-right), clusterNHits (bottom left) and clusterTheta (bottom right) [49].

A.2.1 Continuum suppression

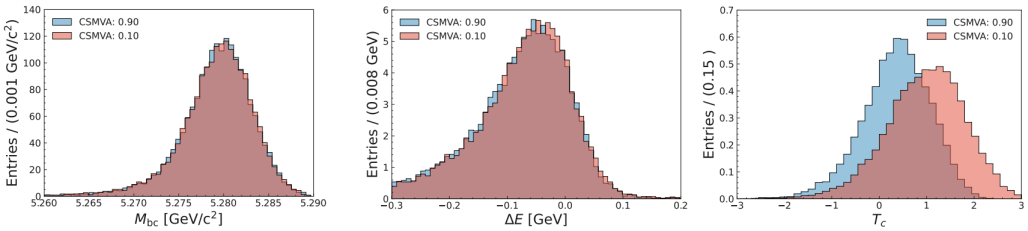


Figure A.3: Distributions of M_{bc} (left), ΔE (middle) and T_c (right) for a continuum suppression selection of 0.10 *(red) and* 0.90 *(blue) [49].*

Appendix B

PhotonMVA variables

- clusterAbsZernikeMoment40: returns absolute value of Zernike moment 40 ($|Z_{40}|$), a shower shape variable.
- clusterAbsZernikeMoment51: returns absolute value of Zernike moment 51 ($|Z_{51}|$), a shower shape variable.
- clusterE1E9: returns ratio of energies of the central crystal, E1, and 3×3 crystals, E9, around the central crystal. Since E1$\leq$E9, this ratio is ≤ 1 and tends towards larger values for photons and smaller values for hadrons.
- clusterE9E21: returns ratio of energies in inner 3×3 crystals, E9, and 5×5 crystals around the central crystal without corners. Since E9$\leq$E21, this ratio is ≤ 1 and tends towards larger values for photons and smaller values for hadrons.
- clusterHighestE: returns energy of the highest energetic crystal in ECL cluster after reweighting.
- clusterSecondMoment: returns second moment S, defined as:

$$S = \frac{\sum_{i=0}^{n} \omega_i E_i r_i^2}{\sum_{i=0}^{n} \omega_i E_i} \tag{B.1}$$

  where $E_i = (E_0, E_1, ...)$ are the single crystal energies sorted by energy, ω_i is the crystal weight, and r_i is the distance of the i-th digit to the shower center projected to a plane perpendicular to the shower axis.

- clusterZernikeMVA: returns output of a MVA using eleven Zernike moments of the cluster. Zernike moments are calculated per shower in a plane perpendicular to the shower direction via:

$$|Z_{nm}| = \frac{n+1}{\pi} \frac{1}{\sum_i \omega_i E_i} \left| \sum_i R_{nm}(\rho_i) e^{-im\alpha_i} \omega_i E_i \right| \tag{B.2}$$

  where n, m are the integers, i runs over the crystals in the shower, E_i is the energy of the i-th crystal in the shower, R_{nm} is a polynomial od degree n, ρ_i is the radial distance of the i-th

crystal in the perpendicular plane. As a crystal can be related to more than one shower, ω_i is the fraction of the energy of the i-th crystal associated with the shower.

- minC2TDist: Returns distance between ECL cluster and nearest track hitting the ECL. A cluster comprises the energy depositions of several crystals. All these crystals have slightly different orientations in space. A shower direction can be constructed by calculating the weighted average of these orientations using the corresponding energy depositions as weights. The intersection (more precisely the point of closest approach) of the vector with this direction originating from the cluster center and an extrapolated track can be used as reference for the calculation of the track depth. It is defined as the distance between this intersection and the track hit position on the front face of the ECL.

Appendix C

Neutral pion reconstruction efficiency

C.1 Selection of the $D^0 \to K^-\pi^+\pi^0$ decay

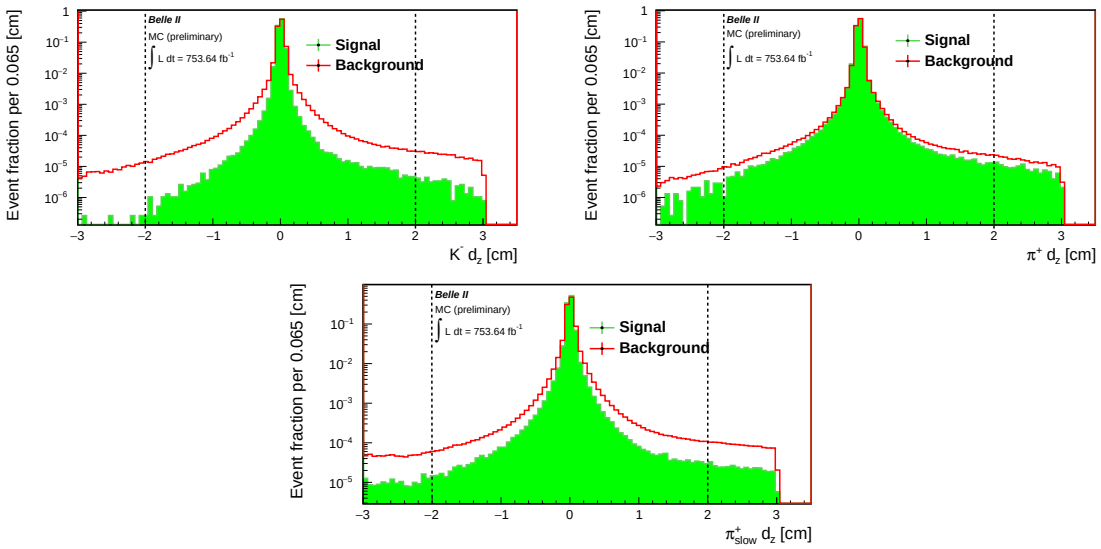


Figure C.1: Longitudinal displacement in cm *from the interaction point for kaon (left), the slow pion (center) and pion (right). The histogram filled in green is the signal, the background is represented as a red line. The two vertical dashed lines correspond to the chosen cut values.*

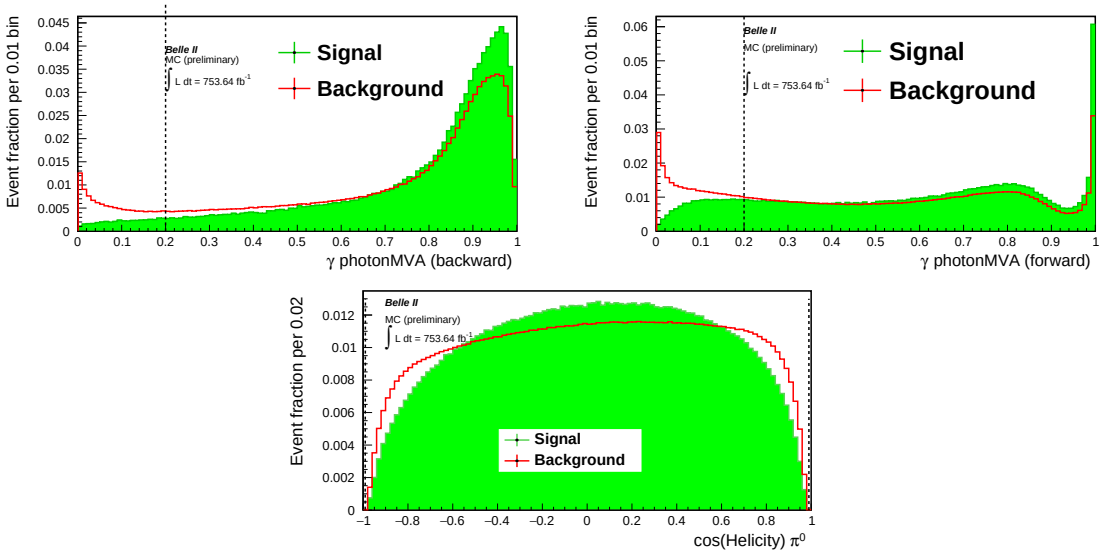


Figure C.2: Distribution of the cosine of the helicity angle (middle), the photonMVA output in the forward (left) and backward region of the ECL (right). The histogram filled in green is the signal, the background is represented as a red line and the vertical dashed line corresponds to the chosen cut value.

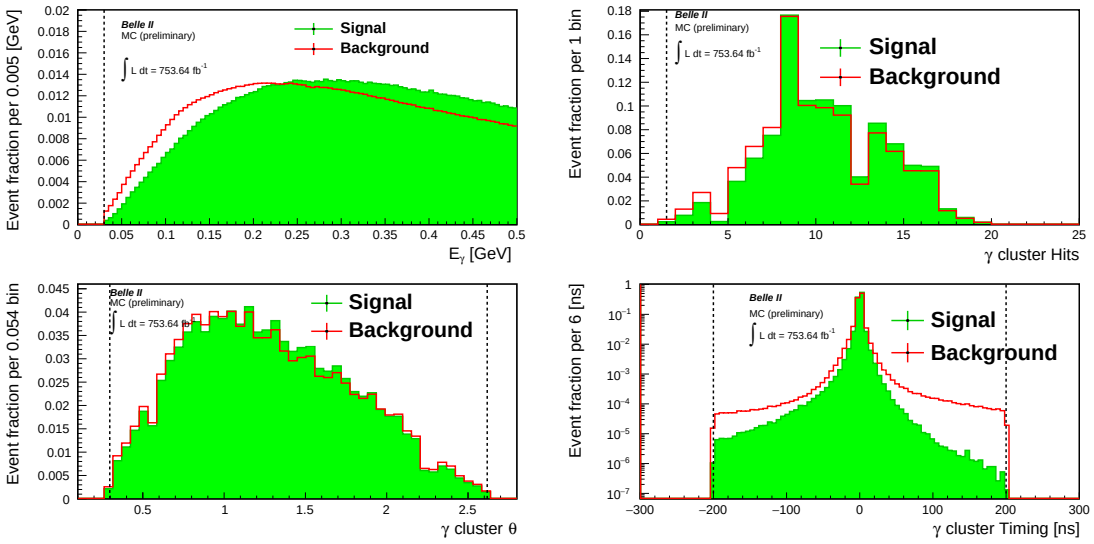


Figure C.3: Distributions of the photon energy (top-left), clusterHits (top-right), clusterTheta (bottom-left) and clusterTiming (bottom-right). See Section 3.2.1 for more details on these variables. The histogram filled in green is the signal, the background is represented as a red line and the vertical dashed line corresponds to the chosen cut value.

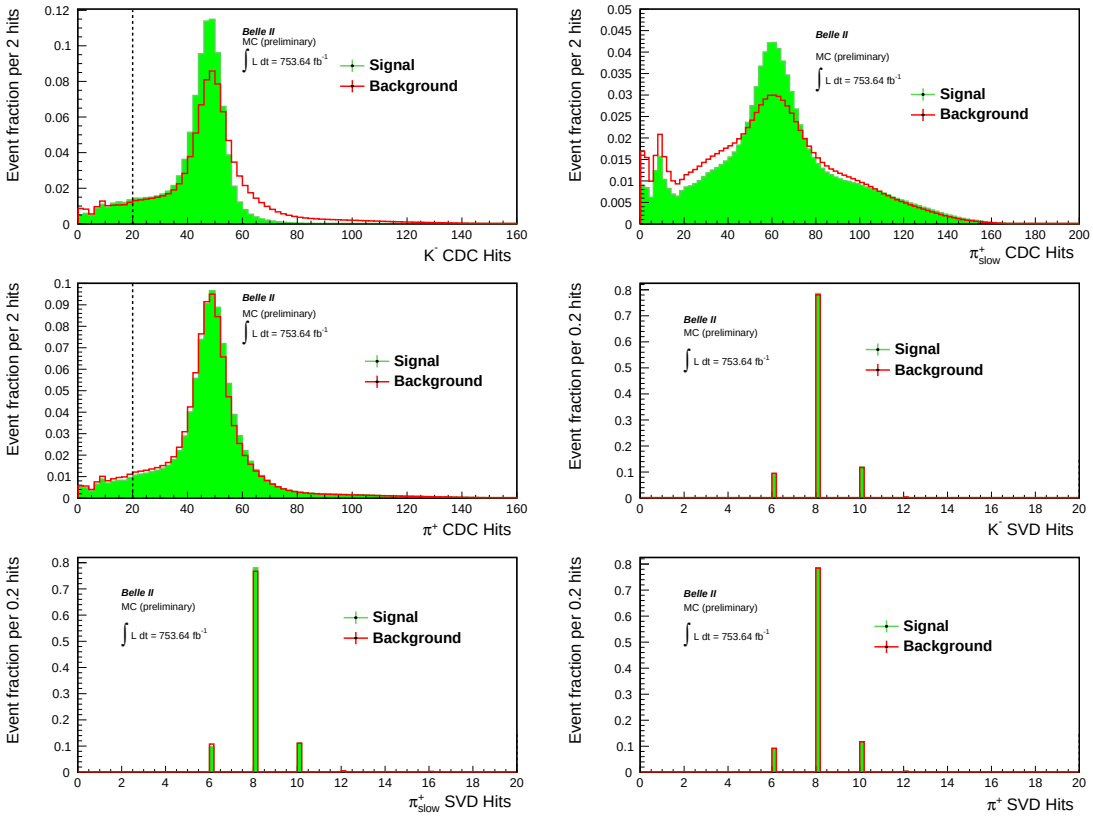


Figure C.4: Hits in the central drift chamber for kaon (top-left), the slow pion (top-right) and pion (middle-left). For the slow pion the cut value is set to 0*. Hits in the first SVD layer for kaon (middle-right), the slow pion (bottom-left) and pion (bottom-right); the cut value is set to* 0*. The histogram filled in green is the signal, the background is represented as a red line and the vertical dashed line corresponds to the chosen cut value.*

C.2 Selection for the $D^0 \to K^-\pi^+$ decay

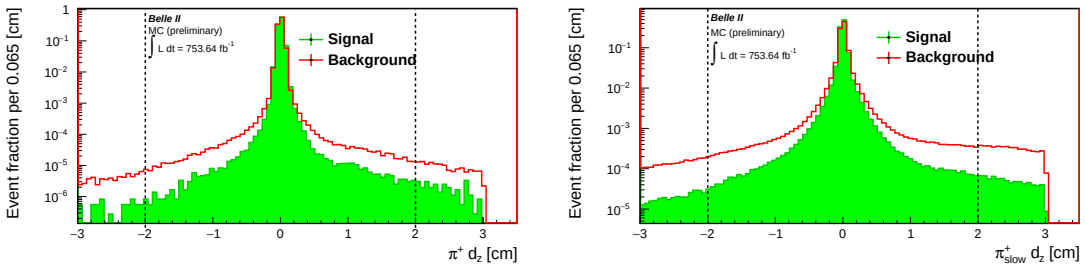


Figure C.5: Longitudinal displacement from the interaction point for π^+ (left) and the π^+_{slow} (right). The histograms filled in green are the signal, the background is represented as a red line and the vertical dashed lines corresponds to the chosen cut values.

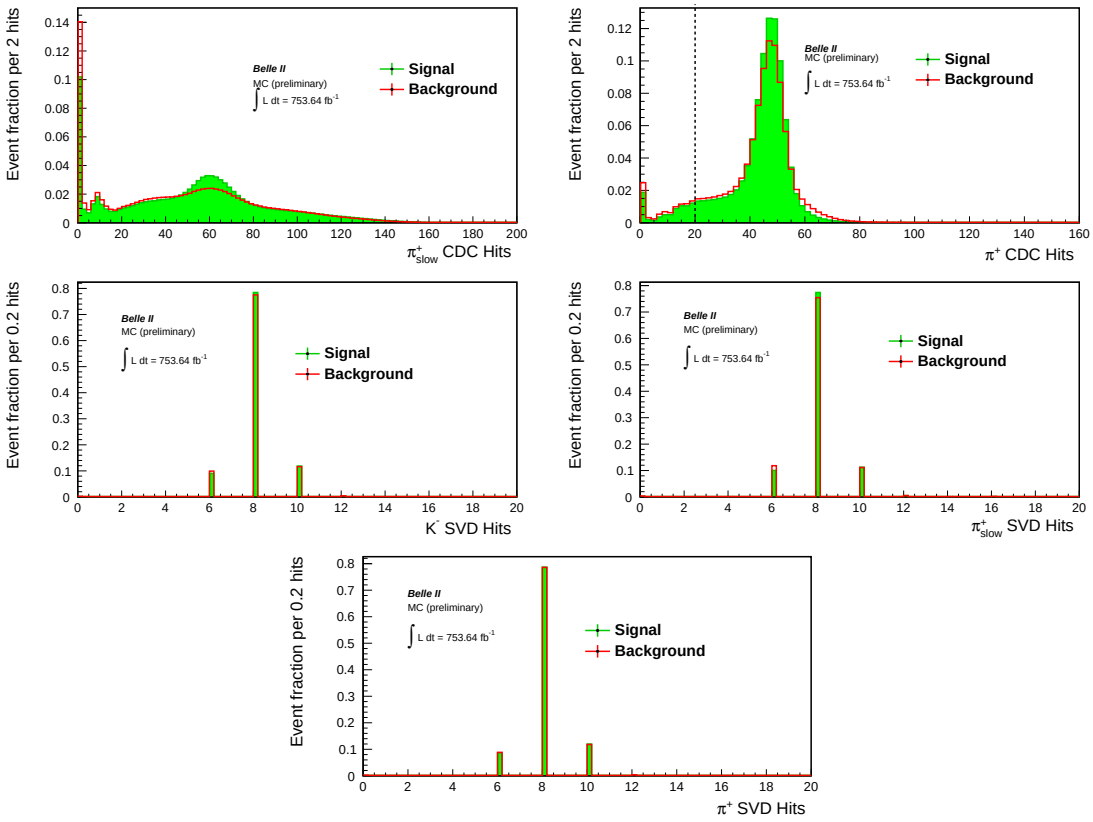


Figure C.6: Number of hits in the CDC for π^+_{soft} (top-left), the π^+ (top-right). Number of hits in the first SVD layer for K^- (middle-left), π^+_{slow} (middle-right) and the π^+ (bottom). The histograms filled in green are the signal, the background is represented as a red line. The cut values are set to 0 for the SVD hits distributions.

C.3 Data-MC correction factor

The Johnson probability density function is defined as:

$$PDF[Johnson\,S_U] = \frac{\delta}{\lambda\sqrt{2\pi}} \frac{1}{\sqrt{1+\left(\frac{x-\mu}{\lambda}\right)^2}} \exp\left[-\frac{1}{2}\left(\gamma + \delta \sinh^{-1}\left(\frac{x-\mu}{\lambda}\right)\right)^2\right] \tag{C.1}$$

It is often used to fit the reconstructed invariant mass or the mass difference in charm decays. For this reason x is called mass; δ, λ, μ and γ are free parameters.
The crystal ball function is defined as:

$$f = N \cdot \begin{cases} \exp\left(-\frac{(x-m_0)^2}{2\sigma^2}\right), & \text{for } \frac{x-m_0}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-m_0}{\sigma})^{-n}, & \text{for } \frac{x-m_0}{\sigma} \leq -\alpha \end{cases} \tag{C.2}$$

where α, n, m_0, σ are free parameters and:

$$\begin{aligned} A &= \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(-\frac{|\alpha|^2}{2}\right), \\ B &= \frac{n}{|\alpha|} - |\alpha|, \\ N &= \frac{1}{\sigma(C+D)} \\ C &= \frac{n}{|\alpha|} \cdot \frac{1}{n-1} \cdot \exp\left(-\frac{|\alpha|^2}{2}\right) \\ D &= \sqrt{\frac{\pi}{2}}\left(1 + \operatorname{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right)\right) \end{aligned} \tag{C.3}$$

Appendix D

Systematic uncertainty

D.1 Common selections

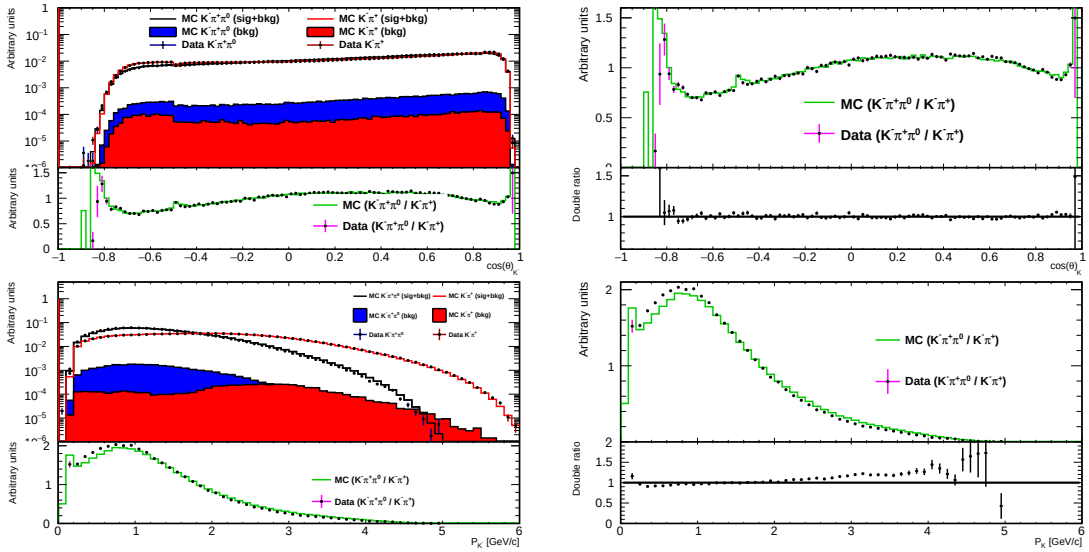


Figure D.1: On the left side, MC and data distributions of $cos(\theta)$ (top) and momentum (bottom) of the kaon, for both the decays, are presented. The histograms are normalized, to the same area. The blue and red-filled histograms are the background distributions. The ratio of the MC and data distributions are shown in the lower panels. On the right side, the same ratios are shown in the upper panels, double ratios are reported in the lower panels.

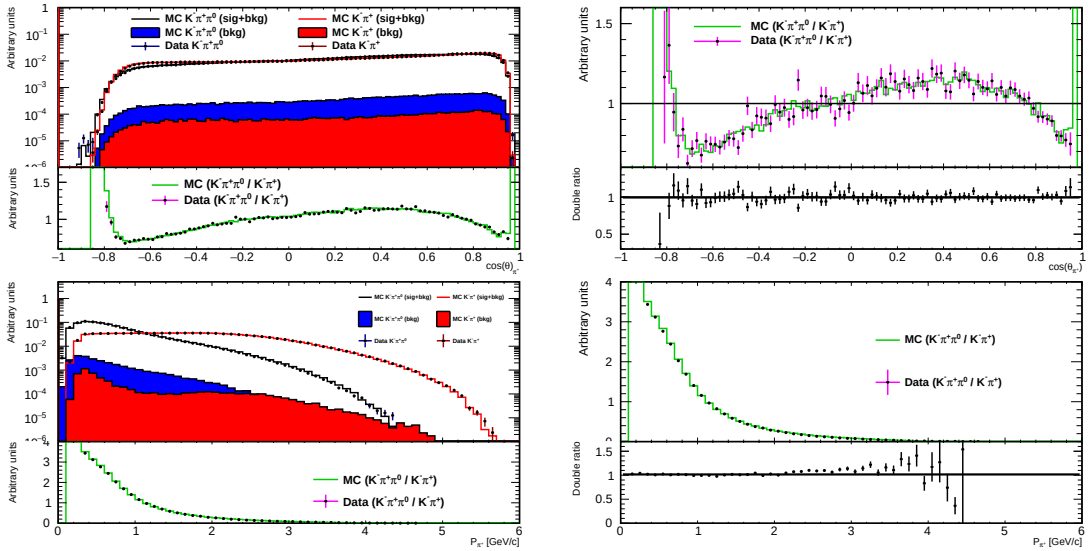


Figure D.2: On the left side, MC and data distributions of $\cos(\theta)$ (top) and momentum (bottom) of the π^+, for both the decays, are presented. The histograms are normalized, to the same area. The blue and red-filled histograms are the background distributions. The ratio of the MC and data distributions are shown in the lower panels. On the right side, the same ratios are shown in the upper panels, double ratios are reported in the lower panels.

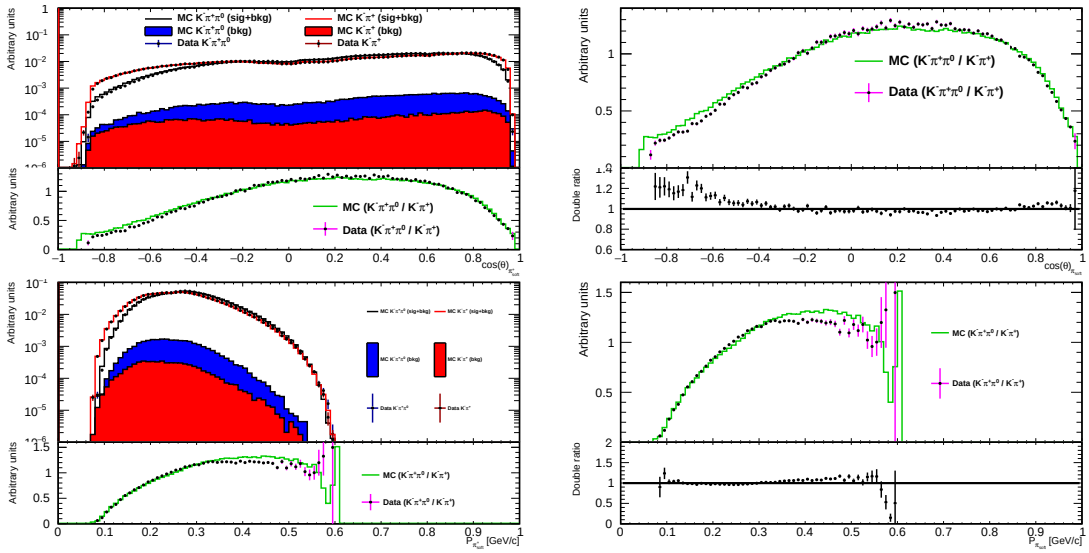


Figure D.3: On the left side, MC and data distributions of $\cos(\theta)$ (top) and momentum (bottom) of the slow-pion, for both the decays, are presented. The histograms are normalized, to the same area. The blue and red-filled histograms are the background distributions. The ratio of the MC and data distributions are shown in the lower panels. On the right side, the same ratios are shown in the upper panels, double ratios are reported in the lower panels.

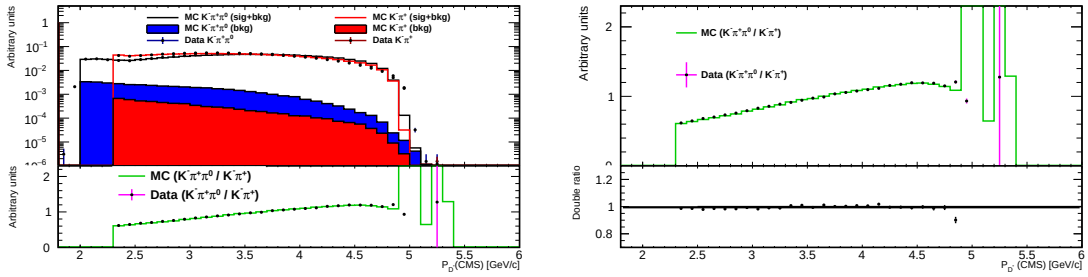


Figure D.4: On the left side, MC and data distributions of the momentum in the center of mass reference frame of the D^{*+}, for both the decays, are presented. The histograms are normalized, to the same area. The blue and red-filled histograms are the background distributions. The ratio of the MC and data distributions are shown in the lower panels. On the right side, the same ratios are shown in the upper panels, double ratios are reported in the lower panels.

Bibliography

[1] Belle II Collaboration, F. Abudinén, et al. *Measurement of the branching fraction for $B^0 \to \pi^0\pi^0$ decays reconstructed in 2019-2020 Belle II data*. 2021. DOI: 10.48550/ARXIV.2107.02373. URL: https://arxiv.org/abs/2107.02373.

[2] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 1967), p. 1264. DOI: 10.1103/PhysRevLett.19.1264.

[3] Sheldon L. Glashow. "Partial-symmetries of weak interactions". In: *Nuclear Physics* 22 (1961), p. 579. ISSN: 0029-5582. DOI: https://doi.org/10.1016/0029-5582(61)90469-2.

[4] Abdus Salam. "Weak and Electromagnetic Interactions". In: *Conf. Proc. C* 680519 (1968), p. 367. DOI: 10.1142/9789812795915_0034.

[5] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J.C. Taylor, p. 321. DOI: 10.1103/PhysRevLett.13.321.

[6] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J.C. Taylor, p. 508. DOI: 10.1103/PhysRevLett.13.508.

[7] G.S. Guralnik, C.R. Hagen, and T.W.B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J.C. Taylor, p. 585. DOI: 10.1103/PhysRevLett.13.585.

[8] P.A. Zyla et al. "Review of Particle Physics". In: *Prog. Theor. Exp. Phys.* 2020 (3 2020), p. 083C01. DOI: 10.1103/PhysRevD.98.030001.

[9] LHCb collaboration. *LHCb pubblications*. 2022. URL: https://lhcbproject.web.cern.ch/Publications/LHCbProjectPublic/Summary_all.html.

[10] Gerhart Lüders. "Proof of the TCP theorem". In: *Annals of Physics* 2 (1957), p. 1. DOI: 10.1016/0003-4916(57)90032-5.

[11] Wolfgang Pauli. "Exclusion Principle, Lorentz Group, and reversal of space-time and charge". In: *Niels Bohr and the Development of Physics* (1955), p. 1.

[12] J. H. Christenson et al. "Evidence for the 2π Decay of the K_2^0 Meson". In: *Phys. Rev. Lett.* 13 (4 July 1964), pp. 138–140. DOI: 10.1103/PhysRevLett.13.138. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.138.

[13] Nicola Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Phys. Rev. Lett.* 10 (1963), p. 531. DOI: 10.1103/PhysRevLett.10.531.

[14] Makoto Kobayashi and Toshihide Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". In: *Prog. Theor. Phys.* 49 (1973), p. 652. DOI: 10.1143/PTP.49.652.

[15] Lincoln Wolfenstein. "Parametrization of the Kobayashi-Maskawa Matrix". In: *Phys. Rev. Lett.* 51 (21 1983), p. 1945. DOI: 10.1103/PhysRevLett.51.1945.

[16] C. Jarlskog. "Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Violation". In: *Phys. Rev. Lett.* 55 (1985), p. 1039. DOI: 10.1103/PhysRevLett.55.1039.

[17] C. Jarlskog. "A Basis Independent Formulation of the Connection Between Quark Mass Matrices, CP Violation and Experiment". In: *Z. Phys. C* 29 (1985), p. 491. DOI: 10.1007/BF01565198.

[18] I. Dunietz, O. W. Greenberg, and Dan-Di Wu. "A priori definition of maximal CP nonconservation". In: *Phys. Rev. Lett.* 55 (1985), p. 2935. DOI: 10.1103/PhysRevLett.55.2935.

[19] G. Isidori. "B physics in the LHC era". In: (Jan. 2010). DOI: 10.1201/b11865-5.

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

[21] A. Ceccucci et al. "CKM Quark-Mixing Matrix". In: (2020). URL: https://pdg.lbl.gov/2021/reviews/rpp2020-rev-ckm-matrix.pdf.

[22] J. Charles et al. "CP violation and the CKM matrix: assessing the impact of the asymmetric B factories". In: *The European Physical Journal C* 41.1 (May 2005), pp. 1–131. DOI: 10.1140/epjc/s2005-02169-1. URL: https://doi.org/10.1140%2Fepjc%2Fs2005-02169-1.

[23] Belle II Collaboration and B2TiP theory Community. *The Belle II Physics Book*. 1st. Prog. Theor. Exp. Phys., 2018.

[24] Ganiev Eldar. "Measurement of the branchin fraction, longitudinal polarization fraction, and charge-parity violating asymmetry in $B^+ \to \rho^+\rho^0$ decays at Belle". PhD thesis. Università degli studi di Trieste, 2020. URL: http://hdl.handle.net/11368/2995661.

[25] Y. Amhis et al. "Averages of b-hadron, c-hadron, and tau-lepton properties as of 2021". In: (June 2022). arXiv: 2206.07501 [hep-ex].

[26] Michael Gronau and David London. "Isospin analysis of CP asymmetries in B decays". In: *Phys. Rev. Lett.* 65 (27 Dec. 1990), pp. 3381–3384. DOI: 10.1103/PhysRevLett.65.3381. URL: https://link.aps.org/doi/10.1103/PhysRevLett.65.3381.

[27] J. Charles et al. "Isospin analysis of charmless B-meson decays". In: *The European Physical Journal C* 77.8 (Aug. 2017). DOI: 10.1140/epjc/s10052-017-5126-9. URL: https://doi.org/10.1140%2Fepjc%2Fs10052-017-5126-9.

[28] Yukiyoshi Ohnishi et al. "Accelerator design at SuperKEKB". In: *Prog. Theo. Exp. Phys.* (2013), 03A011. DOI: 10.1093/ptep/pts083.

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

[30] Paolo Raimondi et al. "Beam-beam issues for colliding schemes with large Piwinski angle and crabbed waist". In: (2007). arXiv: physics/0702033.

[31] T. Abe et al. "Belle II Technical Design Report". In: (2010). arXiv: 1011.0352.

[32] C. Marinas. "The Belle II pixel detector: High precision with low material". In: *Nucl. Instrum. Meth.* A731 (2013), p. 31. DOI: 10.1016/j.nima.2013.03.025.

[33] K. Adamczyk et al. "The Belle II silicon vertex detector assembly and mechanics". In: *Nucl. Instrum. Meth.* A845 (2017), p. 38. DOI: 10.1016/j.nima.2016.03.100.

[34] N. Taniguchi. "Central Drift Chamber for Belle-II". In: *JINST* 12 (2017), p. C06014. DOI: 10.1088/1748-0221/12/06/C06014.

[35] Valerio Bertacchi et al. "Track finding at Belle II". In: *Computer Physics Communications* 259 (Feb. 2021), p. 107610. DOI: 10.1016/j.cpc.2020.107610. URL: https://doi.org/10.1016%2Fj.cpc.2020.107610.

[36] A. Abashian et al. "The K_L/μ detector subsystem for the BELLE experiment at the KEK B factory". In: *Nucl. Instrum. Meth.* A449 (2000), p. 112. DOI: 10.1016/S0168-9002(99)01383-2.

[37] T. Abe, Adachi, et al. *Belle II Technical Design Report*. 2010. DOI: 10.48550/ARXIV.1011.0352. URL: https://arxiv.org/abs/1011.0352.

[38] E Kou et al. "The Belle II Physics Book". In: *Progress of Theoretical and Experimental Physics* 2019.12 (Dec. 2019). DOI: 10.1093/ptep/ptz106. URL: https://doi.org/10.1093%2Fptep%2Fptz106.

[39] Francesco Forti. "Snowmass Whitepaper: The Belle II Detector Upgrade Program". In: *2022 Snowmass Summer Study*. Mar. 2022. arXiv: 2203.11349 [hep-ex].

[40] Belle II collaboration. *Belle II Luminosity*. 2022. URL: https://confluence.desy.de/display/BI/Belle+II+Luminosity#BelleIILuminosity-Integratedluminosity(perweek/day)forthepastrunperiod2022a-b.

[41] T. Kuhr et al. "The Belle II Core Software". In: *Computing and Software for Big Science* 3.1 (Nov. 2018). DOI: 10.1007/s41781-018-0017-9. URL: https://doi.org/10.1007%2Fs41781-018-0017-9.

[42] David J. Lange. "The EvtGen particle decay simulation package". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1 (2001). BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines, pp. 152–155. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(01)00089-4. URL: https://www.sciencedirect.com/science/article/pii/S0168900201000894.

[43] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Computer Physics Communications* 191 (2015), pp. 159–177. ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2015.01.024. URL: https://www.sciencedirect.com/science/article/pii/S0010465515000442.

[44] S. Jadach et al. "The precision Monte Carlo event generator for two-fermion final states in collisions". In: *Computer Physics Communications* 130.3 (Aug. 2000), pp. 260–325. DOI: 10.1016/s0010-4655(00)00048-5. URL: https://doi.org/10.1016%2Fs0010-4655%2800%2900048-5.

[45] S. Jadach et al. "Coherent exclusive exponentiation for precision Monte Carlo calculations". In: *Phys. Rev. D* 63 (11 May 2001), p. 113009. DOI: 10.1103/PhysRevD.63.113009. URL: https://link.aps.org/doi/10.1103/PhysRevD.63.113009.

[46] Stanisław Jadach et al. "TAUOLA - a library of Monte Carlo programs to simulate decays of polarized tau leptons". In: *Computer Physics Communications* 64.2 (1991), pp. 275–299. ISSN: 0010-4655. DOI: https://doi.org/10.1016/0010-4655(91)90038-M. URL: https://www.sciencedirect.com/science/article/pii/001046559190038M.

[47] N. Davidson et al. "Universal interface of TAUOLA: Technical and physics documentation". In: *Computer Physics Communications* 183.3 (2012), pp. 821–843. ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2011.12.009. URL: https://www.sciencedirect.com/science/article/pii/S0010465511003973.

[48] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[49] Francis Pham. "Measurement of $B^0 \to \pi^0\pi^0$ branching fraction and $\mathcal{A}_{CP}$". In: (Nov. 2021).

[50] Geoffrey C. Fox and Stephen Wolfram. "Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes". In: *Phys. Rev. Lett.* 41 (23 1978), p. 1581. DOI: 10.1103/PhysRevLett.41.1581.

[51] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[52] Justin Skorupa. "Recent Belle II results on hadronic B decays". In: (July 2022). URL: https://docs.belle2.org/record/3126?ln=en.

[53] F. Abudinén et al. "B-flavor tagging at Belle II". In: *The European Physical Journal C* 82.4 (Apr. 2022). DOI: 10.1140/epjc/s10052-022-10180-9. URL: https://doi.org/10.1140%2Fepjc%2Fs10052-022-10180-9.

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

[55] T. Julius and others. "Measurement of the branching fraction and CP asymmetry in $B^0 \to \pi^0\pi^0$ decays, and an improved constraint on ϕ_2". In: *Phys. Rev. D* 96 (3 Aug. 2017), p. 032007. DOI: 10.1103/PhysRevD.96.032007. URL: https://link.aps.org/doi/10.1103/PhysRevD.96.032007.

[56] J. P. Lees et al. "Measurement of CP asymmetries and branching fractions in charmless two-body B-meson decays to pions and kaons". In: *Phys. Rev. D* 87 (5 Mar. 2013), p. 052009. DOI: 10.1103/PhysRevD.87.052009. URL: https://link.aps.org/doi/10.1103/PhysRevD.87.052009.

[57] F. Rossi et al. "Measurement of π^0 reconstruction efficiency for $B^0 \to \pi^0\pi^0$ decays BELLE2-NOTE-TE-2022-010". In: (June 2022).

[58] Belle II Collaboration. *Measurement of the branching fraction for $B^0 \to \pi^0\pi^0$ decays reconstructed in 2019-2020 Belle II data*. 2021. DOI: 10.48550/ARXIV.2107.02373. URL: https://arxiv.org/abs/2107.02373.

[59] A. J. Bevan et al. "The Physics of the B Factories". In: *The European Physical Journal C* 74.11 (Nov. 2014). DOI: 10.1140/epjc/s10052-014-3026-9. URL: https://doi.org/10.1140%2Fepjc%2Fs10052-014-3026-9.

[60] Particle Data Group. 2022. URL: https://pdglive.lbl.gov/BranchingRatio.action?desig=8&parCode=S032&home=MXXX035.

[61] Belle II collaboration. *Belle II tracks performance*. 2022. URL: https://confluence.desy.de/pages/viewpage.action?spaceKey=BI&title=Tracking+and+Vertexing+Performance.

[62] Valerio Bertacchi et al. "Track finding at Belle II". In: *Computer Physics Communications* 259 (Feb. 2021), p. 107610. DOI: 10.1016/j.cpc.2020.107610. URL: https://doi.org/10.1016%2Fj.cpc.2020.107610.

[63] G. Bonvicini et al. "Updated measurements of absolute D^+ and D^0 hadronic branching fractions and $\sigma(e^+e^- \to D\overline{D})$ at $E_{\rm cm} = 3774$ MeV". In: *Phys. Rev. D* 89.7 (2014). [Erratum: Phys.Rev.D 91, 019903 (2015)], p. 072002. DOI: 10.1103/PhysRevD.89.072002. arXiv: 1312.6775 [hep-ex].