Università degli Studi di Pisa

Dipartimento di Fisica "E. Fermi"
PhD Program in Physics – XXXVII Cycle

PhD Thesis

Search for CP violation in the $D^0 \to \pi^+\pi^-\pi^0$ decay at the Belle II experiment

Candidate
Ludovico Massaccesi

Supervisors
Francesco Forti
Giulia Casarosa

August 28, 2025

Search for CP violation in the $D^0 \to \pi^+\pi^-\pi^0$ decay at the Belle II experiment, PhD thesis in Physics
Ludovico Massaccesi – August 28, 2025

ABSTRACT

The Standard Model of particle physics is the most successfully tested theory describing nature at a fundamental level. Still, it is unable to provide answers to several fundamental questions, for instance the origin of the matter-antimatter asymmetry of the universe: the amount of the charge-parity (CP) symmetry violation necessary to explain it is, within the assumptions of the Standard Model, more than what has been observed so far. There must exist, therefore, undiscovered sources of CP violation.

CP violation was first observed in the K system in 1964 [1], then in the B system in 2001 [2] and in the B_s system in 2007 [3], and finally in the D system only in 2019 [4]. The measurement in the D system is particularly interesting because the *charm* quark is the only up-type quark that exhibits mixing and CP violation.

In this thesis, I measure the time-integrated CP asymmetry of the D^0 meson decay to three pions $\pi^+\pi^-\pi^0$. The Standard Model predicts, with large theoretical uncertainties, that the CP asymmetry in this decay is at or below the level of current experimental sensitivity; however, extensions of the Standard Model may enhance it. To perform the measurement, I use the sample of electron-positron collisions collected by the Belle II experiment between 2019 and 2022 with a collision energy at or around the $\Upsilon(4S)$ resonance, corresponding to an integrated luminosity of $428\,\mathrm{fb}^{-1}$.

My result, $\mathcal{A}_{CP} = (0.29 \pm 0.27(\mathrm{stat}) \pm 0.13(\mathrm{sys}))\%$, is compatible with conservation of the CP symmetry, and improves the precision of the current world's best determination by 34%. With an increase in integrated luminosity of only 10%, most of the improvement is due to the superior Belle II detector performance compared to previous B-factory experiments, and the improved selection criteria and analysis strategy. At the time of writing, this result is considered preliminary by the Belle II collaboration.

CONTENTS

ACRONYMS

ARICH	aerogel ring-imaging Čerenkov detector
basf2	Belle II analysis software framework
BCS	best candidate selection
BWD	backward
CDC	central drift chamber
CKM	Cabibbo-Kobayashi-Maskawa
CM	center of mass
DAQ	data acquisition
DP	Dalitz plot
ECL	electromagnetic calorimeter
FSP	final-state particle
FWD	forward
GIM	Glashow-Iliopoulos-Maiani
HER	high-energy ring
HLT	high-level trigger
IP	interaction point
IR	interaction region
KLM	K_L^0 and μ detector
L1	level 1
LER	low-energy ring
LINAC	linear accelerator
LS1	long shutdown 1
MC	Monte Carlo
MPV	most probable value
MVA	multivariate analysis
NN	neural network
PID	particle identification
PDF	probability density function
PoCA	point of closest approach
PXD	pixel detector
QCD	quantum chromodynamics

QED	quantum electrodynamics
RF	radio frequency
SM	Standard Model
SVD	silicon vertex detector
TOP	time of propagation detector

INTRODUCTION

The Standard Model of particle physics is the most successfully tested theory describing nature at a fundamental level; still, it is unable to provide answers to several fundamental questions. Among these, the origin of the matter-antimatter asymmetry of the universe: the amount of the charge-parity (*CP*) symmetry violation that is necessary to explain the observations at the cosmological scale is, within the assumptions of the Standard Model, more than what was observed so far at the elementary particle level. There must exist, therefore, undiscovered sources of *CP* violation.

While *CP* violation in the *strange* and *beauty* quark sectors has been established long ago, and extensively studied, significant results in the *charm* sector are much fewer and more recent. This is due to the much smaller magnitude of *CP* asymmetry and neutral meson mixing in this sector, which are suppressed by the Glashow-Iliopoulos-Maiani mechanism and by the small couplings to the third generation of quarks. Indeed, mixing was only established in 2007 [5, 6], and *CP* violation was only observed in one channel, only by the LHCb experiment, in 2019.

The LHCb collaboration first observed a non-zero difference between the *CP* asymmetries of $D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$ decays [4]; then, only in 2023, restricted the majority of the *CP* violation to $D^0 \to \pi^+\pi^-$ by measuring a $D^0 \to K^+K^-$ *CP* asymmetry that is compatible with zero [7]. It is not yet completely clear whether these measurements are compatible with the Standard Model, as the predictions for the charm sector are plagued by large theoretical uncertainties. Other observations have been conducted to restrict the possible origin of this *CP* asymmetry, including measurements of the *CP* asymmetry in $D^0 \to \pi^0\pi^0$ and $D^+ \to \pi^+\pi^0$ decays at the Belle II experiment [8, 9]. In this context, both improving existing *CP* asymmetry measurements and searching for *CP* violation in unexplored channels become very important.

In this thesis, I measure the time-integrated *CP* asymmetry of the D^0 meson decay to three pions $\pi^+\pi^-\pi^0$. The Standard Model predicts, with large theoretical uncertainties, that the *CP* asymmetry in this decay is at or below the level of current experimental sensitivity; however, extensions of the Standard Model may enhance it, even by relatively large factors. Therefore, the purpose of this measurement is both testing the Standard Model, and searching for possible new physics contributions.

To perform the measurement, I use the sample of electron-positron collisions collected by the Belle II experiment between 2019 and 2022 with a collision energy at or around the $\Upsilon(4S)$ resonance, corresponding to an integrated luminosity of $428\,\mathrm{fb}^{-1}$. With only a 10% increase in integrated luminosity compared to the current world's best determination, which was performed by the *BABAR* collaboration in 2007, I adopt a different analysis strategy to improve the precision. I employ much looser selection criteria, exploiting the superior performance of the Belle II detector, but also allowing my sample of signal decays to have much lower purity compare to previous works. I then use a more advanced fit, that models several background contributions, to extract the signal asymmetry, and employ $D^0 \to K\pi$ decays as a control channel to correct for detector-induced asymmetry biases.

I find no evidence of *CP* violation, nor any tension with previous measurements, and improve the precision of the current world's best determination of the *CP* asymmetry of this channel by 34%. At the time of writing, this result is yet to be approved by the Belle II collaboration, therefore it may only be shown within the context of this thesis.

The thesis is organized as follows:

- in Chapter 1 I briefly introduce the Standard Model of particle physics, as well as CP violation and mixing in the charm sector from the theoretical and experimental points of view;
- in Chapter 2 I briefly introduce B-factory experiments, I describe the Belle II experiment and the SuperKEKB particle collider, and I discuss particle reconstruction in the Belle II software;
- in Chapter 3 I introduce the general strategy of my work, with a particular focus on experimentally-induced asymmetries;
- in Chapter 4 I describe the reconstruction of the sample of signal $D^0 \to \pi^+\pi^-\pi^0$ decays, detailing the selection criteria I use and their motivation;
- in Chapter 5 I describe the various signal and background components of the signal sample, the models I use to describe them, and the fit procedure I use to extract the signal asymmetries;
- in Chapter 6 I describe how I reconstruct the $D^0 \to K\pi$ control samples, and how I use them to compute the corrections for the experimentally-induced asymmetries;
- in Chapter 7 I show the final result, and discuss the consistency checks I perform to validate my strategy; and
- in Chapter 8 I describe the various source of systematic uncertainty that I consider, and the method I use to estimate them

Finally, Chapter 9 concludes the thesis, with an outlook on future perspective of this measurement at the Belle II experiment, and possible extensions.

In addition to this analysis work, during the three years of this PhD course I actively took part in the operation and commissioning of the Belle II detector, particularly for what concerns the silicon vertex detector which I describe in Chapter 2. Furthermore, I took part in the development of the vertex detector upgrade from the point of view of both hardware characterization and performance studies. Finally, I contributed to the development of the Belle II software, which is used for data-taking, simulation, and analysis in this and all other Belle II works, especially for what concerns the reconstruction of charged particles.

1 THEORETICAL OVERVIEW AND MOTIVATION

The Standard Model (SM) of particle physics is, so far, the most successfully tested theory describing nature at a fundamental level. Still, despite its tremendous accuracy in predicting the interactions among particles, it is unable to provide answers to several fundamental questions. Among these, at the cosmological scale, is the origin of the matter-antimatter asymmetry of the universe: the violation of the charge-parity (*CP*) symmetry is necessary to explain it, but the amount of *CP* violation observed so far is insufficient, within the assumptions of the SM, to match the observations at the cosmological scale. There must exist, therefore, undiscovered sources of *CP* violation.

In this thesis, I search for *CP* violation in charmed particle decays, specifically in the $D^0 \to \pi^+\pi^-\pi^0$ decay channel[1].

In Section 1.1 I briefly introduce the SM; in Section 1.2 I discuss the *CP* symmetry and its violation within the SM, introducing the formalism needed to describe it; in Section 1.3 I focus on *CP* violation in the charm sector, recalling recent experimental results; finally, in Section 1.4 I review the present experimental status of *CP* asymmetry measurements in the $D^0 \to \pi^+\pi^-\pi^0$ decay channel.

1.1 THE STANDARD MODEL OF PARTICLE PHYSICS

The SM is a theory describing the elementary constituents of matter, and three of the four fundamental interactions among them: the electromagnetic, the weak, and the strong interactions (leaving out the gravitational interaction). In this framework, matter is made of elementary particles following the Fermi-Dirac statistics, or *fermions*, while their interactions are mediated by the exchange of elementary particles following the Bose-Einstein statistics, or *bosons*. Presently, the known particles include 12 fermions, each with its corresponding anti-fermion, and five types of bosons [10]. This is summarized in Figure 1.1.

FERMIONS The fermions can be divided into *leptons* and *quarks*. Both are organized in three *generations*, each generation made of a doublet of particles. Fermions in the three generations have very different masses, as shown in Figure 1.1; the reason for this *mass hierarchy* is one of the open questions that the SM has left unanswered so far. All known elementary fermions have *spin* $\frac{1}{2}\hbar$.

Lepton doublets are each made of a charged lepton (in the three generations: electron e^-, muon μ^-, and tauon τ^-) and its corresponding *neutrino* (ν_e, ν_μ, ν_τ). Charged leptons have electric charge $-e$ and interact electromagnetically and weakly, while neutrinos have zero electric charge and therefore only interact weakly. A different lepton family number ($L_e = 1$, $L_\mu = 1$, $L_\tau = 1$) is associated with each lepton: these are empirically observed to be conserved, however the discovery of neutrino oscillations provided indirect evidence of their violation.

1 In this work, *CP*-conjugate decays are implied unless otherwise noted.

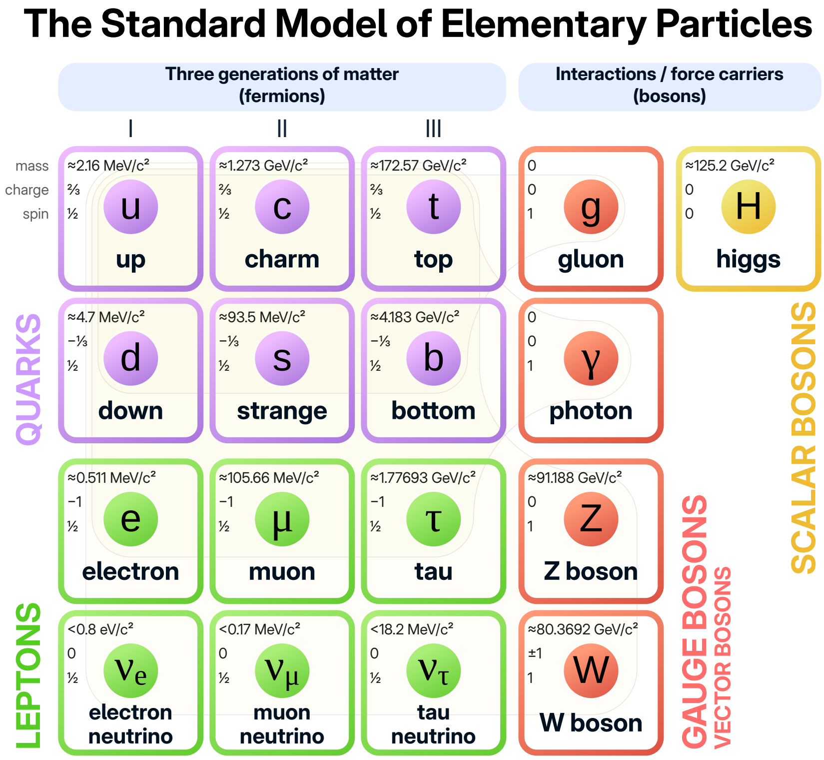


Figure 1.1: The particles of the SM.

Each lepton has a corresponding anti-lepton (e^+, μ^+, τ^+, and $\overline{\nu}_e$, $\overline{\nu}_\mu$, $\overline{\nu}_\tau$) with the same mass and opposite quantum numbers.

Quark doublets are each made of an *up-type* quark (up u, charm c, and top or truth t) and a *down-type* quark (down d, strange s, and bottom or beauty b). Up-type quarks have electric charge $\frac{2}{3}e$, while down-type quarks have $-\frac{1}{3}e$. All quarks interact through all three interactions. Each quark has its own *flavor*, which is conserved by the strong and electromagnetic interactions, but not by the weak interaction. Quarks also have an additional quantum number, the *color*, which is the charge of the strong interaction, and can assume three values called *red*, *green*, and *blue*. Each quark has a corresponding anti-quark ($\overline{u}$, $\overline{c}$, $\overline{t}$ and $\overline{d}$, $\overline{s}$, $\overline{u}$) with the same mass and opposite quantum numbers.

Unlike leptons, quarks are never observed isolated due to the phenomenon dubbed *confinement*, which is a consequence of their color charge and the strong interaction. Confinement demands that the color charge be neutralized by forming bound states of quarks that result in a color singlet under the strong interaction gauge symmetry $SU(3)_C$, which I discuss in Section 1.1.1. There are mainly two kinds of such states: *mesons*, made of a quark and an anti-quark; and *baryons*, made of three quarks (with the corresponding anti-baryons, made of three anti-quarks). More recently, a number of exotic four- or five-quark bound states have been established.

BOSONS Bosons act as *mediators* of the interactions by being exchanged between fermions. As such, they are strictly related to the fundamental interactions, and originate from local gauge group invariance, as discussed in Section 1.1.1.

The photon γ is the vector (*i.e.* spin 1) boson mediating the electromagnetic interaction. It is massless and chargeless, which allows it to propagate at any distance without decaying.

The W^+, W^-, and Z^0 vector bosons are the mediators of the weak interaction. They are massive, which gives them the possibility to decay, causing the weak interaction to happen only at short distance.

The gluons g are the vector bosons mediating the strong interaction. They exist in eight states carrying different color charge. They are massless; however, due to their charge, they quickly interact with quarks or other gluons, causing the strong interaction to happen only at very short distance. This is also the reason behind quark confinement.

Finally, the scalar (*i.e.* spin-zero) Higgs boson H is responsible for the mechanism that generates the masses of all massive particles, including itself.

1.1.1 *The SM Lagrangian*

The SM is a quantum gauge theory based on the $SU(3)_C \otimes SU(2)_I \otimes U(1)_Y$ symmetry group. $SU(3)_C$ represents the color symmetry, related to the strong interaction, while $SU(2)_I$ and $U(1)_Y$ represent the *weak isospin* and *weak hypercharge* symmetries, related to the *electroweak* theory (the unified description of electromagnetic and weak interactions). In this framework, bosons arise from the gauge fields associated with the symmetry groups, while fermions can be interpreted as irreducible representations of the groups.

Before delving into the SM Lagrangian, let us introduce the notation. The fields associated with the symmetry groups, which correspond directly to the generators of the group transformations, are:

- B_μ (where μ refers to the space and time components) for $U(1)_Y$;
- W^a_μ (with $a = 1, 2, 3$) for $SU(2)_I$; and
- G^A_μ (with $A = 1, \ldots, 8$) for $SU(3)_C$.

G^A_μ is directly related to the eight gluons. We shall see later, instead, how the other two relate to the photon, and the $W^\pm$ and Z^0 bosons.

The weak interaction only couples to left-handed fermions. Therefore, it is necessary to group them into multiplets of *spinors* of defined *chirality*. The notation for this is presented in Table 1.1. Note that this notation refers to massless interaction eigenstates: the masses of the fermions, and their mass eigenstates, shall be introduced later.

		Weak isospin I_3	Weak hypercharge Y	Color C
$Q_L^{\text{int}} = \begin{pmatrix} U_L^{\text{int}} \\ D_L^{\text{int}} \end{pmatrix}$	$U_L^{\text{int}} = (u_L^{\text{int}}, c_L^{\text{int}}, t_L^{\text{int}})$ $D_L^{\text{int}} = (d_L^{\text{int}}, s_L^{\text{int}}, b_L^{\text{int}})$	Doublet $\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$	$1/3$	Triplet
	$U_R^{\text{int}} = (u_R^{\text{int}}, c_R^{\text{int}}, t_R^{\text{int}})$	Singlet (0)	$+4/3$	
	$D_R^{\text{int}} = (d_R^{\text{int}}, s_R^{\text{int}}, b_R^{\text{int}})$	Singlet (0)	$-2/3$	
$L_L^{\text{int}} = \begin{pmatrix} \nu_L^{\text{int}} \\ \ell_L^{\text{int}} \end{pmatrix}$	$\nu_L^{\text{int}} = (\nu_{e,L}^{\text{int}}, \nu_{\mu,L}^{\text{int}}, \nu_{\tau,L}^{\text{int}})$ $\ell_L^{\text{int}} = (e_L^{\text{int}}, \mu_L^{\text{int}}, \tau_L^{\text{int}})$	Doublet $\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$	-1	Singlet
	$\ell_R^{\text{int}} = (e_R^{\text{int}}, \mu_R^{\text{int}}, \tau_R^{\text{int}})$	Singlet (0)	-2	

Table 1.1: Fermion representations as spinors of defined chirality. The L and R subscripts indicate the left or right chirality of the field ($\psi_{R,L} = (1 \pm \gamma_5)\psi$), while the int superscript indicates interaction eigenstates. Right-handed neutrinos are not included, since there is only indirect evidence of their existence.

The SM lagrangian can be organized in five parts:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{kin}} + \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}} + \mathcal{L}_H + \mathcal{L}_Y \,. \tag{1.1}$$

$\mathcal{L}_{\text{kin}}$ This is the *kinetic* part of the gauge fields

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4} G_{\mu\nu}^A G_A^{\mu\nu} \tag{1.2}$$

with

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \,, \tag{1.3a}$$
$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + 2g\varepsilon^{abc} W_\mu^b W_\nu^c \,, \tag{1.3b}$$
$$G_{\mu\nu}^A = \partial_\mu G_\nu^A - \partial_\nu G_\mu^A + g_s f^{ABC} G_\mu^B G_\nu^C \,, \tag{1.3c}$$

where g and g_s are the weak and strong coupling constants, ε^{abc} is the Levi-Civita tensor, which gives the structure constants of $SU(2)_I$, and f^{ABC} are the structure constants of $SU(3)_C$. This part also contains the self-interaction terms of the gluons (through f^{ABC}). The kinetic terms of the fermions are instead customarily included in the electroweak interaction part.

$\mathcal{L}_{\text{QCD}}$ This is the quantum chromodynamics (QCD) part, which describes quark-gluon interactions

$$\mathcal{L}_{\text{QCD}} = \sum_\psi g_s \overline{\psi} G_\mu^a T^a \gamma^\mu \psi \quad \text{for} \quad \psi = Q_L^{\text{int}}, U_R^{\text{int}}, D_R^{\text{int}} \,, \tag{1.4}$$

where T^a are the generators of the $SU(3)_C$ group, and γ^μ are the Dirac matrices. I will not describe QCD in detail, as it is not so relevant for this work.

$\mathcal{L}_{\text{EW}}$ This is the electroweak part, which describes the interaction between the B_μ and W_μ^a fields and the fermions

$$\mathcal{L}_{\text{EW}} = \sum_\psi \overline{\psi} i \gamma^\mu \mathcal{D}_\mu \psi \quad \text{for} \quad \psi = Q_L^{\text{int}}, U_R^{\text{int}}, D_R^{\text{int}}, L_L^{\text{int}}, \ell_R^{\text{int}} \,, \tag{1.5}$$

where I introduced the covariant derivative

$$\mathcal{D}_\mu = \partial_\mu - i\frac{g'}{2}YB_\mu - i\frac{g}{2}\tau_L^a W_\mu^a\,. \tag{1.6}$$

τ_L^a are the Pauli matrices, *i.e.* the generators of $SU(2)_I$, which only act on the left-handed chiral fermions. g' is the electromagnetic coupling constant. Y is the weak hypercharge, *i.e.* the generator of $U(1)_Y$, which is related to the electric charge Q and weak isospin I by

$$Q = I_3 + \frac{1}{2}Y\,. \tag{1.7}$$

THE HIGGS MECHANISM So far, there is no term in the Lagrangian accounting for the masses of the fermions and the bosons. It is not possible to introduce such terms while respecting gauge symmetry without introducing an external field. This is what happens under the Higgs mechanism: a new *ad-hoc* scalar field Φ, the Higgs fields, is added to the Lagrangian

$$\Phi = \begin{pmatrix}\phi_+\\ \phi_0\end{pmatrix} \qquad \mathcal{L}_H = (\mathcal{D}_\mu\Phi)^\dagger(\mathcal{D}^\mu\Phi) - V(\Phi) \qquad V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2\,. \tag{1.8}$$

Φ is a weak isospin doublet with weak hypercharge $Y = 1$. If we choose $\mu^2 < 0$ and $\lambda > 0$, the potential $V(\Phi)$ reduces to

$$V(\Phi) = \lambda\left(\Phi^\dagger\Phi - \frac{v}{2}\right)^2, \tag{1.9}$$

which has a minimum for $\Phi^\dagger\Phi = \mu^2/(2\lambda)$. This means that the *vacuum expectation value* is non-zero: $\langle\Phi\rangle = v/\sqrt{2} = -\sqrt{\mu^2/(2\lambda)}$. This is called *spontaneous symmetry breaking*, as the ground state is not symmetric under the symmetry groups. The magnitude of Φ is fixed, but its direction is not: infinite equivalent ground states exist. However, it is always possible to apply a gauge transformation so that only the real part of ϕ_0 is non-zero; thus

$$\langle\Phi\rangle = \begin{pmatrix}0\\ v/\sqrt{2}\end{pmatrix} \quad\Rightarrow\quad \Phi = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v + h(x)\end{pmatrix}, \tag{1.10}$$

and the Higgs boson arises from h. v is the only dimensional parameter of the SM, and its value is measured to be $(G_F\sqrt{2})^{-1/2} \sim 246\,\mathrm{GeV}$, where G_F is the Fermi coupling constant. The quadratic term in the potential generates the mass of the Higgs boson, $m_H = v\sqrt{2\lambda} \sim 125\,\mathrm{GeV}/c^2$ [11].

Spontaneous symmetry breaking reduces $SU(2)_I \otimes U(1)_Y$ to the electromagnetic $U(1)_Q$. This happens through the covariant derivative in the $(\mathcal{D}_\mu\Phi)^\dagger(\mathcal{D}^\mu\Phi)$ terms of the Higgs Lagrangian, where the B_μ and W_μ^a appear. Diagonalizing, the fields associated with the $W^\pm$ bosons ($W_\pm^\mu$), Z^0 boson (Z^μ), photon (A^μ) appear:

$$W_\pm^\mu = \frac{W_1^\mu \mp iW_2\mu}{2}, \qquad \begin{pmatrix}Z^\mu\\ A^\mu\end{pmatrix} = \begin{pmatrix}\cos\theta_W & -\sin\theta_W\\ \sin\theta_W & \cos\theta_W\end{pmatrix}\begin{pmatrix}W_3^\mu\\ B^\mu\end{pmatrix}, \tag{1.11}$$

where the θ_W is the Weinberg angle, related to the weak and electromagnetic coupling constants by $e = g\sin\theta_W = g'\cos\theta_W$, where e is the elementary charge. Also, the boson masses can be predicted by

$$m_W = \frac{1}{2}vg, \qquad m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}\,, \tag{1.12}$$

while the photon remains massless because of the residual $U(1)_Q$ symmetry.

After spontaneous symmetry breaking, $\mathcal{L}_{\text{EW}}$ can be rewritten in terms of the interaction between physical bosons and fermions:

$$\mathcal{L}_{\text{EW}} = \mathcal{L}_{\text{free}} + \mathcal{L}_{\text{CC}} + \mathcal{L}_{\text{NC}}\,, \tag{1.13a}$$

$$\mathcal{L}_{\text{free}} = \sum_{\psi} \overline{\psi} i\gamma^{\mu}\partial_{\mu}\psi \quad \text{for} \quad \psi = Q_L^{\text{int}}, U_R^{\text{int}}, D_R^{\text{int}}, L_L^{\text{int}}, \ell_R^{\text{int}}\,, \tag{1.13b}$$

$$\mathcal{L}_{\text{CC}} = \frac{g}{\sqrt{2}}(J_{\mu}^{+} W_{-}^{\mu} + \text{h.c.}), \qquad \mathcal{L}_{\text{NC}} = eJ_{\mu}^{\text{em}} A^{\mu} + \frac{g}{2\cos\theta_W} J_{\mu}^{0} Z^{\mu}\,. \tag{1.13c}$$

$\mathcal{L}_{\text{free}}$ is the free motion kinetic term of the fermions. $\mathcal{L}_{\text{CC}}$ and $\mathcal{L}_{\text{NC}}$ are the charged and neutral current terms, where I introduced the weak charged current J_{μ}^{+} (associated with $W^{\pm}$), the electromagnetic current J_{μ}^{em} (associated with the photon), and the weak neutral current J_{μ}^{0} (associated with the Z^0). The charged current only interacts with the left-handed doublets, coupling particles with opposite weak isospin

$$J_{\mu}^{+} = i\overline{U}_L^{\text{int}}\gamma_{\mu} D_L^{\text{int}} + i\overline{\nu}_L^{\text{int}}\gamma_{\mu}\ell_L^{\text{int}}\,, \tag{1.14}$$

and corresponds to the charged $V - A$ interaction. The electromagnetic and neutral currents, instead, couple particles of both left and right chirality with particles of the same chirality and weak isospin

$$J_{\mu}^{\text{em}} = \frac{2}{3} i\overline{U}^{\text{int}}\gamma_{\mu} U^{\text{int}} - \frac{1}{3} i\overline{D}^{\text{int}}\gamma_{\mu} D^{\text{int}} + i\overline{\ell}^{\text{int}}\gamma_{\mu}\ell^{\text{int}}\,, \tag{1.15a}$$

$$J_{\mu}^{0} = i\overline{U}^{\text{int}}(c_V^u - c_A^u\gamma^5)\gamma_{\mu} U^{\text{int}} + i\overline{D}^{\text{int}}(c_V^d - c_A^d\gamma^5)\gamma_{\mu} D^{\text{int}} \tag{1.15b}$$

$$+\, i\overline{\ell}^{\text{int}}(c_V^{\ell} - c_A^{\ell}\gamma^5)\gamma_{\mu}\ell^{\text{int}} + i\overline{\nu}_L^{\text{int}}(1-\gamma^5)\gamma_{\mu}\nu_L^{\text{int}}\,, \tag{1.15c}$$

where the spinors without L or R suffix contain both chiral components, and the coefficients c_V and c_A represent the vector and pseudovector couplings of the fermions f:

$$c_V^f = I_3^f - 2Q^f \sin^2\theta_W, \qquad c_A^f = I_3^f\,. \tag{1.16}$$

Finally, $\mathcal{L}_{\text{Y}}$ contains the Yukawa couplings between the Higgs field and the fermions, which are responsible for the masses of the latter:

$$\mathcal{L}_{\text{Y}} = -\sum_{i,j=1}^{3} \left(Y_{ij}^{u}\overline{Q}_{L,i}^{\text{int}}\Phi^{*} U_{R,j}^{\text{int}} + Y_{ij}^{d}\overline{Q}_{L,i}^{\text{int}}\Phi D_{R,j}^{\text{int}} + Y_{ij}^{\ell}\overline{L}_{L,i}^{\text{int}}\Phi\ell_{R,j}^{\text{int}} + \text{h.c.} \right), \tag{1.17}$$

where i and j loop over the generations, and $Y_{ij}^{u,d,\ell}$ are the 3×3 Yukawa couplings matrices. This Yukawa part does not contain any term for neutrino masses, despite neutrino oscillations being evidence that neutrino are massive. There are two possible solutions for this problem (the Dirac and Majorana mass terms). However, it presently remains an open question. I will not discuss this in any more detail, as it is not relevant for this work.

$Y^{u,d,\ell}$ can be diagonalized using six unitary matrices $V_{L,R}^{u,d,\ell}$:

$$Y^{u,d,\ell} \to V_L^{u,d,\ell} Y^{u,d,\ell} (V_R^{u,d,\ell})^{\dagger}\,. \tag{1.18}$$

By doing this and substituting Φ with its vacuum expectation value in $\mathcal{L}_{\text{Y}}$, we find the masses of the fermions. The fact that $Y^{u,d,\ell}$ are not diagonal in the interaction basis is symptom of the fact that the interaction eigenstates U^{int} , D^{int}, ℓ^{int}, and ν^{int} are *not* eigenstates of the free-motion Lagrangian that governs the propagation of particles through space. Let us call the latter eigenstates, *i.e.* the mass eigenstates, U, D, ℓ and ν: this will be defined by the relations

$$D_{L,i} = V^d_{L,ij} D^{\text{int}}_{L,j}, \qquad D_{R,i} = V^d_{R,ij} D^{\text{int}}_{R,j}, \tag{1.19a}$$

$$U_{L,i} = V^u_{L,ij} U^{\text{int}}_{L,j}, \qquad U_{R,i} = V^u_{R,ij} U^{\text{int}}_{R,j}, \tag{1.19b}$$

$$\ldots \tag{1.19c}$$

This change of basis has direct impact on the charged currents, but not on the neutral ones: for instance, for the electromagnetic current of the up-type quarks

$$J^{\text{em},u}_\mu = \sum_{i=1}^{3} \frac{2}{3} i\overline{U}^{\text{int}}_i \gamma_\mu U^{\text{int}}_i = \sum_{i=1}^{3} \frac{2}{3} i\overline{U}_i \gamma_\mu \left(\sum_{j=1}^{3} V^u_{ij} (V^u_{ji})^\dagger \right) U_i = \sum_{i=1}^{3} \frac{2}{3} i\overline{U}_i \gamma_\mu U_i \,, \tag{1.20}$$

where the summation cancels due to the unitary of V^u. A similar cancellation happens for the weak neutral current, but not for the charged current: for instance

$$J^{q+}_\mu = \sum_{i,j=1}^{3} i\overline{U}^{\text{int}}_{L,i} \gamma_\mu D^{\text{int}}_{L,j} = \sum_{i,j=1}^{3} i\overline{U}_{L,i} \gamma_\mu \left(\sum_{k=1}^{3} V^u_{L,ik} (V^d_{L,kj})^\dagger \right) D_{L,j} \,. \tag{1.21}$$

The $V_{ij} = V^u_{L,ik}(V^d_{L,kj})^\dagger$ matrix is known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix. It allows to predict charged weak interactions while working with the mass eigenstates of the quarks, which are customarily used. An analogous matrix, the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix, exists for the lepton sector; I will not describe it any further though, as it is not relevant for this work.

1.1.2 *The CKM matrix*

A 3×3 complex matrix has, in general, 18 free parameters. The unitarity condition of the CKM matrix reduces its parameters to 9: three angles and six complex phases. Furthermore, five of these phases can be absorbed in the definition of the quark fields, leaving only three angles and one phase.

The CKM matrix can be written as

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{1.22}$$

where $V_{q_1 q_2}$ indicates the term relative to the coupling between the quarks q_1 and q_2. Many parametrization of the CKM matrix exist. A "standard" choice is to express it as the product of three rotation matrices, with the complex phase added in one of these [12]:

$$V_{\text{CKM}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & \sin\theta_{23} \\ 0 & -\sin\theta_{23} & \cos\theta_{23} \end{pmatrix} \begin{pmatrix} \cos\theta_{13} & 0 & e^{-i\delta}\sin\theta_{13} \\ 0 & 1 & 0 \\ -e^{i\delta}\sin\theta_{13} & 0 & \cos\theta_{13} \end{pmatrix} \begin{pmatrix} \cos\theta_{12} & \sin\theta_{12} & 0 \\ -\sin\theta_{12} & \cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1.23}$$

where θ_{ij} are the mixing angles among the generations (which can all be chosen in the first quadrant for ease of notation), and δ is the irreducible complex phase. Using the experimental result that $\theta_{13} \ll \theta_{23} \ll \theta_{12} \ll 1$, Lincoln Wolfenstein defined the parametrization [13, 14]

$$V_{\text{CKM}} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta + i\eta\lambda^2/2) \\ -\lambda & 1-\lambda^2/2 - i\eta A^2\lambda^4 & A\lambda^2(1+i\eta\lambda^2) \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^5)\,, \quad (1.24)$$

which is unitary to all orders in λ. The four independent parameters λ, A, η and ρ are defined so that they are all of $\mathcal{O}(1)$ with the exception of the sine of the Cabibbo angle $\lambda \sim 0.23$ [15], which acts as the small parameter for the series expansion:

$$\lambda = \sin\theta_{12} = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}, \quad A = \frac{\sin\theta_{23}}{\lambda^2} = \frac{1}{\lambda}\left|\frac{V_{cb}}{V_{us}}\right|, \quad \rho + i\eta = \frac{e^{i\delta}\sin\theta_{13}}{A\lambda^3} = \frac{V_{ub}^*}{A\lambda^3}\,. \quad (1.25)$$

It is also common to replace ρ and η with the CKM-phase-independent

$$\bar{\rho} + i\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\,. \quad (1.26)$$

UNITARITY TRIANGLES Precise measurement of the CKM parameters is of fundamental importance for testing the SM. The unitarity of the CKM matrix imposes two strong requirements

$$\sum_i V_{ij}V_{ik}^* = \delta_{jk}, \qquad \sum_i V_{ji}V_{ki}^* = \delta_{jk}\,. \quad (1.27)$$

The six vanishing sums can be interpreted as triangular relations, and represented as triangles in the complex plane. The least degenerate triangle is given by

$$\frac{V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0\,, \quad (1.28)$$

and is shown in Figure 1.2. Many different measurements are performed to test the relations represented by these triangles in terms of its angles and its free sides; they are then combined to over-constrain the free vertex: see, for example, [16, 17].

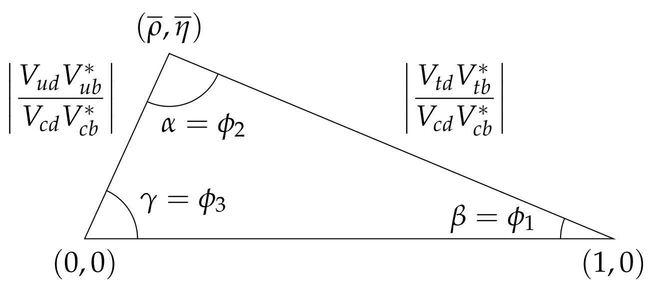


Figure 1.2: The least degenerate unitarity triangle, with the two naming conventions for the angles.

The triangle relevant for the charm sector is given by

$$V_{ud}V_{cb}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* = 0\,. \quad (1.29)$$

It is close to degeneracy: two sides are of $\mathcal{O}(\lambda)$, while the third is of $\mathcal{O}(\lambda^5)$. This is related to the small magnitude of *CP* violation in the charm sector, which I discuss in Section 1.3.

1.2 *CP* VIOLATION AND NEUTRAL MESON MIXING

Parity (P) is the operation of spatial inversion through the origin $\vec{x} \to -\vec{x}$. Even if they are point-like, elementary particles exhibit an intrinsic parity. Parity conservation arises naturally in quantum electrodynamics (QED) and in QCD due to the vector nature of such interactions. However, the weak interaction does *not* conserve P, due to the simultaneous presence of vector and pseudovector terms.

Charge conjugation (C) is the operation that inverts all quantum numbers. Essentially, it replaces each particle with its antiparticle. While both QED and QCD conserve C, the weak interactions do not.

When P and C violations were first observed in the β decay [18–21], it was proposed that the combination CP of the two broken symmetries remains unbroken [22]. Alas, also CP is only approximately conserved, though to a much better degree than C and P [1].

CP violation is actually a fundamental ingredient for a convincing explanation of the prevalence of matter over antimatter at the cosmological scale [23]. One can identify three possible sources of CP violation: one in the quark sector, which is discussed later; one in the strong interaction, which, however, is observed to be vanishingly small; and one in the lepton sector, for which only upper limits exist so far due to the elusiveness of neutrinos. The observed CP violation in the quark sector is insufficient, within the SM framework, to explain the matter-antimatter asymmetry of the universe. This could be, however, resolved by the presence of sufficient, yet-to-be-observed CP violation in the lepton sector. If this, too, is found by experiment to be insufficient, new sources of CP violation will have to be introduced by new physics beyond the SM.

CP VIOLATION IN THE QUARK SECTOR Under the SM, CP violation in the quark sector can be attributed to a single parameter: the irreducible complex phase δ of the CKM matrix. The reason for this may be seen by considering the amplitude for a decay process $M \to f$, and its CP conjugate $\overline{M} \to \bar{f}$

$$\mathcal{A} = \mathcal{A}(M \to f) = |\mathcal{A}|\, e^{i\theta} e^{i\phi}, \qquad \bar{\mathcal{A}} = \mathcal{A}(\overline{M} \to \bar{f}) = |\mathcal{A}|\, e^{i\theta} e^{-i\phi}, \tag{1.30}$$

where I used $|\mathcal{A}|$ for the magnitude (invariant under CP), and I split the phase in non-CKM part θ (invariant under CP, often called "strong phase") and CKM part ϕ (which changes sign under CP, and is often called "weak phase"). Physical rates depend on $|\mathcal{A}|^2$, so there is no difference in this scenario. However, consider a process that can happen through two different contributions, $M \xrightarrow{1} f$ and $M \xrightarrow{2} f$, each with its own amplitude

$$\mathcal{A} = |\mathcal{A}_1|\, e^{i\theta_1} e^{i\phi_1} + |\mathcal{A}_2|\, e^{i\theta_2} e^{i\phi_2}, \tag{1.31a}$$

$$\bar{\mathcal{A}} = |\mathcal{A}_1|\, e^{i\theta_1} e^{-i\phi_1} + |\mathcal{A}_2|\, e^{i\theta_2} e^{-i\phi_2}, \tag{1.31b}$$

$$\Rightarrow |\mathcal{A}|^2 - |\bar{\mathcal{A}}|^2 = 4|\mathcal{A}_1|\,|\mathcal{A}_2|\, \sin(\theta_1 - \theta_2) \sin(\phi_1 - \phi_2)\,. \tag{1.31c}$$

A difference arises, thus CP is violated.

CP violation is also linked to the area of the unitarity triangles. All triangles have the same area, which is half the Jarlskog invariant

$$J = \cos\theta_{12}\cos^2\theta_{13}\cos\theta_{23}\sin\theta_{12}\sin\theta_{13}\sin\theta_{23}\sin\delta \tag{1.32a}$$

$$= \mathrm{Im}[V_{ij}V_{kl}V_{il}^*V_{kj}^*] \quad \text{for} \quad (i,k) \text{ and } (j,l) = (1,2),(2,3),(3,1) \tag{1.32b}$$

$$\sim 3\times 10^{-5}\,. \tag{1.32c}$$

The area thus becomes a measurement for *CP* violation.

1.2.1 *Neutral meson mixing*

There are four known neutral meson systems that undergo the phenomenon known as *mixing*. These are the K^0–$\bar{K}^0$, B^0–$\bar{B}^0$, B_s–$\bar{B}_s$, and D^0–$\bar{D}^0$ systems[2]. The phenomenon arises because these mesons are flavor, *i.e.* interaction, eigenstates, but not mass eigenstates: they are produced and decay as the flavor eigenstates, but these states mix as the particles propagate in space as mass eigenstates.

Let us consider a system made of only two states, the neutral meson P^0 and its antiparticle $\bar{P}^0$, which are distinguished only by their flavor. At time $t=0$ the state will be, in general, a superposition of the two

$$|\psi(0)\rangle = a(0)\left|P^0\right\rangle + b(0)\left|\bar{P}^0\right\rangle\,. \tag{1.33}$$

This evolves in time as dictated by the Schrödinger equation

$$i\partial_t\,|\psi(t)\rangle = \mathbf{H}\,|\psi(t)\rangle \quad\Rightarrow\quad |\psi(t)\rangle = a(t)\left|P^0\right\rangle + b(t)\left|\bar{P}^0\right\rangle + c_1(t)\,|f_1\rangle + \ldots \tag{1.34}$$

where $\mathbf{H}$ is the Hamiltonian of the system, and f_i are the possible final states in which the neutral (anti-)meson can decay. For simplicity, let us forget the decay final states and restrict ourselves only to the P^0 and $\bar{P}^0$ states: $\mathbf{H}$ becomes a 2×2 matrix that is, however, no longer Hermitian (because the decay final states are now outside of our model). Let us rewrite it as

$$\mathbf{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} \tag{1.35}$$

where both $\mathbf{M}$ and $\mathbf{H}$ are Hermitian. The former is associated with the masses of the states, while the latter with their decay widths. On-diagonal terms are associated with flavor-conserving transitions $P^0 \to P^0$ and $\bar{P}^0 \to \bar{P}^0$, while off-diagonal terms are associated with flavor-changing $P^0 \leftrightarrow \bar{P}^0$ transitions.

The eigenvectors of $\mathbf{H}$, *i.e.* the mass eigenstates, have well-defined masses and widths. Let us define the two mass eigenstates P_1 and P_2 as a function of the flavor eigenstates by introducing p and q, two complex parameters such that

$$|P_1\rangle = p\left|P^0\right\rangle + q\left|\bar{P}^0\right\rangle, \quad |P_2\rangle = p\left|P^0\right\rangle - q\left|\bar{P}^0\right\rangle \quad \text{with} \quad |p|^2 + |q|^2 = 1\,. \tag{1.36}$$

Note that, by adding a third parameter, *CPT* violation can be described as well; however, I will restrict myself to the cases where *CPT* is conserved: this implies $\mathbf{M}_{11} = \mathbf{M}_{22}$ and $\mathbf{\Gamma}_{11} = \mathbf{\Gamma}_{22}$. Then, diagonalizing $\mathbf{H}$, q and p are given by

$$\frac{q}{p} = \pm\sqrt{\frac{\mathbf{M}_{12}^* - \frac{i}{2}\mathbf{\Gamma}_{12}^*}{\mathbf{M}_{12} - \frac{i}{2}\mathbf{\Gamma}_{12}}}\,. \tag{1.37}$$

2 K^0 is defined as a $d\bar{s}$ state; B^0 as a $d\bar{b}$ state; B_s as an $s\bar{b}$ state; and D^0 as a $c\bar{u}$ state.

Here I choose the plus sign for q/p (choosing the minus corresponds to swapping P_1 and P_2). Note that the eigenstates have, in general, different masses and decay widths[3]:

$$\Delta m \equiv m_2 - m_1 = -2\,\mathrm{Re}\left[\frac{q}{p}\left(\mathbf{M}_{12} - \frac{i}{2}\mathbf{\Gamma}_{12}\right)\right], \tag{1.38a}$$

$$\Delta\Gamma \equiv \Gamma_2 - \Gamma_1 = 4\,\mathrm{Im}\left[\frac{q}{p}\left(\mathbf{M}_{12} - \frac{i}{2}\mathbf{\Gamma}_{12}\right)\right], \tag{1.38b}$$

where $m_{1,2}$ and $\Gamma_{1,2}$ are the masses and widths of the eigenstates. If CP is conserved, then the eigenstates are orthogonal, thus $\langle P_1|P_2\rangle = |p|^2 - |q|^2 = 0$.

The time evolution of the mass eigenstates is, as one can expect, trivial: each eigenstate decays exponentially, and never mixes with the other:

$$|P_1(t)\rangle = e^{-i\lambda_1 t}\,|P_1\rangle\,, \quad |P_2(t)\rangle = e^{-i\lambda_2 t}\,|P_2\rangle\,, \tag{1.39a}$$

$$\lambda_{1,2} = m_{1,2} - \frac{i}{2}\Gamma_{1,2} = \left(m \pm \frac{\Delta m}{2}\right) - \frac{i}{2}\left(\Gamma \pm \frac{\Delta\Gamma}{2}\right) = \mathbf{M}_{11} - \frac{i}{2}\mathbf{\Gamma}_{11} \pm \frac{q}{p}\left(\mathbf{M}_{12} - \frac{i}{2}\mathbf{\Gamma}_{12}\right), \tag{1.39b}$$

$$\mathrm{Probability}(P_1 \to P_1, t) = |\langle P_1(t)|P_1\rangle|^2 \propto e^{-\Gamma_1 t}, \tag{1.39c}$$

$$\mathrm{Probability}(P_1 \to P_2, t) = |\langle P_2(t)|P_1\rangle|^2 = 0\,, \tag{1.39d}$$

where $\lambda_{1,2}$ are the eigenvalues of $\mathbf{H}$, and m and Γ are the average mass and decay width. Changing basis to the flavor eigenstates, the equations become

$$\mathrm{Probability}(P^0 \to P^0, t) = |\langle P^0(t)|P^0\rangle|^2 = |g_+(t)|^2, \tag{1.40a}$$

$$\mathrm{Probability}(\bar{P}^0 \to \bar{P}^0, t) = |\langle \bar{P}^0(t)|\bar{P}^0\rangle|^2 = |g_+(t)|^2, \tag{1.40b}$$

$$\mathrm{Probability}(P^0 \to \bar{P}^0, t) = |\langle \bar{P}^0(t)|P^0\rangle|^2 = |g_-(t)|^2|q/p|^2, \tag{1.40c}$$

$$\mathrm{Probability}(\bar{P}^0 \to P^0, t) = |\langle P^0(t)|\bar{P}^0\rangle|^2 = |g_-(t)|^2|p/q|^2, \tag{1.40d}$$

$$\text{with} \quad g_\pm(t) \equiv \frac{e^{-i\lambda_1 t} \pm e^{-i\lambda_2 t}}{2}\,. \tag{1.40e}$$

If $|q/p| = 1$ (this is the case if either CP or the time-inversion symmetry T is conserved), then

$$\mathrm{Probability}(P^0 \to P^0, t) = \frac{1}{2}e^{-\Gamma t}\left(\cosh\left(\frac{\Delta\Gamma}{2}t\right) + \cos(\Delta m\, t)\right), \tag{1.41a}$$

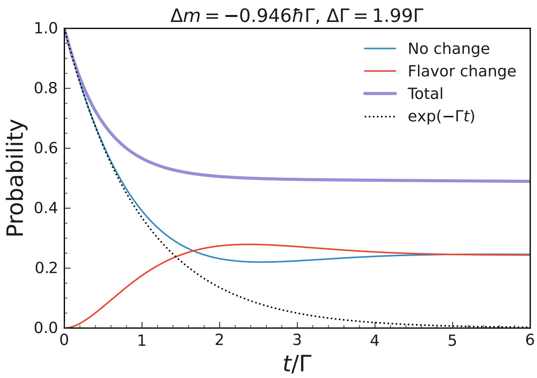
$$\mathrm{Probability}(P^0 \to \bar{P}^0, t) = \frac{1}{2}e^{-\Gamma t}\left(\cosh\left(\frac{\Delta\Gamma}{2}t\right) - \cos(\Delta m\, t)\right). \tag{1.41b}$$

This is shown in Figure 1.3 for the sets of parameters roughly corresponding to the four mixing systems, which are reported in Table 1.2. The Δm parameter is responsible for the so-called *flavor oscillations*, which are particularly evident for the B_s–$\bar{B}_s$ system. The $\Delta\Gamma$ parameter, instead, can subvert the exponential decay: this happens because large $\Delta\Gamma$ means that the mass eigenstates have very different decay widths and, therefore, mean lifetimes.

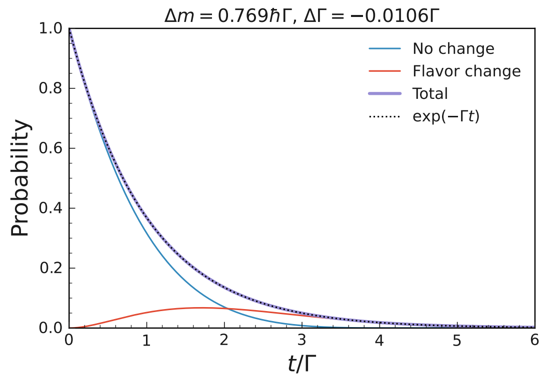
Let us now consider the decay into a final state f. I will assume that flavor is not conserved by the decay, and therefore that it must happen through the weak interaction. Let use define the decay amplitudes for the flavor eigenstates

$$\mathcal{A}_f = \mathcal{A}(P^0 \to f) = \langle f|H_w|P^0\rangle\,, \qquad \bar{\mathcal{A}}_f = \mathcal{A}(\bar{P}^0 \to f) = \langle f|H_w|\bar{P}^0\rangle\,, \tag{1.42}$$

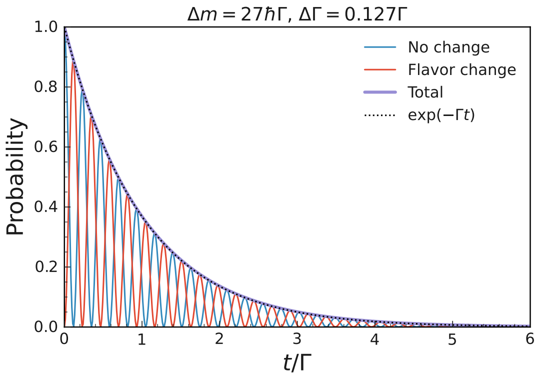
3 In this theoretical overview, I use natural units $c = 1 = \hbar$.



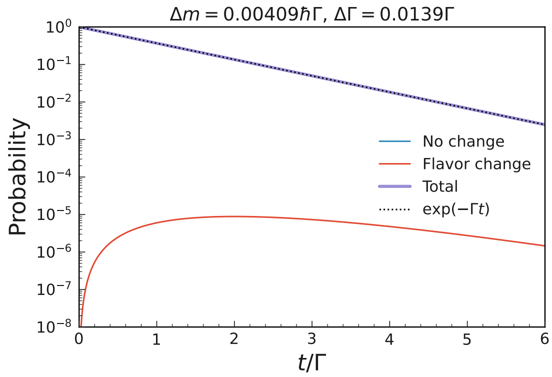
(a) K^0–$\bar{K}^0$ system.



(b) B^0–$\bar{B}^0$ system.



(c) B_s–$\bar{B}_s$ system.



(d) D^0–$\bar{D}^0$ system.

Figure 1.3: Rough plots of the neutral meson oscillations for the four mixing systems. Parameter values from [15]. The blue line is for the probability of keeping the producing flavor; the red line for changing flavor; the thick purple line is their sum; and the black dotted line is the simple exponential we would have without mixing.

where H_w is the Hamiltonian of the weak interaction. Then the decay rates are

$$\frac{\mathrm{d}\Gamma(P^0(t)\to f)}{\mathrm{d}t} \propto |\langle f|H_w|P^0(t)\rangle|^2 = \left|g_+(t)\langle f|H_w|P^0\rangle + \frac{q}{p}g_-(t)\langle f|H_w|\bar{P}^0\rangle\right|^2, \tag{1.43a}$$

$$= \left|g_+(t)\mathcal{A}_f + \frac{q}{p}g_-(t)\bar{\mathcal{A}}_f\right|^2 = |\mathcal{A}_f|^2|g_+(t)+\lambda_f g_-(t)|^2, \tag{1.43b}$$

$$\frac{\mathrm{d}\Gamma(\bar{P}^0(t)\to f)}{\mathrm{d}t} \propto \left|g_+(t)\bar{\mathcal{A}}_f + \frac{p}{q}g_-(t)\mathcal{A}_f\right|^2 = |\bar{\mathcal{A}}_f|^2\left|g_+(t)+\frac{1}{\lambda_f}g_-(t)\right|^2, \tag{1.43c}$$

$$\text{where } \lambda_f \equiv \frac{q}{p}\frac{\bar{\mathcal{A}}_f}{\mathcal{A}_f}, \tag{1.43d}$$

System	K^0–$\bar{K}^0$	B^0–$\bar{B}^0$	B_s–$\bar{B}_s$	D^0–$\bar{D}^0$
$x \equiv \Delta m/\Gamma$	-0.9462 ± 0.0017	0.7697 ± 0.0035	26.99 ± 0.09	0.0041 ± 0.0005
$y \equiv \Delta\Gamma/(2\Gamma)$	$0.996\,506 \pm 0.000\,014$	-0.005 ± 0.017	0.064 ± 0.004	$0.006\,47 \pm 0.000\,24$

Table 1.2: Parameter values for the four mixing systems [15].

or, more explicitly, and highlighting in red the terms that differ [24],

$$\frac{d\Gamma(P^0(t) \to f)}{dt} \propto \frac{1}{2}|\mathcal{A}_f|^2 e^{-\Gamma t}\Big((1+|\lambda_f|^2)\cosh\frac{\Delta\Gamma}{2}t + (1-|\lambda_f|^2)\cos\Delta m\, t \tag{1.44a}$$

$$- 2\,\mathrm{Re}[\lambda_f]\sinh\frac{\Delta\Gamma}{2}t + 2\,\mathrm{Im}[\lambda_f]\sin\Delta m\, t\Big), \tag{1.44b}$$

$$\frac{d\Gamma(\bar{P}^0(t) \to f)}{dt} \propto \frac{1}{2}|\bar{\mathcal{A}}_f|^2 e^{-\Gamma t}\Big((1+|1/\lambda_f|^2)\cosh\frac{\Delta\Gamma}{2}t + (1-|1/\lambda_f|^2)\cos\Delta m\, t \tag{1.44c}$$

$$- 2\,\mathrm{Re}[1/\lambda_f]\sinh\frac{\Delta\Gamma}{2}t + 2\,\mathrm{Im}[1/\lambda_f]\sin\Delta m\, t\Big), \tag{1.44d}$$

and, for the *CP*-conjugate final state $\bar{f}$,

$$\frac{d\Gamma(P^0(t) \to \bar{f})}{dt} \propto \frac{1}{2}|\mathcal{A}_{\bar{f}}|^2 e^{-\Gamma t}\Big((1+|\lambda_{\bar{f}}|^2)\cosh\frac{\Delta\Gamma}{2}t + (1-|\lambda_{\bar{f}}|^2)\cos\Delta m\, t \tag{1.44e}$$

$$- 2\,\mathrm{Re}[\lambda_{\bar{f}}]\sinh\frac{\Delta\Gamma}{2}t + 2\,\mathrm{Im}[\lambda_{\bar{f}}]\sin\Delta m\, t\Big), \tag{1.44f}$$

$$\frac{d\Gamma(\bar{P}^0(t) \to \bar{f})}{dt} \propto \frac{1}{2}|\bar{\mathcal{A}}_{\bar{f}}|^2 e^{-\Gamma t}\Big((1+|1/\lambda_{\bar{f}}|^2)\cosh\frac{\Delta\Gamma}{2}t + (1-|1/\lambda_{\bar{f}}|^2)\cos\Delta m\, t \tag{1.44g}$$

$$- 2\,\mathrm{Re}[1/\lambda_{\bar{f}}]\sinh\frac{\Delta\Gamma}{2}t + 2\,\mathrm{Im}[1/\lambda_{\bar{f}}]\sin\Delta m\, t\Big), \tag{1.44h}$$

$$\text{with } \lambda_{\bar{f}} \equiv \frac{q}{p}\frac{\bar{\mathcal{A}}_{\bar{f}}}{\mathcal{A}_{\bar{f}}}, \quad \mathcal{A}_{\bar{f}} = \left\langle \bar{f}\middle|H_w\middle|P^0\right\rangle, \quad \bar{\mathcal{A}}_{\bar{f}} = \left\langle \bar{f}\middle|H_w\middle|\bar{P}^0\right\rangle. \tag{1.44i}$$

1.2.2 *CP violation in mixing systems*

In neutral meson mixing systems, *CP* violation can occur in three different ways: in the decay, in mixing, or in the interference between decay amplitudes with and without mixing. The experimental observable I measure in this work is the *time-integrated CP asymmetry*

$$\mathcal{A}_{CP}(P^0 \to f) \equiv \frac{\Gamma(P^0 \to f) - \Gamma(\bar{P}^0 \to \bar{f})}{\Gamma(P^0 \to f) + \Gamma(\bar{P}^0 \to \bar{f})} \quad \text{with} \quad \Gamma(a \to b) \equiv \int_0^\infty \frac{d\Gamma(a(t) \to b)}{dt}\, dt\,. \tag{1.45}$$

All three ways may contribute to *CP* violation in a specific decay at the same time, and all will contribute to $\mathcal{A}_{CP}$.

CP VIOLATION IN THE DECAY This occurs when the probability of $P^0 \to f$ is different from that of $\bar{P}^0 \to \bar{f}$. This implies a difference in the modulo of the amplitudes $\mathcal{A}_f$ and $\bar{\mathcal{A}}_{\bar{f}}$:

$$\left|\frac{\mathcal{A}_f}{\bar{\mathcal{A}}_{\bar{f}}}\right| \neq 1 \quad \Rightarrow \quad \mathcal{A}_{CP} = \frac{\left|\mathcal{A}_f/\bar{\mathcal{A}}_{\bar{f}}\right|^2 - 1}{\left|\mathcal{A}_f/\bar{\mathcal{A}}_{\bar{f}}\right|^2 + 1}. \tag{1.46}$$

This kind of *CP* violation can only occur if the amplitudes $\mathcal{A}_f$ and $\bar{\mathcal{A}}_{\bar{f}}$ contain at least two different contributions with different "strong" and "weak" phases, as shown in Equation 1.31.

CP violation in the decay is a type of *direct CP* violation, as it involves an irreducible weak phase in the physical decay amplitudes $\mathcal{A}_f$ and $\bar{\mathcal{A}}_{\bar{f}}$. This is also the only kind of *CP* violation that can occur in charged mesons (which do not form a mixing system).

CP VIOLATION IN THE MIXING This occurs when the probability of a flavor change $P^0 \to \bar{P}^0$ is different from that of $\bar{P}^0 \to P^0$. This implies that the mixing parameters p and q have different magnitude:

$$\left|\frac{q}{p}\right| \neq 1 \qquad \Rightarrow \qquad \mathcal{A}_{CP} = \frac{1 - |q/p|^4}{1 + |q/p|^4}. \tag{1.47}$$

Note that this kind of *CP* violation is actually time-independent:

$$\mathcal{A}_{\mathrm{SL}}(t) \equiv \frac{\frac{\mathrm{d}}{\mathrm{d}t}\Gamma(P^0(t) \to f) - \frac{\mathrm{d}}{\mathrm{d}t}\Gamma(\bar{P}^0(t) \to \bar{f})}{\frac{\mathrm{d}}{\mathrm{d}t}\Gamma(P^0(t) \to f) + \frac{\mathrm{d}}{\mathrm{d}t}\Gamma(\bar{P}^0(t) \to \bar{f})} = \frac{1 - |q/p|^4}{1 + |q/p|^4}, \tag{1.48}$$

where the SL subscript comes from the fact that this effect is best measured in charged-current semileptonic decays $P^0 \to \ell^+ X^-$, as this is the only contribution at the lowest SM order.

CP violation in the mixing is a type of *indirect CP* violation, as it does not involve any irreducible weak phase in the physical decay amplitudes $\mathcal{A}_f$ and $\bar{\mathcal{A}}_{\bar{f}}$.

CP VIOLATION IN THE INTERFERENCE When the same final state f (including all *CP* eigenstates) is accessible by both flavors P^0 and $\bar{P}^0$, the decay amplitude $P^0 \to f$ always includes two contributions:

1. the direct decay $P^0 \to f$, and
2. the mixing-then-decay $P^0 \to \bar{P}^0 \to f$.

The amplitude from these two ways can interfere, giving rise to *CP* violation. This happens when λ_f has a "weak" phase, *i.e.* a phase that changes sign under *CP* conjugation

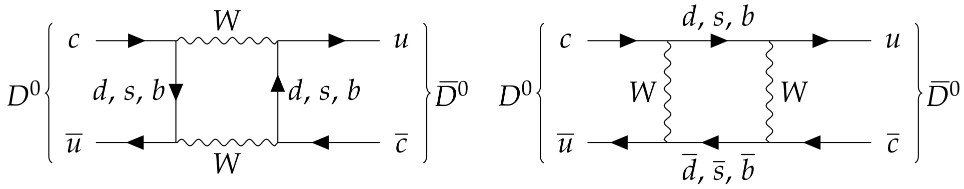
$$\arg[\lambda_f] + \arg[\lambda_{\bar{f}}] \neq 0, \tag{1.49}$$

or just $\mathrm{Im}[\lambda_f] \neq 0$ if f is a *CP* eigenstate.

Note that this kind of *CP* violation may occur even when both other kinds do not, *i.e.* when both $|q/p| = 1$ and $|\bar{\mathcal{A}}_f/\mathcal{A}_f| = 1$.

1.2.3 *Mixing in the Standard Model*

Contributions to the mixing amplitudes can be classified by the travel distance of intermediate states, compared to the typical strong interaction scale. Figure 1.4 shows the Feynman diagrams



(a) Short-distance contribution. (b) Long-distance contribution.

Figure 1.4: Leading-order contributions to mixing amplitudes for the D^0–$\overline{D}^0$ system.

for the leading-order mixing contributions (for the D^0–$\overline{D}^0$ system, but analogous diagrams can be drawn for the other systems).

In Figure 1.4a the intermediate states are the massive W bosons. Since $m_W \sim 80\,\text{GeV}/c^2$ while $m_D^0 \sim 1.9\,\text{GeV}/c^2$, these intermediate states will always be off-shell virtual particles. Their travel distance will be vanishingly small: in fact, this interaction can be written as a Fermi four-point interaction. Contributions of this kind are called *short-distance* contributions, and contribute mainly to Δm, the mixing parameter responsible for the oscillations.

In Figure 1.4b the intermediate states are made of quarks, which may hadronize into on-shell intermediate states (especially the light ones). Besides introducing QCD contributions, these states may travel for distances comparable to the typical strong interaction scale, and are therefore called *long-distance* contributions. They mainly contribute to $\Delta\Gamma$.

The mixing parameters for the four systems are reported in Table 1.2. Short-distance contributions dominate in the B^0–$\overline{B}^0$ system thanks to the contribution for the very massive top quark. They are, instead, very suppressed in the D^0–$\overline{D}^0$ system both because the bottom quark is not nearly as heavy as the top, and because inter-generation interactions are suppressed by the CKM matrix. In this case, also the contributions from the light quarks (u and d) are very suppressed, since the Glashow-Iliopoulos-Maiani (GIM) mechanism is more effective when there are only two generations involved [25].

The amount of long-distance contribution depends on the amount of final states that are common between the meson and the anti-meson, since only these can act as intermediate hadronic states. The K^0 mesons almost exclusively decay in two pions (charged or neutral), meaning that almost all the possible final states are in common: this leads to the very large $\Delta\Gamma$. The B^0 system is in the opposite situation. For D^0 mesons, long-distance contributions are dominant, but only because of the strong suppression of the short-distance diagrams.

1.3 *CP* VIOLATION IN THE CHARM SECTOR

CP violation and mixing were first discovered in th K^0–$\overline{K}^0$ system in the 1960s. The B^0–$\overline{B}^0$ system has been under study since the 1990s, both in terms of mixing and *CP* asymmetry, as it is our main gateway to measure most SM parameters. *CP* violation and mixing of D^0 mesons, instead, have a much more recent history.

The SM predicts much smaller mixing and *CP* asymmetries in charm compared to the other systems:

- mixing is suppressed by the relatively small mass of the b quark and the GIM mechanism, as discussed in Section 1.2.3 and shown in Table 1.2 and Figure 1.3;

- *CP* violation at the tree level is introduced only by V_{cs}, which has a very small imaginary part $\text{Im}[V_{cs}] = -\eta A^2 \lambda^4 \sim -6 \times 10^{-4}$; and
- *CP* violation at the loop level could take a large phase from V_{ub}, but this is always paired to V_{cb}, suppressing the effect to $|V_{ub}V_{cb}| \sim \mathcal{O}(\lambda^5)$.

In particular, *CP* asymmetries are typically predicted to be $\mathcal{O}(10^{-3})$ at most. This makes their measurement much harder from the experimental point of view: On top of that, the presence of long-distance contributions makes the computation of theoretical predictions particularly cumbersome, as these contributions are essentially non-perturbative due to the involvement of low-energy QCD.

The D^0–$\bar{D}^0$ mixing, which is now well-established, was actually only observed in 2007 at the first-generation B-factory experiments *BABAR* and Belle [5, 6]. Evidence of *CP* violation was only observed indirectly in one decay channel, $D^0 \to \pi^+\pi^-$, through the measurement in 2019 of the *CP* asymmetry difference

$$\Delta\mathcal{A}_{CP} = \mathcal{A}_{CP}(D^0 \to K^+K^-) - \mathcal{A}_{CP}(D^0 \to \pi^+\pi^-) = (-1.54 \pm 0.29) \times 10^{-3}\,, \tag{1.50}$$

where the uncertainty includes both statistical and systematic contributions, and of the *CP* asymmetry $\mathcal{A}_{CP}(D^0 \to K^+K^-) = (0.68 \pm 0.54(\text{stat}) \pm 0.16(\text{sys})) \times 10^{-3}$ in 2023, both at the LHCb experiment [4, 7]. These measurements allow to compute the *direct* $D^0 \to \pi^+\pi^-$ *CP asymmetry* $a^d_{\pi^+\pi^-} = (2.3 \pm 0.6) \times 10^{-3}$. It is not completely clear whether these measurements are compatible with SM predictions, due to the aforementioned non-perturbative QCD effects affecting the prediction itself. Several attempts are being made to reduce these uncertainties, both from the theoretical and the experimental side [26–30].

Among these efforts, the Belle II collaboration is publishing measurements of the *CP* asymmetry in the isospin-related $D^0 \to \pi^0\pi^0$ and $D^+ \to \pi^+\pi^0$ decay modes [8, 9]. A sum rule based on the asymmetries and branching fractions of these three modes, and on the mean lives of the D^0 and D^+ mesons, allows to determine whether *CP* violation arises from $\Delta I = 1/2$ amplitudes, allowed within the SM, or $\Delta I = 3/2$ amplitudes, forbidden by the SM [30]. The Belle II experiment is particularly suited for this kind of measurements involving neutral final-state particles, thanks to its low-particle-multiplicity events (compared to hadronic collider experiments), completely known initial state kinematic, and excellent photon reconstruction performance. To complete the picture, other notable charm *CP* violation measurements at Belle II include those of the $D^0 \to K^0_S K^0_S$ *CP* asymmetry [31, 32], and the first model-independent measurement of the D^0–$\bar{D}^0$ mixing parameters $x = \Delta m/\Gamma$ and $y = \Delta\Gamma/(2\Gamma)$ at a B factory experiment, using the $D^0 \to K^0_S\pi^+\pi^-$ decay channel [33].

Now that the experimental sensitivity is approaching the level of SM predictions, there are several reasons for studying *CP* violation in the charm sector:

- to test the predictions of the SM in a sector, as of yet, almost unexplored;
- to search for new physics contributions to *CP* violation; and
- to test our quantitative understanding of non-perturbative QCD and or techniques for computing its contributions (*e.g.* lattice QCD).

Concerning the first point, it is important to stress the necessity for studying *CP* violation in different channels, and independently with different experiments.

The second point deserves more discussion. Here, the charm sector presents a unique opportunity in two ways. From the experimental point of view, the fact that SM predictions are so small provides a very clean environment to probe for contributions that are not accounted for. From the theoretical point of view, it is conceivable for a new physics model to introduce effects with a natural preference for up-type quarks [34, 35]. Flavor-changing neutral currents are a common example of such an additional contribution. Note that probing models where the up sector plays a special role is almost exclusively possible with charm: the light u quarks produce only a few neutral hadrons that can not mix (π^0 and η are their own antiparticle), and have so few decay channels that *CP* symmetry is essentially implied by *CPT* invariance; and the t quark decays before hadronizing, again preventing mixing and denying many options for *CP* violation searches. While the top and up quarks still offer opportunities (*e.g.* searches for electric dipole moments), the charm sector offers a much more fertile ground [36].

CABIBBO-SUPPRESSED CHARM DECAYS It is convenient to classify the D^0 decays depending on the *strangeness*[4] S of the final state:

- $D^0 \to [S = -1]$ decays, like $D^0 \to K^-\pi^+$, are called *Cabibbo favored*, as they only involve the large, on-diagonal elements of the CKM matrix $V_{ud} \simeq V_{cs} \simeq \cos\theta_C \sim 0.97$;
- $D^0 \to [S = 0]$ decays, like $D^0 \to \pi^+\pi^-$ or $D^0 \to K^+K^-$, are called *singly Cabibbo suppressed*, as they involve one off-diagonal element $V_{cd} \simeq -V_{us} \simeq \sin\theta_C \sim 0.23$; and
- $D^0 \to [S = 1]$ decays, like $D^0 \to K^+\pi^-$, are called *doubly Cabibbo suppressed*, as they involve two off-diagonal elements;

where $\theta_C \equiv \theta_{12} \sim 13°$ is the Cabibbo angle ($\lambda \equiv \sin\theta_C$), and *CP*-conjugate decays are implied. The most striking difference is that decay rates of the latter two are suppressed by a factor of about 20 and 400 with respect to the former (at least, for SM contributions). However, there are also important implications for the amount of *CP* violation predicted by the SM.

In $D \to PP$ and $D \to PV$ decays (where P and V indicate a pseudoscalar and a vector meson), time-integrated *CP* asymmetries arise only when two amplitudes with different "strong" and "weak" phases contribute (see Equation 1.31):

$$\mathcal{A}_{CP} = -2\frac{|\mathcal{A}_2/\mathcal{A}_1|\,\sin(\theta_1-\theta_2)\sin(\phi_1-\phi_2)}{1+|\mathcal{A}_2/\mathcal{A}_1|^2+2|\mathcal{A}_2/\mathcal{A}_1|\,\cos(\theta_1-\theta_2)\cos(\phi_1-\phi_2)}\,. \tag{1.51}$$

In the SM, two substantial amplitudes of this kind exist *only* for singly Cabibbo suppressed channels:

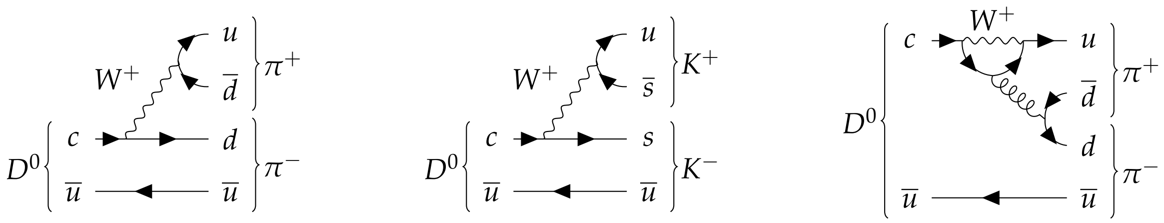
1. the tree-level $c \to d\bar{d}u$, which proceeds through $V_{cd}V_{ud}^* = -\lambda + \lambda^3/2$, or $c \to s\bar{s}u$, which proceeds through $V_{cs}V_{us}^* = \lambda - \lambda^3/2 - i\eta A^2\lambda^5$; and
2. the loop-level "gluonic penguin" diagram, which can proceed through the above and also through the suppressed $V_{cb}V_{ub}^* = A^2\lambda^5(\rho + i\eta) + \mathcal{O}(\lambda^7)$.

Example Feynman diagrams visualizing amplitudes are shown in Figure 1.5.

Let us consider the relative "weak" phase between these amplitudes. I neglect the $V_{cb}V_{ub}^*$ contributions, since they are suppressed by a factor $\lambda^4 \sim 3\times10^{-3}$ with respect to the others, which means a similar suppression factor would act on $\mathcal{A}_{CP}$ as shown by Equation 1.51. $V_{cd}V_{ud}^*$ has no "weak" phase, so that the relative phase between $c \to d\bar{d}u$ and $c \to s\bar{s}u$ is

$$\arg(V_{cs}V_{us}^*) = \arctan\frac{\mathrm{Im}[V_{cs}V_{us}^*]}{\mathrm{Re}[V_{cs}V_{us}^*]} \simeq -\eta A^2\lambda^4 \sim -6\times10^{-4} \tag{1.52}$$

4 The flavor associated to the s quark: $S(s) = -1$ and $S(\bar{s}) = 1$.



(a) $D^0 \to \pi^+\pi^-$ tree-level diagram ($c \to d\bar{d}u$ transition).

(b) $D^0 \to K^+K^-$ tree-level diagram ($c \to s\bar{s}u$ transition).

(c) $D^0 \to \pi^+\pi^-$ gluonic penguin diagram. The looping fermion may be a d, s, or b quark.

Figure 1.5: Some Feynman diagrams for the $D^0 \to \pi^+\pi^-$ and $D^0 \to K^+K^-$ processes, showing the most substantial amplitudes mentioned in the text.

modulo spin effects. The relative "strong" phases can be taken for granted: charm decays proceed through many hadronic resonances, so suppression from strong phase shifts is extremely unlikely [37]. With this, we can expect $\mathcal{A}_{CP}$ approximately of order 10^{-3}.

CP VIOLATION BEYOND THE SM As introduced, current experimental results allow for yet-undetected new physics contributions to the *CP* asymmetry of D^0 decays. There are little constraints on the possible origin of this additional *CP* violation. I report here two different possibilities as an example.

The first is the arise of relative phases between decay amplitudes with and without D^0–$\overline{D}^0$ mixing. The present experimental constraints leave these phases with enough freedom to produce significant effects. In this case, *CP* violation would be produced in the interference between decays with and without mixing [35].

The second possibility involves the extension of the SM with a Z' boson whose couplings are not diagonal in the flavor representation. Besides enhancing *CP* violation in charm, such a model may be used to explain both the measured value of $\Delta\mathcal{A}_{CP}$ (see Equation 1.50), and other anomalies such as the lepton flavor universality violation in $B \to K^*\ell^+\ell^-$ [38].

THE $D^0 \to \pi^+\pi^-\pi^0$ DECAY The focus of this work is the singly-Cabibbo-suppressed $D^0 \to \pi^+\pi^-\pi^0$ decay channel. It is almost exclusively a decay of the form $D \to PV$, as it proceeds mainly through the three $D^0 \to \rho^{\pm,0}\pi^{\mp,0}$ resonances. Each of these may produce *CP* violation alone, even without considering the interference among the three. Other resonances (excited ρ, f_0, and f_2 states) provide only minor contributions. The same can be said of non-resonant amplitudes [39]. Within the framework of the SM, it is reasonable to expect *CP* asymmetries of $\mathcal{O}(10^{-3})$ [40], although theoretical predictions are complicated by non-perturbative QCD contributions which lead to large uncertainties.

While the final state appears identical for the two flavors, there are actually kinematic differences between D^0 and $\overline{D}^0$ mesons. Among the dominant $D^0 \to \rho\pi$ intermediate resonant states, $D^0 \to [\rho^+ \to \pi^+\pi^0]\pi^-$ is significantly more frequent than the others. Since $m_\rho \sim 0.77\,\mathrm{GeV}/c^2$ and $m_{D^0} \sim 1.87\,\mathrm{GeV}/c^2$, low $\pi^+\pi^0$ pair invariant masses are favored; in turn, this means that the π^+ tends to have lower momentum than the π^-. For the $\overline{D}^0$ meson, the situation is reversed by the *CP* operator: $\overline{D}^0 \to [\rho^- \to \pi^-\pi^0]\pi^+$ is the most frequent intermediate state, favoring low $\pi^-\pi^0$ masses and, thus, lower π^- momenta.

1.4 EXPERIMENTAL STATUS OF $D^0 \to \pi^+\pi^-\pi^0$

There are four recent searches for *CP* violation in the $D^0 \to \pi^+\pi^-\pi^0$ decay which are relevant to date, none of which found any evidence for *CP* violation. These are summarized in Table 1.3.

Experiment	Measurement	Luminosity	Candidates	Ref.
BABAR	$\mathcal{A}_{CP} = (0.31 \pm 0.41 \pm 0.17)\%$	$385\,\mathrm{fb}^{-1}$ (e^+e^-)	82×10^3	[41]
Belle	$\mathcal{A}_{CP} = (0.43 \pm 0.41 \pm 1.23)\%$	$532\,\mathrm{fb}^{-1}$ (e^+e^-)	123×10^3	[42]
LHCb	Energy test p-value $= 62\%$	$6\,\mathrm{fb}^{-1}$ (pp)	2.5×10^6	[43]
LHCb	$\Delta Y = (-1.3 \pm 6.3 \pm 2.4) \times 10^{-4}$	$7.7\,\mathrm{fb}^{-1}$ (pp)	3.8×10^6	[44]

Table 1.3: The most recent searches for *CP* violation in $D^0 \to \pi^+\pi^-\pi^0$.

The measurement by the *BABAR* experiment is the most precise $\mathcal{A}_{CP}$ determination to date. It was performed by reconstructing a very pure (98%) sample of D^0 mesons, subtracting background (estimated from sideband regions), correcting efficiency effects (estimated on simulation) and counting candidates for the two flavors. Contextually, Dalitz-plot-dependent and helicity-angle-dependent $\mathcal{A}_{CP}$ were measured, and a Dalitz plot fit was performed (allowing for different amplitudes and phases in the two flavors and looking for differences).

The measurement by the Belle experiment adopted a similar approach as *BABAR*.

The LHCb experiment performed two different searches. In the first one, the energy test method [45] was used to compare the Dalitz plot distributions on an unbinned basis. In the second, the ΔY parameter [46], related to time-dependent *CP* violation, was measured.

Finally, the resonant contributions to the decay amplitude are known from the amplitude analysis performed by *BABAR* experiment in [39], where *CP* symmetry was assumed. This is used as input for the Belle II simulations.

2 THE BELLE II EXPERIMENT

Following the discovery of the b quark, B-factory collider experiments were envisioned for measuring the Standard Model (SM) parameters in the beauty sector, with a particular focus on CP violation, and for verifying the predictive power of the Kobayashi-Maskawa mechanism. This kind of experiment also provides an ideal environment for studies in the charm and tau sectors, for searches for beyond-SM particles, and more. The first-generation B-factory experiments *BABAR* and Belle ran from 1999 to 2010; afterwards, the Belle II experiment was conceived as a major upgrade of Belle and its collider KEKB.

In this chapter, I first introduce B-factory experiments in Section 2.1; then, I describe the SuperKEKB collider in Section 2.2, and the Belle II detector in Section 2.3 more in detail; finally, in Section 2.4 I explain how the various particles are reconstructed using the detector information, highlighting the features that are most relevant for this work.

2.1 INTRODUCTION TO B-FACTORY EXPERIMENTS

In 1973, Kobayashi and Maskawa showed that CP violation could be incorporated into the SM through what we now call the CKM matrix, but only as long as the number of quark flavors was (at least) six [47]. The following year, the c quark was discovered [48, 49], followed by the b quark in 1977 [50], and along with their bare-flavored D and B mesons.

In the early 1980s, the possibility of using B meson decays to test the Kobayashi-Maskawa mechanism was explored from the theoretical point of view. However, for this to be possible, it was necessary that the CKM matrix fell within a narrow range of the allowed parameter space, so as to produce large CP violation in neutral B decays; this range also corresponds to having large B^0–$\bar{B}^0$ mixing. Furthermore, the "golden" channels to observe this CP violation would be those to CP eigenstates, such as $B^0 \to J/\psi K_S^0$, which were never observed at the time: we now know that their branching fractions are below 0.1%.

In the same years, a large B lifetime was measured, and evidence of substantial B^0–$\bar{B}^0$ mixing rate emerged: this restricted the CKM parameter space to the aforementioned range. Additionally, drastic improvements in e^+e^- storage rings were taking place: this allowed the $B\bar{B}$ pair production rate to increase from tens to millions per day in a couple of decades. Under these circumstances, the observation of CP violation in $B^0 \to CP$ eigenstate channels became feasible, and experiments were designed to perform such measurements [51].

ANATOMY OF A B-FACTORY The design choices for B-factories are completely dictated by the requirements for studying the CP asymmetry of the "golden" channels $B^0 \to CP$ eigenstate. These asymmetries are produced by the interference of different amplitudes, with and without mixing, and cancel when integrated over time. Therefore, one has to measure them as a function of the decay time with respect to a moment when the meson's flavor state is known.

In 1980, the $\Upsilon(4S)$ bound $b\bar{b}$ state was discovered [52, 53]: this resonance is just above the $B\bar{B}$ pair production threshold,

$$m_{\Upsilon(4S)} \sim 10.58\,\mathrm{GeV}/c^2, \quad m_{B^0} \sim 5.28\,\mathrm{GeV}/c^2 \quad \Rightarrow \quad q = m_{\Upsilon(4S)} - 2m_{B^0} \sim 20\,\mathrm{MeV}/c^2\,, \quad (2.1)$$

and essentially decays only to $B\bar{B}$ pairs. Furthermore, in the $\Upsilon(4S) \to B^0\bar{B}^0$ decay, the B mesons are produced in an entangled state. This prompted the following strategy: reconstruct events where one B meson decays to a flavor-specific final state, and the other to a CP eigenstate. When the first meson decays, its flavor is known from the final state, and the flavor of the other is known to be opposite thanks to the entanglement. Then, one can disentangle the mixing contributions by measuring the decay time difference between the two mesons. This technique is called "flavor tagging", and is quite efficient thanks to the large number of flavor-specific decay channels available to neutral B mesons.

To make this kind of measurement, one needs the following.

1. A collider working at the $\Upsilon(4S)$ resonance; this naturally suggests an e^+e^- collider, where the collision energy is known.
2. High luminosity, given the small branching fractions of B^0 decays to CP eigenstates.
3. Boosted B meson production: since the mean B lifetime is $\sim 1.5\,\mathrm{ps}$, the decay time must be inferred from the flight distance; to have large enough flight distance, a minimum B momentum is needed.
4. Nearly-hermetic detector able to perform particle identification (PID) and reconstruct neutrals: the former is paramount for good tagging efficiency, while the latter is required to reconstruct many channels of interest.

The third point has two important consequences.

Fist: since the $\Upsilon(4S)$ is just above threshold for $B\bar{B}$ pair production, the B mesons are produced almost at rest in the collision center of mass (CM) frame. To boost them, the whole collision CM frame must be boosted in the laboratory frame. This implies having an asymmetric-energy collider, with severe impact on the design of the machine, *e.g.* the necessity for separate storage rings. I discuss this more in detail in Section 2.2.

Second: since the decay time difference must be measured from the flight distance, decay vertex reconstruction accuracy is paramount. This imposes strict requirements on the design of the inner part of the tracking detector, such as very high granularity and very low material budget. I discuss this more in detail in Section 2.3.

NON-B PHYSICS AT B-FACTORIES While the design of B-factory experiments is dictated by time-dependent CP violation measurements in neutral B decays, these conditions are also very favorable for many other kinds of measurements. e^+e^- colliders provide a much cleaner and controlled environment with respect to hadronic machines: each bunch crossing rarely counts more than one $e^+e^- \to X$ event, and the initial state is completely and accurately known. This makes it possible to study channels with many neutrals in the final state, or even without any charged particle, and allows to constrain the kinematic of decays with undetected particles (such as neutrinos).

With a collision CM energy equal to the mass of the $\Upsilon(4S)$ resonance, several other processes take place just as frequently as the $\Upsilon(4S)$ production. Cross sections for these are reported in Table 2.1. Among them, those of greatest physical interest include $e^+e^- \to c\bar{c}$ and $\tau^+\tau^-$.

Process	Cross sec. [nb]	Process	Cross sec. [nb]	Process	Cross sec. [nb]
$\Upsilon(4S)$	1.11	$c\bar{c}$	1.30	$\tau^+\tau^-$	0.92
$s\bar{s}$	0.38	$d\bar{d}$	0.40	$u\bar{u}$	1.61
$e^+e^-(\gamma)$	300	$\mu^+\mu^-(\gamma)$	1.15	$\gamma\gamma(\gamma)$	4.99
$e^+e^-e^+e^-$	39.7	$e^+e^-\mu^+\mu^-$	18.9	$\nu\bar{\nu}$	0.25×10^{-3}

Table 2.1: Cross sections of the main $e^+e^- \to X$ processes at the $\Upsilon(4S)$ energy.

The light-quark final states, while considered background by most analyses, allow for measurements of hadronic form factors through processes such as $e^+e^- \to \pi^+\pi^-\pi^0$. Moreover, the abundant Bhabha scattering $e^+e^- \to e^+e^-$, which is a purely QED process and thus is well-understood from the theoretical point of view, allows for accurate measurements of the integrated luminosity, which are fundamental for absolute cross section and branching ratio measurements.

Additionally, searches for beyond-SM particles can be carried out, not just indirectly (by searching for inconsistencies with the SM predictions), but also from direct production. The clean e^+e^- environment allows very good sensitivity to rare events. Furthermore, the controlled initial conditions allow searching for undetected final states.

2.2 THE SUPERKEKB COLLIDER

The SuperKEKB collider [54] is the result of a major upgrade and redesign of the KEKB circular collider, which operated from 1999 to 2010 providing collisions to the Belle experiment. KEKB reached the (at the time) world-record instantaneous luminosity of $2.1 \times 10^{34}\,\mathrm{cm^{-2}\,s^{-1}}$, and integrated almost $1\,\mathrm{ab^{-1}}$ over a decade of activity, corresponding to about one billion $B\bar{B}$ pairs. SuperKEKB aims to surpass its predecessor by more than a factor 30, targeting instantaneous luminosities in excess of $6 \times 10^{35}\,\mathrm{cm^{-2}\,s^{-1}}$ and an integrated luminosity of $50\,\mathrm{ab^{-1}}$. As of today, the world-record peak luminosity of $5.1 \times 10^{34}\,\mathrm{cm^{-2}\,s^{-1}}$ (over twice KEKB's record) was reached, accumulating a sample of $575\,\mathrm{fb^{-1}}$ [55].

SuperKEKB collides 7 GeV electrons with 4 GeV positrons at a single interaction point (IP), where the Belle II detector is located. The choice of a single IP is to maximize luminosity.

A schematic view of the facility is shown in Figure 2.1. The asymmetric-energy requirement forces to have separate rings for electrons and positrons, dubbed high-energy ring (HER) and low-energy ring (LER). Electrons are provided by a linear accelerator (LINAC), which already brings them to their nominal energy before injecting them into the main storage ring: this on-energy injection is necessary to optimize the storage ring for luminosity. The LINAC also produces positrons as secondary particles from the collision of electrons on a target. Since these positrons have high emittance, they are first sent into a 1.1 GeV damping ring, then re-injected into the LINAC, which brings them to their nominal energy before injecting them into the main storage ring. The main parameters of SuperKEKB are reported in Table 2.2.

Although the particles are injected in the main rings already at their nominal energy, they quickly lose energy to synchrotron radiation. This is a characteristic of electron and positron rings, and is due to the small mass of the particles: the power lost to synchrotron radiation

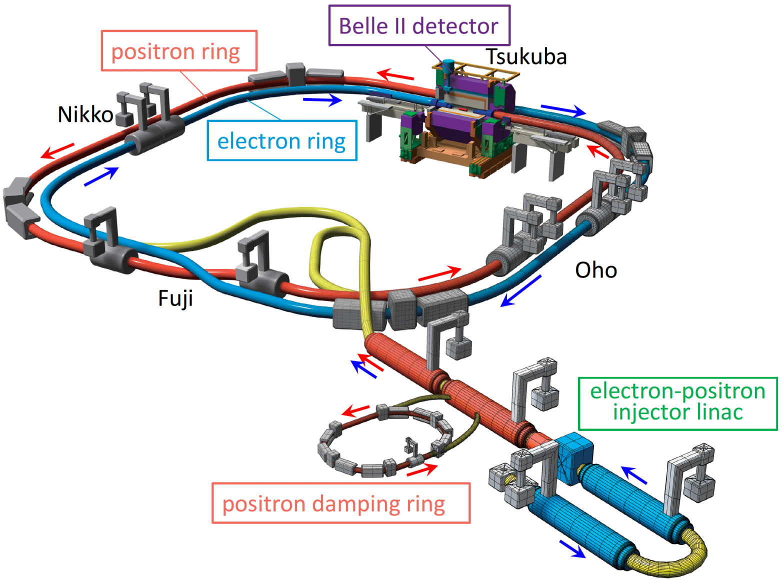


Figure 2.1: Schematic view of the SuperKEKB collider complex. Blue arrows are for electrons, while red arrows are for positrons.

is $\propto E^4/m^3$, where E is the particle energy and m is its mass [57]. Therefore, accelerating RF cavities are needed to maintain the beam energies constant. This mechanism is instrumental in damping oscillations around the nominal orbit: synchrotron radiation decreases momentum in all three directions, but the cavities only restore longitudinal momentum, resulting in a reduction of transverse momentum. The typical damping time is $\mathcal{O}(1\,\text{ms})$. This principle is also exploited in the positron damping ring.

Another noteworthy characteristic of these machines is the relatively short beam lifetime. Electrons and positrons in high-intensity beams are easily diverted from the nominal beam orbit by several effects, which cause relatively large losses of beam particles at every turn. The typical beam lifetimes at SuperKEKB are 5 to 10 minutes. To counteract this, the "top-up" injection scheme is adopted, where a new bunch is injected "on top" of an already present bunch, thus "refilling" its charge and maintaining beam currents almost constant.

The design of SuperKEKB is dictated by the challenges that reaching high instantaneous luminosity poses. The luminosity of an e^+e^- collider is given by

$$L = \frac{N_\text{b} n_+ n_- f}{A_\text{eff}} \quad \text{with} \quad A_\text{eff} = \frac{\sigma_x^* \sigma_y^*}{R_L}\,, \tag{2.2}$$

where N_b is the number of bunches circulating in each ring, n_+ and n_- are the number of positrons and electrons per bunch, f is the revolution frequency, and A_eff is the effective transverse overlap area of the bunches at the IP; the effective area is given by the product of σ_x^* and σ_y^* (the transverse bunches sizes at the IP), and is reduced by an effective factor R_L, which limits the luminosity increase at high beam currents[1].

The reduction factor R_L is mainly due to the "beam-beam" interaction effects. When two bunches cross each other, the particles in one bunch are deflected by the electromagnetic

1 The beam currents are given by $I_\pm = eN_b n_\pm f$, where e is the elementary charge.

	Design		Achieved on Dec. 27, 2024	
Ring	LER	HER	LER	HER
Beam energy [GeV]	4	7		
Crossing angle [mrad]	83			
Luminosity [$\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$]	6×10^{35}		5.1×10^{34}	
Beam current [A]	3.6	2.6	1.632	1.259
Number of bunches	2560		2346	
Bunch length [mm]	6	5		
Horizontal bunch size at IP [µm]	10	11	15.5	16.6
Vertical bunch size at IP [nm]	48	62	265	265

Table 2.2: The main parameters of SuperKEKB from [56].

field of the other bunch, which acts as a focusing lens. This is very non-linear, causing the operating point of the machine (in the betatron oscillation tune plane) to spread and shift. Besides limiting luminosity, this effect is also hard to predict from first principles, complicating the operation of the machine and forcing the final operating parameters to diverge from their design values.

There are two main ways to increase the instantaneous luminosity: increasing the beam currents (thus the number of interacting particles), and decreasing the effective overlap area. Furthermore, the currents may be increased both by increasing the number of bunches N_b, and by increasing the bunch charge $n_\pm$. All of these are pursued at SuperKEKB.

INCREASING THE BEAM CURRENTS Increasing the number of bunches N_b is challenging because, as the separation between successive bunches becomes smaller, the residual electromagnetic fields produced by one bunch in the beam chambers begin to sensibly affect the following bunch. This leads to “coupled bunch-bunch” instabilities, which can push the beams out of their nominal orbits, potentially leading to beam losses.

The RF cavities of SuperKEKB operate at a frequency of $\sim 508\,\mathrm{MHz}$, meaning that the separation between bunches is a multiple of $\sim 2\,\mathrm{ns}$ (0.6 m). However, to keep the bunch-bunch instability manageable, bunches can not be placed closer than $\sim 4\,\mathrm{ns}$ (1.2 m).

Increasing N_b also requires a more powerful injection, since a larger number of bunches would have to be “refilled” with the top-up injection. The same difficulty is encountered by attempting to increase the bunch charge $n_\pm$. A more powerful injection can be obtained in several ways, such as: increasing the amount of charge injected at a time; reducing the time between successive injections; increasing the injection efficiency (the fraction of the injected charge that actually reaches the nominal beam orbit); and injecting multiple bunches at a time. To these ends, the SuperKEKB design included major upgrades of the KEKB LINAC, the beam-transport line from the LINAC to the main rings, and the injection magnets. Also, more upgrades are being considered for the future. I will not discuss this any further, as it is beyond the scope of this work.

Lastly, increasing the beam currents is also challenging for the hardware of the machine itself. First, a very hard vacuum must be maintained to maximize beam lifetime. Second, the constant bombardment of synchrotron radiation and stray beam particles forces every component to meet very strict radiation tolerance criteria. Third, larger beam currents correspond to large image currents along the beam chambers, which means that a large a mount of heat is dissipated in the various components, and must be effectively removed; fluctuations in temperature must also be avoided, as they can shift the orbit parameters. Finally, these amounts of synchrotron radiation and heat add further stress on the vacuum system, as they increase the amount of gas and ions released into the vacuum by the beam pipe walls.

REDUCING THE EFFECTIVE CROSSING AREA A two-ring machine with small bunch spacing requires a system that separates the two beams around the IP. PEP-II (*BABAR*'s collider) used dipolar magnets close to the IP, which allow head-on collisions. KEKB, instead, allowed for a small crossing angle (22 mrad) between the beams at the IP: this allows for smaller bunch spacing and leaves more room for the detector, however it introduces transverse-longitudinal coupled instabilities and can reduce the luminosity through R_L.

SuperKEKB inherited the crossing-angle strategy from its predecessor, improving on it through the Nano-Beam scheme, which was initially proposed for the cancelled SuperB project [58, 59]. The scheme is exemplified in Figure 2.2. In short, the idea is to adopt a relatively large (but still $\ll 1$) crossing angle, together with a relatively great bunch length ($\sim 2\sigma_z$) and a very small horizontal bunch size ($\sim 2\sigma_x$). This way, the transverse overlap area (the segment d in the figure) can be kept small, which has several advantages.

- A smaller horizontal overlap d reflects in a smaller A_{eff}, which implies higher luminosity.
- It naturally solves the "hourglass" effect.
  The vertical size of the bunch σ_y is determined by the vertical betatron function $\beta_y(s)$, where s is the position along the beam orbit. Ideally, one would like to minimize this function at the IP, so that A_{eff} is also minimized, thus maximizing the luminosity. However, β_y is an oscillating function, thus the minimum can be kept only for a finite distance. In head-on collision schemes, this distance must be larger than the bunch length; otherwise, the tails of the bunches are larger than the centers, as shown in Figure 2.3, and their contribution to luminosity is reduced. With the Nano-Beam scheme, one only needs to maintain β_y close to the minimum for the overlap distance $\sim d \ll \sigma_z$, which allows to reach smaller minima, thus further reducing A_{eff} and increasing the luminosity.
- It naturally reduces the parasitic collisions.
  These are the collisions along the enlarged tails, which contribute very little to luminosity, but still produce beam-beam instability effects. In the Nano-Beam scheme, the tails are naturally separated by the geometry, so that parasitic collisions are almost absent.

The drawback of schemes with a large crossing angle is the introduction of horizontal-vertical instability coupling, which can limit the choice of operating points and, potentially, the maximum achievable luminosity. These are addressed by the "crab-waist" scheme, also proposed for SuperB, and tested at KEKB during its final run years. The coupling originates from the fact that, due to the crossing angle, the positrons that are displaced along the transverse horizontal direction x meet the center of the electron bunch at different longitudinal coordinates z, and *vice versa* for the electrons. This is shown in Figure 2.4. Since the minimum of β_y occurs at a fixed z, both the luminosity and the magnitude of the beam-beam effects depend

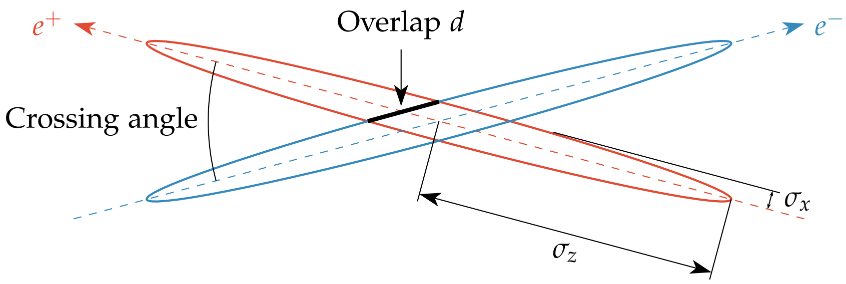


Figure 2.2: Schematic of a collision in the Nano-Beam scheme. Sizes and angles are not realistic, but chosen for readability.

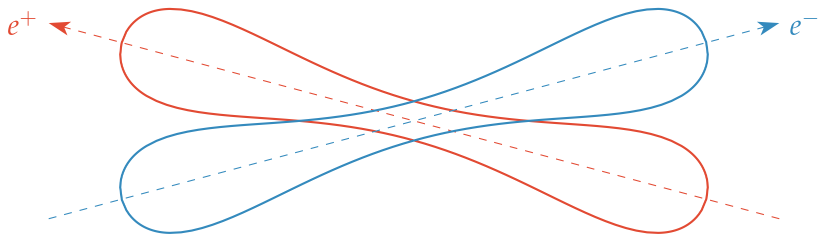


Figure 2.3: Visualization of the hourglass effect. Sizes and angles are not realistic, but chosen for readability.

on x, thus increasing the horizontal spread and coupling vertical effects to the horizontal position. The crab waist scheme resolves this by using a sextupolar magnet near the IP to introduce non-linear effects that can shift the z position of the β_y minimum as a function of x.

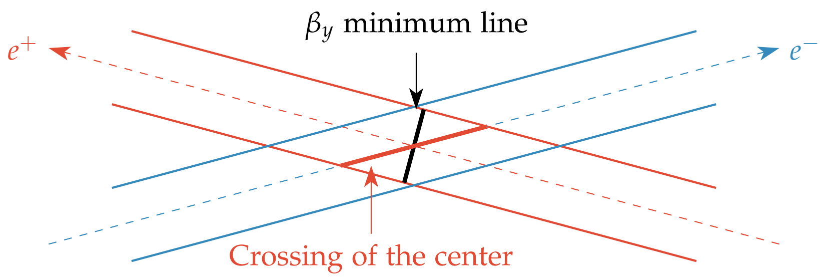


Figure 2.4: Schematic explanation of the crab waist scheme. Sizes and angles are not realistic, but chosen for readability. The thick black line is the location of the minimum of the positron β_y when the crab-waist is *not* used, while the thick red line shows the points where horizontally-displaced positrons meet the center of the electron bunch.

BEAM-INDUCED BACKGROUNDS An important characteristic of high-luminosity e^+e^- machines is the large amount of particles produced or lost by the beams through various effects. These reach the accelerator components and the detector, causing radiation damage and polluting the otherwise very clean e^+e^- events. Many of these effects happen all along the storage ring, and are very difficult to simulate accurately; as a consequence, all three B-factories had to adopt, during detector design, safety factors of one order of magnitude with respect to the simulated backgrounds.

Some of these effects are produced within each beam, without interaction with the other: these are called "single-beam" backgrounds. Synchrotron radiation photons belong to this category. Beam particles can also change their momentum and stray out of the beam, reaching the walls of the beam chambers, due to Coulomb scattering with the other particles of the bunch: this is called Touschek scattering. Beam particles may also scatter against residual gas and ions released into the vacuum by the walls of the beam chambers: this is called "beam-gas" scattering. While these effects happen all around the storage rings, a substantial number of particles reach the beam pipe around the IP, and is detected by Belle II, polluting the events. From the accelerator side, these backgrounds are suppressed by placing movable collimators in strategic positions to scrape the beam halo. From the event reconstruction side, charged particle tracks can be filtered by imposing a maximum distance to the IP at the point of closest approach (PoCA), or a common vertex with other tracks; this is not possible for photons, however.

Residual ions are also responsible for bunch-bunch instabilities and for enlarging the transverse size of the electron beam. The latter effect is called "beam size blow-up". Similar effects happen for the positron beam due to clouds of photoelectrons produced by the interaction of synchrotron radiation with the walls of the beam chambers. Three dedicated countermeasures are in place for the electron clouds: coating the beam chambers with a material with low secondary emission yield; using grooved surfaces that further suppresses the yield from synchrotron radiation; and placing "clearing electrodes" in suitable locations to deflect electrons using an electric field.

Other effects happen at the collision of the bunches, and are called "luminosity-dependent" backgrounds: these are e^+e^- processes of no physical interest, but with large cross sections, so that they are likely to happen close in time to interesting events and pollute them. The main ones are the radiative Bhabha scattering ($e^+e^- \to e^+e^-\gamma$), and the two-photon process $e^+e^- \to e^+e^-e^+e^-$, which are mentioned in Section 2.1 and Table 2.1. The filtering methods discussed for the single-beam backgrounds are ineffective for these; however, these effects are much better understood and simulated.

Finally, an important background comes from the top-up injection. A schematic of the procedure is shown in Figure 2.5. The bunch to be injected enters the main ring at an angle, and is deflected on the nominal orbit direction by a septum magnet, *i.e.* a dipolar magnet with a very thin wall on the side of the existing beam. Before reaching the injection point, the pre-existing bunch is moved close to the wall of the septum magnet by a pair of kickers (magnets or capacitors that can be activated for a very short time with a very strong field). Then, the new bunch is injected in phase with the pre-existing bunch. Finally, the two are brought back on the nominal orbit by a second pair of kickers. However, because of the finite size of the septum magnet wall, the newly-injected bunch will be slightly out of the nominal orbit, and will oscillate around it until synchrotron radiation losses dampens them, in a few ms.

Before the damping is complete, a large number of synchrotron radiation photons will be produced. Also, a substantial fraction of the newly injected particles will never reach the nominal orbit, instead straying away and reaching the beam chamber walls. This background behaves like the single-beam backgrounds, but has a much more defined structure in time and is much more severe. Belle II adopts a specially-designed trigger veto to avoid most of this; however, a significant portion still reaches the detector [61].

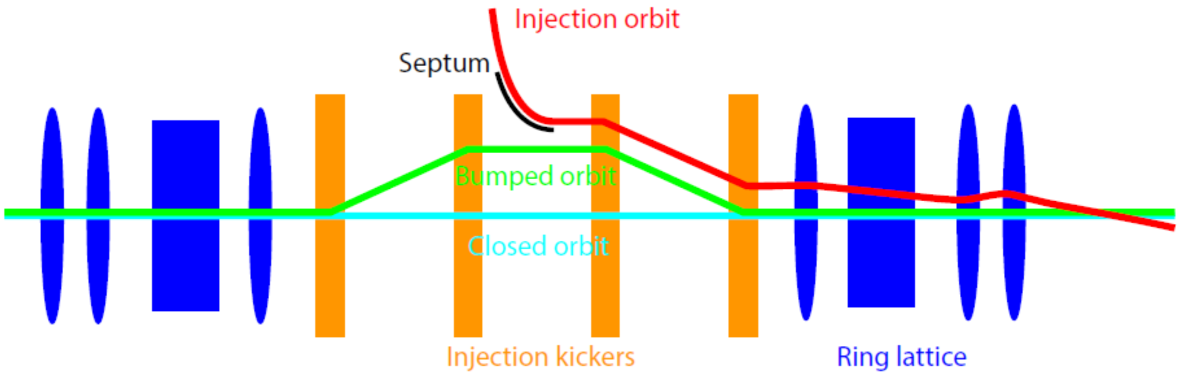


Figure 2.5: Schematic of the injection system from [60].

2.3 the Belle II detector

The Belle II detector [62] is a major redesign of the Belle detector. It reuses part of the Belle components and mechanical structures (*e.g.* the superconducting solenoid magnet, its iron return yoke, and the calorimeter crystals), while upgrading or redesigning other parts (*e.g.* the inner detectors, the readout electronics, and the data acquisition systems).

As for the colliders, the requirements for detectors at B-factories are set by what is needed to study CP violation in $B^0 \to CP$ eigenstate channels:

- reconstruction of as much as possible of the event, *i.e.* near-hermeticity;
- reconstruction of both charged and neutral particles;
- excellent decay vertex position resolution to infer the decay time of B mesons from their flight distance; and
- excellent PID capability, to achieve efficient B flavor tagging.

The first point, together with the boosted CM frame resulting from asymmetric collision energies, results in detector geometries with very wide and asymmetric acceptance. The detector develops in a series of subdetector layers, starting from the IP and going outwards:

- first are the *tracking* subdetectors, responsible for reconstructing charged particles, and divided into an inner *vertex detector*, primarily responsible for the vertex position resolution, and an outer *drift chamber*, primarily responsible for the momentum resolution;
- then, subdetectors dedicated to charged particle identification (PID), based either on Čerenkov light production or time of flight;
- afterwards, an electromagnetic calorimeter, mainly responsible for photon reconstruction;
- then, a superconducting solenoid, producing the 1.5 T magnetic field required by the tracking subdetectors to measure the tracks' momenta; and
- finally, chambers dedicated to muon identification and detection of very-long-lived K^0_L mesons, implemented by instrumenting the return iron yoke of the solenoid magnet.

Figure 2.6 shows a schematic view of these systems, which I describe in detail the following sections, together with their contribution to the event reconstruction, which is explained more in detail in Section 2.4.

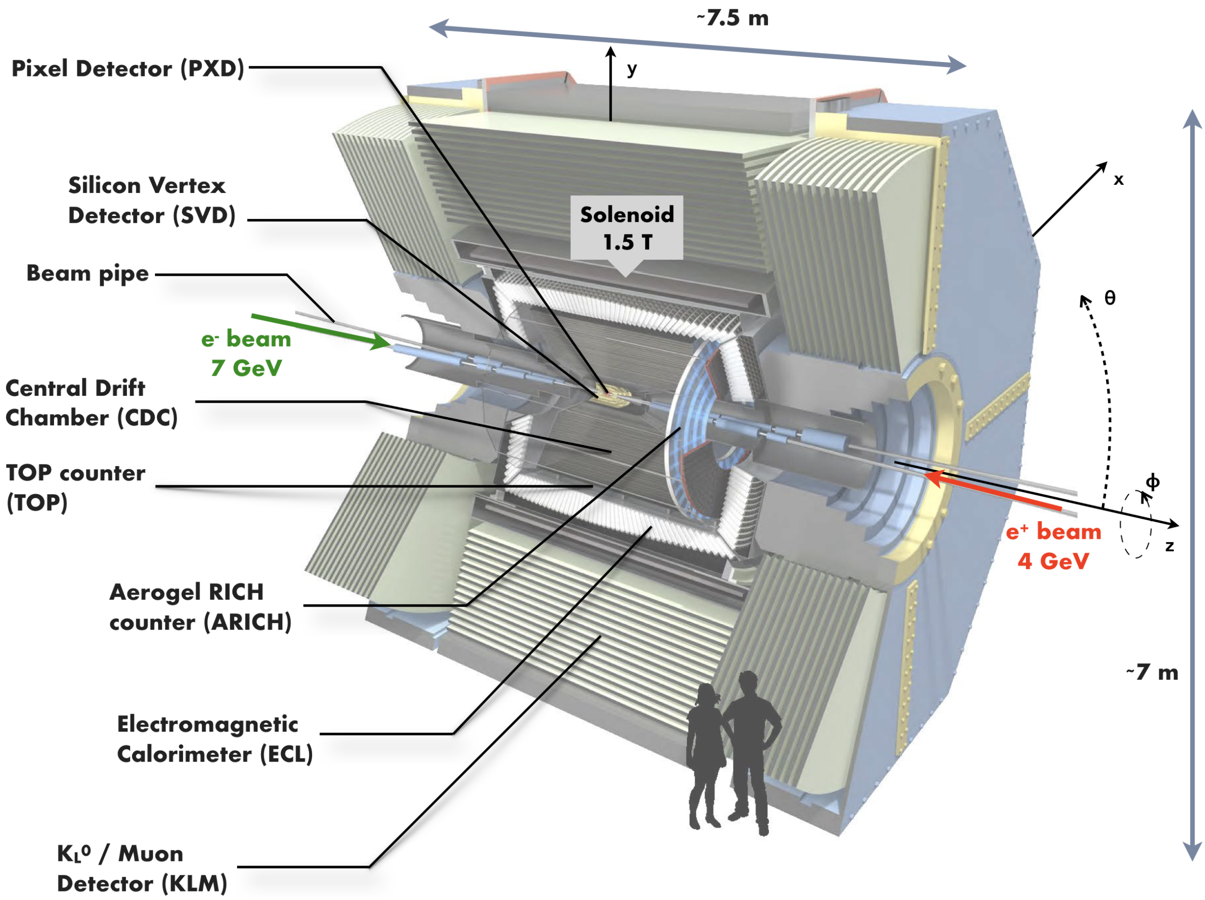


Figure 2.6: A schematic 3D view of the Belle II detector (people silhouettes for scale).

2.3.1 *The vertex detector*

Decay vertices are primarily reconstructed by searching for the geometric intersection of charge particle tracks. Therefore, the decay vertex position resolution is dominated by the resolution of the point of origin of said tracks. To achieve the necessary spatial resolution, two silicon-based subdetectors are placed as close as possible to the IP: the pixel detector (PXD), which makes up the two innermost layers, and the silicon vertex detector (SVD), which makes up the four outermost layers, for a total of six layers of silicon sensors.

To achieve excellent momentum resolution and PID performance, it is paramount to minimize the effect of detector material on the particle trajectories (multiple Coulomb scattering, *etc.*). This is accomplished choosing materials with high radiation length (X_0), and by minimizing the thickness of material crossed by the particles. This has a particularly important impact on the vertex detector, given the relatively high density and low X_0 of silicon, and on the beam chamber walls around the IP.

The beam chamber at the interaction region (IR) is 1 cm in diameter, and made of a double-walled beryllium pipe, for a total of 1 mm of beryllium thickness. The 1 mm interstice between the two walls is used to flow liquid paraffin as a coolant. Finally, the inner wall of the beam pipe is coated in 10 µm of gold to shield the detector from synchrotron radiation, for a total material budget of $\sim 0.8\,\%\,X_0$.

The PXD layers are placed just outside the beam pipe, at 1.4 cm and 2.2 cm radii. Until the long shutdown 1 (LS1), which begun in July 2022, only two ladders (one sixth) of the second layer were present: this is the condition for all the data used in this work. The PXD is made of depleted FET (DEPFET) pixel sensors thinned to 75 µm and supported by a 400 µm thick silicon frame, resulting in a material budget of $\sim 0.2\,\%\,X_0$ per layer. They have a pixel pitch of 50 µm in the azimuthal angle (ϕ) direction, and variable between 56 µm in the longitudinal (z) direction (77 µm for the second layer). This excellent spatial resolution comes at the cost of a longer integration time of about 20 µs (corresponding to about two beam revolutions). A schematic rendering of the PXD is shown in Figure 2.7

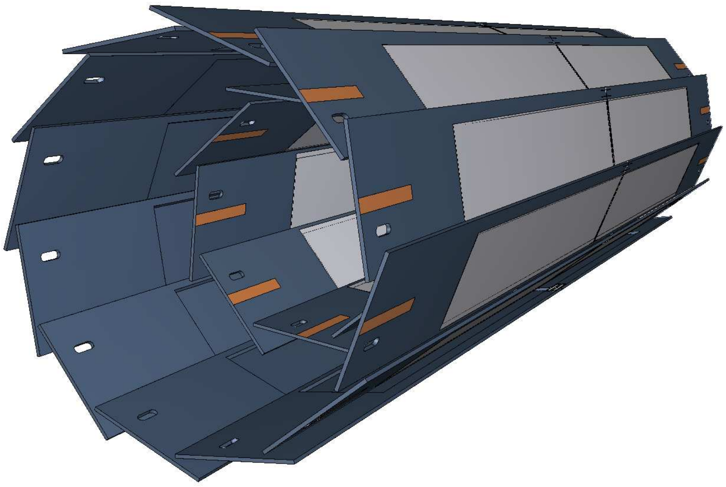


Figure 2.7: A schematic 3D rendering of the PXD.

The SVD layers are placed at 3.9 cm, 8.0 cm, 10.4 cm, and 13.5 cm radii. It is made of double-sided silicon strip (DSSD) sensors, 300 µm thick and supported by a carbon fiber structure. In the three outermost layers, part of the readout ICs are placed within the detector acceptance on flexible printed circuit boards, together with their stainless steel CO_2 pipes for cooling; in the innermost layers, instead, all of these are placed outside of the acceptance in the forward and backward regions. The total material budget is $\sim 0.6\,\%\,X_0$ per layer. The forward sensors of the three outer layers are tilted into a lantern-like shape to further reduce the amount of material crossed by particles with small polar angle. A schematic rendering of the SVD is shown in Figure 2.8.

The innermost SVD layer has a readout pitch of 50 µm in ϕ and 160 µm in z, while the outermost layers have a pitch of 75 µm in ϕ and 240 µm in z; the tilted forward sensors, which are trapezoidal in shape, have a z-dependent pitch in ϕ ranging from 50 µm to 75 µm. The readout ICs operate on a 127 MHz clock, allowing to reach a cluster time resolution of $\mathcal{O}(2\,\text{ns})$ [63]. Also, the amount of deposited charge in the cluster is measured, and used for PID.

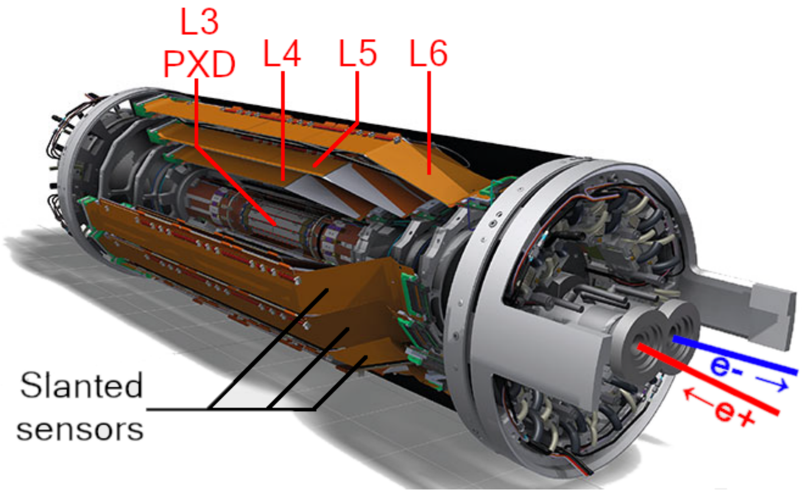


Figure 2.8: A schematic 3D rendering of the four SVD layers (L3-6) and PXD.

2.3.2 *The drift chamber*

The central drift chamber (CDC) design is largely based on that of Belle, with updated geometry and readout electronics. It covers the radial range from 17 cm to 113 cm, and has a polar angle (θ) acceptance ranging from 17° (forwards) to 150° (backwards), which is the smallest of all subdetectors. In the collision CM frame, this range corresponds to the symmetric θ_{CM} range from $\sim 22°$ to $\sim 158°$. The forward and backward endcap walls are tapered to make room for the final-focusing quadrupole magnets of SuperKEKB, responsible for the minimization of β_y at the IP, and the other subdetectors. A cross section of the CDC is schematized in Figure 2.9

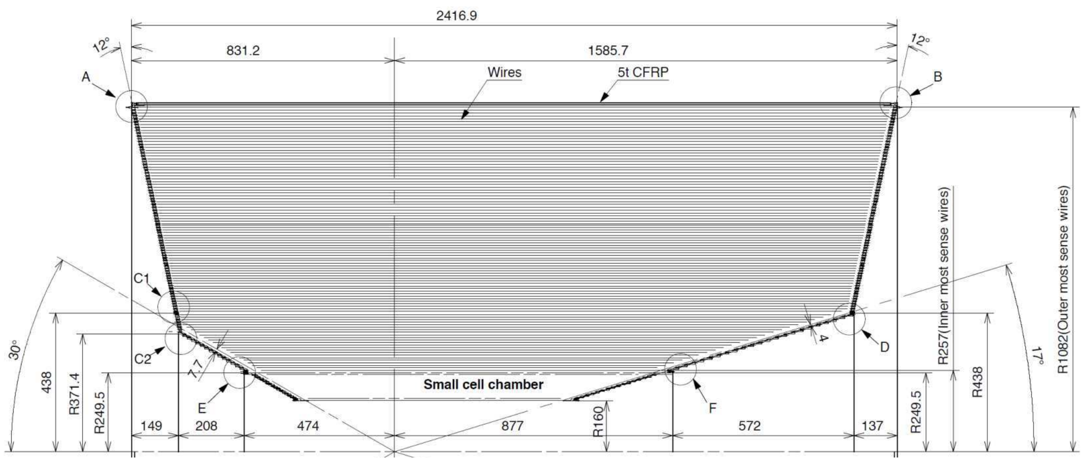


Figure 2.9: A schematic cross-sectional view of the CDC.

The CDC employs a 50% helium, 50% ethane gas mixture, and is instrumented with 14 thousand tungsten-gold sense wires (30 µm in diameter), and 42 thousand aluminum field wires (126 µm in diameter). The choice of a gas-filled drift chamber instrumented with thin wires is, again, dictated by the necessity to minimize material budget, while still instrumenting a relatively large volume.

The sense wires are distributed in 56 layers, grouped into nine "superlayers", as shown in Figure 2.10. The innermost superlayer (the "small cell chamber") includes eight layers, with a cell size of ~ 7 mm, while the other superlayers include six layers each, with a cell size ranging from 10 mm to 18 mm. Superlayers are alternating between the "axial" configuration (wires parallel to the z axis, essentially insensitive to the longitudinal position) and "stereo" configuration (wires tilted by ~ 50 mrad to provide some longitudinal position information).

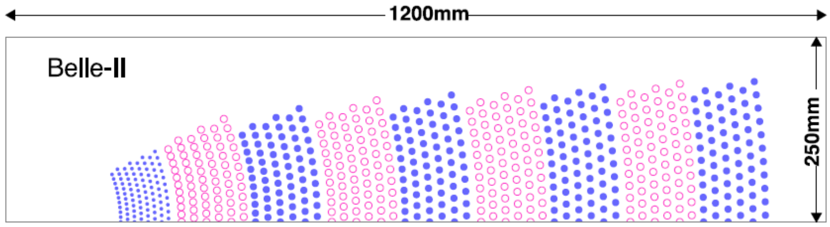


Figure 2.10: Scheme of the CDC layers and superlayers, from innermost (left) to outermost (right). Filled blue circles are for "axial" wires, while pink hollow circles are for "stereo" wires.

The front-end electronics measures both drift time and deposited charge for each wire. The former is necessary to infer the track position, and the resulting resolution is ~ 100 µm. The latter is used for PID, and has a resolution of about 5%.

2.3.3 *The particle identification systems*

There are two subdetectors dedicated to PID, both located just outside the drift chamber: the time of propagation detector (TOP), located in the barrel section, and the aerogel ring-imaging Čerenkov detector (ARICH) located in the forward-endcap section.

TOP (BARREL) The TOP barrel system is made of 16 quartz bars (2.5 m long, 45 cm wide, and 2 cm thick), distributed in ϕ around the CDC outer wall. It covers the θ range from 33° to 120°. The schematic of a bar is shown in Figure 2.11.

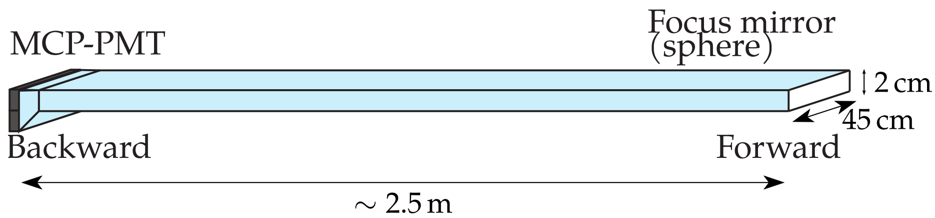


Figure 2.11: A schematic of a TOP quartz bar.

Charged particles interact with the quartz by emitting photons at a Čerenkov angle, which depends on the particle's velocity. Such Čerenkov photons are reflected inside the quartz bar and reach the detector at its backward end, which are micro-channel plate photo-multiplier tubes (MCP-PMT). These measure the 2D transverse position of the photons (pitch ~ 5 mm), as well as their precise time of arrival (31 ps resolution). This is shown in Figure 2.12.

The Čerenkov angle can then be determined from the pattern of photons in the (x, y, t) space, combined with the incidence angle of the track, which is measured by the tracking

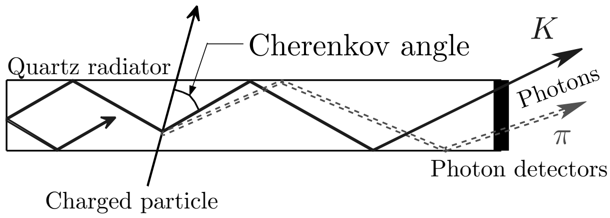


Figure 2.12: A schematic of the TOP working principle.

subdetectors. Then, hypotheses for the particle identity can be tested against each other using the likelihood given by the Čerenkov angle and the track's momentum (also measured by the tracking subdetectors).

ARICH (ENDCAP) The ARICH is located in front of the CDC forward end plate. It is made of a proximity-focusing aerogel Čerenkov radiator, followed by an empty expansion volume, and an array of photon detectors. This is schematically shown in Figure 2.13

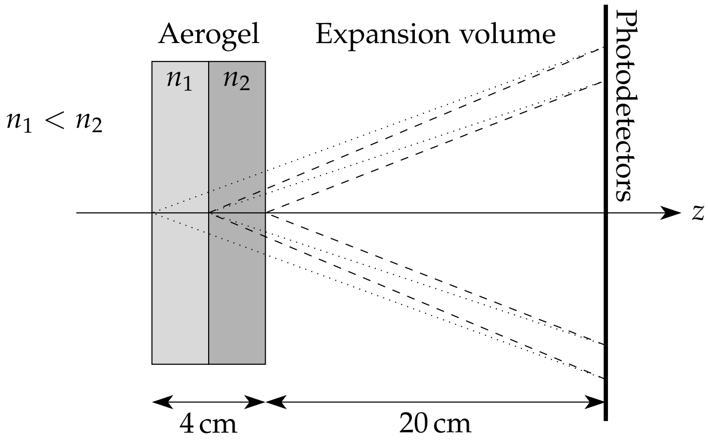


Figure 2.13: Schematic side view of the ARICH.

To ensure that particles crossing the radiator produce enough Čerenkov photons, a relatively thick (few cm) radiator is necessary. However, a thicker radiator implies a larger dispersion of the photon's position on the photodetector plane. To limit this dispersion, two slabs of silica aerogel with slightly different refraction indices ($n_1 = 1.046$ and $n_2 = 1.056$) are used. This way, the rings produced by the two slabs overlap on the photodetector plane, reducing the dispersion by a factor two.

The photodetectors are pixelated hybrid avalanche photo-diodes (HAPD) with a pitch of 4.9 mm. This gives a single-photon angular resolution of about 14 mrad.

Aerogel and HAPDs are arranged in nine concentric rings of square tiles. They cover the radial range from 43.5 cm to 113.7 cm, corresponding to the θ range from 15° to 31°.

The computation of the identity likelihoods is done similarly as for the TOP. The particle's intercept point and angle on the aerogel plane are provided by the tracking subdetectors; the ARICH then searches for a ring around this point, from which the Čerenkov angle is measured; then, using the particle momentum measured by the tracking subdetectors, the likelihood for the various identity hypotheses are computed.

2.3.4 *The electromagnetic calorimeter*

The electromagnetic calorimeter (ECL) envelops the CDC and PID detectors on all sides, covering the θ region from 12.3° to 155.1° (except for two 1° gaps between the barrel and endcap sections, at 32° and 130°). It is designed to reconstruct photons from energies as small as 20 MeV; it also aids in electron identification, luminosity measurement, and K_L^0 detection.

The ECL is made of 8736 CsI(Tl) scintillator crystals with an average cross section of about $6\,\text{cm} \times 6\,\text{cm}$ and a length of 30 cm, corresponding to about $16\,X_0$. The barrel crystals are truncated pyramids, while for the endcap 69 different crystals shapes are used.

Scintillation light is detected by PIN photodiodes (two per crystal), and both the amount of light and the time of arrival are measured. This results in a cluster energy resolution approximately given by

$$\frac{\sigma_E}{E} = \sqrt{\left(\frac{0.066\%}{E}\right)^2 + \left(\frac{0.81\%}{\sqrt[4]{E}}\right)^2 + (1.34\%)^2}\,, \tag{2.3}$$

with E in GeV, a cluster position resolution of about $5\,\text{mm}/\sqrt{E} \sim 0.25°/\sqrt{E}$, and a cluster time resolution of below 100 ns.

2.3.5 *The K_L^0 and μ detector*

The ECL barrel section is enveloped in the superconducting solenoid that produces the 1.5 T magnetic field needed by the tracking subdetectors to measure particle momenta. Outside the solenoid is a return yoke for the magnetic flux, made of 4.7 cm thick iron plates. The K_L^0 and μ detector (KLM) is made by instrumenting this yoke, creating an alternating sandwich of iron plates and active detector elements.

The KLM covers the θ range from 20° to 155°. The barrel section is made of 15 detector layers and 14 iron plates, while in the endcap sections only 14 detector layers are present. The iron provides a total of 3.9 hadronic interaction lengths, to be added to the 0.8 interaction lengths of the ECL. The detector layers are made partly of resistive plate chambers (RPC) (for the barrel section), and partly of plastic scintillator (for the endcap section, where more severe beam background is expected).

The KLM primarily exists to detect hadronic showers produced by scarcely-interacting, very-long-lived K_L^0 mesons. It also provides information for muon identification. Its granularity allows for an angular resolution of about 0.5°.

2.3.6 *The trigger system*

Because of the limited data bandwidth at the various data acquisition (DAQ) steps, and of the limited storage space available for acquired data, only events that are likely to contain interesting physics are stored. To efficiently select such events while discarding non-interesting background events, two levels of "triggers" are used: the level 1 (L1) trigger, and the high-level trigger (HLT).

LEVEL 1 TRIGGER Each subdetector's readout electronics stores the information about all events in a small local buffer: the L1 trigger is responsible for indicating which information to transfer to the global DAQ system for further processing before the buffer is overwritten. For this reason, it is required to have a fixed latency of 5 µs, and a time resolution of about 10 ns. Due to limitations of DAQ bandwidth and online processing power, the maximum average rate is further limited to 30 kHz.

Through dedicated lines, the trigger gathers coarse but quickly-available information from the CDC, TOP, ARICH, ECL, and KLM. It then processes this information in a series of field-programmable gate arrays (FPGA), allowing fast processing and ease of reconfiguration.

Many trigger lines are present for different kinds of events. Most are mainly based on coarse track reconstruction from CDC information, and cluster energy and topology information from the ECL; the TOP detector is instead mainly used for its fine time resolution.

A very high hadronic event efficiency is paramount for B physics measurements: events of this kind are easily identified by the presence of a relatively large number of tracks. Several lines are provided for calibration and luminosity measurement, mainly designed for events such as Bhabha, $\mu^+\mu^-$, and $\gamma\gamma$; given the large cross sections of these processes compared to that of the $\Upsilon(4S)$ and other physically-interesting events, many of these are prescaled by randomly selecting or rejecting events with some fixed probability. Finally, dedicated lines are used for low-multiplicity events, such as $\tau^+\tau^-$ and those with production of possible dark matter candidates; due to the similarities with calibration lines, low-multiplicity lines tend to have lower efficiencies, and several dedicated selections were optimized for these specific cases.

Finally, the L1 trigger receives a veto signal from the SuperKEKB injection system. This is tuned to reject events around the passage time of a recently-injected bunch, as the beam background is expected to be much more severe in this case, and the information would be hardly useful for physics.

HIGH-LEVEL TRIGGER Once detector data selected by the L1 trigger is acquired, it is sent to a farm of computers that perform a full event reconstruction and further tighten the selection criteria: this farm is the HLT. The HLT uses the same Belle II software framework used for offline data reprocessing and analysis (described in Section 2.3.7). The idea is to perform the same reconstruction during data-taking and offline to minimize inconsistencies. The main differences between online and offline processing are:

- the lack of data from the PXD on the HLT (due to bandwidth limitations); and
- the calibrations which are produced *a posteriori* from data, which of course are not available before the data has been acquired.

The HLT selections are based on a number of alternative sets of requirements on the fully-reconstructed tracks and clusters, aimed at selecting the different types of events that are interesting for analyses or calibration, with appropriate prescaling if needed. The HLT reduces the event rate by almost one order of magnitude compared to the L1 trigger.

The HLT is also responsible for transmitting the selected events to the PXD DAQ system, so that PXD information may be available during offline processing. This forces the PXD to buffer its data for a much longer time (~ 1 s), but allows for a much reduced data rate. Also, the HLT transmits regions of interest (ROI) to the PXD DAQ, *i.e.* the position of intercepts between tracks (reconstructed with CDC and SVD information) and PXD sensors, which allow to only transfer PXD hits close to these intercepts, further reducing the PXD data size.

Finally, the HLT provides a coarse classification of events based on a number of selections, which is stored in the saved data. When a specific type of event is needed (*e.g.* hadronic or Bhabha), this classification allows faster offline processing by skimming, *i.e.* skipping events that lack the precomputed classification one is looking for.

2.3.7 *The Belle II software and computing infrastructure*

The Belle II software revolves around the Belle II analysis software framework (basf2) [64], which contains almost all of the libraries and programs required for data-taking, reprocessing, simulation and analysis. This framework relies on a number of external libraries, including: ROOT [65] for data storage and processing; EvtGen [66], PYTHIA8 [67], KKMC [68], and PHOTOS [69, 70] for event generation; Geant4 [71] for the simulation of particle-matter interactions; and more.

basf2 itself is based on a number of *modules* organized in several *packages* (*e.g.* for each subdetector or reconstruction step), plus auxiliary tools, libraries, and scripts. Each module performs one specific task (from data input/output to simulation, reconstruction, selection, analysis and monitoring), and most modules are highly configurable. The vast majority of the modules are written in C++ to allow fast and efficient data processing. Data is processed by using Python scripts called *steering files*, which first set up a sequence of configured modules called a *path*, and then initiate the execution of the path by the basf2 core framework: this allows to exploit the flexibility of Python to construct complex analysis workflows, while still relying on the velocity and efficiency of C++ for the actual processing. The modules in a path are called in sequence, once per input event, and exchange data with each other through a common interface called *data store*.

An additional infrastructure, called the *conditions database*, is used to store and retrieve information about specific data-taking *runs* and periods. This is a centralized online database, allowing for consistent use of the information among all processes and users. Information in the conditions database typically include detector condition, configuration, and calibrations.

Finally, Belle II uses a distributed GRID-based approach for both data processing and storage. Many affiliated institutions worldwide provide the computational resources and storage space for this model. A central management service allows to distribute the workload of data transfer, reprocessing, calibration, production of simulated samples, and analysis among the many available centers and machines, exploiting the huge available resources by running as many processes as possible at the same time.

2.4 EVENT RECONSTRUCTION AT BELLE II

In this section, I briefly outline the reconstruction process for the different kinds of particles used in this analysis, highlighting the features that must be taken into account in this *CP* violation measurement. For each event, *i.e.* for each bunch crossing that resulted in a positive trigger decision, basf2 produces lists of particle *candidates* of different types. These are all final-state particles (FSPs), *i.e.* particles that live long enough to actually reach the detector; there is only a limited number of such particles:

- charged π, K, p, deuterons, electrons and muons, which are reconstructed using the tracking subdetectors;

- photons, which are reconstructed using the ECL; and
- K^0_L and neutrons, which are detected by the ECL or the KLM.

Two other kinds of particles are able to reach the detector, but are not reconstructed directly: V^0s (*i.e.* K^0_S mesons and Λ baryons), which are reconstructed from the pair of charged daughters they decay into (when such a decay occurs); and neutrinos, which remain undetected. All other particles are reconstructed (sometimes only partially) from the combination of their daughters.

Since the detector and reconstruction algorithms have finite resolution, limited efficiency, noise, backgrounds, *etc.*, these candidates may or may not be real particles; candidates that do not correspond to any real particle are called *fakes*. Moreover, a real particle may be missed entirely, or, in some cases, be reconstructed twice (*i.e.* two candidates may actually correspond to the same particle). Candidates from the latter case are called *clones*; clones mostly occur for charged particles, *i.e.* for tracks, and their rate is $\lesssim 1\%$.

Of course, it is also possible to reconstruct a particle but assign it the incorrect type. This occurrence is called *misidentification*, and may occur in several ways depending on the type of particles and the subdetectors involved.

Finally, it is possible to reconstruct particles that did not originate from the collision of interest. These may be produced in a nearby collision: the interval between bunch crossings can be as small as 4 ns, and most subdetectors' time resolution is much worse than this. Alternatively, the particle may be from beam-induced background, which does not necessarily originate from the IR. Time-based selections are generally useful to reduce these backgrounds, if necessary. Additionally, beam-induced backgrounds may be suppressed by selections on the particle's point of origin; this is only possible for charged particles, however, since the point of origin of neutrals is not measured.

In the following sections, I describe the three key reconstruction processes for this work: track reconstruction (Section 2.4.1), particle identification (Section 2.4.2), and photon reconstruction (Section 2.4.3).

2.4.1 *Track reconstruction*

Track reconstruction, *i.e.* the reconstruction of charged particles, is performed in two steps: pattern finding and track fitting. Pattern finding involves identifying groups of detector hits that likely belong to the same particle. Track fitting, as the name suggests, consists in finding the most-likely parameters of the track trajectory given the hits of each group, and taking into account energy losses, interaction with material, and magnetic field non-uniformities.

At Belle II, only the CDC and the SVD participate to pattern finding. Due to the fundamental differences in the detectors, dedicated pattern recognition algorithms exists for each subdetector; tracks found by one subdetector are fitted, then their trajectories are extrapolated to the other detector, nearby hits are attached and the tracks are fitted again; finally, dedicated merging algorithms are used to minimize the number of clones. Last, the tracks are extrapolated to the PXD, hits are attached, and the tracks are fitted again: this allows to exploit the high granularity and proximity to the IP of the PXD to achieve the necessary position resolution [72].

The model for a track's trajectory is the helix (or a series of helix arcs when material effects, *etc.* need to be taken into account). It can be written as

$$\vec{x}(s) = \begin{pmatrix} \cos\phi_0 & -\sin\phi_0 & 0 \\ \sin\phi_0 & \cos\phi_0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sin(\omega s)}{\omega} \\ \frac{\cos(\omega s)-1}{\omega} - d_0 \\ s\tan\lambda + z_0 \end{pmatrix}, \tag{2.4}$$

where s is the arc length travelled from the point of closest approach (PoCA) to the z axis along the transverse (xy) plane, $\lambda = \pi/2 - \theta$ is the dip angle, $\omega \propto q/p_\mathrm{T}$ is the inverse curvature radius (p_T is the transverse momentum, and q is the signed charge), and the PoCA position is determined by d_0, ϕ_0, and z_0. The helix function and its parameters are shown in Figure 2.14.

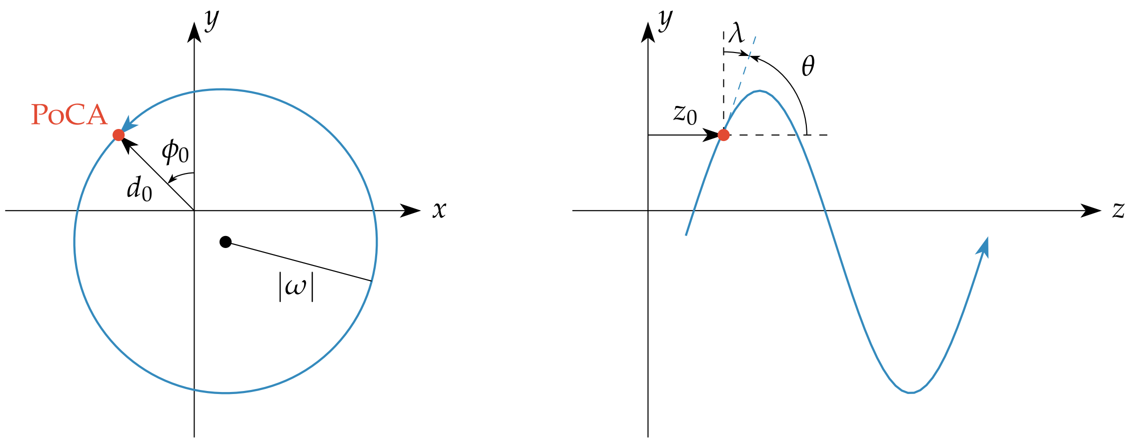


Figure 2.14: The track trajectory model, the helix, with its parameters.

Note that, in this model, both the momentum and the point of origin of the particle are reconstructed. In Belle II, the typical resolution of the *impact parameters* (d_0 and z_0) is $\sim 20\,\mu\mathrm{m}$, while the transverse momentum resolution is given by

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} \simeq \sqrt{(0.1\% \cdot p_\mathrm{T})^2 + \left(\frac{0.3\%}{\beta}\right)^2}, \tag{2.5}$$

where β is the velocity in units of c [62]. Also the charge sign is deduced from the rotation direction in the transverse plane. The mass, instead, must be given as an assumption when using the track in the analysis process: for this reason, it is called *mass hypothesis*.

TRACK RECONSTRUCTION EFFICIENCY The track reconstruction efficiency at Belle II is $\sim 99\%$ for high-transverse-momentum tracks, *i.e.* $p_\mathrm{T} \gtrsim 255\,\mathrm{MeV}/c$. This is the minimum transverse momentum required so that the particle leaves the CDC volume through the barrel wall, *i.e.* so that the helix diameter is greater than $\sim 1.2\,\mathrm{m}$. Once the particle exits the CDC, it either stops in the ECL material, or it encounters the fringes of the magnetic field, allowing it to escape the detector. The tracking subdetectors therefore only capture one "arm" (about half a turn of the helix) of the particle's trajectory: this is the condition for the majority of the tracks of interest at Belle II, and the track reconstruction algorithms are optimized for this case.

Tracks with lower transverse momentum, instead, curl back into the CDC volume. If their longitudinal momentum is low enough (*i.e.* if $\cos\theta \sim \pi/2$), they may complete one or more full turns into the CDC volume. The track reconstruction chain begins from the pattern recognition in the CDC axial layers, which are essentially only sensitive to the position in the transverse plane. In this plane, curling tracks are circles (barring energy losses): there is, therefore, a relatively high probability that the wrong assumption for the charge sign is used. The estimate of the parameters of tracks reconstructed with wrong charge assumption are largely incorrect; it is safe to assume that these tracks will never contribute to decay reconstruction, and treat this effect as an inefficiency.

It is important to note that this inefficiency is charge-sign-dependent. Since it has a large impact on this work, it is worth to describe this effect in detail. Let us define the asymmetry of the track reconstruction efficiency as

$$\mathcal{A}_{\text{tracking}} = \frac{\varepsilon^+ - \varepsilon^-}{\varepsilon^+ + \varepsilon^-}, \tag{2.6}$$

where $\varepsilon^\pm$ are the efficiencies for positive- and negative-charge particles. Figure 2.15 shows the asymmetry of the track reconstruction efficiency computed on simulation[2] as a function of the transverse momentum and polar angle of the track. The figure highlights several regions.

- $p_\text{T} \gtrsim 255\,\text{MeV}/c$, *i.e.* particles that exit the CDC through its barrel wall before completing half a turn of the helix: here the asymmetry is negligible.
- $p_\text{T} \lesssim 255\,\text{MeV}/c$ but large $|\cos\theta|$, *i.e.* particles that exit the CDC though either of its endcap walls before completing a full turn of the helix: also here the asymmetry is negligible.
- $p_\text{T} \lesssim 255\,\text{MeV}/c$ and small $|\cos\theta|$, *i.e.* particles that complete at least one full turn within the tracking subdetector: here asymmetries are large, up to $\mathcal{O}(10\%)$. Eight p_T bands can be seen in this region, corresponding to the last superlayer of the CDC reached by the particle before curling back:
  - if the last superlayer is "stereo" (*i.e.* the CDC has some sensitivity to longitudinal position), the asymmetry is lower;
  - if the last superlayer is "axial" (*i.e.* the CDC has no longitudinal position sensitivity), the asymmetry is larger.

To reconstruct a track, at least three hits are necessary. Tracks with transverse momentum below $\sim 30\,\text{MeV}/c$ are unable to reach the third SVD layer, thus can never be reconstructed[3].

The above is valid for tracks that originate from around the IP. Tracks that originate elsewhere need different treatment. I will not discuss this, however, since all the tracks needed in this work originate within a few mm from the IP, and I apply appropriate selections on d_0 and z_0 to reject tracks that do not.

IN-FLIGHT DECAYS Pions, kaons and muons, while relatively long-lived, have a non-negligible chance of decaying while inside the tracking subdetectors. This occurrence is called *in-flight decay*.

The final states of these decays are almost exclusively made of a long-lived charged particle, plus one or more neutrinos ($\pi \to \mu\nu_\mu$, $K \to \mu\nu_\mu$, $\mu \to e\nu_e\nu_\mu$). Depending on the amount of momentum taken by the neutrinos, a more or less evident *kink* appears on the particle's

2 The efficiencies are computed as $\varepsilon^\pm = n^\pm_\text{reconstructed}/n^\pm_\text{generated}$.
3 Because such tracks suffer heavily from energy losses, the efficiency already begins to drop at $p_\text{T} \sim 100\,\text{MeV}/c$.

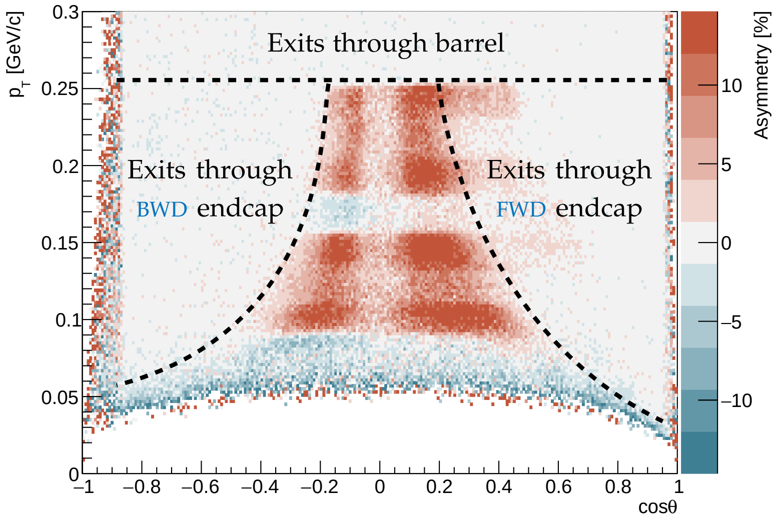


Figure 2.15: Charge sign asymmetry of the track reconstruction efficiency as a function of the transverse momentum p_{T} and polar angle θ of the track, in simulation, for tracks originating near the IP. The three regions where the asymmetry is negligible are highlighted.

trajectory. When the kink is small enough, the track is reconstructed using hits from both the mother and the daughter at once. This causes a degradation of the resolution of the track parameters; still, for many analyses, the track is still usable.

In this work, signal candidates built using tracks with in-flight decays are considered together with the rest of the signal candidates, without distinction or special treatment.

2.4.2 *Particle identification*

Information from the SVD, CDC, TOP, ARICH, ECL and KLM contribute to the identification of charged particles. For each subdetector, the information is summarized by a likelihood $\mathcal{L}^{\text{subdetector}}_{\text{hypothesis}}$, with hypotheses running over the six long-lived charged particle species considered. A global likelihood may then be computed as

$$\mathcal{L}_h = \prod_{s \in \text{subdetectors}} \mathcal{L}^s_h\,, \tag{2.7}$$

and a particle identification "probability" may be defined as

$$p(h \textit{ vs.} \text{ other hypotheses}) = \frac{\mathcal{L}_h}{\mathcal{L}_h + \sum_{h' \in \text{other hypotheses}} \mathcal{L}_{h'}}\,. \tag{2.8}$$

In this work, I only use "binary" identification probabilities of the form "π *vs.* K".

The particle identification may be used in a selection criterion of the form $p(h \textit{ vs.} \text{ others}) \gtrless c$, with $c \in (0, 1)$. This will give some identification efficiency (*i.e.* probability to identify a particle of type h as h) and fake rate (*i.e.* probability to identify a particle of other type as h), which depends on the hypotheses considered, on the cut c, and on how the particle interacts with the detectors.

The dependence of these quantities on the interactions with the detectors are rather complex: in principle, one would have to know which subdetectors are reached by the particle, which parts of the subdetector, with which momentum, incidence angle, and so on. However, for tracks that originate near the IP, all these dependencies may be empirically reduced to the kinematic variables p_t and θ, and the sign of the charge. The angle ϕ of the particle's momentum could also be considered, however dependencies on the azimuthal angle are typically very small, so I will not take them into account.

NEURAL-NETWORK BASED PID Since the likelihoods for the various subdetectors are modeled empirically, they fail, in some cases, to fully exploit the complexity of the particle-detector interaction. To improve the PID performance, a neural network (NN) has been trained on simulation to improve the identification probability. This NN-based PID uses as input the likelihoods from the single subdetectors, as well as the particle's momentum, angles, and ECL cluster information (if a cluster is associated to the particle). Note that the simulation, while unable to completely model particle-detector interactions in all their details, takes into account the detector conditions of the various data-taking periods.

This work mainly uses the NN-based PID.

CONTRIBUTIONS OF THE SINGLE SUBDETECTORS The tracking subdetectors (SVD and CDC) provide energy loss $(\mathrm{d}E/\mathrm{d}x)$ measurements. These are particularly useful for low-momentum (sub-GeV$/c$) π–K discrimination, as shown in Figure 2.16.

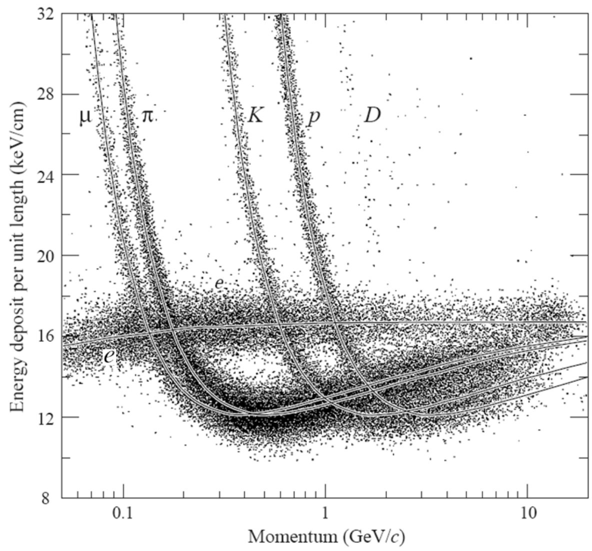


Figure 2.16: Energy loss $(\mathrm{d}E/\mathrm{d}x)$ for various particle species as a function of the particle's momentum.

The dedicated PID subdetectors (TOP and ARICH), instead, provide measurements of the Čerenkov angle, which are useful, for instance, for discriminating pions from kaons at higher momenta.

When a track intersects the ECL, a cluster that is sufficiently close to the intercept is associated to it. The ratio between cluster energy and track momentum is useful for PID, particularly for identifying electrons, which tend to lose all of their energy in the calorimeter through electromagnetic showers, unlike other particles.

Finally, the same cluster-track association is performed with KLM clusters. Electrons tend to stop in the ECL; hadrons tend to stop in the first KLM layers; muons, instead, tend to pass through all detector layers loosing very little energy. Therefore, this information is particularly useful for identifying high-momentum muons.

2.4.3 *Photon reconstruction*

Photon candidates are made by taking an ECL cluster with no track associated to it, *i.e.* for which no charged particle passes close enough. Since the ECL acceptance is wider than that of the tracking subdetectors (the tightest being that of the CDC, 17° to 150°), one typically selects only photons within the CDC acceptance: this discards clusters that may be related to an undetected charged particle.

Since an ECL cluster provides a single point to constrain the photon trajectory, the point of origin must be assumed as an hypothesis. In basf2, photons are assumed to originate at the coordinates origin, which roughly coincides with the IP; essentially, only the angular coordinates of the photon momentum (θ and ϕ) are measured. The point-of-origin hypothesis may be adjusted *a posteriori*, typically through a decay vertex fit that uses other information (*e.g.* track trajectories) to constrain the vertex position.

The lack of a measurement of the origin of the photons precludes the possibility of rejecting beam-induced background clusters based on the fact that they do not originate from the IP (which is the main filter for beam-background tracks). This is particularly relevant when looking for low-energy (order of 100 MeV or less) photons, as the calorimeter is bombarded by low-energy showers from stray beam particles, particularly in the endcap sections. To reject these clusters, other selections must be used, *e.g.* based on the time of arrival or cluster shape variables.

The ECL is almost 100% efficient for photons of energy above 2 GeV; below this energy, the efficiency gradually drops until it reaches about 80% in the $\mathcal{O}(100\,\mathrm{MeV})$ energy range. The efficiency is almost uniform throughout the acceptance. Some of the photons interact with the material in front of the calorimeter (TOP, ARICH, and mechanical structures of the CDC) and are detected as showers. These, however, tend to be rejected by selections on the cluster shape variables, as they become nearly indistinguishable from the showers of stray beam particles.

π^0 RECONSTRUCTION Since π^0 mesons decay to two photons almost 99% of the time [15], π^0 candidates are made by combining pairs of photon candidates. A selection on the invariant mass of the pair is used to reject most of the random combinations: as the mass resolution is about $6.5\,\mathrm{MeV}/c^2$, this window typically spans from about $100\,\mathrm{MeV}/c^2$ to about $150\,\mathrm{MeV}/c^2$. However, a large number of random combination inevitably falls within the window, especially when looking for low-momentum π^0 mesons, for which the low-energy beam-induced background clusters become very relevant.

Depending on the necessary efficiency and purity, several solutions may be adopted to suppress this background. Selection on the aforementioned cluster properties (time of arrival and shape variables) are a possibility. Additionally, fits using geometric and kinematic information may be used; this is especially useful when precise vertex position information can be reconstructed from other particles involved in the decay, especially charged particles for which a point-of-origin measurement is available.

3 ANALYSIS OVERVIEW AND STRATEGY

The target of this work is to measure the time-integrated CP asymmetry $\mathcal{A}_{CP}$ in the $D^0 \to \pi^+\pi^-\pi^0$ decay channel. In Section 3.1 I introduce the strategy I employ to extract this observable from the data. The greatest challenge is accounting for the asymmetries introduced by the detector and by the D^0 production process, which I describe in detail in Section 3.2 and Section 3.3. Finally, a schematic summary of the various steps can be found in Section 3.4.

3.1 GENERAL STRATEGY

The $\mathcal{A}_{CP}$ observable, defined in Equation 1.45, is a function of the partial decay widths Γ of the decay; however, the partial decay widths of neutral meson decays are not directly measurable. Therefore, in order to access $\mathcal{A}_{CP}$, I define the "raw asymmetry" as

$$\mathcal{A}_{\text{raw}} = \frac{n(D^0 \to \pi^+\pi^-\pi^0) - n(\overline{D}^0 \to \pi^+\pi^-\pi^0)}{n(D^0 \to \pi^+\pi^-\pi^0) + n(\overline{D}^0 \to \pi^+\pi^-\pi^0)}\,, \tag{3.1}$$

where $n(i \to f)$ represents the yield, *i.e.* the number of observed $i \to f$ decays, which is more readily accessible from the sample of reconstructed $D^0 \to \pi^+\pi^-\pi^0$ decays.

I reconstruct D^0 and $\overline{D}^0$ meson candidates from their decay products: this is described in detail in Section 4.2. Since the final states for the two flavors appear identical from the experimental point of view, additional information is needed to determine (to "tag") the D^0 candidate's flavor at production. I reconstruct D^* candidates in the strong $D^{*+} \to D^0\pi^+_{\text{tag}}$ and $D^{*-} \to \overline{D}^0\pi^-_{\text{tag}}$ decays: this allows me to exploit the sign of the charge of the "tag" pion to infer the flavor of the D^0.

At this point, I perform a simultaneous fit of the D^0 candidate mass and D^*–D^0 mass difference to extract the raw asymmetry $\mathcal{A}_{\text{raw}}$: this is described in Chapter 5. In principle, instead of fitting, I could simply count the number of reconstructed candidates, assuming that my selection criteria produce a sample pure enough, and that I have a way to reliably estimate the residual background: this is the strategy employed by previous B-factory measurements such as [41]. However, by fitting variables that discriminate signal from background, I can accept a less pure sample, which allows me employ less stringent selection criteria, significantly increasing the reconstruction efficiency and the sample size.

The raw asymmetry that I measure is due to several flavor-asymmetric effects. Assuming that all of these asymmetries are small, I can write (truncating at the first order)

$$\mathcal{A}_{\text{raw}} \simeq \mathcal{A}_{\text{det}} + \mathcal{A}_{\text{prod}} + \mathcal{A}_{CP}\,. \tag{3.2}$$

Proof of this approximation is given in Section A.1. Besides the CP asymmetry $\mathcal{A}_{CP}$ that I aim to measure, $\mathcal{A}_{\text{raw}}$ includes two other asymmetries, which are typically dubbed "experimental asymmetries", that I need to remove: $\mathcal{A}_{\text{det}}$, which is due to detection and reconstruction effects; and $\mathcal{A}_{\text{prod}}$, which is due to the production process of the charmed mesons. Section 3.2 and Section 3.3 describe these experimental asymmetries and the methods I use to remove them.

3.2 DETECTION ASYMMETRIES

Detection asymmetries are caused by the geometry and functioning of the detector components, by the reconstruction algorithms, and by the selection criteria that I impose on the signal candidates. When the final reconstruction efficiency is different for D^0 and $\overline{D}^0$ mesons, a detection asymmetry is present. However, to understand these effects, we must separately consider the reconstruction efficiency of each final-state particle of the decay chain, *i.e.* those particles that reach the detector: all the charged pions, and the two photons from the decay of the neutral pion.

The reconstruction efficiency of charged particles depends on two terms: the tracking efficiency, *i.e.* the probability to identify the particle's trajectory in the tracking detectors and correctly reconstruct its parameters, which was introduced in Section 2.4.1; and the particle identification (PID) efficiency, *i.e.* the probability to correctly identify the particle type, which depends on the PID probability criterion imposed, and was introduced in Section 2.4.2. Both of these terms depend on the specific subdetector layers crossed by the particle; additionally, the PID efficiency depends on the particles' momentum; and, finally, the subdetectors' performance is different for particles of different charge sign, since the trajectories and interactions with material are different. While these dependencies are quite complex, all this information ultimately depends only on three variables, assuming our particles originate near the interaction point (IP): the particle's transverse momentum p_{T}, its polar angle θ, and its charge.

The reconstruction efficiency of photons, which was introduced in Section 2.4.3, depends on the region of the electromagnetic calorimeter (ECL) that the photon points to, and on the material that the photon crosses before reaching the ECL. As for the charged particles, for photons originating near the IP, this is completely determined by the photon's energy E and polar angle θ.

A detection asymmetry is produced when the reconstruction efficiency of the D^0 decay products is different from that of the $\overline{D}^0$ decay products, *i.e.* when the distribution of p_{T} (E) and θ of the final-state tracks (photons) is different for the two flavors, or some of the final-state tracks have opposite charge for opposite flavors. The latter condition is particularly relevant for the tag pion. For practical reasons, I treat as one the asymmetry induced by the reconstruction of all D^0 daughters $\mathcal{A}_{\varepsilon}^{D^0}$, but I treat separately the tag pion asymmetry $\mathcal{A}_{\varepsilon}^{\pi\,\mathrm{tag}}$: these two sources of asymmetry are discussed in Section 3.2.1 and Section 3.2.2.

3.2.1 *Asymmetry from the reconstruction of the D^0*

The largest difference among D^0 daughters is found in the transverse momentum distributions of the charged pions, which are shown in Figure 3.1 for simulated[1] signal decays. The difference between π^+ and π^- distributions are a result of the interference between the various resonant and non-resonant contributions to the $D^0 \to \pi^+\pi^-\pi^0$ decay amplitude. Even assuming absence of *CP* violation, the two charged pions are swapped under *CP* conjugation: since the track reconstruction efficiency depends on both charge and transverse momentum, this results in a flavor-dependent D^0 reconstruction efficiency.

1 The simulated samples are described in detail in Section 4.1.

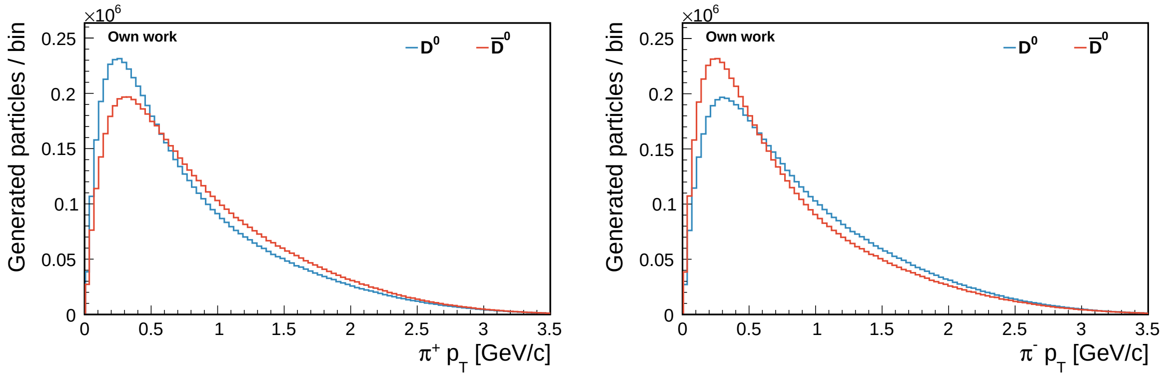


Figure 3.1: Simulated distributions of the transverse momenta of the positive (left) and negative (right) pions produced by the signal D^0 decay, split by D^0 flavor.

REMOVAL METHOD The track reconstruction efficiency only shows sensible asymmetry for low transverse momenta ($p_\mathrm{T} \lesssim 255\,\mathrm{MeV}/c$, as shown in Section 3.2.2). D^0 daughters with such low p_T constitute a small fraction of the total, implying that the induced detection asymmetry must also be relatively small. Therefore, I do not apply any correction for this effect. Instead, I evaluate its magnitude on simulation, and assign a systematic uncertainty for its neglection: this is detailed in Section 8.1.

3.2.2 *Asymmetry from the tagging*

The tag pion transverse momentum and polar angle distributions have very small differences between the two flavors, as shown in Figure 3.2. However, the majority of the tag pions have transverse momentum below $255\,\mathrm{MeV}/c$, *i.e.* within the range where the track reconstruction efficiency depends more heavily on the particle's charge sign, as discussed in detail in Section 2.4.1.

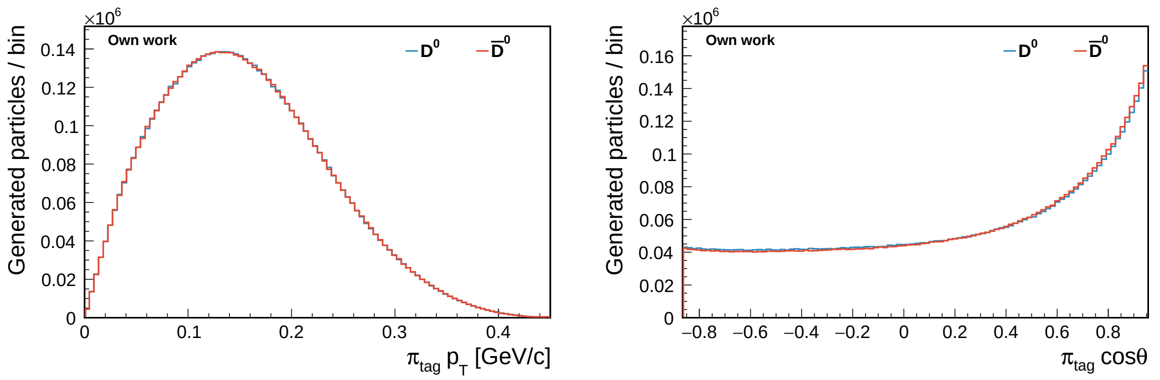


Figure 3.2: Simulated distributions of the transverse momentum and polar angle of the tag pion, split by D^0 flavor. The polar angle range is limited to the detector's acceptance.

REMOVAL METHOD The asymmetry induced by reconstructing the $D^* \to D^0\pi$ decay is much larger than that coming from the D^0 daughters, and of the uncertainty I expect on

$\mathcal{A}_{CP}$. While the simulation reproduces the p_T and θ dependence relatively well, it can not be assumed able to reproduce the magnitude of the asymmetry accurately enough for correcting the raw asymmetry. Instead, I use $D^0 \to K\pi$ decays as a control channel to estimate this effect. $D^0 \to K^-\pi^+$ ($\overline{D}^0 \to K^+\pi^-$) is a Cabibbo-favored decay with a large branching fraction; the presence of only charged particles in the final state makes it easier to reconstruct the decays with both high efficiency and high purity, resulting in a small statistical uncertainty on $\mathcal{A}_{\text{raw}}$; and the different identities of the daughters allow me to infer the D^0 flavor without any tagging. The process is the following.

I reconstruct a "tagged" sample of $D^0 \to K\pi$: the tagging is performed by reconstructing $D^* \to D^0\pi$ decays, imposing on the tag pion the same criteria used for the signal $D^0 \to \pi^+\pi^-\pi^0$ channel. As for the signal channel, I use a fit (to the D^*–D^0 mass difference, see Section 6.3.2) to extract the raw asymmetry

$$\mathcal{A}_{\text{raw}}^{\text{tagged}} \simeq \mathcal{A}_{\varepsilon}^{\pi\,\text{tag}} + \mathcal{A}_{\varepsilon}^{D^0}(D^0 \to K\pi) + \mathcal{A}_{\text{prod}} + \mathcal{A}_{CP}(D^0 \to K\pi)\,, \tag{3.3}$$

which comprises several terms: the tag pion reconstruction asymmetry $\mathcal{A}_{\varepsilon}^{\pi\,\text{tag}}$; the D^0 reconstruction asymmetry for $D^0 \to K\pi$; the production asymmetry $\mathcal{A}_{\text{prod}}$; and the $D^0 \to K\pi$ *CP* asymmetry $\mathcal{A}_{CP}$.

Since I only need $\mathcal{A}_{\varepsilon}^{\pi\,\text{tag}}$, I also reconstruct a second, "untagged" sample of $D^0 \to K\pi$, for which I do not reconstruct the D^*. Also in this case, I use a fit (to the D^0 candidate invariant mass, see Section 6.3.1) to extract the $\mathcal{A}_{\text{raw}}$, which will contain all the terms in $\mathcal{A}_{\text{raw}}^{\text{tagged}}$ except $\mathcal{A}_{\varepsilon}^{\pi\,\text{tag}}$:

$$\mathcal{A}_{\text{raw}}^{\text{untagged}} \simeq \mathcal{A}_{\varepsilon}^{D^0}(D^0 \to K\pi) + \mathcal{A}_{\text{prod}} + \mathcal{A}_{CP}(D^0 \to K\pi)\,. \tag{3.4}$$

Therefore, I can compute the tag pion asymmetry correction as

$$\mathcal{A}_{\varepsilon}^{\pi\,\text{tag}} = \mathcal{A}_{\text{raw}}^{\text{tagged}} - \mathcal{A}_{\text{raw}}^{\text{untagged}}\,, \tag{3.5}$$

which I will then subtract from $\mathcal{A}_{\text{raw}}$ of $D^0 \to \pi^+\pi^-\pi^0$.

In all these operations I assumed that the $\mathcal{A}_{\varepsilon}^{\pi\,\text{tag}}$ of $D^0 \to K\pi$ is the same as the $\mathcal{A}_{\varepsilon}^{\pi\,\text{tag}}$ of $D^0 \to \pi^+\pi^-\pi^0$, and that the other asymmetries are the same in the tagged and untagged $D^0 \to K\pi$ samples. However, this is not guaranteed by simply using the same selection criteria for the relevant particles (tag pion for the first case, D^0 meson and its daughters for the second). The detection asymmetries indeed depend on the kinematic distributions of the involved particles. To make my assumption true, I use per-candidate weights to equalize the distributions of the kinematic variables that the asymmetry depends on.

$\mathcal{A}_{\varepsilon}^{\pi\,\text{tag}}$ is empirically known to depend on the p_T and θ of the tag pion, as discussed in Section 2.4.1. To ensure that the tagged sample has the same $\mathcal{A}_{\varepsilon}^{\pi\,\text{tag}}$ as the $D^0 \to \pi^+\pi^-\pi^0$ sample, I assign weights to the tagged sample so that its tag pion p_T–θ 2D distribution matches that of the $D^0 \to \pi^+\pi^-\pi^0$ sample.

Similarly, to ensure that the untagged sample has the same $\mathcal{A}_{\text{prod}}$ and $\mathcal{A}_{\varepsilon}^{D^0}$ as the tagged sample, I assign weights to the untagged sample so that its distributions match those of the (already weighted) tagged sample. In this case, the relevant variables are the p_T and θ of the D^0 daughters (on which $\mathcal{A}_{\varepsilon}^{D^0}$ depends), plus the $\cos\theta_{\text{CM}}$ of the D^0 (on which $\mathcal{A}_{\text{prod}}$ depends, as explained in Section 3.3). $\mathcal{A}_{CP}$, of course, does not depend on the kinematic of the decay. To ensure the correlations among these five variables are accounted for, I compute the weights iteratively on one variable at a time (see Section 6.4.3).

The $\mathcal{A}_{\varepsilon}^{\pi\,\text{tag}}$ determination procedure is explained in detail in Chapter 6.

3.3 PRODUCTION ASYMMETRY

As the name suggests, the production asymmetry is induced by the process that produces the head of the decay chain, *i.e.* the D^* meson. The selection criteria I illustrate in Section 4.2 ensure that my sample only contains "prompt" D^* mesons, *i.e.* D^* mesons produced in the $e^+e^- \to c\bar{c}$ process and the subsequent hadronization, rather than in the decay of a B meson (or heavier beauty particle).

The production asymmetry $\mathcal{A}_{\text{prod}}$ arises from the interference of the γ and Z^0 contributions to the $e^+e^- \to c\bar{c}$ process. In the collision center of mass (CM) frame, the momentum of the c quark is more likely to point towards the direction of the incoming positrons (*i.e.* backwards, towards negative $\cos\theta$), while the $\bar{c}$ antiquark momentum is more likely to point towards the direction of the incoming electrons (*i.e.* forwards, towards positive $\cos\theta_{\text{CM}}$). For this reason, the production asymmetry is often referred to as "forward-backward asymmetry".

The differential $c\bar{c}$ cross section can be written, at the Z^0 pole, as a function of θ_{CM} (the polar angle of the quark's momentum in the collision CM frame):

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta_{\text{CM}}} \propto 1 + \cos^2\theta_{\text{CM}} + \frac{8}{3}\mathcal{A}_{\text{fb}}\cos\theta_{\text{CM}}\,, \tag{3.6}$$

where $\mathcal{A}_{\text{fb}}$ is the forward-backward asymmetry proper (a constant between -1 and 1) [73]. The resulting c–$\bar{c}$ asymmetry is given by

$$\mathcal{A}_{\text{prod}}(\cos\theta_{\text{CM}}) = \frac{8}{3}\mathcal{A}_{\text{fb}}\frac{\cos\theta_{\text{CM}}}{1+\cos^2\theta_{\text{CM}}}\,, \tag{3.7}$$

which is an odd function of $\cos\theta_{\text{CM}}$.

Since the momentum of the c quark is not reconstructible, I use the momentum of the D^0 instead. I could use the momentum of the D^* as well, however this is not available in the untagged $D^0 \to K\pi$ sample, therefore I use the D^0 momentum everywhere for consistency. The functional shape, as well as the magnitude of the asymmetry as a function of the D^0 $\cos\theta_{\text{CM}}$ differs from Equation 3.7; however, since both the hadronization process and the $D^* \to D^0\pi$ decay are caused by the *CP*-conserving strong interaction, the asymmetry is still an odd function of the D^0 $\cos\theta_{\text{CM}}$.

Figure 3.3 shows the production asymmetry computed on simulation as a function of the D^0 $\cos\theta_{\text{CM}}$. Since no parametrization exists for the production asymmetry as a function of the D^0 $\cos\theta_{\text{CM}}$, the simulation simply uses Equation 3.7 as the best available approximation. From other analyses in the charm sector, the actual asymmetry on data is known to be over a factor two larger.

Being an odd function, the integral of the production asymmetry is zero, so it would seem that it should not contribute to the raw asymmetry in any way. However, for the sample of reconstructed decays, this is only correct if the reconstruction efficiency is an even function of $\cos\theta_{\text{CM}}$, which is not the case, as shown in Chapter 5 (Figure 5.1).

REMOVAL METHOD Even if the reconstruction efficiency is not $\cos\theta_{\text{CM}}$-even, I can still exploit the $\cos\theta_{\text{CM}}$-oddness of the production asymmetry to cancel it. I split my sample in a number of $\cos\theta_{\text{CM}}$ bins that are:

- narrow enough that the reconstruction efficiency is approximately constant within each bin; and

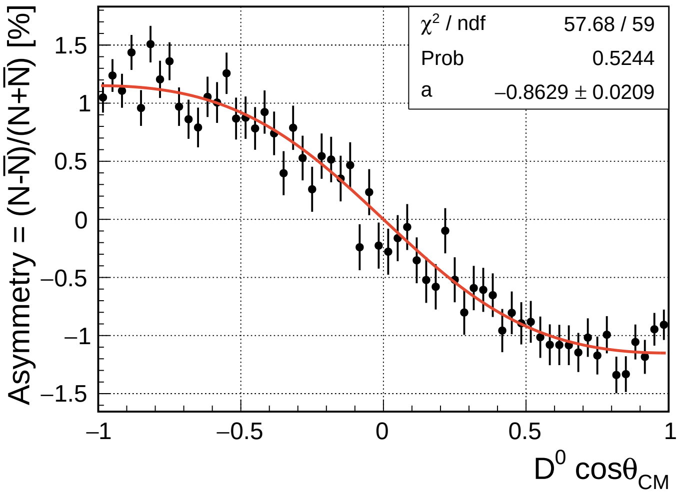


Figure 3.3: D^0 production asymmetry in simulated events, as a function of the true $\cos\theta_{\mathrm{CM}}$. The red line shows a fit with Equation 3.7.

- symmetric with respect to $\cos\theta_{\mathrm{CM}} = 0$, *i.e.* for each bin $\cos\theta_{\mathrm{CM}} \in [a, b)$ there is a corresponding symmetric bin $\cos\theta_{\mathrm{CM}} \in [-b, -a)$.

Consider a pair of symmetric bins, denoted by "+" (for the bin in $\cos\theta_{\mathrm{CM}} > 0$) and "−" (for the bin in $\cos\theta_{\mathrm{CM}} < 0$). Let us measure the raw asymmetry in each bin, and assume that the detection asymmetries have already been removed: only $\mathcal{A}_{\mathrm{prod}}$ and $\mathcal{A}_{CP}$ remain. Therefore, the asymmetries in the two bins $\mathcal{A}^+$ and $\mathcal{A}^-$ are, in average,

$$\mathcal{A}^+ = \mathcal{A}_{CP} + \mathcal{A}^{\mathrm{bin}}_{\mathrm{prod}}, \qquad \mathcal{A}^- = \mathcal{A}_{CP} - \mathcal{A}^{\mathrm{bin}}_{\mathrm{prod}}\,. \tag{3.8}$$

Because the efficiency is constant within each bin, the $\mathcal{A}_{\mathrm{prod}}$ of the two bins are opposite. Note that the efficiency simplifies in the ratio of Equation 3.1 (assuming that the bias due to flavor-asymmetric efficiencies has already been corrected): as this happens independently for each bin, different efficiencies in the two bins do not affect the result.

I can now compute the unweighted average of the asymmetries from the two bins:

$$\mathcal{A}^{\mathrm{avg}} = \frac{\mathcal{A}^+ + \mathcal{A}^-}{2} = \frac{\mathcal{A}_{CP} + \cancel{\mathcal{A}^{\mathrm{bin}}_{\mathrm{prod}}} + \mathcal{A}_{CP} - \cancel{\mathcal{A}^{\mathrm{bin}}_{\mathrm{prod}}}}{2} = \mathcal{A}_{CP}\,. \tag{3.9}$$

The production asymmetry correctly cancels, while the *CP* asymmetry remains.

I will repeat this procedure for every pair of bins, "folding" the $\cos\theta_{\mathrm{CM}} < 0$ half of the sample over the $\cos\theta_{\mathrm{CM}} > 0$ half. For each pair, I will obtain a measurement of the same $\mathcal{A}_{CP}$: these measurements should all be consistent with each other, and I will compute the final $\mathcal{A}_{CP}$ measurement by fitting the bin-pair measurements with a constant function.

3.4 SUMMARY

Once I reconstruct my sample of $D^0 \to \pi^+\pi^-\pi^0$ decays, I use a fit to measure its raw asymmetry $\mathcal{A}_{\mathrm{raw}}$. This observable contains the *CP* asymmetry $\mathcal{A}_{CP}$ that I want to measure, the asymmetry induced by the D^* production process $\mathcal{A}_{\mathrm{prod}}$, and the asymmetry induced by the reconstruction efficiency of the tag pion $\mathcal{A}^{\pi\,\mathrm{tag}}_{\varepsilon}$. The latter two must be removed.

I can measure $\mathcal{A}_{\varepsilon}^{\pi\,\text{tag}}$ as the raw asymmetry difference of the tagged and untagged $D^0 \to K\pi$ control samples. Measuring this effect directly on data is necessary, since it depends on the detector itself, and the simulation might not be able to reproduce all involved effects accurately enough.

To remove $\mathcal{A}_{\text{prod}}$, I split my sample in bins of $\cos\theta_{\text{CM}}$ of the D^0 that are symmetric with respect to $\cos\theta_{\text{CM}} = 0$. I measure $\mathcal{A}_{\text{raw}}$ in each bin, and subtract $\mathcal{A}_{\varepsilon}^{\pi\,\text{tag}}$ bin-wise. Since $\mathcal{A}_{\text{prod}}$ is an odd function of $\cos\theta_{\text{CM}}$, and the binning evens out efficiency effects, I compute the average asymmetry of each pair of symmetric bins. In this operation, called "folding", $\mathcal{A}_{\text{prod}}$ cancels out, while $\mathcal{A}_{CP}$, which does not depend on $\cos\theta_{\text{CM}}$, remains.

In summary:

1. I measure the $\mathcal{A}_{\text{raw}}$ of the $D^0 \to \pi^+\pi^-\pi^0$ sample in each $\cos\theta_{\text{CM}}$ bin;
2. I measure the $\mathcal{A}_{\text{raw}}$ of the $D^0 \to K\pi$ control samples in each $\cos\theta_{\text{CM}}$ bin, and compute each bin's $\mathcal{A}_{\varepsilon}^{\pi\,\text{tag}}$ as the difference between tagged and untagged sample asymmetry;
3. I subtract $\mathcal{A}_{\varepsilon}^{\pi\,\text{tag}}$ from the $D^0 \to \pi^+\pi^-\pi^0$ $\mathcal{A}_{\text{raw}}$ bin-wise;
4. I "fold" the $\cos\theta_{\text{CM}}$ bins, *i.e.* compute the average asymmetry in each pair of symmetric bins; and
5. I fit the "folded" asymmetries, which are all $\mathcal{A}_{CP}$ measurements, with a constant function to obtain the final measurement.

CLOSED-BOX ANALYSIS To avoid experimenter's biases, the analysis is first performed with closed box. This is achieved by shifting the data $\mathcal{A}_{\text{raw}}$ of each $\cos\theta_{\text{CM}}$ bin by the same, unknown amount between -10% and 10%. Signal, tagged, and untagged sample asymmetries are shifted by three different unknown amounts to avoid cancellations when computing the corrected asymmetries. The whole analysis procedure, including the systematic uncertainties and consistency checks, is finalized before opening the box. In this thesis I show preliminary open-box results (not yet approved by the Belle II collaboration).

4 SIGNAL RECONSTRUCTION AND SELECTION

In this chapter I describe the data that I use, and the I take to produce the sample of signal decays from which I extract the asymmetry (the asymmetry extraction process is described later, in Chapter 5). Section 4.1 introduces the Belle II samples that I use, including both real-world data and Monte Carlo (MC) simulation. Section 4.2 described the candidate reconstruction process, along with all the selection criteria I employ to reject backgrounds, and compares the expectations from MC to the real data distributions.

4.1 DATA AND SIMULATED SAMPLES

The Belle II experiment begun taking data in March 2019, and stopped in June 2022 as the long shutdown 1 (LS1) begun. During this 2019-2022 period, dubbed *Run 1*, $428\,\mathrm{fb}^{-1}$ of collisions were recorded [55]. Data-taking resumed in January 2024, and an additional $\sim 150\,\mathrm{fb}^{-1}$ was recorded so far; however, due to the time necessary for calibration and offline reprocessing, my work is only based on Run 1 data.

The majority of the data ($365\,\mathrm{fb}^{-1}$) was taken with a collision energy $\sqrt{s}$ equal to the $\Upsilon(4S)$ mass (about $10.58\,\mathrm{GeV}/c^2$). About one tenth of the data ($44\,\mathrm{fb}^{-1}$) was taken with $\sqrt{s}$ below the $B\bar{B}$ pair production threshold (at $\sim 10.52\,\mathrm{GeV}$): these are called *off-resonance* runs, and are mainly intended for calibration (*e.g.* $B\bar{B}$ pair counting). Finally, about one twentieth of the data ($19\,\mathrm{fb}^{-1}$) was taken while scanning $\sqrt{s}$ above the $\Upsilon(4S)$ mass (up to $10.8\,\mathrm{GeV}$): these are referred to as $\Upsilon(5S)$ *energy scans*, and are mainly intended for quarkonium physics analyses and searches for hadronic resonances. The $e^+e^- \to c\bar{c}$ process is almost unaffected by these changes in $\sqrt{s}$, therefore in this work I use all $428\,\mathrm{fb}^{-1}$ of Run 1.

Besides the real-world data, I also use large simulated samples. The optimization of the selection criteria and study of the signal and background components, as well as the estimation of some systematic uncertainties, are based on these samples. Belle II MC samples are centrally produced, and correspond to about four times the data luminosity. Also, these samples are *run dependent*: events are produced for each data-taking *run* (periods of no more than 8 h), in an amount proportional to the run's integrated luminosity, taking into account the actual detector conditions and calibrations *for that specific period* as much as possible. Moreover, beam-induced background hits, which are notoriously difficult to simulate accurately, are taken from data (using a dedicated random trigger line) and overlaid onto the simulated events.

The simulation process is the following.

1. The $e^+e^- \to X$ process is simulated by an appropriate generator; for instance, PYTHIA8 [67] and KKMC [68] simulate $e^+e^- \to q\bar{q}$ processes, using a combination of models and experimental inputs (*e.g.* fragmentation factors).
2. EvtGen [66] simulates the decay of the particles X from the previous step, and of their daughters, recursively, until only long-lived particles that may reach the detector remain; the inputs to EvtGen are tables of decay modes with their branching fractions and models for the daughters' kinematic distributions.

3. Geant4 [71] simulates the flight of the particles, their interaction with the detector material (active and non), and their decay (if appropriate); the inputs to Geant4 are the geometry and material properties of the detector and its mechanical structure, a map of the solenoidal magnetic field, and the mean life and decay modes of the particles, as well as their interaction cross section with the various materials.
4. Dedicated basf2 modules for each Belle II subdetector simulate the detector response to the particle-detector interactions simulated by Geant4 (including secondary particles generated by Geant4 itself); this simulation includes both the detector physics and its electronics, and is often parametrized to achieve a reasonably fast simulation.
5. Detector hits from beam-induced backgrounds (taken from real-data random-trigger events) are added to the simulated hits.
6. The level 1 (L1) trigger logic is simulated, using the detector hits as input; if the response of the trigger to the event is negative, the event is discarded.
7. The simulated detector information is passed to the reconstruction process; this is, at every step, the same process used for real data, including the high-level trigger (HLT) decision and flagging.
8. Reconstructed particle candidates are matched to the corresponding generated particles, if any; depending on the subdetectors, the matching also takes into account the quality of the reconstructed particle and its measured parameters.

For my work, I use *generic* MC samples of $e^+e^- \to B^0\overline{B}^0$, B^+B^-, $c\overline{c}$, $s\overline{s}$, $u\overline{u}$, $d\overline{d}$, and $\tau^+\tau^-$ events, accounting for $1712\,\mathrm{fb}^{-1}$ (exactly four times the data luminosity). I also use dedicated *signal* MC samples of $e^+e^- \to c\overline{c}$ events where at least one D^{*+} (or D^{*-}) is present, and is forced to decay to $D^0\pi^+$ ($\overline{D}^0\pi^-$), with the D^0 ($\overline{D}^0$) decaying to $\pi^+\pi^-\pi^0$. The signal samples contain 13×10^6 events, corresponding to $1806\,\mathrm{fb}^{-1}$ (a bit more than four times the data luminosity).

The signal samples use a modified EvtGen model for the $D^0 \to \pi^+\pi^-\pi^0$ decay kinematics, which fixes some inconsistencies between the generated Dalitz plot (DP) distribution, and that observed by the *BABAR* experiment in [39]. The amplitude analysis results from this paper are used as input to the EvtGen DP model. More details about the modification may be found in [74]. To avoid double-counting, $e^+e^- \to c\overline{c}$ events with a signal decay ($D^* \to D^0\pi$ with $D^0 \to \pi^+\pi^-\pi^0$) are removed from the generic MC sample.

4.1.1 *Sample skim*

To allow faster sample processing and minimize the usage of computational resources, *skimmed* versions of the data and MC samples are produced during the centralized offline reprocessing. These *physics skims* are different from the HLT skims described in Section 2.3, as they are dedicated to specific channels or groups of similar channels.

Physics skims are produced by reconstructing candidates for specific decay chains with loose selections, and retaining only events with at least one candidate. The physics skim I use for the $D^0 \to \pi^+\pi^-\pi^0$ signal sample requires:

1. that the event is flagged as "hadronic" at the HLT skim level; and
2. the presence of at least one $D^* \to D^0\pi$, $D^0 \to h^+h'^-\pi^0$ candidate in the event, where h and h' are either pions or kaons in any combination.

The first requirement is computed during data taking, and translates in the presence, inside the event, of at least three tracks with relatively large transverse momentum and originating from the interaction point (IP): $p_T > 200\,\mathrm{MeV}/c$, $|d_0| < 2\,\mathrm{cm}$, and $|z_0| < 4\,\mathrm{cm}$. Events consistent with Bhabha scattering are also loosely vetoed.

The second requirement applies to *reconstructed* D^* candidates. Physics skim employ the same candidate reconstruction strategy and tools as actual analyses. Since I describe this in detail in Section 4.2, I defer this explanation (and the definition of the variables used), and I only report the skim's selection criteria in Table 4.1. The only noteworthy difference with respect to my selection criteria is the requirement that a kinematic fit, performed with KFit of the π^0 meson parameters converges. This fit is performed using KFit, the kinematic fitting algorithm of the Belle experiment, which is based on [75]. All other criteria are either unchanged or tightened in the final set of selections.

Variable	Selection
Photons	
`clusterNHits`	> 1.5
θ	within CDC acceptance (17° to 150°)
E	$> 25\,\mathrm{MeV}$ for forward and barrel ECL
E	$> 40\,\mathrm{MeV}$ for backward ECL
Neutral pion	
M	between $105\,\mathrm{MeV}/c^2$ and $150\,\mathrm{MeV}/c^2$
KFit	must converge
Charged particles	
d_r	$< 1\,\mathrm{cm}$
$\lvert d_z \rvert$	$< 3\,\mathrm{cm}$
θ	within CDC acceptance (17° to 150°)
π`IDNN`	> 0.1 (for pions)
K`IDNN`	> 0.1 (for kaons)
D^0 **meson**	
M	between $1.7\,\mathrm{MeV}/c^2$ and $2.1\,\mathrm{MeV}/c^2$
D^* **meson**	
ΔM	$< 160\,\mathrm{MeV}/c^2$
p_{CM}	$> 2\,\mathrm{GeV}/c$

Table 4.1: The selection criteria used by the $D^* \to D^0\pi, D^0 \to h^+h'^-\pi^0$ physics skim.

It is important to note that I apply the skim selection to all the samples I use. This reflects on the distributions and values I show throughout this thesis.

4.2 SIGNAL CANDIDATES RECONSTRUCTION

Signal candidates are reconstructed using the basf2 tools for analysis. A candidate for a *mother* particle is built by grouping candidates corresponding to the *daughter* particles, and computing the mother's four-momentum as the sum of the four-momenta of the daughters. Daughter candidates may be both final-state particles (FSPs), or particles reconstructed previously using the same process; this allows to reconstruct a decay chain starting from the FSPs and going backwards.

The basf2 tools build *lists* of candidates by considering all possible combinations of daughters that are physically meaningful (*e.g.* the same FSP is not used twice in the same candidate, the charges are correct, the correct mass hypotheses are used for tracks, *etc.*). To keep the number of combinations manageable, some loose pre-selections are applied already when forming the candidates: those I use for my signal channel are listed in Appendix B; it is important to note that these selection are already applied in all distributions and values I show in this thesis.

The first step in reconstructing my signal channel, $D^0 \to \pi^+\pi^-\pi^0$, is to reconstruct the neutral pion. Then, I combine π^0 candidates with two oppositely-charged tracks to form D^0 candidates. Finally, I combined D^0 candidates with tracks of any charge to form $D^* \to D^0\pi$ candidates from which the flavor of the D^0 can be inferred. In the following, I describe in detail each step, the selections I impose, and the reason behind them; the selection criteria are also summarized in Table 4.2.

I reconstruct π^0 mesons from pairs of photons: this covers $\sim 98.8\%$ of the π^0 decays [15]. A substantial background is produced by random combinations of unrelated photons or beam-induced background clusters: to suppress this, I use several selection criteria on the photon candidates themselves, which I report in the following list.

- `clusterNHits` > 1.5: this variable counts the number of ECL crystals that take part in the cluster, which can be fractional when the energy detected by a single crystal is subdivided among two nearby clusters; this is a standard selection to reject noise-induced photon candidates.
- θ within CDC acceptance: as explained in Section 2.4.3, this criterion ensures that, if the cluster is produced by a charged particle, said particle can be detected by the tracking subdetectors, and thus the cluster can be excluded from the photon candidates.
- $E > 100\,\text{MeV}$: the vast majority of beam-induced background clusters have very low energy in the laboratory frame, synchrotron radiation being one of the main culprits; thus, minimum energy requirements are effective at suppressing them, with small impact on the reconstruction efficiency since my D^* mesons are boosted in the laboratory frame (see later selections).
- $|$`clusterTime`$| < 200\,\text{ns}$: this variable is the time of arrival of the cluster, measured with respect to the event time; this criterion is meant to suppress beam-induced background clusters, which are uniformly distributed in time; the 200 ns value is a standard selection based on the electromagnetic calorimeter (ECL) time resolution measured on data.
- $|$`clusterTimeSignificance`$| < 2$: this variable is the ratio between `clusterTime` and its uncertainty, parametrized so that the ± 1 significance range contains 99% of the correctly-reconstructed photons in simulation; the rationale behind this selection is the same as

Variable	Selection
Photons	
`clusterNHits`	> 1.5
θ	within CDC acceptance (17° to 150°)
E	$> 100\,\mathrm{MeV}$
|`clusterTime`|	$< 200\,\mathrm{ns}$
|`clusterTimeSignificance`|	< 2
E_1/E_9	> 0.3
E_9/E_{21}	> 0.8
`clusterZernikeMVA`	> 0.1
Neutral pion	
M	between $116\,\mathrm{MeV}/c^2$ and $150\,\mathrm{MeV}/c^2$
p	$> 500\,\mathrm{MeV}/c$
Charged particles	
d_r	$< 0.5\,\mathrm{cm}$
$\lvert d_z \rvert$	$< 1\,\mathrm{cm}$
θ	within CDC acceptance (17° to 150°)
π`IDNN`	> 0.2 (D^0 daughters)
π`IDNN`	> 0.1 (tag pion)
D^0 meson	
M	between $1.785\,\mathrm{GeV}/c^2$ and $1.95\,\mathrm{GeV}/c^2$
$M(\pi^+\pi^-)$	**not** between $470\,\mathrm{MeV}/c^2$ and $530\,\mathrm{MeV}/c^2$ (K^0_S veto)
D^* meson	
TreeFit	must converge
ΔM	between $140\,\mathrm{MeV}/c^2$ and $150\,\mathrm{MeV}/c^2$
X	> 1.06 (veto on D^* mesons from B meson decays)
Best candidate selection	Keep candidate with highest TreeFit probability

Table 4.2: The selection criteria used for signal candidates.

for `clusterTime`; the value of 2 is a standard selection based on the comparison of data and simulated distributions.

- $E_1/E_9 > 0.3$: this variable is the ratio between the energy deposited in the central crystal of the cluster, and that deposited in the 3×3 crystals around the central one (see Figure 4.1); since photon clusters tend to be narrower than clusters produced by other particles, this variable provides some discriminating power against non-photon clusters.
- $E_9/E_{21} > 0.8$: this variable is the ratio between the energy deposited in the 3×3 central crystals of the cluster, and that deposited in the 5×5 crystals around the central block, excluding the four corners (see Figure 4.1); the rationale behind this selection is the same as for E_1/E_9.
- `clusterZernikeMVA` > 0.1: this variable is the output of a multivariate analysis (MVA) (specifically, a boosted decision tree [76]) that uses the Zernike moments of the cluster's transverse energy distribution to distinguish photon clusters from K^0_L clusters [77]; like E_1/E_9 and E_9/E_{21}, it provides some discriminating power against non-photon clusters.

Plots of the distributions of some of these variables can be found in Figure 4.2.

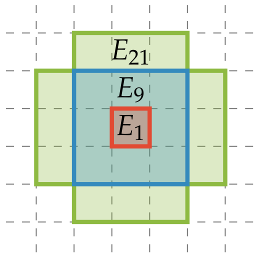
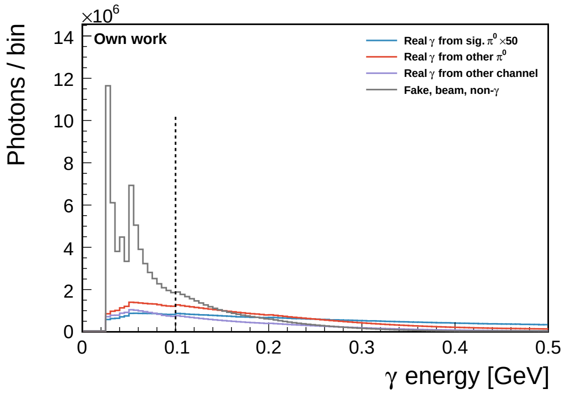


Figure 4.1: Visualization of the E_1, E_9, and E_{21} cluster variables.

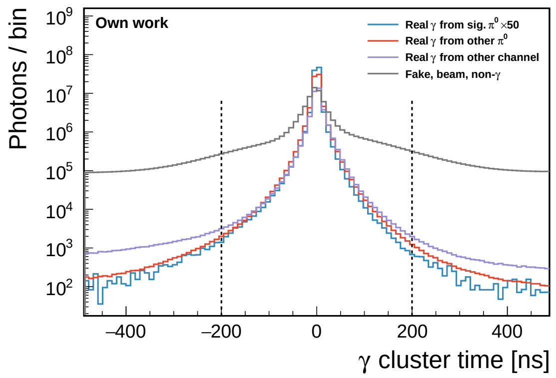
Besides the selections on the photon candidates, I also required the reconstructed π^0 mass to be between $116\,\mathrm{MeV}/c^2$ and $150\,\mathrm{MeV}/c^2$ (window of about 2.5 standard deviations), and the minimum π^0 momentum to be $500\,\mathrm{MeV}/c$ (further suppresses combinatorial candidates). The distributions of these variables are shown in Figure 4.3.

I select tracks that originate reasonably close to the IP by requiring that $d_r < 0.5\,\mathrm{cm}$ and $|d_z| < 1\,\mathrm{cm}$; d_r and d_z differ from d_0 and z_0 in that they are computed with respect to to the measured position of the collision point in each run, rather than the nominal one (the origin of the coordinates): this allows to use tighter selection criteria. Also note that d_r is always positive, being defined as the transverse distance between point of closest approach (PoCA) and IP. I also require the θ angle of the track's momentum to be within the acceptance of the tracking subdetectors: this is a standard selection to reject badly-reconstructed track candidates. I use all of the tracks that pass my selection assuming they are pions. Distributions for the impact parameters are shown in Figure 4.4.

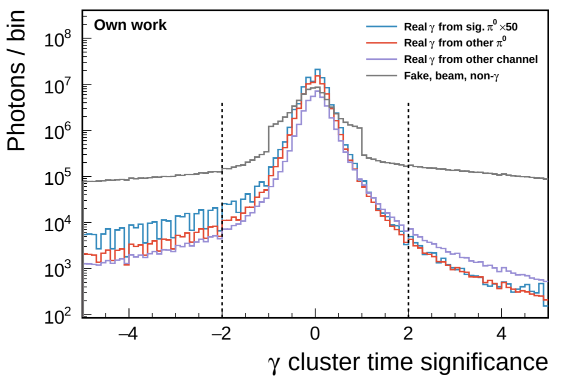
I reconstruct D^0 meson candidates by combining π^0 candidates with pairs of oppositely-charged tracks. Such tracks must have particle identification (PID) probability greater than 0.2; I use the pion-*vs.*-kaon identification probability based on neural network (NN) (π`IDNN`), which I introduced in Section 2.4.2. This is a loose requirement meant to suppress backgrounds from $D^0 \to K\pi\pi^0$ and similar channels, where I misidentify a kaon as a pion, without introducing strong experimental asymmetry from the PID requirement; such background can



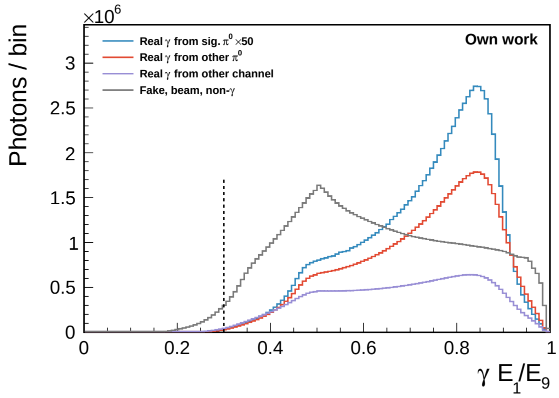
(a) Photon energy.



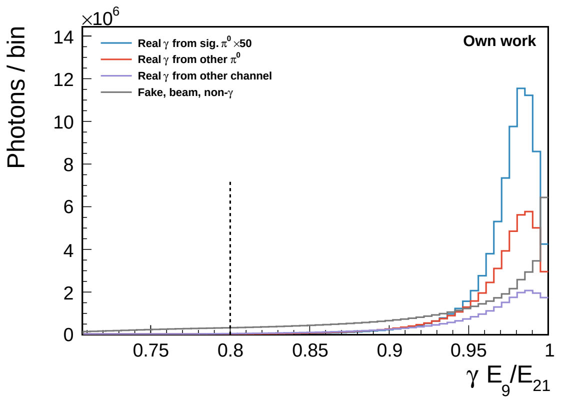
(b) Photon time.



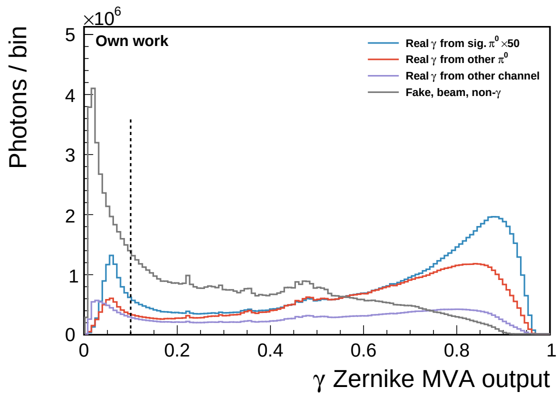
(c) Photon time significance.



(d) E_1/E_9 ratio.

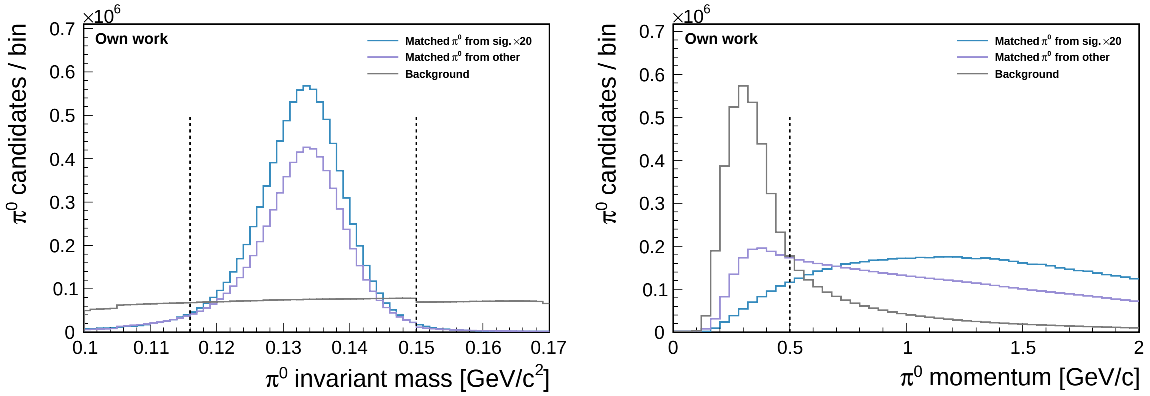


(e) E_9/E_{21} ratio.



(f) Zernike MVA output.

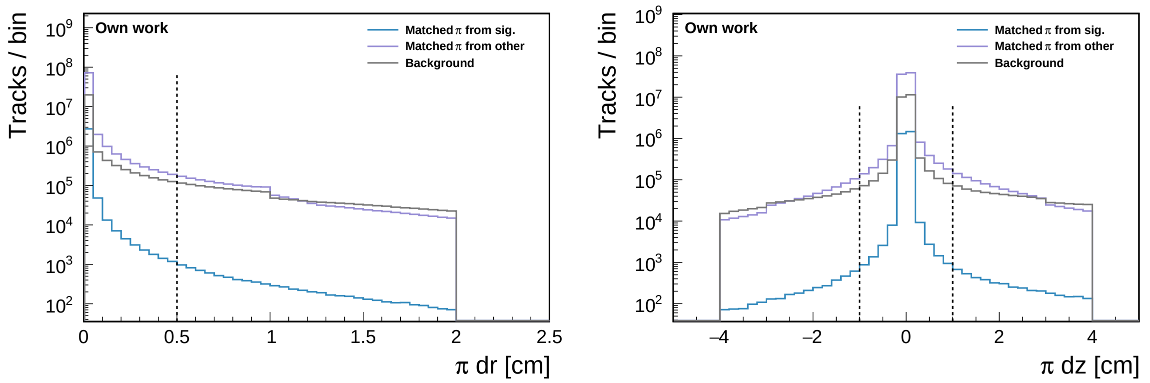
Figure 4.2: Distributions of photon variables used for selections, in simulation (1712 fb^{-1} scaled to 428 fb^{-1}). The blue lines are for correctly-reconstructed photons from $D^0 \to \pi^+\pi^-[\pi^0 \to \gamma\gamma]$ decay chains, scaled by a factor 50 for readability; the red lines are for correctly-reconstructed photons from other $\pi^0 \to \gamma\gamma$ decays; the purple lines are for other correctly-reconstructed photons; and the gray lines are for all other ECL clusters.



(a) Invariant mass.

(b) Momentum.

Figure 4.3: Distributions of π^0 variables used for selections, in simulation (1712 fb^{-1} scaled to 428 fb^{-1}) and after applying photon selections. The blue lines are for correctly-reconstructed π^0 mesons from $D^0 \to \pi^+\pi^-\pi^0$ decays, scaled by a factor 20 for readability; the purple lines are for other correctly-reconstructed π^0 mesons; and the gray lines are for all other π^0 candidates.



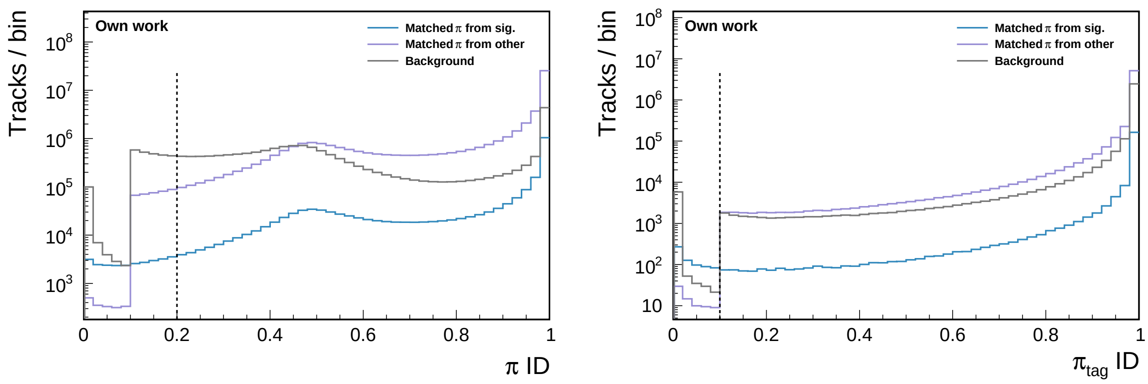
(a) Transverse.

(b) Longitudinal.

Figure 4.4: Distributions of charged pion impact parameters, in simulation (1712 fb^{-1} scaled to 428 fb^{-1}). The blue lines are for correctly-reconstructed pions from $D^* \to [D^0 \to \pi^+\pi^-\pi^0]\pi$ decay chains; the purple lines are for other correctly-reconstructed pions; and the gray lines are for all other track candidates.

be discerned from the signal from their D^0 mass distribution, as I explain in Chapter 5. The πIDNN distribution can be found in Figure 4.5a.

At this point, it is not possible to univocally assign the flavor of the D^0 meson. I therefore, reconstruct D^* meson candidates by combining D^0 candidates with tracks of any charge sign. Such tracks are required to have $\pi\texttt{IDNN} > 0.1$ to ensure consistency with the physics skim selection (see Table 4.1). The πIDNN distribution can be found in Figure 4.5b.



(a) Tracks used as D^0 daughters.

(b) Tracks used as tag pions.

Figure 4.5: Distributions of charged pion PID probabilities, in simulation (1712 fb^{-1} scaled to 428 fb^{-1}). The blue lines are for correctly-reconstructed pions from $D^0 \to \pi^+\pi^-\pi^0$ or $D^* \to D^0\pi$ decays; the purple lines are for other correctly-reconstructed pions; and the gray lines are for all other track candidates.

I perform a *vertex fit* on the reconstructed decay chain using TreeFitter [78, 79]: this tool allows to simultaneously fit the kinematic and geometry of the decay chain using appropriate constraints (*e.g.* four-momentum conservation, strong decays happening with negligible flight distance, *etc.*), determining the position of the decay vertices and updating the kinematic variables, improving their resolution. In this step, I assume the D^* meson to be produced at the IP, which adds an additional constraint to the fit. I reject candidates for which this fit does not converge.

After the vertex fit, I require the invariant mass of the $\pi^+\pi^-$ pair, $M(\pi^+\pi^-)$, *not* to be between 470 MeV/c^2 and 530 MeV/c^2. This selection is meant to suppress the $D^0 \to K^0_S\pi^0$ background with early-decaying K^0_S, which mimics the signal in all its characteristics. $D^0 \to K^0_S\pi^0$ decays with a K^0_S decaying far from the D^0 vertex are suppressed by the selections on the tracks' impact parameters. Figure 4.6 shows the distribution of $M(\pi^+\pi^-)$.

In Section 3.3 I introduced the necessity to ensure that the D^* mesons I reconstruct are "prompt", *i.e.* produced by $e^+e^- \to c\bar{c}$ processes: this implies rejecting D^* mesons produced in the decay of B mesons. Note that I also used this assumption in imposing the production-at-IP constraint to the vertex fit.

Since B mesons are produced in pairs, their energy in the collision center of mass (CM) frame is half of the collision energy ($\sqrt{s}/2$). From this information, a limit on the CM momentum of

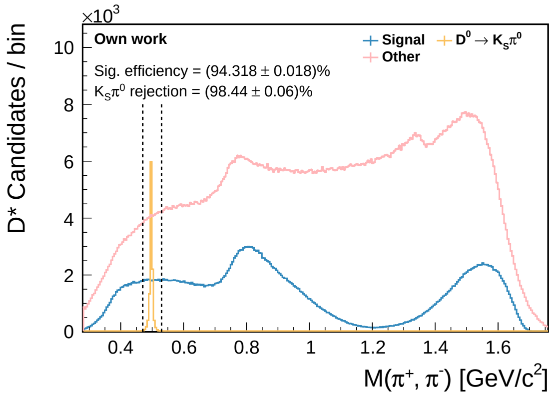


Figure 4.6: Distribution of $M(\pi^+\pi^-)$, in simulation (1712 fb^{-1} scaled to 428 fb^{-1}), after having applied all aforementioned selections. The blue line is for correctly-reconstructed $D^* \to [D^0 \to \pi^+\pi^-\pi^0]\pi$ decay chains; the yellow line is for the $D^0 \to K^0_S\pi^0$ background; and the pink line is for all other candidates.

the D^* produced as $B \to D^*X$ can be computed: let this be $p^{\max}_{\text{CM}}$. I compute the value of this limit as a function of $\sqrt{s}$ in Section A.2. I then define the X variable as

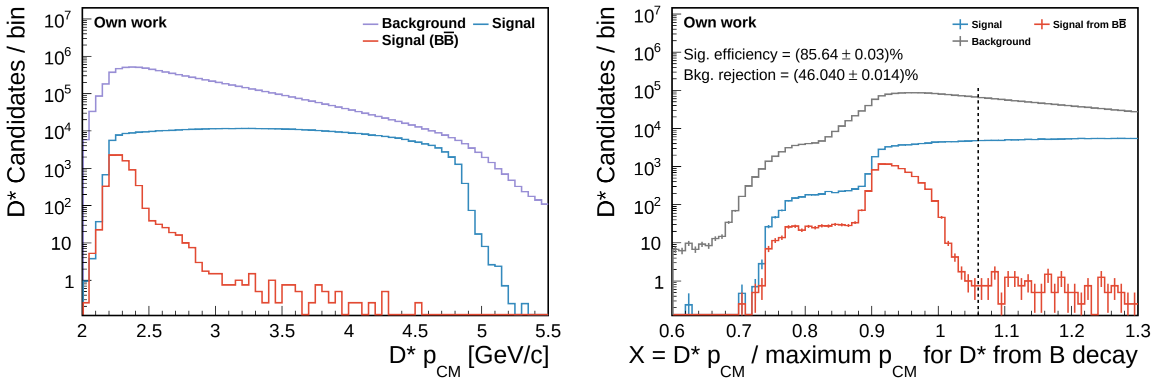
$$X = \frac{p_{\text{CM}}}{p^{\max}_{\text{CM}}(\sqrt{s})}, \tag{4.1}$$

where I use $2M_B$ instead of $\sqrt{s}$ when the latter is below the $B\bar{B}$ pair production threshold (*i.e.*, in off-resonance runs). I then require that $X > 1.06$ to reject D^* mesons from B meson decays; for $\Upsilon(4S)$ runs, this corresponds to $p_{\text{CM}} \gtrsim 2.6\,\text{GeV}/c$. The p_{CM} and X distributions are shown in Figure 4.7. Note that this selection has the positive side-effect of suppressing over 50% of the (mostly combinatorial) low-momentum background candidates; most of this background comes from $e^+e^- \to B\bar{B}$ events, where the particle multiplicity is particularly high and the average particle momentum is small.

The variables I fit to extract the signal asymmetry are the D^0 invariant mass, and the mass difference between D^* and D^0 mesons, ΔM. I require the D^0 mass to be between $1.785\,\text{GeV}/c^2$ and $1.95\,\text{GeV}/c^2$ (a window corresponding to over four standard deviations), and ΔM to be between $140\,\text{MeV}/c^2$ and $150\,\text{MeV}/c^2$ (a window corresponding to over ten standard deviations, meant to be able to capture the shape of the background components). The value of the lower D^0 mass limit is chosen to ease the handling of the main peaking background component, $D^0 \to K\pi\pi^0$: this is discussed in detail in Chapter 5. The reason for using ΔM instead of the D^* mass is because the latter is strongly correlated with the D^0 invariant mass by construction, while for the former this effect is much smaller.

Finally, since I expect events with two signal decays to be relatively rare (about 0.18% of the events with at least one signal decay), I keep only the candidate with the highest TreeFit probability (*i.e.* p-value from the χ^2 test) in each event. This further suppresses some of the residual background. Events with multiple D^* candidates occur approximately 12% of the time in the signal MC samples, and 10% of the time in the generic MC samples. In events where at least one correctly-reconstructed signal candidate is present, this best candidate selection (BCS) is $(61.03 \pm 0.12)\%$ efficient at selecting the signal candidate.

The total signal efficiency is $(11.275 \pm 0.009)\%$ on MC.



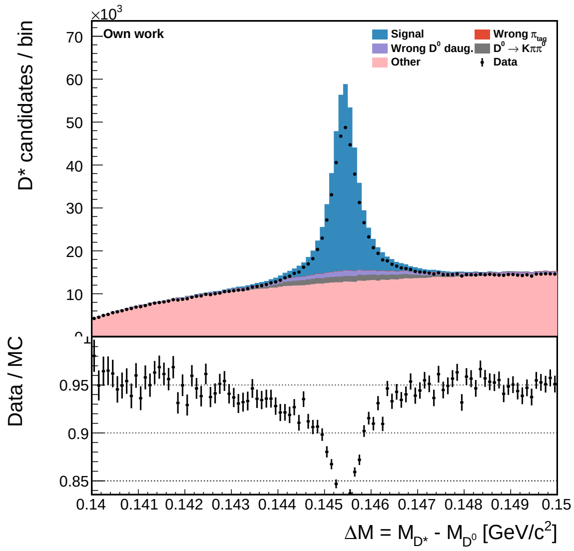
(a) CM momentum.

(b) X variable.

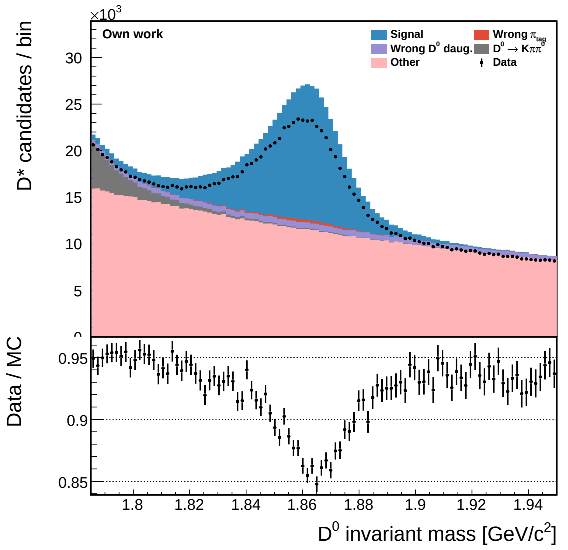
Figure 4.7: Distributions of the D^* CM momentum and X variable, in simulation (1712 fb^{-1} scaled to 428 fb^{-1}), after having applied all aforementioned selections. The blue lines are for correctly-reconstructed, prompt $D^* \to [D^0 \to \pi^+\pi^-\pi^0]\pi$ decay chains; the red lines are for correctly-reconstructed $D^* \to [D^0 \to \pi^+\pi^-\pi^0]\pi$ decay chains from $B \to D^*X$ decays; and the purple/grey lines are for all other candidates.

Figure 4.8 shows the distribution of the fit variables after applying all selections, and compares data and MC distributions. Both signal and background shapes appear to be relatively well-modeled in simulation, although signal abundance is about 20% lower in data with respect to MC. This has been observed in all D^*-tagged analyses at Belle II, and is believed to be due to incorrect D^* fragmentation inputs to the generator (PYTHIA), which is plausible given the large experimental uncertainties on such value.

To improve the agreement between data and MC, which is especially important for the reliability of the systematic uncertainties computed on simulation, I apply an *ad-hoc* rescaling of the various components on MC. I obtain the scale factors via a dedicated least-squares fit based on the histograms of the D^0 invariant mass for the data and the various MC components. I apply a shift of $-1\,\mathrm{MeV}/c^2$ to the simulation *only* for the purpose of computing the scale factors: this shift is probably caused by the uncertainty of the photon energy bias corrections used when reconstructing the data sample. Figure 4.9 shows the D^0 invariant mass and ΔM distributions after the rescaling: the maximum discrepancy is reduced to about 2%.

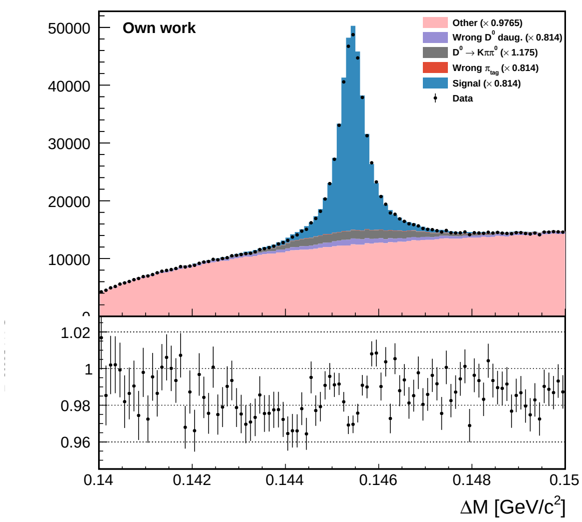


(a) ΔM.

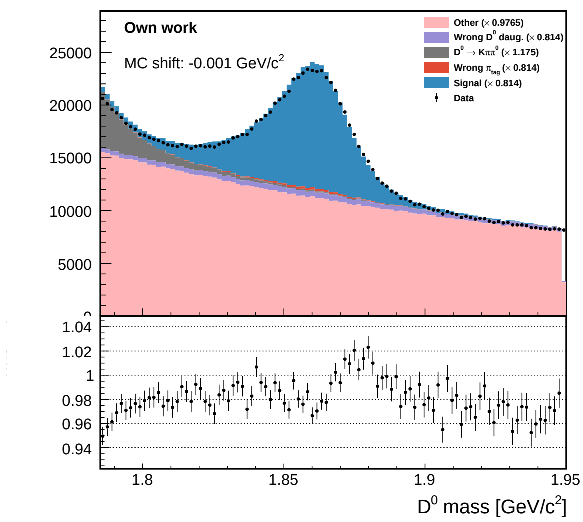


(b) D^0 invariant mass.

Figure 4.8: Distributions of the fit variables in simulation (1712 fb^{-1} scaled to 428 fb^{-1}) and data, with all selections applied. Blue is for correctly-reconstructed signal decay chains; red is for correctly-reconstructed signal D^0 mesons with random tag pion attached; purple is for signal D^0 mesons with one incorrect daughter; grey is for the $D^0 \to K\pi\pi^0$ background; pink is for all other combinatorial background; and the black points are the data distribution. At the bottom: ratio between data and MC.



(a) ΔM.



(b) D^0 invariant mass.

Figure 4.9: Distributions of the fit variables after applying the *ad-hoc* rescaling of the MC.

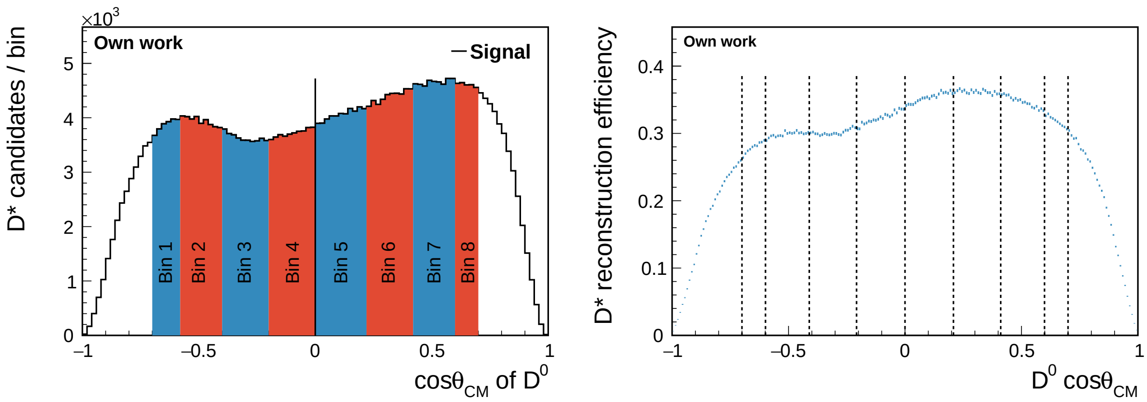
5 SAMPLE MODELING AND ASYMMETRY EXTRACTION

To extract the raw asymmetry of the signal decays, I employ an unbinned likelihood fit of the D^0 invariant mass and ΔM distributions. I split my samples into eight symmetric $\cos\theta_{\mathrm{CM}}$ (of the D^0 meson) bins, as described in Section 3.3, and by flavor; then, I simultaneously fit these 16 subsamples.

The $\cos\theta_{\mathrm{CM}}$ binning is reported in Table 5.1, and shown in Figure 5.1. The latter also shows the signal reconstruction efficiency in Monte Carlo (MC), since one of the assumptions used in Section 3.3 is that the signal efficiency is approximately constant within each bin. As this is not exact, systematic effects from this are evaluated in Section 7.3.2.

$\cos\theta_{\mathrm{CM}}$	-0.7	-0.599	-0.411	-0.208	0	0.208	0.411	0.599	0.7
Bins		Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8

Table 5.1: The $\cos\theta_{\mathrm{CM}}$ binning used throughout this work.



(a) Reconstructed signal decays distribution.

(b) Reconstructed signal decays efficiency.

Figure 5.1: The $\cos\theta_{\mathrm{CM}}$ binning used throughout this work, superimposed on the simulated distribution of reconstructed signal decays.

Splitting the sample into $\cos\theta_{\mathrm{CM}}$ bins also eases the definition of suitable probability density function (PDF), as I can capture the $\cos\theta_{\mathrm{CM}}$ dependence of the distributions by splitting PDF parameters by bin, rather than having to define a PDF that works on the whole sample. However, note that I exclude the candidates with $|\cos\theta_{\mathrm{CM}}| > 0.7$: I do this because the PDFs I define for the "inner" region ($|\cos\theta_{\mathrm{CM}}| \leq 0.7$) are unable to describe the signal and background distributions in the "outer" region ($|\cos\theta_{\mathrm{CM}}| > 0.7$). The "outer" region contains about 20% of the signal candidates, and has a 30% worse signal-to-background ratio: adding the "outer" region would only improve the statistical uncertainty by less than 10% (relative), which I deem insufficient to warrant finding a specific model for the "outer" region.

The PDF I use considers five different components: these are described in Section 5.1. The complete fit and the resulting raw asymmetry measurement, instead are shown in Section 5.2.

5.1 SAMPLE COMPOSITION

There are five prominent components in the reconstructed signal decays sample: correctly-reconstructed signal decay chains, purely combinatorial background, $D^* \to [D^0 \to K\pi\pi^0]\pi$ decay chains (with the kaon misidentified as a pion), signal D^0 decays with random tag pion attached, and partially mis-reconstructed signal D^0 decays. Table 5.2 reports their abundance in MC.

Component	Yield N^x	Ratio N^x/N^{signal}
Signal	$(346.5 \pm 0.3) \times 10^3$	**1**
with decayed-in-flight tag pion	$(7.30 \pm 0.04) \times 10^3$	2.1%
Random-pion signal	$(7.04 \pm 0.04) \times 10^3$	2.0%
with correct flavor	$(3.63 \pm 0.03) \times 10^3$	1.0%
with wrong flavor	$(3.41 \pm 0.03) \times 10^3$	0.98%
Mis-reconstructed signal	$(52.94 \pm 0.12) \times 10^3$	15%
$D^0 \to K\pi\pi^0$	$(47.17 \pm 0.11) \times 10^3$	14%
Other	$(1152.7 \pm 0.5) \times 10^3$	**3.3**
Residual $D^0 \to K^0_S\pi^0$	149 ± 6	4.2×10^{-4}
Total	$(1613.7 \pm 0.6) \times 10^3$	**4.7**

Table 5.2: Abundance of the various components in the simulated sample of reconstructed $D^0 \to \pi^+\pi^-\pi^0$ decays (scaled to $428\,\text{fb}^{-1}$).

In the following sections I describe these components in detail, together with the PDFs I use to describe them in the fit.

5.1.1 *Correctly-reconstructed signal*

This component is made of correctly-reconstructed signal decay chains. The D^0 invariant mass and ΔM distributions peak at the expected values, $\sim 1.865\,\text{GeV}/c^2$ and $\sim 145.4\,\text{MeV}/c^2$ [15].

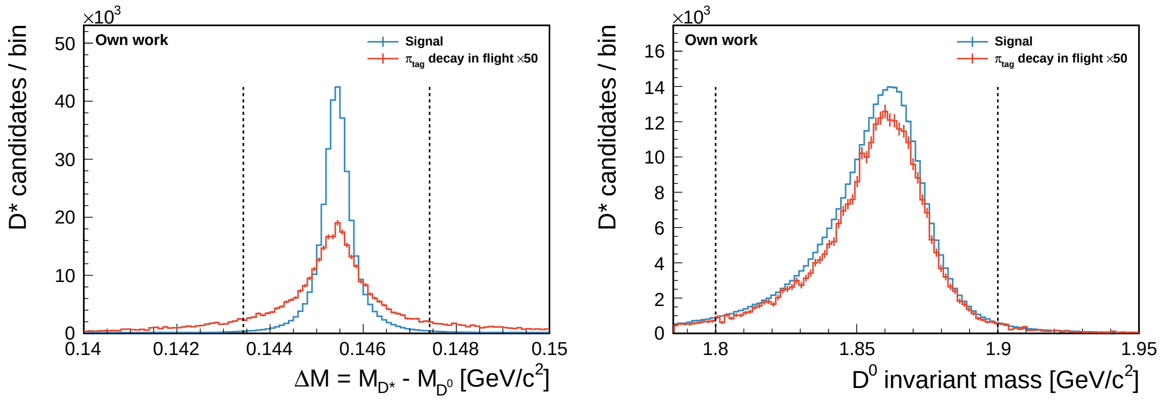
In this component I include those candidates for which the tag pion decayed inside the detector, which are about 2% of the total. This subcomponent has a wider ΔM distribution due to the worse resolution on the track parameters of the tag pion, as shown in Figure 5.2.

To model the distribution of this component I use the product of two independent Johnson's S_U PDFs:

$$p^{\text{sig}}_M(M) = \text{Johnson}(M|\mu^{\text{sig}}_M, \lambda^{\text{sig}}_M, \gamma^{\text{sig}}_M, \delta^{\text{sig}}_M), \tag{5.1a}$$

$$p^{\text{sig}}_{\Delta M}(\Delta M) = \text{Johnson}\left(\Delta M|J_\mu(\mu^{\text{sig}}_{\Delta M}, \lambda^{\text{sig}}_{\Delta M}, \gamma^{\text{sig}}_{\Delta M}, \delta^{\text{sig}}_{\Delta M}), \lambda^{\text{sig}}_{\Delta M}, \gamma^{\text{sig}}_{\Delta M}, \delta^{\text{sig}}_{\Delta M}\right), \tag{5.1b}$$

$$p^{\text{sig}}(M, \Delta M) = p^{\text{sig}}_M(M) \cdot p^{\text{sig}}_{\Delta M}(\Delta M), \tag{5.1c}$$

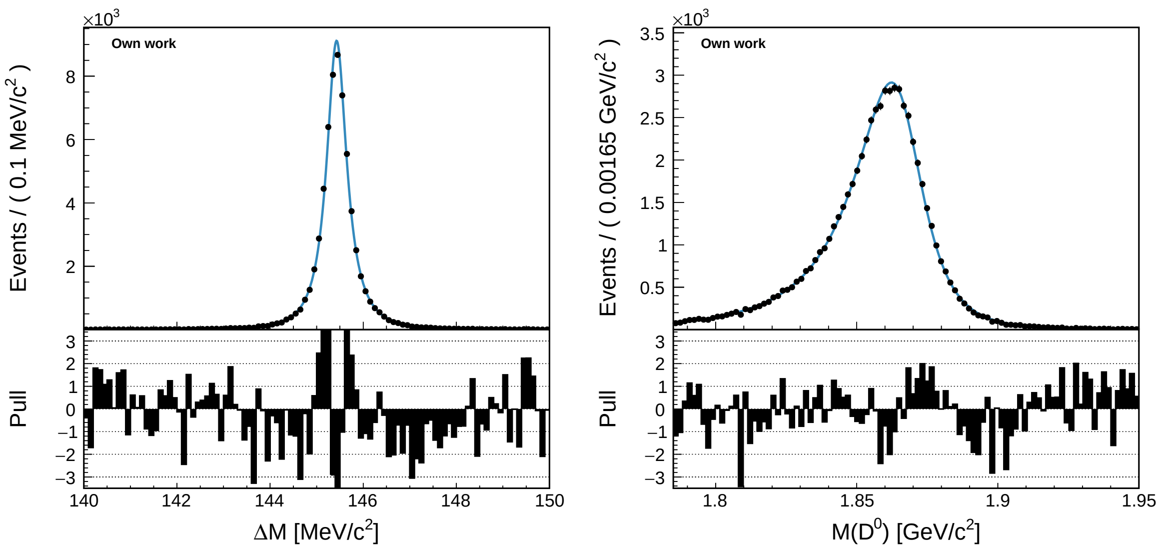


(a) ΔM distribution.

(b) D^0 invariant mass distribution.

Figure 5.2: Fit variables distribution for the signal component in simulation (1712 fb^{-1} scaled to 428 fb^{-1}). The subcomponents without (blue) and with (red) decayed-in-flight tag pion are shown separately, the latter scaled up by a factor 50 for readability.

where the J_μ function is defined so that $\mu^{\text{sig}}_{\Delta M}$ is the mode of the distribution. The definitions of both the Johnson PDF and the J_μ function can be found in Section A.3. Figure 5.3 shows the result of fitting the signal component alone, on MC, for one $\cos\theta_{\text{CM}}$ bin.



(a) ΔM distribution.

(b) D^0 invariant mass distribution.

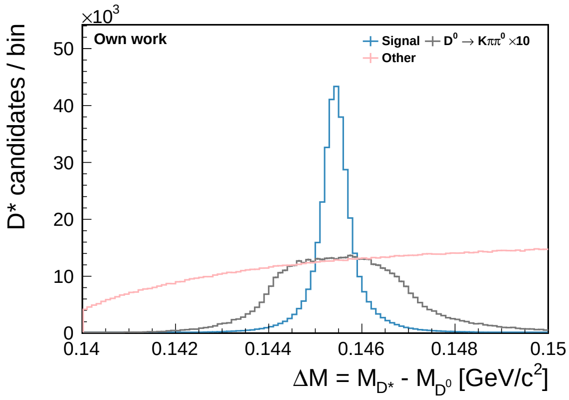
Figure 5.3: Fit of the signal component on MC for (D^0s, bin 4, *i.e.* $\cos\theta_{\text{CM}}$ from -0.208 to 0).

The two λ parameters are split by $\cos\theta_{\text{CM}}$ bin; all other parameters are shared among the bins.

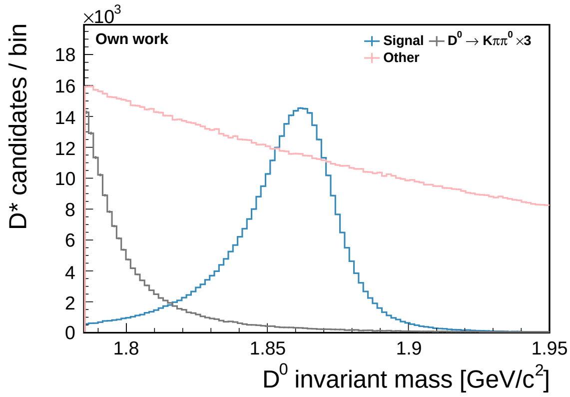
5.1.2 $D^0 \to K\pi\pi^0$ background

This component is made of candidates corresponding to $D^* \to [D^0 \to K\pi\pi^0]\pi$ decays that are correctly reconstructed, except for the mis-identification of the kaon as a pion. The D^0 invariant mass distribution peaks at lower values than the nominal D^0 mass (because of "missing" mass from the wrong hypothesis on the kaon track), while the ΔM distribution peaks around the expected value, but with a distorted peak shape.

To ease the modeling of this distribution, I cut the D^0 invariant mass window at 1.875 GeV/c^2 (as already introduced in Section 4.2), thus leaving only part of the right tail of the D^0 mass peak. This can be seen in Figure 5.4.



(a) ΔM distribution.



(b) D^0 invariant mass distribution.

Figure 5.4: Fit variables distribution for the signal (blue), $D^0 \to K\pi\pi^0$ (gray), and combinatorial (pink) components in simulation (1712 fb^{-1} scaled to 428 fb^{-1}). The $D^0 \to K\pi\pi^0$ component is scaled up by a factor 10 on the left and 3 on the right for readability.

To model the distribution of this component, I use the product of a power law for the D^0 invariant mass, and a fully-asymmetric double-sided Crystal Ball PDF for ΔM:

$$p_M^{\text{pkg}}(M) = (M - x_{0,M}^{\text{pkg}})^{-n_M^{\text{pkg}}}, \tag{5.2a}$$

$$p_{\Delta M}^{\text{pkg}}(\Delta M) = \text{CB}(\Delta M | \mu_{\Delta M}^{\text{pkg}}, \sigma_{L,\Delta M}^{\text{pkg}}, \sigma_{R,\Delta M}^{\text{pkg}}, \alpha_{L,\Delta M}^{\text{pkg}}, n_{L,\Delta M}^{\text{pkg}}, \alpha_{R,\Delta M}^{\text{pkg}}, n_{R,\Delta M}^{\text{pkg}}), \tag{5.2b}$$

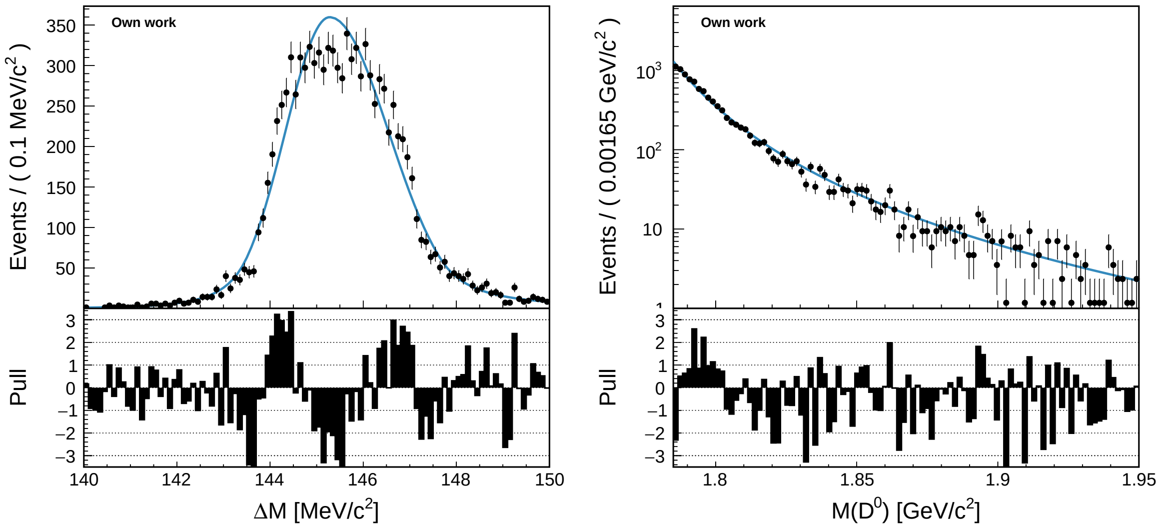
$$p^{\text{pkg}}(M, \Delta M) = p_M^{\text{pkg}}(M) \cdot p_{\Delta M}^{\text{pkg}}(\Delta M)\,. \tag{5.2c}$$

The definition of the Crystal Ball PDF can be found in Section A.3. Figure 5.5 shows the result of fitting the $D^0 \to K\pi\pi^0$ component alone, on MC, for one $\cos\theta_{\text{CM}}$ bin.

The x_0 parameter is fixed at the lower limit of the D^0 invariant mass range. The σ_L, σ_R, and α_R parameters are split by $\cos\theta_{\text{CM}}$ bin; all other parameters are shared among the bins.

5.1.3 Combinatorial background

This component includes all candidates that are not related to signal D^0 decays or $D^0 \to K\pi\pi^0$ decays. The D^0 invariant mass and ΔM distributions are non-peaking. This can be seen in Figure 5.4.



(a) ΔM distribution. (b) D^0 invariant mass distribution.

Figure 5.5: Fit of the $D^0 \to K\pi\pi^0$ component on MC for (D^0s, bin 4, *i.e.* $\cos\theta_{\mathrm{CM}}$ from -0.208 to 0).

To model the distribution of this component, I use the product of a 2$^{\mathrm{nd}}$-degree Chebyshev polynomial for the D^0 invariant mass with a threshold-like function for ΔM:

$$p_M^{\mathrm{npk}}(M) = 1 + p_{1,M}^{\mathrm{npk}} \cdot M + p_{2,M}^{\mathrm{npk}} \cdot (2M^2 - 1), \tag{5.3a}$$

$$p_{\Delta M}^{\mathrm{npk}}(\Delta M) = \begin{cases} \sqrt{\Delta M - x_{0,\Delta M}^{\mathrm{npk}}} + p_{0,\Delta M}^{\mathrm{npk}} + p_{1,\Delta M}^{\mathrm{npk}}(\Delta M - x_{0,\Delta M}^{\mathrm{npk}}) & \Delta M > x_{0,\Delta M}^{\mathrm{npk}} \\ 0 & \Delta M \leq x_{0,\Delta M}^{\mathrm{npk}} \end{cases}, \tag{5.3b}$$

$$p^{\mathrm{npk}}(M, \Delta M) = p_M^{\mathrm{npk}}(M) \cdot p_{\Delta M}^{\mathrm{npk}}(\Delta M)\,. \tag{5.3c}$$

Figure 5.6 shows the result of fitting the combinatorial component alone, on MC, for one $\cos\theta_{\mathrm{CM}}$ bin.

The x_0 parameter is fixed at the nominal mass of the charged pion, $139.57\,\mathrm{MeV}/c^2$ [15]. Both 1D PDFs are normalized to one within the range considered for the fit. The ΔM PDF parameters are split by $\cos\theta_{\mathrm{CM}}$ bin; the other parameters are shared among the bins.

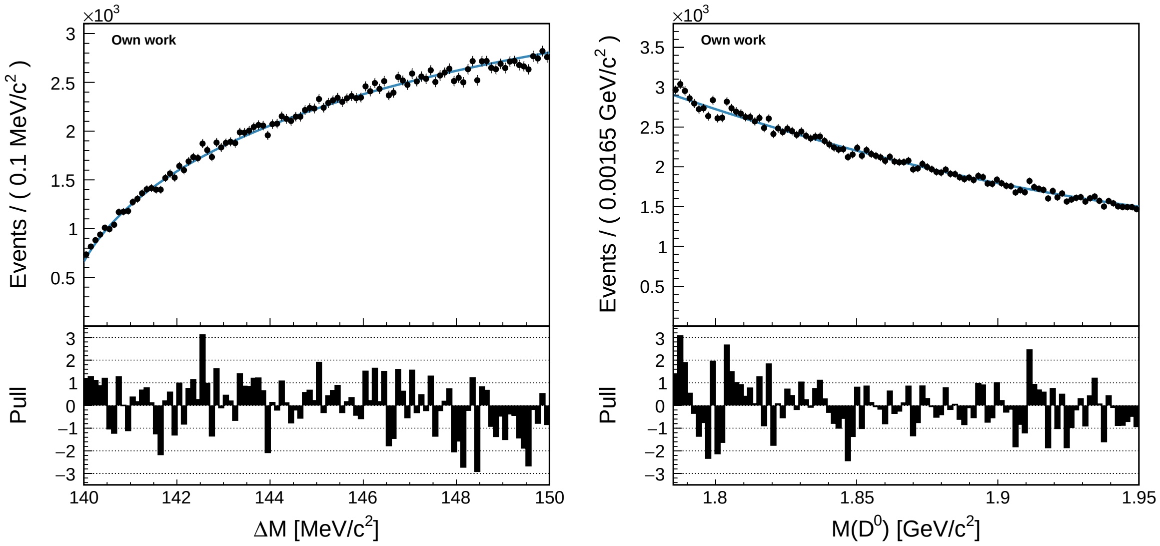
5.1.4 *Signal D^0 decays with random pion*

This component is made of candidates with correctly-reconstructed signal $D^0 \to \pi^+\pi^-\pi^0$ decay, but random track attached as tag pion. Such random track may have any charge sign. The D^0 invariant mass distribution is the same as for the signal component, while the ΔM distribution is the same as for the combinatorial component. This is shown in Figure 5.7.

To model the distribution of this component I use the product of the signal component's D^0 invariant mass PDF, and the combinatorial component's ΔM PDF:

$$p^{\mathrm{R}\pi}(M, \Delta M) = p_M^{\mathrm{sig}}(M) \cdot p_{\Delta M}^{\mathrm{npk}}(\Delta M)\,. \tag{5.4}$$

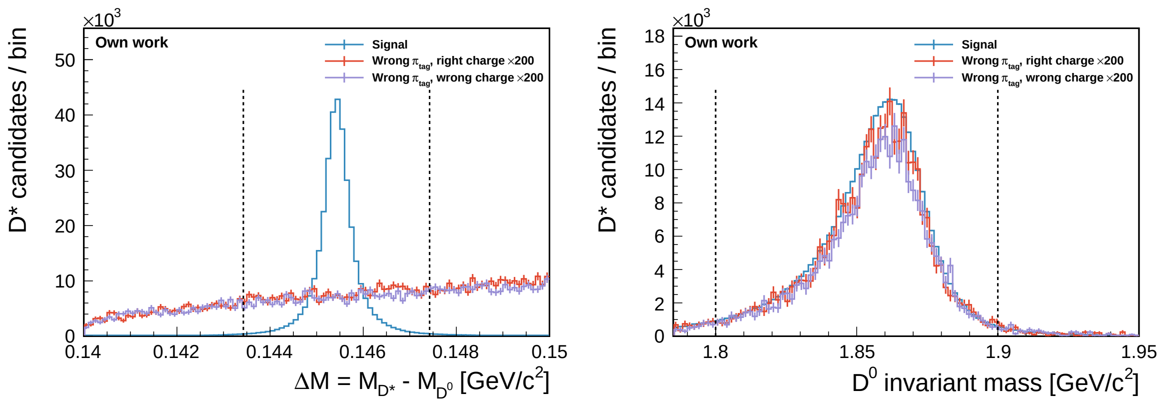
Figure 5.8 shows the result of fitting the random-pion component alone, on MC, for one $\cos\theta_{\mathrm{CM}}$ bin.



(a) ΔM distribution.

(b) D^0 invariant mass distribution.

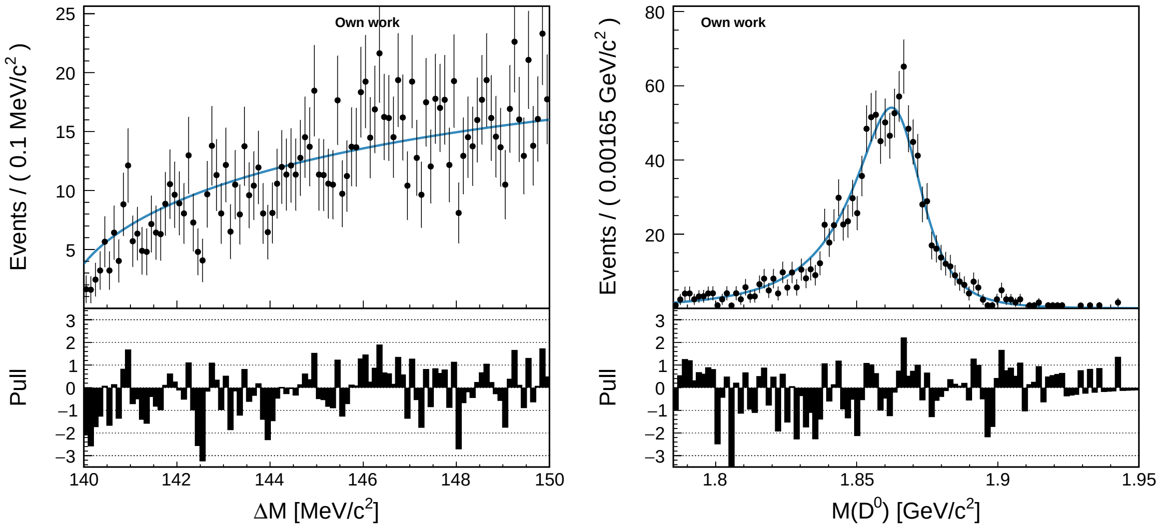
Figure 5.6: Fit of the combinatorial component on MC for (D^0s, bin 4, *i.e.* $\cos\theta_{\text{CM}}$ from -0.208 to 0).



(a) ΔM distribution.

(b) D^0 invariant mass distribution.

Figure 5.7: Fit variables distribution for the random-pion component in simulation (1712 fb^{-1} scaled to 428 fb^{-1}). The signal (blue) component is shown for comparison. The subcomponents with correct (red) and wrong (purple) tag pion charge are shown separately, and scaled up by a factor 200 for readability.



(a) ΔM distribution.

(b) D^0 invariant mass distribution.

Figure 5.8: Fit of the random-pion component on MC for (D^0s, bin 4, *i.e.* $\cos\theta_{\mathrm{CM}}$ from -0.208 to 0).

In the final fit (with all components together), the parameters are shared with the other components' PDFs.

5.1.5 *Partially mis-reconstructed signal*

This component is made of candidates with incorrectly-reconstructed signal $D^0 \to \pi^+\pi^-\pi^0$ decay, where one of the D^0 daughters is not the correct particle. There are two subcomponents: a *non-peaking* one, with distributions similar to those of the combinatorial component; and a *peaking* one, with a very wide D^0 invariant mass peak and a ΔM distribution similar to that of the signal component but wider. This is shown in Figure 5.9.

To model the D^0 invariant mass distribution of this component, I use a 2$^{\text{nd}}$-degree Chebyshev polynomial. For ΔM, I use the signal component's Johnson distribution (with a different λ parameter to allow larger width) for the peaking subcomponent, and the combinatorial component's threshold function for the non-peaking subcomponent.

$$p_M^{\text{misreco}}(M) = 1 + p_{1,M}^{\text{misreco}} \cdot M + p_{2,M}^{\text{misreco}} \cdot (2M^2 - 1), \tag{5.5a}$$

$$p_{\Delta M}^{\text{misreco,pkg}}(\Delta M) = \text{Johnson}\left(\Delta M | J_\mu(\mu_{\Delta M}^{\text{sig}}, \lambda_{\Delta M}^{\text{misreco}}, \gamma_{\Delta M}^{\text{sig}}, \delta_{\Delta M}^{\text{sig}}), \lambda_{\Delta M}^{\text{misreco}}, \gamma_{\Delta M}^{\text{sig}}, \delta_{\Delta M}^{\text{sig}}\right), \tag{5.5b}$$

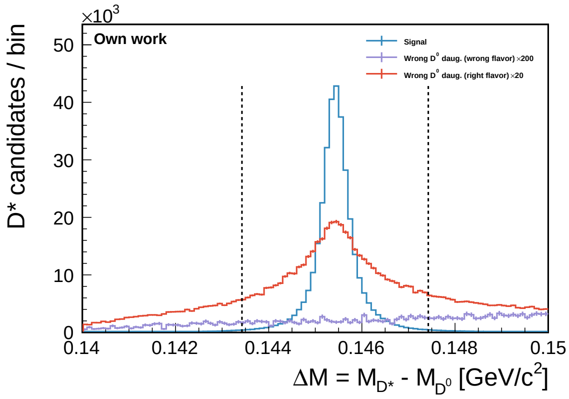
$$p_{\Delta M}^{\text{misreco,npk}}(\Delta M) = \sqrt{\Delta M - x_{0,\Delta M}^{\text{npk}}} + p_{0,\Delta M}^{\text{npk}} + p_{1,\Delta M}^{\text{npk}}(\Delta M - x_{0,\Delta M}^{\text{npk}}), \tag{5.5c}$$

$$p^{\text{misreco,pkg}}(M, \Delta M) = p_M^{\text{misreco}}(M) \cdot p_{\Delta M}^{\text{misreco,pkg}}(\Delta M), \tag{5.5d}$$

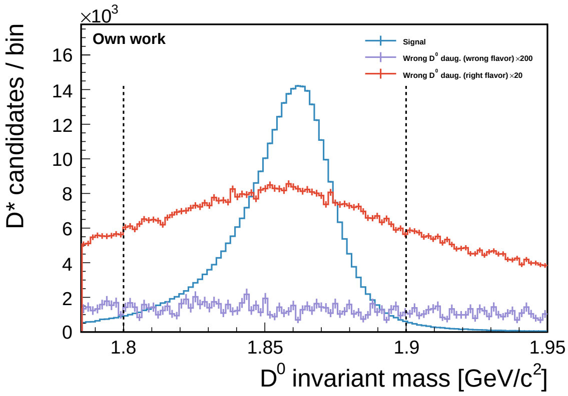
$$p^{\text{misreco,npk}}(M, \Delta M) = p_M^{\text{misreco}}(M) \cdot p_{\Delta M}^{\text{misreco,npk}}(\Delta M), \tag{5.5e}$$

$$p^{\text{misreco}}(M, \Delta M) = n^{\text{pkg}} \cdot p^{\text{misreco,pkg}}(M, \Delta M) + n^{\text{npk}} \cdot p^{\text{misreco,npk}}(M, \Delta M)\,. \tag{5.5f}$$

Figure 5.10 shows the result of fitting the mis-reconstructed signal component alone, on MC, for one $\cos\theta_{\mathrm{CM}}$ bin.

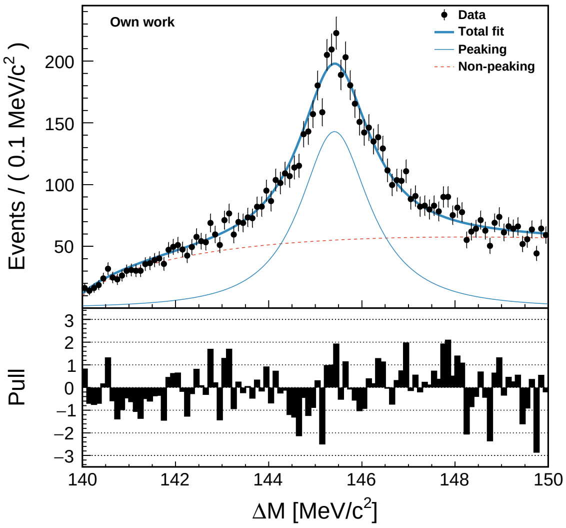


(a) ΔM distribution.

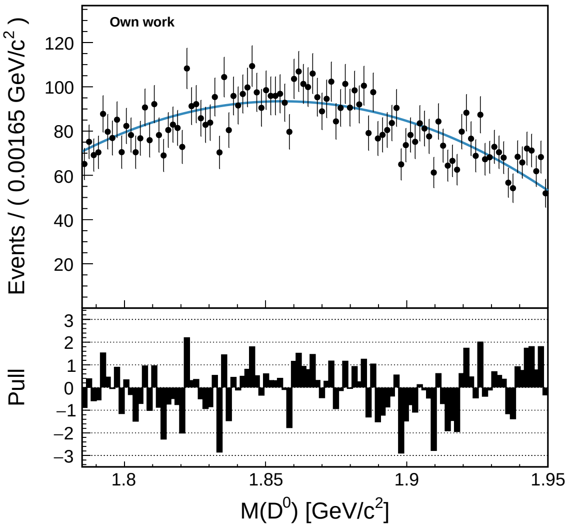


(b) D^0 invariant mass distribution.

Figure 5.9: Fit variables distribution for the mis-reconstructed signal component in simulation (1712 fb^{-1} scaled to 428 fb^{-1}). The signal (blue) component is shown for comparison. The fractions with correct (red, scaled up by a factor 20) and wrong (purple, scaled up by a factor 200) tag pion charge are shown separately.



(a) ΔM distribution.



(b) D^0 invariant mass distribution.

Figure 5.10: Fit of the mis-reconstructed signal component on MC for (D^0s, bin 4, *i.e.* $\cos\theta_{\text{CM}}$ from -0.208 to 0).

The λ parameter is split by $\cos\theta_{CM}$ bin; the D^0 invariant mass PDF parameters are shared among the bins; all other parameters are taken from their respective component's PDF and, in the final fit (with all components together), are shared with the said component. Also, in the final fit, the non-peaking subcomponent is not used, letting the combinatorial component absorb it.

5.2 RAW ASYMMETRY EXTRACTION

To extract the asymmetry from data, all components are fitted at the same time. The yields of each component are fitted independently for each $\cos\theta_{CM}$ bin and for each D^0 flavor by defining

$$n^x = n^x_{tot}\frac{1+\mathcal{A}^x_{raw}}{2}, \qquad \overline{n}^x = n^x_{tot}\frac{1-\mathcal{A}^x_{raw}}{2}, \tag{5.6}$$

where n and $\overline{n}$ are the D^0 and $\bar{D}^0$ yields, n_{tot} is the total yield, and $\mathcal{A}_{raw}$ is the raw asymmetry of the component x. The fit directly extracts n_{tot} and $\mathcal{A}_{raw}$.

Because of the large number of parameters, many are determined by fitting the single component on MC and then fixed when fitting all components together. This mostly includes shape parameters. Several width parameters are fixed to their MC values in the single $\cos\theta_{CM}$ bin, but have a global (*i.e.* shared among the $\cos\theta_{CM}$ bins) scale factor that is left free. This is summarized in Table 5.3

Component		Model	Parameters	
			Split	Shared
Signal	M	Johnson	$\lambda^\dagger$	μ, γ^*, δ^*
	ΔM	Johnson	$\lambda^\dagger$	μ, γ^*, δ^*
$D^0 \to K\pi\pi^0$	M	Power law	–	n^*
	ΔM	Crystal Ball	$\sigma^\dagger_L, \sigma^\dagger_R, \alpha^*_R$	$\mu, \alpha^*_L, n^*_L, n^*_R$
Combinatorial	M	2nd-deg. Chebyshev	–	p_1, p_2
	ΔM	Threshold func.	p_0, p_1	–
Random pion	M	Same as signal	–	–
	ΔM	Same as combinatorial	–	–
Mis-reconstructed (peaking)	M	2nd-deg. Chebyshev	–	p^*_1, p^*_2
	ΔM	Johnson	$\lambda^\dagger$	–

Table 5.3: Summary of the PDF models used in the $D^0 \to \pi^+\pi^-\pi^0$ fit. "Split" and "shared" refer to parameter splitting among $\cos\theta_{CM}$ bins. The * and † indicate fixed parameters (see footnotes).

*Parameter value is fixed from single-component fit on MC.
†Parameter value is fixed, but a global (shared among $\cos\theta_{CM}$ bins) scaling factor s is left floating.

Also, because the abundances of the random pion and mis-reconstructed components are too small to be resolved by the fit, their yields are defined as a fraction of the signal yield, which is fixed from MC. Since the mis-reconstructed component almost always has a tag pion with the correct charge sign, its asymmetry contributes to that of the signal. The random-pion component asymmetry, instead, is fixed from MC.

D^0 and $\overline{D}^0$ PDFs for each component are assumed to be the same in each $\cos\theta_{\mathrm{CM}}$ bin, letting only the yields differ. The total PDF is therefore given by

$$p(M, \Delta M) = n^{\mathrm{sig}} \left[(1 - f^{\mathrm{misreco}}) p^{\mathrm{sig}} + f^{\mathrm{misreco}} \cdot p^{\mathrm{misreco,pkg}} \right] \tag{5.7a}$$

$$+ n^{\mathrm{R}\pi} \cdot p^{\mathrm{R}\pi} + n^{\mathrm{pkg}} \cdot p^{\mathrm{pkg}} + n^{\mathrm{npk}} \cdot p^{\mathrm{npk}}, \tag{5.7b}$$

$$\overline{p}(M, \Delta M) = \overline{n}^{\mathrm{sig}} \left[(1 - f^{\mathrm{misreco}}) p^{\mathrm{sig}} + f^{\mathrm{misreco}} \cdot p^{\mathrm{misreco,pkg}} \right] \tag{5.7c}$$

$$+ \overline{n}^{\mathrm{R}\pi} \cdot p^{\mathrm{R}\pi} + \overline{n}^{\mathrm{pkg}} \cdot p^{\mathrm{pkg}} + \overline{n}^{\mathrm{npk}} \cdot p^{\mathrm{npk}}, \tag{5.7d}$$

where the various n and $\overline{n}$ are defined as in Equation 5.6, while the various p^x are the 2D PDFs defined in the previous sections. The f^{misreco} parameter is the fraction of mis-reconstructed signal component with respect to the correctly-reconstructed signal component, whose values are reported in Table 5.4. The random-pion component yields $n^{\mathrm{R}\pi}$ and $\overline{n}^{\mathrm{R}\pi}$ are also defined as in Equation 5.6, but with $n_{\mathrm{tot}}^{\mathrm{R}\pi} = n_{\mathrm{tot}}^{\mathrm{sig}} f^{\mathrm{R}\pi}$, where the fraction $f^{\mathrm{R}\pi}$ is taken from MC and reported in Table 5.5; also the asymmetry $\mathcal{A}_{\mathrm{raw}}^{\mathrm{R}\pi}$ is fixed from MC and reported in Table 5.6.

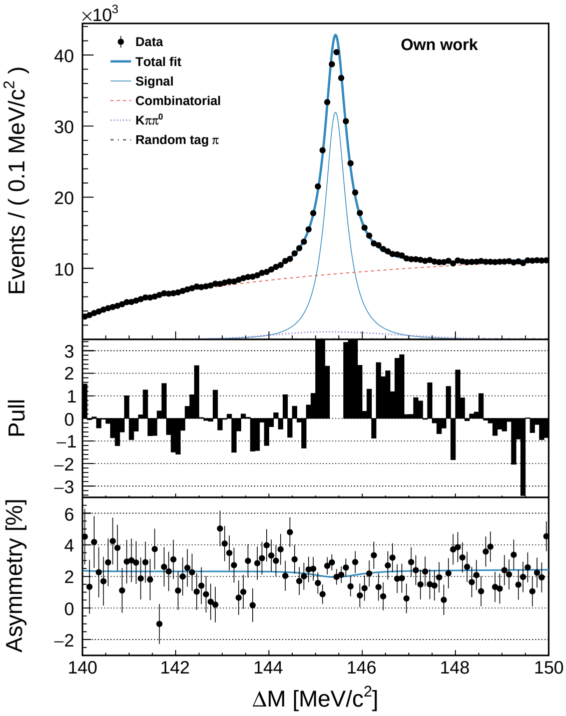
$\cos\theta_{\mathrm{CM}}$ bin	Mis-reconstructed fraction	$\cos\theta_{\mathrm{CM}}$ bin	Mis-reconstructed fraction
1	$(5.1 \pm 0.3)\%$	2	$(4.8 \pm 0.2)\%$
3	$(5.0 \pm 0.2)\%$	4	$(4.9 \pm 0.2)\%$
5	$(5.7 \pm 0.2)\%$	6	$(5.7 \pm 0.2)\%$
7	$(7.0 \pm 0.3)\%$	8	$(7.6 \pm 0.3)\%$

Table 5.4: Mis-reconstructed signal component fractions with respect to the correctly-reconstructed signal component, determined from the MC sample.

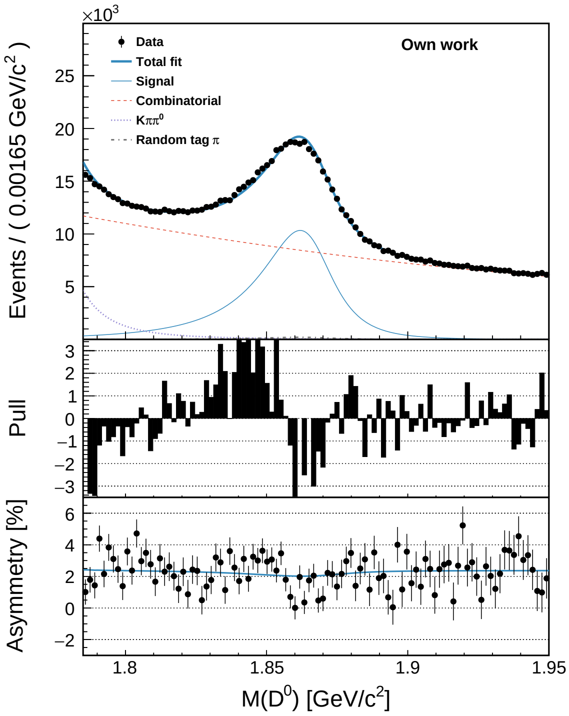
$\cos\theta_{\mathrm{CM}}$ bin	Random π fraction	$\cos\theta_{\mathrm{CM}}$ bin	Random π fraction
1	$(1.80 \pm 0.05)\%$	2	$(1.71 \pm 0.03)\%$
3	$(1.68 \pm 0.03)\%$	4	$(1.67 \pm 0.03)\%$
5	$(1.69 \pm 0.03)\%$	6	$(1.68 \pm 0.03)\%$
7	$(1.73 \pm 0.03)\%$	8	$(1.70 \pm 0.04)\%$

Table 5.5: Random-pion component fractions with respect to the signal component, determined from the MC sample.

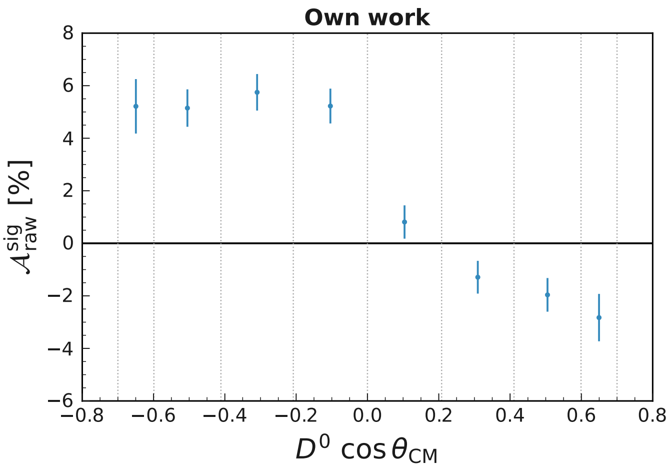
Figure 5.11 shows the fit on data and the resulting raw asymmetries. Figures and numerical results for each $\cos\theta_{\mathrm{CM}}$ bin can be found in Table C.1 and Section C.1.



(a) ΔM distribution.



(b) D^0 invariant mass distribution.



(c) Raw signal asymmetries in percent.

Figure 5.11: Final fit of the data sample. All $\cos\theta_{\mathrm{CM}}$ bins are projected together in the distribution plots. The middle plots show the pulls, while the bottom plots show the asymmetry for data (points) and PDF (line).

$\cos\theta_{\text{CM}}$ bin	Random π asymmetry	$\cos\theta_{\text{CM}}$ bin	Random π asymmetry
1	$(-4.3 \pm 2.5)\%$	2	$(2.7 \pm 1.9)\%$
3	$(-0.1 \pm 1.9)\%$	4	$(-0.8 \pm 1.9)\%$
5	$(1.5 \pm 1.8)\%$	6	$(-0.6 \pm 1.7)\%$
7	$(-3.7 \pm 1.7)\%$	8	$(3.3 \pm 2.3)\%$

Table 5.6: Random-pion component asymmetries, determined from the MC sample.

6 FLAVOR TAGGING ASYMMETRY CORRECTION

In this chapter I explain how I use my control channel, $D^0 \to K\pi$, to measure the asymmetry induced by reconstructing the $D^* \to D^0\pi$ decay, *i.e.* the flavor tagging asymmetry. This is mainly due to the track reconstruction charge asymmetry, which is large at low transverse momentum as explained in Section 2.4.1. Past studies performed by the Belle II tracking group showed that the simulation reproduce the dependence of this asymmetry on p_T and $\cos\theta$ relatively well, but cannot be relied upon where the magnitude of the effect is concerned.

$D^0 \to K^-\pi^+$ ($\overline{D}^0 \to K^+\pi^-$) is a Cabibbo-favored decay with a relatively large branching fraction. Only charged particles are present in the final state: this allows to reconstruction such the decays with both high efficiency and high purity, resulting in a small statistical uncertainty on $\mathcal{A}_\mathrm{raw}$. Also, the D^0 daughters have different identities, allowing to infer the D^0 flavor without any tagging as long as sufficiently hard particle identification (PID) selections are used.

To measure the tagging asymmetry, I reconstruct two $D^0 \to K\pi$ samples: one "tagged" (*i.e.* I reconstruct the $D^* \to [D^0 \to K\pi]\pi$ chain), and the other "untagged" (*i.e.* I only reconstruct $D^0 \to K\pi$ decays). When I extract the raw asymmetries of these samples in each D^0 $\cos\theta_\mathrm{CM}$ bin, I get

$$\mathcal{A}_\mathrm{raw}^\mathrm{tagged} \simeq \mathcal{A}_\varepsilon^{D^0\to K\pi} + \mathcal{A}_\mathrm{prod} + \mathcal{A}_{CP}^{D^0\to K\pi} + \mathcal{A}_\varepsilon^{\pi\,\mathrm{tag}}, \tag{6.1a}$$

$$\mathcal{A}_\mathrm{raw}^\mathrm{untagged} \simeq \mathcal{A}_\varepsilon^{D^0\to K\pi} + \mathcal{A}_\mathrm{prod} + \mathcal{A}_{CP}^{D^0\to K\pi}, \tag{6.1b}$$

where $\mathcal{A}_\varepsilon^{D^0\to K\pi}$ is the D^0 reconstruction asymmetry, $\mathcal{A}_\mathrm{prod}$ is the production asymmetry, $\mathcal{A}_{CP}^{D^0\to K\pi}$ is the $D^0 \to K\pi$ CP asymmetry (negligible at our sensitivity, since this is a Cabibbo-favored decay), and $\mathcal{A}_\varepsilon^{\pi\,\mathrm{tag}}$ is the tagging asymmetry. Assuming that the asymmetries that are present in both samples are the same in each $\cos\theta_\mathrm{CM}$ bin, I compute the tagging asymmetry as

$$\mathcal{A}_\varepsilon^{\pi\,\mathrm{tag}} = \mathcal{A}_\mathrm{raw}^\mathrm{tagged} - \mathcal{A}_\mathrm{raw}^\mathrm{untagged}, \tag{6.2}$$

and I then subtract these values from the $D^0 \to \pi^+\pi^-\pi^0$ raw asymmetries.

My assumptions in this process are that:

- the $\mathcal{A}_\varepsilon^{\pi\,\mathrm{tag}}$ of the tagged $D^0 \to K\pi$ sample is the same (in each $\cos\theta_\mathrm{CM}$ bin) as the $\mathcal{A}_\varepsilon^{\pi\,\mathrm{tag}}$ of the $D^0 \to \pi^+\pi^-\pi^0$ sample; and
- the $\mathcal{A}_\varepsilon^{D^0\to K\pi}$, $\mathcal{A}_\mathrm{prod}$, and $\mathcal{A}_{CP}^{D^0\to K\pi}$ of the untagged sample are identical to those of the tagged sample.

For this to be true, the distributions of all variables on which these asymmetries depend must be the same between the relevant samples. To this end:

- I use the same selection criteria for the tag pion in the signal $D^0 \to \pi^+\pi^-\pi^0$ sample and in the tagged $D^0 \to K\pi$ sample;
- I use the same selection criteria for all other particles in the tagged and untagged samples; and

- since the asymmetries also depend on the kinematic of the relevant particles, I equalize the kinematic distributions of the samples by assigning appropriate weights to the candidates.

The relevant kinematic variables for the $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$ are the p_T and $\cos\theta$ of the tag pion; those for the $\mathcal{A}_\varepsilon^{D^0\to K\pi}$ are the p_T and $\cos\theta$ of the D^0 daughters; and that for the $\mathcal{A}_\text{prod}$ is the $\cos\theta_\text{CM}$ of the D^0.

In summary, the process takes the following steps, which are repeated for the tagged and untagged samples.

1. Reconstruct the sample using common selection criteria.
2. Reweight the sample ("original") to match the "target" sample (tagged must match signal, untagged must match tagged):
   a) obtain signal $_s\mathcal{W}eights$ with a flavor-agnostic 1D fit;
   b) use the $_s\mathcal{W}eights$ to obtain the signal distributions of the relevant variables of the two samples ("original" and "target");
   c) use the ratio of the signal distributions to assign weights to the candidates of the "original" sample (to account for the correlation among the variables, either use the n-dimensional distribution of the n variables directly, or proceed iteratively on one variable at a time).
3. Extract the raw asymmetry with a simultaneous 1D fit of the two flavors.

And finally compute the difference between $\mathcal{A}_\text{raw}^\text{tagged}$ and $\mathcal{A}_\text{raw}^\text{untagged}$, taking into account the correlation between the two samples. All steps are performed independently for each $\cos\theta_\text{CM}$ bin: this prevents any correlation among the bins due to the procedure, thus reducing the systematic effects to consider. The $_s\mathcal{W}eights$ are the weights computed using the $_s\mathcal{P}lot$ technique [80], which is described more in detail later.

The samples used are described in Section 6.1. The reconstruction process and selection criteria are described in Section 6.2. The composition of the samples and the probability density functions (PDFs) used for the fits are described in Section 6.3; the same models are used for both the flavor-agnostic fit that computes the $_s\mathcal{W}eights$, and the simultaneous fit of the two flavors. Finally, Section 6.4 describes the rest of the procedure ($_s\mathcal{P}lots$, reweighting, asymmetry fit, computation of the difference accounting for correlation).

6.1 DATA AND SIMULATED SAMPLES

The data and simulated Monte Carlo (MC) samples used for the $D^0 \to K\pi$ control sample are the same as for the signal $D^0 \to \pi^+\pi^-\pi^0$ sample, described in Section 4.1. The only difference lies in the physics skim used, which requires:

1. that the event is flagged as "hadronic" at the HLT skim level (at least three tracks with $p_\text{T} > 200\,\text{MeV}/c$, $|d_0| < 2\,\text{cm}$, and $|z_0| < 4\,\text{cm}$, vetoing Bhabha events); and
2. the presence of at least one $D^0 \to h^+h'^-$ candidate in the event, where h and h' are either pions or kaons in any combination.

The selection criteria for the $D^0 \to h^+h'^-$ candidates are reported in Table 6.1. The KID variable mentioned in the table is the likelihood-based kaon-*vs.*-pion identification probability (introduced in Section 2.4.2).

Variable	Selection
Charged particles	
d_r	$< 1\,\text{cm}$
$\lvert d_z \rvert$	$< 3\,\text{cm}$
θ	within CDC acceptance ($17°$ to $150°$)
πIDNN	> 0.1 (for pions)
KID	> 0.2 (for kaons)
D^0 **meson**	
M	between $1.7\,\text{MeV}/c^2$ and $2.1\,\text{MeV}/c^2$
p_{CM}	$> 2\,\text{GeV}/c$

Table 6.1: The selection criteria used by the $D^0 \to h^+h'^-$ physics skim.

6.2 CANDIDATES RECONSTRUCTION

I reconstruct the tagged and untagged sample at the same time. First, I reconstruct $D^0 \to K^-\pi^+$ ($\bar{D}^0 \to K^+\pi^-$) candidates using track candidates that satisfy the selection criteria listed in Table 6.2. Then, I further process and select these candidates to build the two samples, as described in the next paragraphs and summarized in Figure 6.1.

To select candidates for the untagged sample, I first perform a vertex fit using TreeFitter [78, 79], and constraining the D^0 production vertex at the interaction point (IP). This improves the resolution of the kinematic variables. I only keep candidates whose fit probability (*i.e.* the p-value obtained from the χ^2 test) is $\leq 10^{-3}$, and whose invariant mass is between $1.8\,\text{GeV}/c^2$ and $1.92\,\text{GeV}/c^2$.

To discard D^0 mesons produced in the decay of a B meson (thus ensuring I have a "prompt" sample), I define the X variable as I do for the signal sample (Equation 4.1)

$$X = \frac{p_{\text{CM}}(D^0)}{\max p_{\text{CM}}(D^0)}, \tag{6.3}$$

where I use $2M_B$ instead of $\sqrt{s}$ when the latter is below the $B\bar{B}$ pair production threshold (*i.e.*, in off-resonance runs). The only difference is that I am computing this on the D^0 meson, rather than the D^* meson, so that I can use the same selection in both the untagged and tagged samples. I then require X to be greater than 1.06 (corresponding to $p_{\text{CM}} \gtrsim 2.65\,\text{GeV}/c$ in $\Upsilon(4S)$ runs).

Since the probability of having both a $D^0 \to K^-\pi^+$ and a $\bar{D}^0 \to K^+\pi^-$ in the same event is relatively small (about 1.8% [15]), I keep only one D^0 candidate in each event. This candidate is chosen randomly to avoid any possible bias in $\mathcal{A}_{\text{raw}}$ that might be different in the tagged sample.

Finally, I apply some loose selections on the p_{T} and $\cos\theta$ of the D^0 daughters. I introduced these *a posteriori* to exclude regions where discrepancies arise between the $_s$*Weighted* and true signal distributions; this is discussed more in detail in Section 6.4.3. The D^0 candidates thus selected constitute the untagged sample.

For the tagged sample, I create D^* candidates by combining the D^0 candidates from before the vertex fit with a track candidate satisfying the criteria for pions given in Table 6.2, and with the expected charge sign ($D^{*+} \to D^0\pi^+$, $D^{*-} \to \bar{D}^0\pi^-$). Then, I perform a vertex fit of the whole decay chain using TreeFitter (which assumes the D^* flight distance to be negligible, as the decay happens through strong interaction), and I impose the same requirements on the fit probability, the D^0 invariant mass, and X. Additionally, I require ΔM to be between 140 MeV/c^2 and 150 MeV/c^2. Finally, as for the untagged sample, I keep only one randomly-chosen candidate per event, and I apply the loose selections on the p_T and $\cos\theta$ of the D^0 daughters.

Variable	Selection
Charged particles	
d_r	< 0.5 cm
$\lvert d_z \rvert$	< 1 cm
θ	within CDC acceptance (17° to 150°)
Pions	
πIDNN	> 0.1 (tag pion)
πIDNN	> 0.8 (D^0 daughter)
p_T	> 0.15 GeV/c (D^0 daughter)
$\cos\theta$	> -0.6 (D^0 daughter)
Kaons	
KID	> 0.2 (re-apply skim selection)
KIDNN	> 0.8
p_T	> 0.15 GeV/c
$\cos\theta$	> -0.75
D^0 meson	
M	between 1.8 GeV/c^2 and 1.92 GeV/c^2
X	> 1.06
TreeFit p-value	$> 10^{-3}$
D^* meson	
ΔM	between 140 MeV/c^2 and 150 MeV/c^2
TreeFit p-value	$> 10^{-3}$
Single candidate selection	
Untagged sample	Choose only one random D^0 candidate per event
Tagged sample	Choose only one random D^* candidate per event

Table 6.2: The selection criteria use for $D^0 \to K\pi$ candidates.

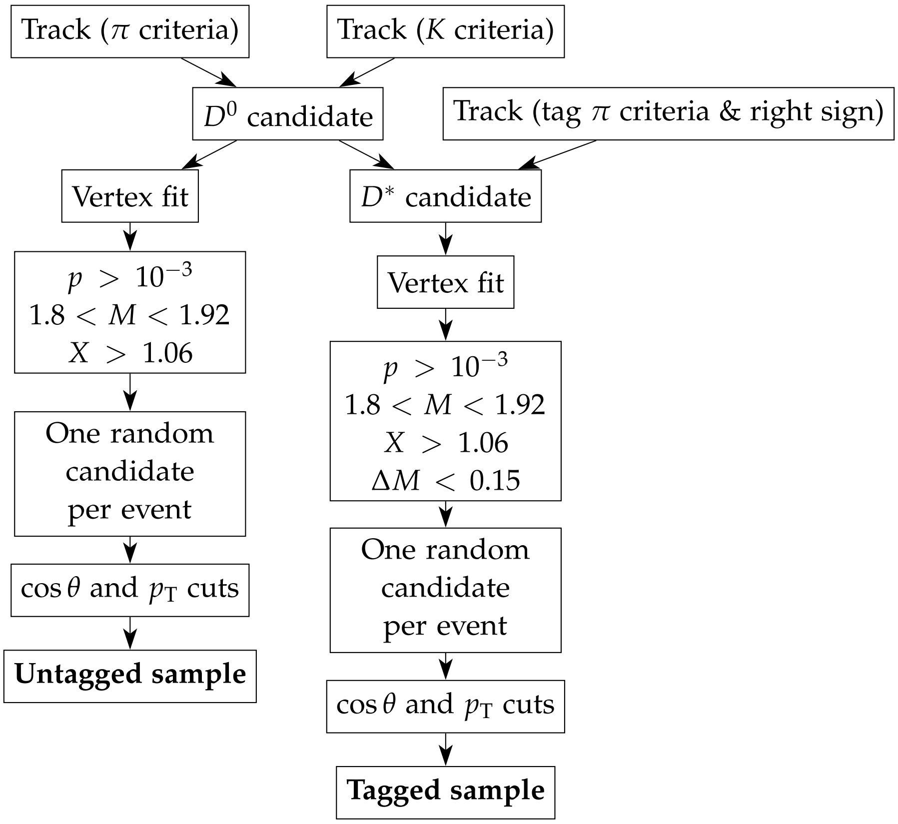


Figure 6.1: Schematic of the $D^0 \to K\pi$ tagged and untagged samples reconstruction.

6.3 SAMPLE COMPOSITION AND MODELING

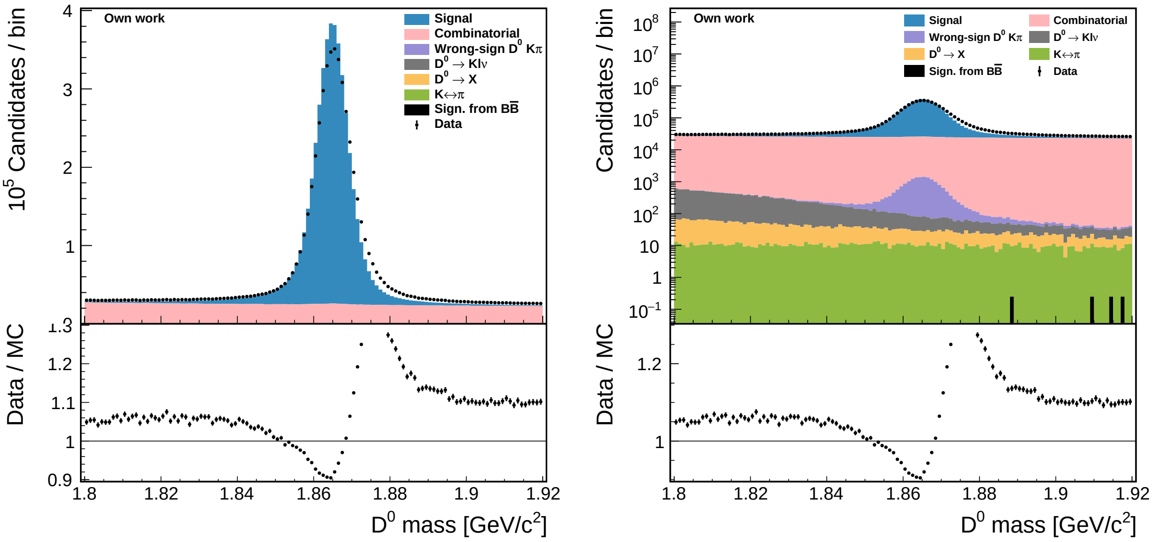
In this section I describe the composition of the two samples, and the PDFs that I use to model signal and background for both the $_s\mathcal{W}eights$ and $\mathcal{A}_{\mathrm{raw}}$ extraction.

6.3.1 *Untagged sample*

The D^0 invariant mass distribution is shown in Figure 6.2. This is the variable used in the fits. The most prominent components are the correctly-reconstructed signal decays, and the purely combinatorial background candidates.

It is necessary to mention the "wrong-sign" $D^0 \to K\pi$ decays, *i.e.* the doubly-Cabibbo-suppressed $D^0 \to K^+\pi^-$ and $\overline{D}^0 \to K^-\pi^+$. Since this sample has no tagging, such candidates perfectly mimic the signal, and the wrong flavor is always assumed for them. Their fraction with respect to "right-sign" decays on MC is $(0.3795 \pm 0.0014)\%$, compatible with the nominal value $(0.379 \pm 0.018)\%$ [15], which results in a negligible dilution of the signal asymmetry, as discussed in Section 8.6.

Other background components do not present any peaking structure, and can therefore be considered together with the combinatorial. It is worth mentioning the small number of candidates where a signal decay is reconstructed with double misidentification of the D^0 daughters, *i.e.* with the kaon swapped with the pion. Also this background always has the



(a) Linear scale.

(b) Logarithmic scale.

Figure 6.2: D^0 invariant mass distribution of the $D^0 \to K\pi$ untagged sample. The components shown are: correctly-reconstructed signal (blue), combinatorial background (pink), wrong-sign $D^0 \to K\pi$ decays (purple), $D^0 \to K\ell\nu$ decays (gray), D^0-to-other decays (yellow), right-sign $D^0 \to K\pi$ decays with kaon-pion double misidentification (green), and correctly-reconstructed signal D^0 mesons produced in the decay of a B meson (black). Black point are for data. The bottom plot is the data/MC ratio.

wrong flavor assumed. However, since it is non-peaking, it is fitted as part of the background component, so it does not taint the signal asymmetry; also, its abundance with respect to the signal is well below 0.1%.

Finally, for what concerns the data/MC agreement, there seems to be about 5% more background on data, and the signal peak has a more prominent right tail. As for the signal channel, I use an *ad-hoc* rescaling of the MC components to improve the agreement between data and MC, so that systematic uncertainties computed on simulation are reliable. The scale factors for signal and background linearly depend on the D^0 invariant mass, and are applied on a per-candidate basis. Also, I convolve the MC distributions with a 1.6 MeV/c^2-wide Gaussian distribution so that its resolution roughly matches that of data; the convolution is performed by shifting each candidate by a random number drawn from the aforementioned Gaussian distribution. Figure 6.3 shows the result: the maximum discrepancy is reduced to about 5%.

SAMPLE MODELING The PDF model for the untagged sample only includes two components: signal and combinatorial background. The signal component is modeled by the sum of a Johnson's S_U distribution and a Gaussian distribution with the same central value μ:

$$p^{\text{sig}}(M) = (1 - f_G)\text{Johnson}(M|\mu, \lambda, \gamma, \delta) + f_G \cdot \text{Gauss}(M|\mu, \sigma_G)\,. \tag{6.4}$$

The signal component also captures the wrong-sign $D^0 \to K\pi$ decays. When I fit for asymmetries, I split the λ and σ_G width parameters by D^0 flavor: this allows the PDF to model the different mass resolutions for D^0 and $\overline{D}^0$ mesons. Such difference stems mainly from the different elastic cross sections for the interaction of positive and negative kaons with nuclei:

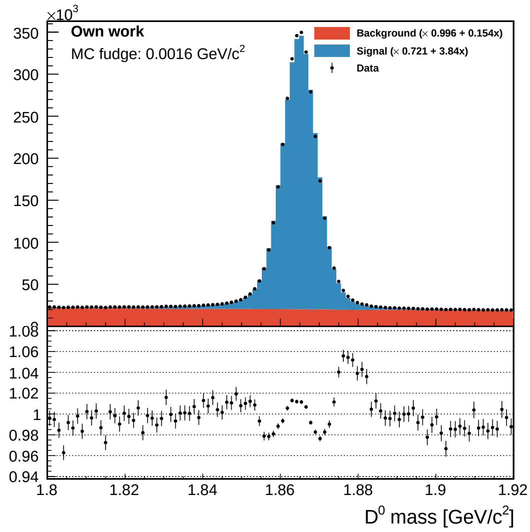


Figure 6.3: D^0 invariant mass distribution of the $D^0 \to K\pi$ untagged sample after applying the *ad-hoc* MC rescaling.

this results in different multiple scattering contributions, especially at low momentum, which in turn produces different momentum and angle resolutions.

The background component is modeled by a 1st-degree polynomial (normalized to one in the considered range):

$$p^{\text{bkg}}(M|p_1) = 1 + p_1 \cdot M\,. \tag{6.5}$$

All parameters are always left floating. The total PDF is defined as

$$p(M) = n^{\text{sig}} \cdot p^{\text{sig}}(M) + n^{\text{bkg}} \cdot p^{\text{bkg}}(M)\,. \tag{6.6}$$

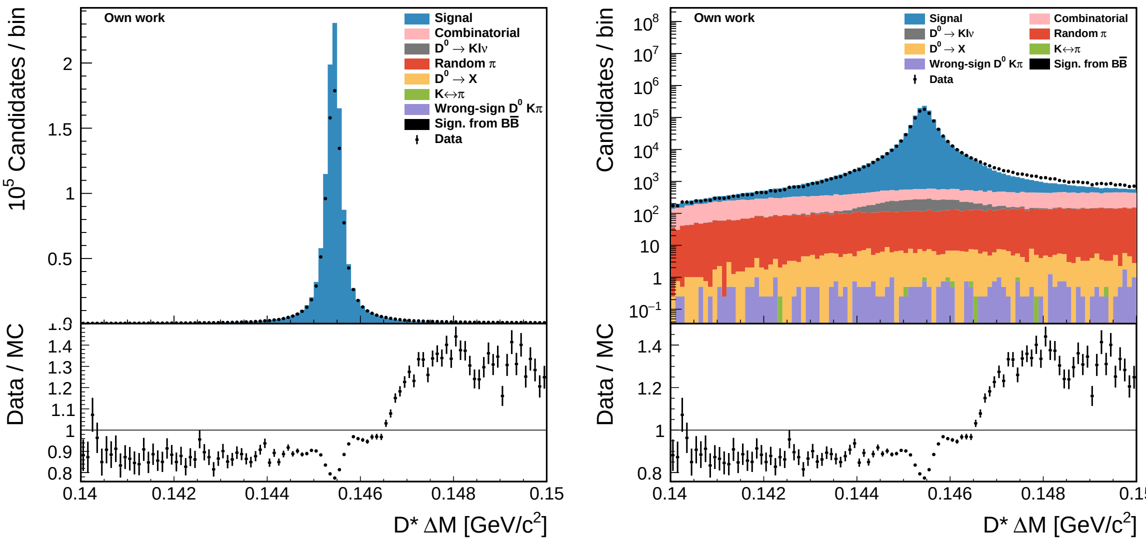
In the fit for the asymmetry extraction (simultaneous fit of the two flavors), the parameters are shared between the two flavors (the same shapes are assumed for D^0 and $\bar{D}^0$ mesons), except for n^{sig} and n^{bkg}; in this case n^{sig} and $\overline{n}^{\text{sig}}$ are defined as a function of the total signal yield $n^{\text{sig}}_{\text{tot}}$ and the raw asymmetry $\mathcal{A}^{\text{untagged}}_{\text{raw}}$, as shown in Equation 5.6 for the signal channel.

6.3.2 *Tagged sample*

The ΔM distribution is shown in Figure 6.4. This is the variable used in the fits. The most prominent components are the correctly-reconstructed signal decays, and the purely combinatorial background candidates, which are only about 5% of the total.

The only peaking background component is made of $D^0 \to K\ell\nu$ decays where the lepton is misidentified as a pion. These present a distorted peak in the ΔM distribution, similarly to the $D^0 \to K\pi\pi^0$ component of the $D^0 \to \pi^+\pi^-\pi^0$ sample. Since the abundance of this component is only $(0.351 \pm 0.003)\%$ of the signal component, I neglect it. Even if the asymmetry of this component was $\pm 100\%$ (which is far from reality), its effect on $\mathcal{A}^{\text{tagged}}_{\text{raw}}$ would still be just as small as for the wrong-sign decays in the untagged sample.

All other background components are non-peaking, and can therefore be considered together with the combinatorial. It is worth mentioning that the wrong-sign decays are non-peaking



(a) Linear scale. (b) Logarithmic scale.

Figure 6.4: ΔM distribution of the $D^0 \to K\pi$ tagged sample. The components shown are: correctly-reconstructed signal (blue), combinatorial background (pink), $D^0 \to K\ell\nu$ decays (gray), random-pion background (red), D^0-to-other decays (yellow), wrong-sign $D^0 \to K\pi$ decays (purple), right-sign $D^0 \to K\pi$ decays with kaon-pion double misidentification (green), and correctly-reconstructed signal D^0 mesons produced in the decay of a B meson (black). Black point are for data. The bottom plot is the data/MC ratio.

in ΔM because they can only be combined with a random tag pion, since I am looking for a pion with the wrong charge because I assumed the wrong flavor; this also contributes to suppressing the component. Also the component with swapped D^0 daughters is suppressed by the same principle.

Finally, for what concerns the data/MC agreement, the background seems to have larger slope in data, and the signal peak has a more prominent right tail, as for the untagged sample. Also, the abundance of signal is about 20% smaller on data with respect to MC, probably because of incorrect D^* fragmentation inputs to the generator as for the signal channel. As for the signal channel and untagged sample, I use an *ad-hoc* rescaling of the MC components to improve the agreement between data and MC, so that systematic uncertainties computed on simulation are reliable. The scale factors for signal and background linearly depend on ΔM, and are applied on a per-candidate basis. Also, I convolve the MC distributions with a 50 keV/c^2-wide Gaussian distribution so that its resolution roughly matches that of data; the convolution is performed by shifting each candidate by a random number drawn from the aforementioned Gaussian distribution. Figure 6.5 shows the result: the maximum discrepancy is reduced to about 5% in the signal peak, and about 20% in the background-dominated tails, which make up anyway a very small fraction of the sample and thus have little impact on the measurement.

SAMPLE MODELING The PDF model for the tagged sample only includes two components: signal and combinatorial background. The signal component is modeled by the sum of a

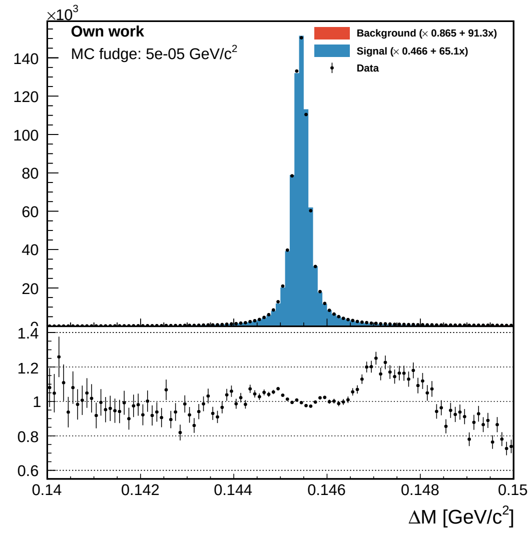


Figure 6.5: ΔM distribution of the $D^0 \to K\pi$ tagged sample after applying the *ad-hoc* MC rescaling.

Johnson's S_U distribution and two Gaussian distributions; the firs two share the same central value μ:

$$p^{\text{sig}}(\Delta M) = (1 - f_1 - f_2)\text{Johnson}(\Delta M|\mu, \lambda, \gamma, \delta) + f_1 \cdot \text{Gauss}(\Delta M|\mu, \sigma_1) \tag{6.7a}$$
$$+ f_2 \cdot \text{Gauss}(\Delta M|\mu_2, \sigma_2)\,. \tag{6.7b}$$

As for the untagged sample, I split the λ and σ_G width parameters by D^0 flavor when I fit for asymmetries, so that the PDF may account for the different ΔM resolutions for D^0 and $\overline{D}^0$ mesons.

The background component is modeled by a threshold-like function:

$$p^{\text{bkg}}(\Delta M|x_0, p_1) = \begin{cases} \sqrt{\Delta M - x_0} + p_1(\Delta M - x_0) & \Delta M > x_0 \\ 0 & \Delta M \leq x_0 \end{cases}\,. \tag{6.8}$$

All parameters are always left floating, except for x_0 which is fixed to the nominal mass of the charged pion, $139.57\,\text{MeV}/c^2$ [15]. The total PDF is defined as

$$p(\Delta M) = n^{\text{sig}} \cdot p^{\text{sig}}(\Delta M) + n^{\text{bkg}} \cdot p^{\text{bkg}}(\Delta M)\,. \tag{6.9}$$

In the fit for the asymmetry extraction (simultaneous fit of the two flavors), the parameters are shared between the two flavors (the same shapes are assumed for D^0 and $\overline{D}^0$ mesons), except for n^{sig} and n^{bkg}; in this case n^{sig} and $\overline{n}^{\text{sig}}$ are defined as a function of the total signal yield $n^{\text{sig}}_{\text{tot}}$ and the raw asymmetry $\mathcal{A}^{\text{tagged}}_{\text{raw}}$, as shown in Equation 5.6 for the signal channel.

6.4 CORRECTION EXTRACTION

In this section I describe in detail the steps I take to measure $\mathcal{A}^{\pi\,\text{tag}}_{\varepsilon}$. In those figures where comparison with the true distributions is relevant, I show results on the MC samples ($1712\,\text{fb}^{-1}$, four times the data). Otherwise, I show the results on data.

First, I discuss the procedure on the tagged sample (Section 6.4.1 and Section 6.4.2). Then, I move on to the untagged sample (Section 6.4.3 and Section 6.4.4) focusing mainly on the differences with respect to the tagged sample. And finally, I explain how I take into account the correlation between the two samples when computing the uncertainty on the $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$ (Section 6.4.5).

6.4.1 *Kinematic equalization of the tagged sample*

The purpose of kinematic equalization, introduced at the beginning of this Chapter, is to make sure that the $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$ of the tagged sample is equal to that of the $D^0 \to \pi^+\pi^-\pi^0$ sample. This is achieved by reweighting the tagged sample ("original" sample) so that the 2D distribution of p_T and $\cos\theta$ of the tag pion (on which the $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$ depends) matched that of the $D^0 \to \pi^+\pi^-\pi^0$ sample ("target" sample).

Because I need the $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$ *of the signal components* to be the same in the two samples, the weights must be computed on the distributions of the signal components only. The first step is therefore to compute the background-subtracted distributions of the two variables of interest in both the original and the target sample. Since I need to use the procedure on data, I must do this without truth information. This is accomplished using the $_s\mathcal{P}lot$ technique [80].

The $_s\mathcal{P}lot$ technique uses a fit with multiple components to define per-candidate weights, dubbed $_s\mathcal{W}eights$. With these $_s\mathcal{W}eights$, one can fill an histogram with the whole sample, and obtain the histogram of the distribution of one specific component. This works as long as the variables whose distribution is sought are independent of the fit variable(s).

The PDFs used for the $_s\mathcal{P}lot$ fits are the same used for the asymmetry extraction: the one described in Section 6.3.2 for the tagged sample, and the 2D PDF described in Chapter 5 for the $D^0 \to \pi^+\pi^-\pi^0$ sample. Note that $_s\mathcal{P}lot$ fit for the signal channel is performed independently for each $\cos\theta_\text{CM}$ bin, as is done for the $D^0 \to K\pi$ samples, in order to avoid correlations among the $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$ measured in the single bins. Also note that these whole procedure is flavor-agnostic, *i.e.* performed without distinguishing D^0 from $\overline{D}^0$ candidates.

Figure 6.6 shows the fit on the tagged data sample (all $\cos\theta_\text{CM}$ bins projected together), while Figure 6.7 compares the $_s\mathcal{W}eighted$ distributions to the true ones on the MC tagged sample, proving the effectiveness of this technique. Figure 6.8 and Figure 6.9 shows the same for the $D^0 \to \pi^+\pi^-\pi^0$ sample.

Weights for the asymmetry extraction of the original (tagged sample) distribution are computed by filling the $_s\mathcal{W}eighted$ 2D histograms of the tag pion p_T and $\cos\theta$, and computing the bin-by-bin ratio of these histograms. For each variable, 50 bins are used, resulting in an uncertainty of $\sim 3\%$ for the weight of each bin. The bin edges are chosen so that the 1D bins are equally populated for the target ($D^0 \to \pi^+\pi^-\pi^0$) sample. The histogram is then smoothed using a 2D Gaussian filter, and the weights are assigned to each candidate depending on the bin that it falls into.

The results of the reweighting procedure are shown in Figure 6.10 and Figure 6.11. Some discrepancies are present in the least-populated regions of the distributions; nevertheless, systematic effects due to the imperfect reweighting are evaluated in Section 8.5.

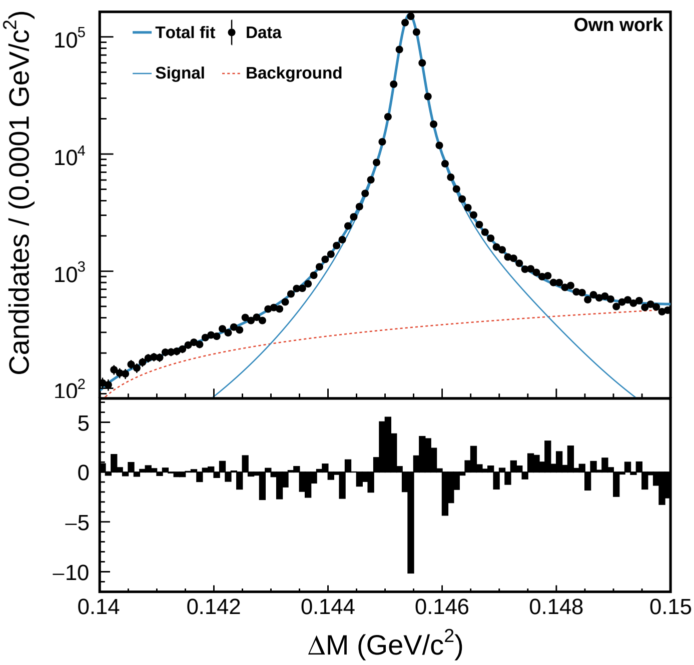
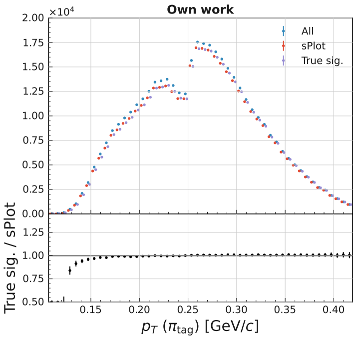
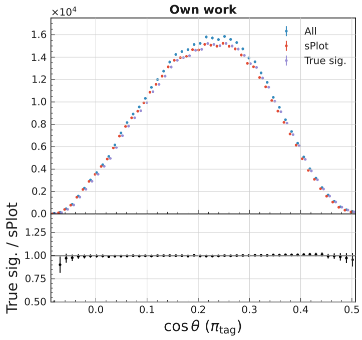


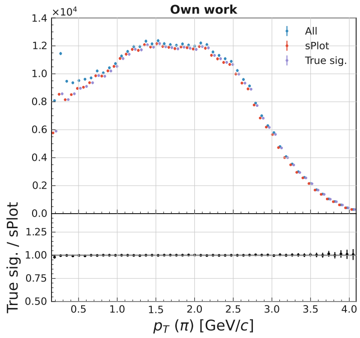
Figure 6.6: ${}_s\mathcal{P}lot$ fit of the tagged $D^0 \to K\pi$ sample on data. All $\cos\theta_{\mathrm{CM}}$ bins are projected together. The bottom plot shows the pulls.



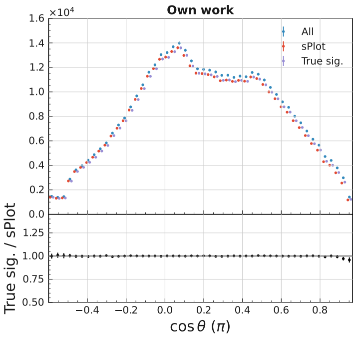
(a) π_{tag} p_{T}.



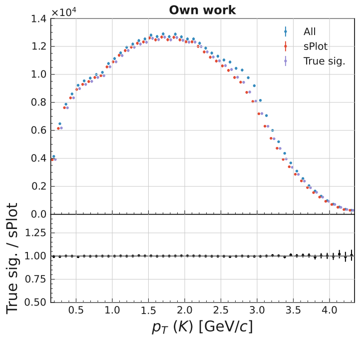
(b) π_{tag} $\cos\theta$.



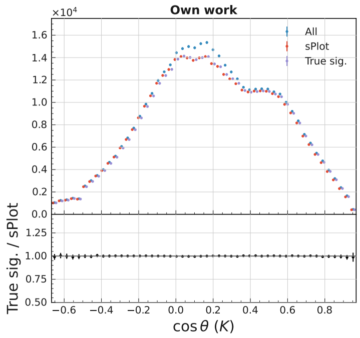
(c) π p_{T}.



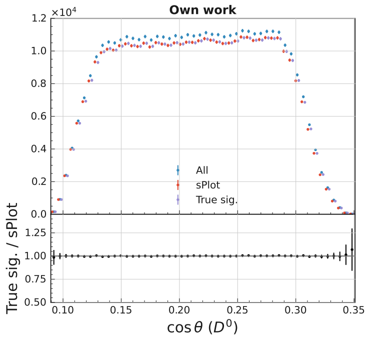
(d) π $\cos\theta$.



(e) K p_{T}.

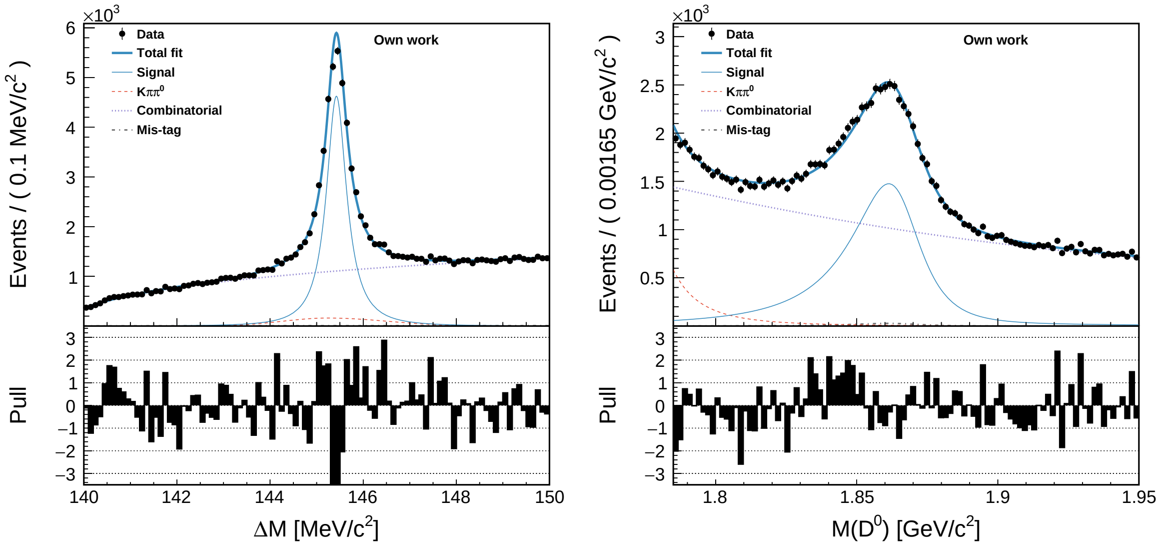


(f) K $\cos\theta$.



(g) D^0 $\cos\theta_{\text{CM}}$.

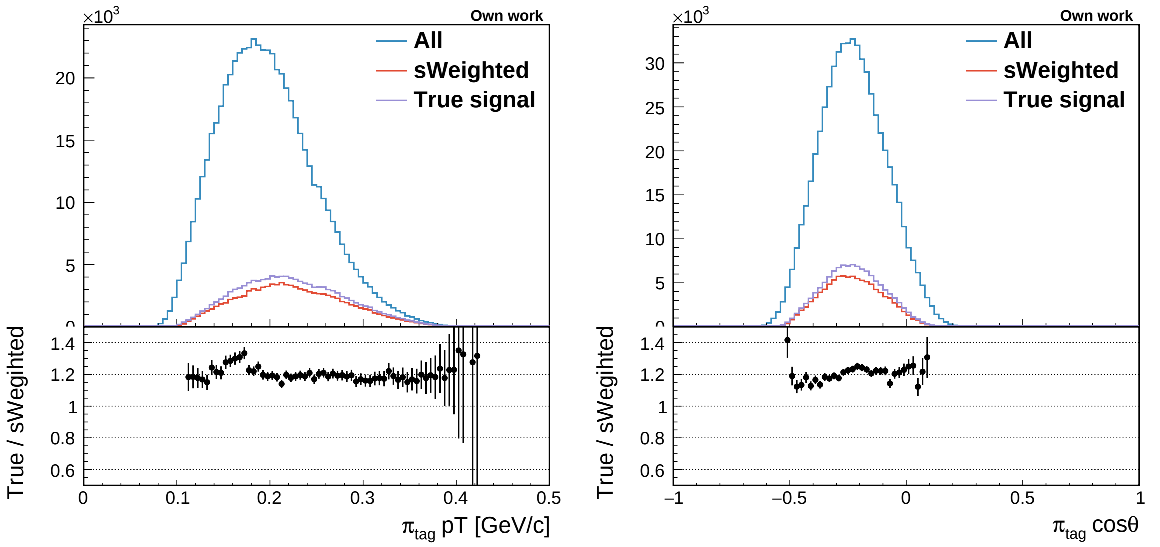
Figure 6.7: Comparison of the whole sample, true signal, and $_s\mathcal{W}eighted$ distributions of the seven tagged sample variables used in the kinematic equalizations, on MC, for bin 4, *i.e.* $\cos\theta_{\text{CM}}$ between -0.208 and 0. The bottom plot shows the ratio between true and $_s\mathcal{W}eighted$ distributions, which should match.



(a) ΔM.

(b) D^0 invariant mass.

Figure 6.8: ${}_s\mathcal{P}lot$ fit of the $D^0 \to \pi^+\pi^-\pi^0$ sample, on data, for bin 4, *i.e.* $\cos\theta_{\text{CM}}$ between -0.208 and 0. The bottom plot shows the pulls.



(a) π_{tag} p_{T}.

(b) π_{tag} $\cos\theta$.

Figure 6.9: Comparison of the whole sample, true signal, and ${}_s\mathcal{W}eighted$ distributions of the two $D^0 \to \pi^+\pi^-\pi^0$ sample variables used in the kinematic equalization, on MC, for bin 4, *i.e.* $\cos\theta_{\text{CM}}$ between -0.208 and 0. The bottom plot shows the ratio between true and ${}_s\mathcal{W}eighted$ distributions, which should match.

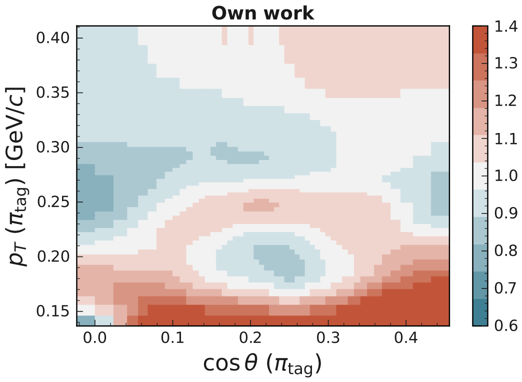


Figure 6.10: Weights for the tagged $D^0 \to K\pi$ sample as a function of the p_T and $\cos\theta$ of the tag pion, on MC, for bin 4, *i.e.* $\cos\theta_\mathrm{CM}$ between -0.208 and 0.

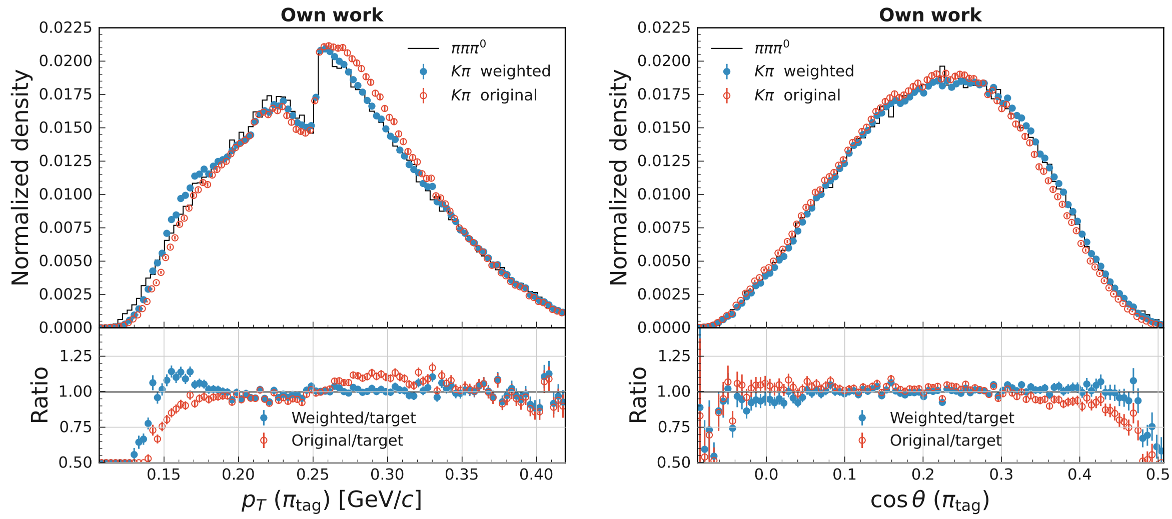
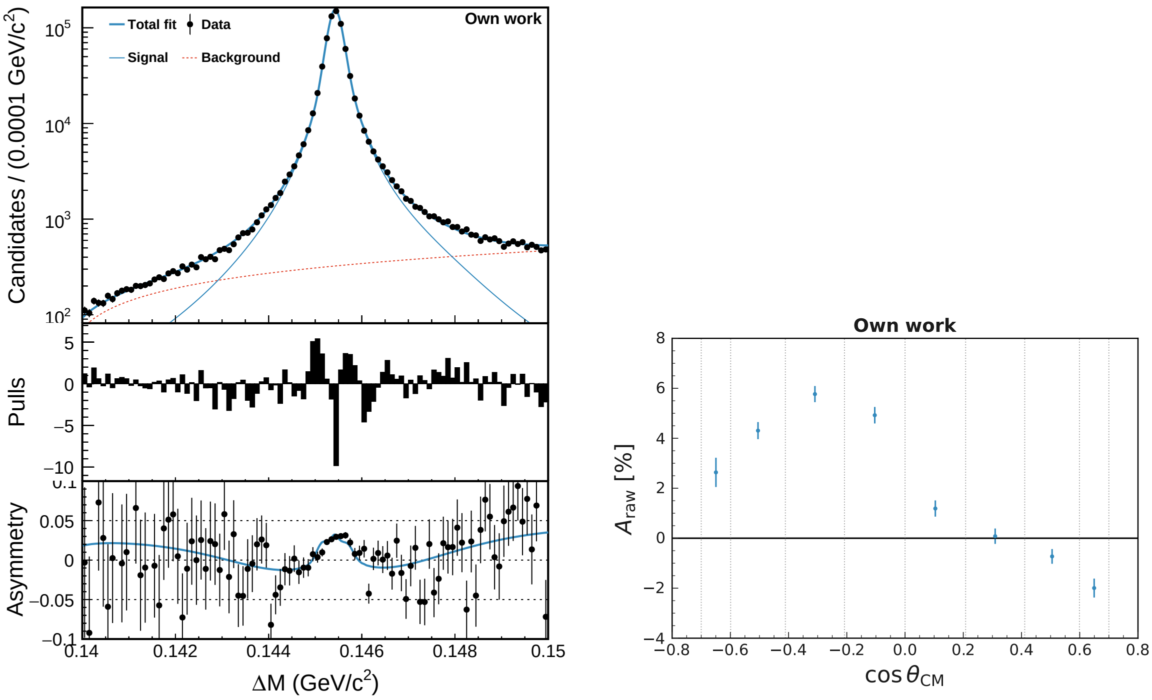


Figure 6.11: Comparison of the original (red empty points) and reweighted (blue filled points) tagged sample distributions to the true signal $D^0 \to \pi^+\pi^-\pi^0$ distributions (black line), for the reweighting variables, on MC, for bin 4, *i.e.* $\cos\theta_\mathrm{CM}$ between -0.208 and 0. The bottom plots show the ratios to the $D^0 \to \pi^+\pi^-\pi^0$ distributions.

6.4.2 *Asymmetry extraction in the tagged sample*

The $\mathcal{A}_{\text{raw}}^{\text{tagged}}$ is extracted by a simultaneous fit of the reweighted D^0 and $\bar{D}^0$ samples, using the same PDF as the ${}_s\mathcal{P}lot$ fit, as anticipated in Section 6.3.2. Again, all parameters are left floating. Figure 6.12 shows the results on data. The $\mathcal{A}_{\text{raw}}^{\text{tagged}}$ extracted as a function of the $\cos\theta_{\text{CM}}$ bin roughly follows the pattern expected from MC. Numerical results and plots for each $\cos\theta_{\text{CM}}$ bin can be found in Table C.1 and Section C.2.



(a) Fit.

(b) Resulting asymmetries.

Figure 6.12: Asymmetry fit of the tagged $D^0 \to K\pi$ sample on data. All $\cos\theta_{\text{CM}}$ bins and both flavors are projected together. The bottom plot shows the pulls.

6.4.3 *Kinematic equalization of the untagged sample*

In this case, the purpose of the kinematic equalization is to make sure that the $\mathcal{A}_{\varepsilon}^{D^0\to K\pi}$ and $\mathcal{A}_{\text{prod}}$ of the untagged sample ("original" sample) are the same as those of the already-reweighted tagged sample. These two asymmetries depend on five variables in total: the p_{T} and $\cos\theta$ of the D^0 daughters ($\mathcal{A}_{\varepsilon}^{D^0\to K\pi}$), and the $\cos\theta_{\text{CM}}$ of the D^0 ($\mathcal{A}_{\text{prod}}$).

The ${}_s\mathcal{P}lot$ procedure is the same as for the tagged sample. Figure 6.13 shows the flavor-agnostic fit on data. Figure 6.14 compares the ${}_s\mathcal{W}eighted$ distributions to the true ones on the MC tagged sample. Because the untagged sample has a much larger background than the tagged sample, the difference between full and ${}_s\mathcal{W}eighted$ distributions is much more evident. Some residual discrepancies can be noticed between these. The largest discrepancies arise at very low $\cos\theta$ and transverse momentum of the D^0 daughters: the bulk of these discrepancies was actually already excluded by introducing selections in these variables, as mentioned in Section 6.2. As the remaining discrepancies may affect the reweighting and, thus, the measurement, systematic effects due to the imperfect reweighting are evaluated in Section 8.5.

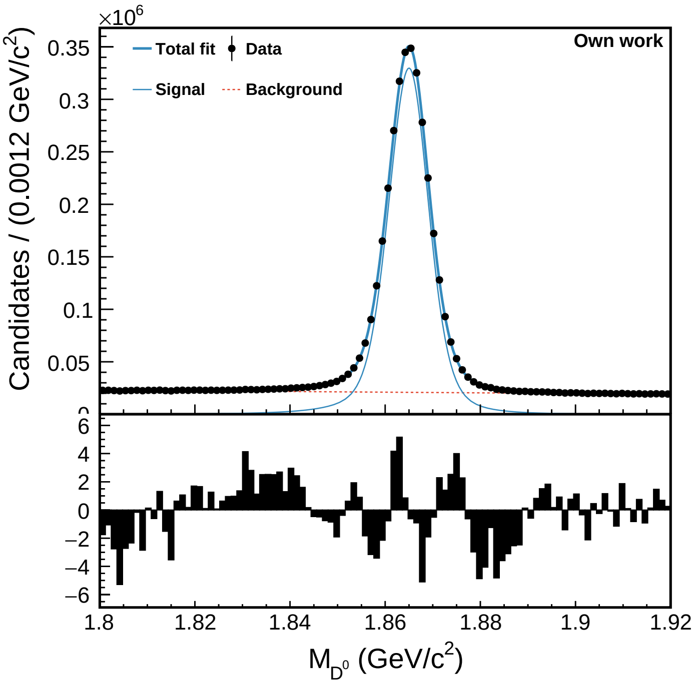
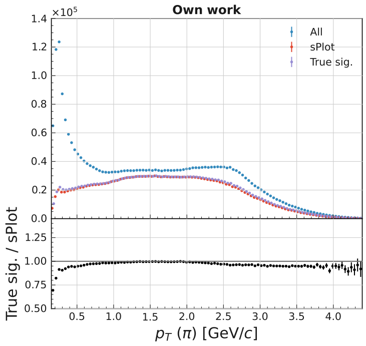


Figure 6.13: ${}_{S}\mathcal{P}lot$ fit of the untagged $D^0 \to K\pi$ sample on data. All $\cos\theta_{\mathrm{CM}}$ bins are projected together. The bottom plot shows the pulls.

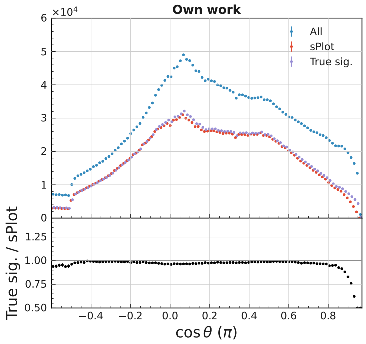
Since the available statistics is insufficient to fill a 5D histogram with fine enough binning, the weight computation procedure is different from the one used for the tagged sample. I fill ${}_{S}\mathcal{W}eighted$ 1D histograms of one variable at a time, compute the bin-by-bin ratio of these histograms, and assign weights to the original (untagged) sample candidates. When I move on to the following variable, I use the product of the ${}_{S}\mathcal{W}eights$ and the weights I just computed for filling the histogram, and I use the ratio of the histograms to update the weights (*i.e.* I compute the product of the weights from the previous step, and those of the current step). Once I compute the weights for the last variable, I start over from the first, and continue until the procedure converges (*i.e.* the weights change by less than their statistical uncertainty).

For each variable, I use 100 bins. The bin edges are chosen so that the 1D bins are equally populated for the target (tagged) sample. In the end, this results in weight uncertainties of $\sim 3\%$.

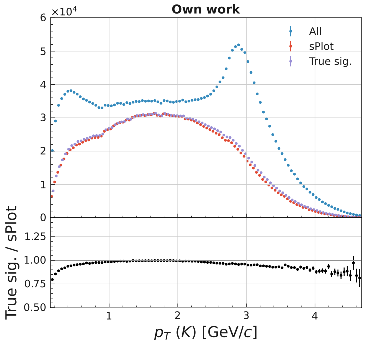
The results of the reweighting procedure are shown in Figure 6.15 and Figure 6.16. The differences in distributions between tagged and untagged samples are much larger than those between tagged and $D^0 \to \pi^+\pi^-\pi^0$ samples: this is because of the impact of the D^* reconstruction efficiency on the kinematic distributions. Some discrepancies are present: systematic effects due to the imperfect reweighting are evaluated in Section 8.5.



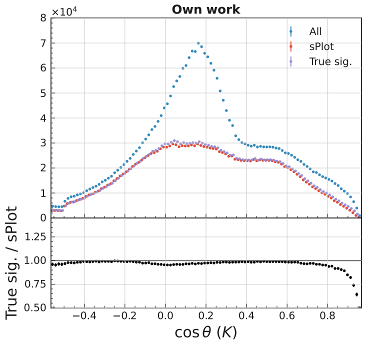
(a) π p_{T}.



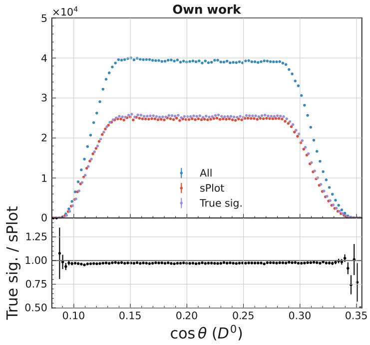
(b) π $\cos\theta$.



(c) K p_{T}.



(d) K $\cos\theta$



(e) D^0 $\cos\theta_{\mathrm{CM}}$.

Figure 6.14: Comparison of the whole sample, true signal, and $_s$*Weighted* distributions of the five untagged sample variables used in the kinematic equalization, on MC, for bin 4, *i.e.* $\cos\theta_{\mathrm{CM}}$ between -0.208 and 0. The bottom plot shows the ratio between true and $_s$*Weighted* distributions, which should match.

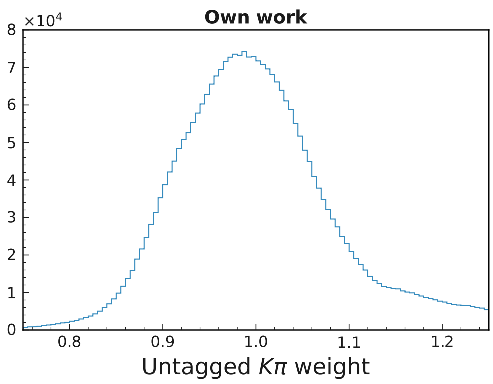
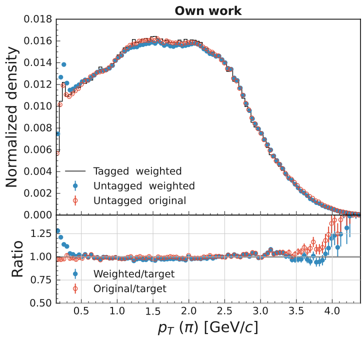
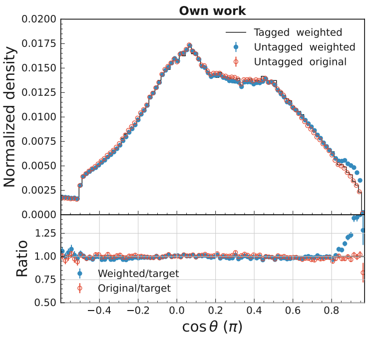


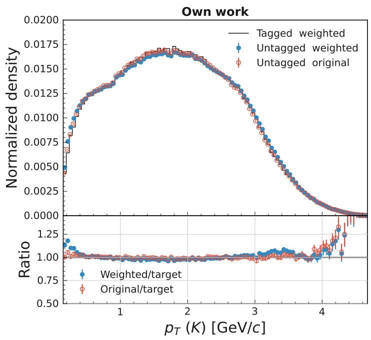
Figure 6.15: Weights distribution for the untagged $D^0 \to K\pi$ sample, on MC, for bin 4, *i.e.* $\cos\theta_{\mathrm{CM}}$ between -0.208 and 0.



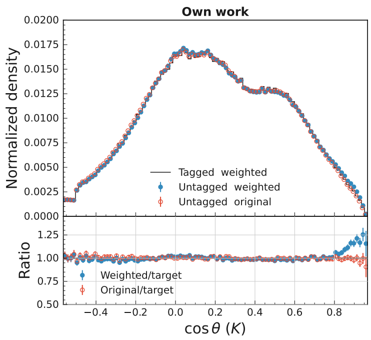
(a) π p_{T}.



(b) π $\cos\theta$.



(c) K p_{T}.



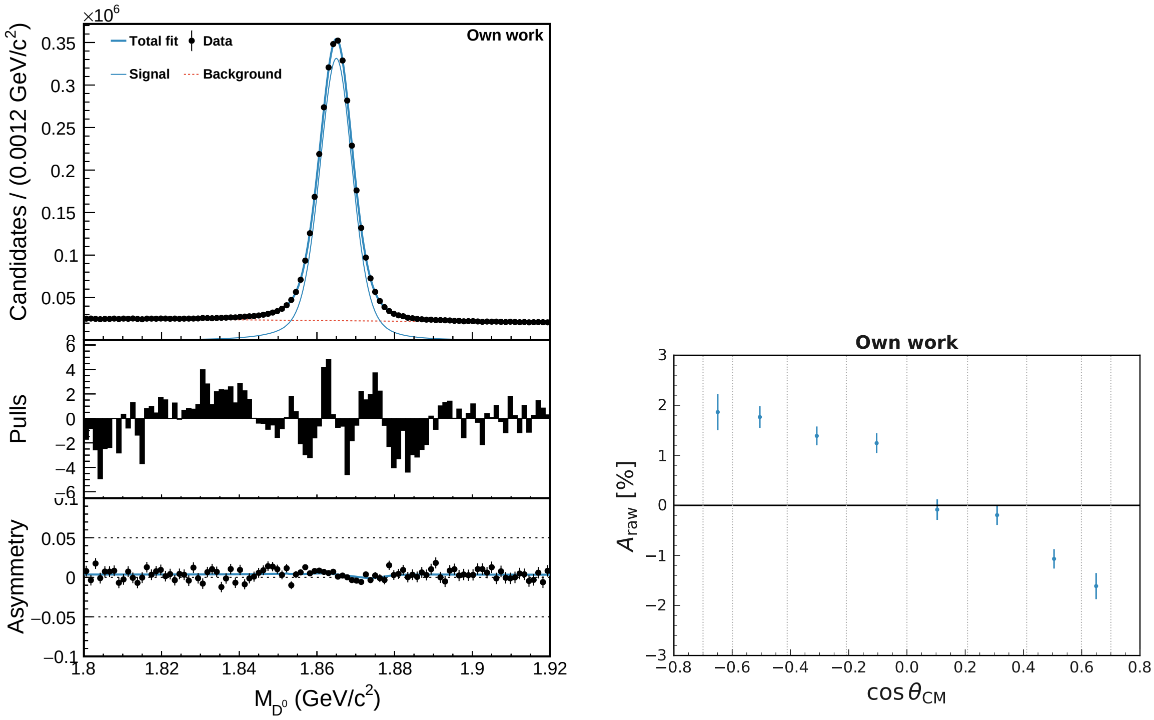
(d) K $\cos\theta$.

(e) D^0 $\cos\theta_{\mathrm{CM}}$.

Figure 6.16: Comparison of the original (red empty points) and reweighted (blue filled points) untagged sample distributions to the true signal tagged sample distributions (black line), for the reweighting variables, on MC, for bin 4, *i.e.* $\cos\theta_{\mathrm{CM}}$ between -0.208 and 0. The bottom plots show the ratios to the tagged sample distributions.

6.4.4 *Asymmetry extraction in the untagged sample*

The $\mathcal{A}_{\text{raw}}^{\text{untagged}}$ is extracted by a simultaneous fit of the reweighted D^0 and $\overline{D}^0$ samples, using the same PDF as the $_{S}\mathcal{P}lot$ fit, as for the tagged sample. Again, all parameters are left floating. Figure 6.17 shows the results on data. The $\mathcal{A}_{\text{raw}}^{\text{untagged}}$ extracted as a function of the $\cos\theta_{\text{CM}}$ bin roughly follows the pattern expected from MC. Numerical results and plots for all bins can be found in Table C.1 and Section C.3.



(a) Fit. (b) Resulting asymmetries.

Figure 6.17: Asymmetry fit of the untagged $D^0 \to K\pi$ sample on data. All $\cos\theta_{\text{CM}}$ bins and both flavors are projected together. The bottom plot shows the pulls.

6.4.5 *Correction computation*

I computed the $\mathcal{A}_{\varepsilon}^{\pi\,\text{tag}}$ simply as the difference $\mathcal{A}_{\text{raw}}^{\text{tagged}} - \mathcal{A}_{\text{raw}}^{\text{untagged}}$ in each $\cos\theta_{\text{CM}}$ bin. However, to correctly compute its statistical uncertainty, I must take into account the correlation between the tagged and untagged samples. All D^0 candidates that appear in the untagged sample also appear in the tagged sample, except for a small fraction of the events in which the single candidate selection chooses two different candidates in the two samples. The overlap between the two samples in quantified in Table 6.3.

In the following, I neglect the $\sim 1.5\%$ of the candidates that only appear in the tagged sample, *i.e.* I assumed that the signal candidates of the tagged sample are a subset of those of the untagged sample. Using a lighter notation, let:

- n_o be the number of signal D^0s that appear *only* in the untagged sample;

	Signal D^0 candidates in the samples			Common fraction	
	Tagged	Untagged	Both	Tagged sample	Untagged sample
D^0	2.270×10^6	8.749×10^6	2.231×10^6	98.3%	25.5%
$\overline{D}^0$	2.282×10^6	8.901×10^6	2.243×10^6	98.3%	25.2%
Total	4.554×10^6	17.66×10^6	4.486×10^6	98.5%	25.4%

Table 6.3: Overlap between the tagged and untagged $D^0 \to K\pi$ samples on MC.

- $\overline{n}_o$ be the number of signal $\overline{D}^0$s that appear *only* in the untagged sample;
- $t_o = n_o + \overline{n}_o$ be the total signal candidates *only* in the untagged sample;
- $\mathcal{A}_o = \frac{n_o - \overline{n}_o}{n_o + \overline{n}_o}$ be the asymmetry of the candidates that are *only* in the untagged sample;

and n_t, $\overline{n}_t$, t_t, and $\mathcal{A}_t$ be the corresponding quantities for the tagged sample, *i.e.* for the candidates that appear in *both* samples.

The asymmetry I measure on the untagged sample $\mathcal{A}_u$ stems both from the candidates that appear only in the untagged sample (n_o and $\overline{n}_o$) and from the candidates that appear in both (n_t and $\overline{n}_t$). Therefore

$$\mathcal{A}_u = \frac{n_t + n_o - \overline{n}_t - \overline{n}_o}{n_t + n_o + \overline{n}_t + \overline{n}_o} = \frac{t_t}{t_t + t_o}\mathcal{A}_t + \frac{t_o}{t_t + t_o}\mathcal{A}_o \,. \tag{6.10}$$

The variance of $\mathcal{A}_\pi = \mathcal{A}_t - \mathcal{A}_u$ is

$$\mathrm{Var}(\mathcal{A}_\pi) = \mathrm{Var}(\mathcal{A}_u) + \mathrm{Var}(\mathcal{A}_t) - 2\,\mathrm{Cov}(\mathcal{A}_u, \mathcal{A}_t) \,. \tag{6.11}$$

The variances of $\mathcal{A}_u$ and $\mathcal{A}_t$ are the statistical uncertainties on $\mathcal{A}_{\mathrm{raw}}$ estimated by the fits, so I only need to compute the covariance:

$$\mathrm{Cov}(\mathcal{A}_u, \mathcal{A}_t) = \mathrm{Cov}(\frac{t_t}{t_t + t_o}\mathcal{A}_t, \mathcal{A}_t) + \mathrm{Cov}(\frac{t_o}{t_t + t_o}\mathcal{A}_o, \mathcal{A}_t) \,. \tag{6.12}$$

The second term cancels since $\mathcal{A}_o$ is independent from $\mathcal{A}_t$ (they are the asymmetries of two non-overlapping subsets of the untagged sample), and $\mathcal{A}_t$ is independent from t_t (see Section A.4). Therefore

$$\mathrm{Cov}(\mathcal{A}_u, \mathcal{A}_t) = \mathrm{E}\left[\frac{t_t \mathcal{A}_t^2}{t_t + t_o}\right] - \mathrm{E}\left[\frac{t_t \mathcal{A}_t}{t_t + t_o}\right]\mathrm{E}[\mathcal{A}_t] \tag{6.13a}$$

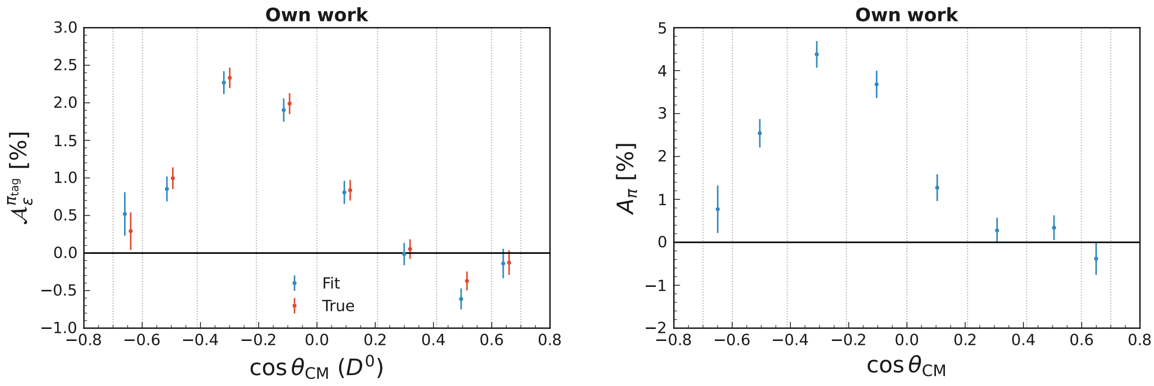
$$= \mathrm{E}\left[\frac{t_t}{t_t + t_o}\right]\left(\mathrm{E}[\mathcal{A}_t^2] - \mathrm{E}^2[\mathcal{A}_t]\right) = \mathrm{E}\left[\frac{t_t}{t_t + t_o}\right]\mathrm{Var}(\mathcal{A}_t) \,, \tag{6.13b}$$

and so I estimate the statistical uncertainty on $\mathcal{A}_\varepsilon^{\pi\,\mathrm{tag}}$ as the square root of (going back to the original notation)

$$\mathrm{Var}(\mathcal{A}_\varepsilon^{\pi\,\mathrm{tag}}) \simeq \mathrm{Var}(\mathcal{A}_{\mathrm{raw}}^{\mathrm{untagged}}) + \left(1 - 2\frac{n_{\mathrm{tot,tagged}}^{\mathrm{sig}}}{n_{\mathrm{tot,untagged}}^{\mathrm{sig}}}\right)\mathrm{Var}(\mathcal{A}_{\mathrm{raw}}^{\mathrm{tagged}}) \,. \tag{6.14}$$

Figure 6.18 shows the $\mathcal{A}_\varepsilon^{\pi\,\mathrm{tag}}$ and its uncertainty thus computed as a function of the $\cos\theta_{\mathrm{CM}}$ bin, for both data and MC. Numerical results are reported in Table 6.4.

The $\mathcal{A}_\varepsilon^{\pi\,\mathrm{tag}}$ on data roughly follows the pattern expected found on MC, but its magnitude is about two times larger. This is compatible with the findings of previous Belle II analyses, and is probably due to overly optimistic resolutions on simulation allowing a better reconstruction of low momentum tracks than what happens on real data. Since the final correction is estimated on data, the discrepancy with MC has no impact on the measurement.



(a) Simulation.

(b) Data.

Figure 6.18: Tag pion asymmetry as a function of the $\cos\theta_{\mathrm{CM}}$ bin for the MC and data samples. The red points in the MC plot are the true asymmetries.

	$\cos\theta_{\mathrm{CM}}$ bin	Tag pion asymmetries [%]		
		MC (true)	MC (fit)	Data
1	$[-0.7, -0.599)$	0.29 ± 0.25	0.52 ± 0.29	0.8 ± 0.6
2	$[-0.599, -0.411)$	1.00 ± 0.15	0.85 ± 0.17	2.54 ± 0.33
3	$[-0.411, -0.208)$	2.33 ± 0.14	2.27 ± 0.15	4.38 ± 0.31
4	$[-0.208, 0)$	1.99 ± 0.14	1.90 ± 0.15	3.68 ± 0.32
5	$[0, 0.208)$	0.84 ± 0.14	0.81 ± 0.15	1.27 ± 0.31
6	$[0.208, 0.411)$	0.05 ± 0.13	-0.01 ± 0.15	0.28 ± 0.30
7	$[0.411, 0.599)$	-0.37 ± 0.13	-0.61 ± 0.14	0.34 ± 0.29
8	$[0.599, 0.7)$	-0.13 ± 0.16	-0.14 ± 0.20	-0.4 ± 0.4

Table 6.4: Tag pion asymmetry in the simulated and data samples.

7 RESULTS

In this chapter I show the final *CP* asymmetry measurement on simulation and data. Summarizing, this measurement is obtained by:

1. measuring the raw asymmetry $\mathcal{A}_{\text{raw}}$ of $D^0 \to \pi^+\pi^-\pi^0$ in each $\cos\theta_{\text{CM}}$ bin, as described in Chapter 5;
2. measuring the tagging-induced asymmetry $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$ from the $D^0 \to K\pi$ control channel in each $\cos\theta_{\text{CM}}$ bin, as described in Chapter 6;
3. subtracting the $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$ from the signal channel $\mathcal{A}_{\text{raw}}$ bin by bin;
4. "folding" the corrected asymmetries, *i.e.* computing the arithmetic average of the asymmetries in each symmetric pair of $\cos\theta_{\text{CM}}$ bins; and
5. fitting the *CP* asymmetry measurements (averages) thus obtained with a constant function to obtain the final result.

I first show the results on Monte Carlo (MC) where the generated asymmetry is known to be zero (Section 7.1), then show the results on data (Section 7.2). Finally, I present the consistency checks I perform to validate my results (Section 7.3).

7.1 FINAL *CP* ASYMMETRY MEASUREMENT ON SIMULATION

Figure 7.1 summarizes the results from the signal $D^0 \to \pi^+\pi^-\pi^0$ channel and the $D^0 \to K\pi$ control channel, and shows the final $\mathcal{A}_{CP}$ measurement. The four arithmetic means from the folding (Figure 7.1d, which are all $\mathcal{A}_{CP}$ measurements themselves) are compatible with each other: this acts as a consistency check.

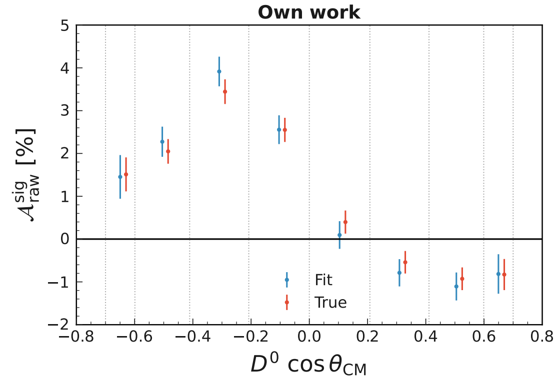
The final measurement, obtained by fitting the means with a constant, is $\mathcal{A}_{CP} = (0.22 \pm 0.14)\%$, where the uncertainty is statistical and includes both the contribution from the $D^0 \to \pi^+\pi^-\pi^0$ sample and from the control samples. Correlations among the signal $\mathcal{A}_{\text{raw}}$ measurements are neglected because the fit estimates them to be compatible with zero; those among the $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$ measurements are zero by construction, as each $\cos\theta_{\text{CM}}$ bin is treated independently of the others.

The final measurement is compatible with zero, the $\mathcal{A}_{CP}$ value assumed by the generator of the simulation. This acts as a second consistency check for the $\mathcal{A}_{CP}$ extraction strategy.

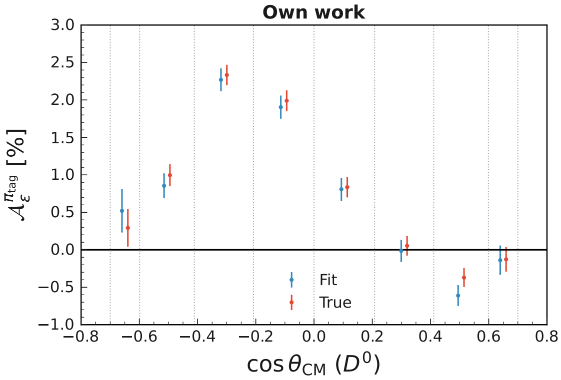
7.2 FINAL *CP* ASYMMETRY MEASUREMENT ON DATA

Figure 7.2 summarizes the results from the signal $D^0 \to \pi^+\pi^-\pi^0$ channel and the $D^0 \to K\pi$ control channel, and shows the final $\mathcal{A}_{CP}$ measurement.

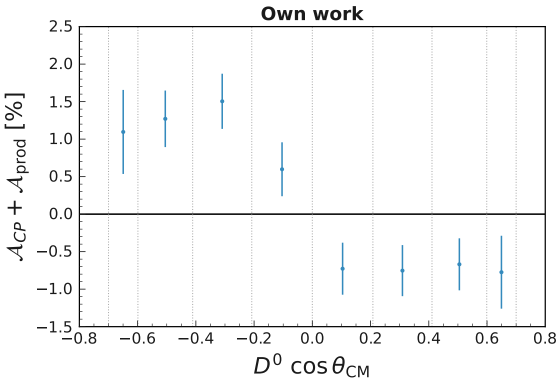
The four arithmetic means from the folding (Figure 7.2d, which are all $\mathcal{A}_{CP}$ measurements themselves) are compatible with each other: this acts as a consistency check. The final result (not yet approved by the Belle II collaboration) is $\mathcal{A}_{CP} = (0.29 \pm 0.27)\%$, where the uncertainty is statistical and is the quadrature sum of the uncertainty from the signal raw asymmetry



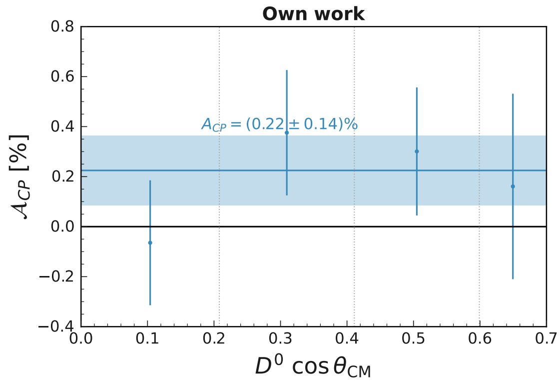
(a) Raw signal asymmetries in percent.



(b) Tagging asymmetry from $D^0 \to K\pi$.

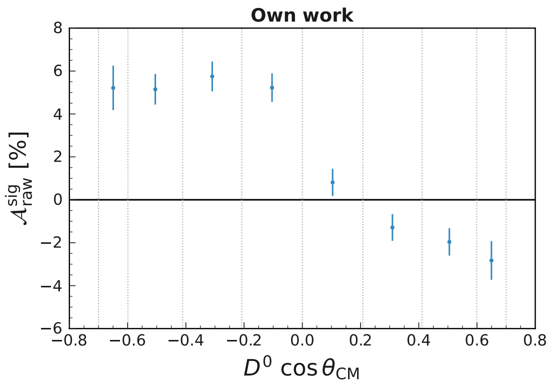


(c) $\mathcal{A}_{\varepsilon}^{\pi\,\rm tag}$-corrected signal asymmetry.

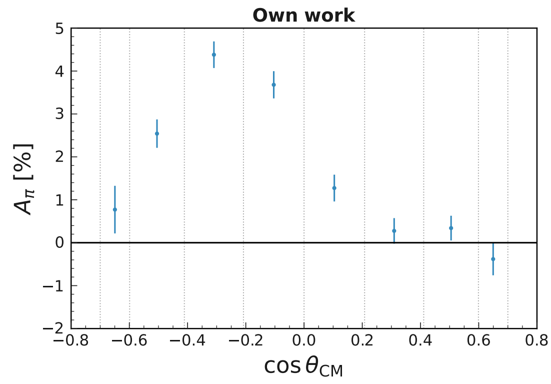


(d) Folded signal asymmetry, *i.e.* $\mathcal{A}_{CP}$. The uncertainty is statistical and includes the contribution from the control samples.

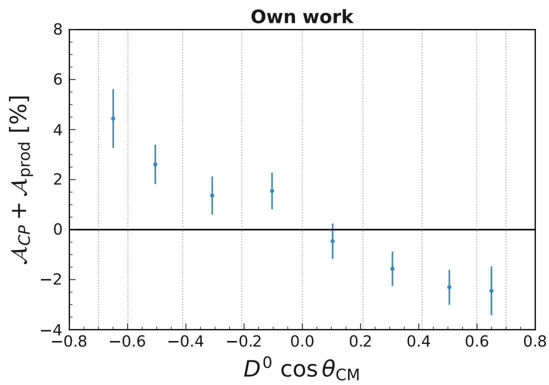
Figure 7.1: CP asymmetry results on simulation (1712 fb^{-1}). The red points show the true asymmetries.



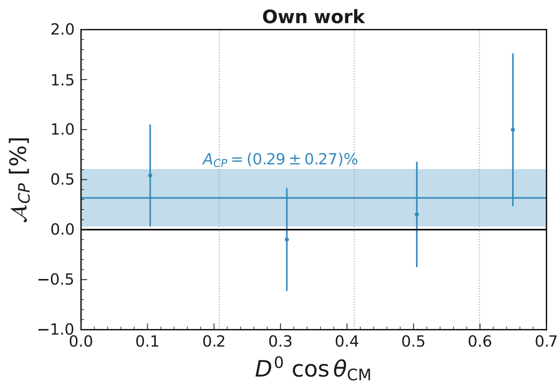
(a) Raw signal asymmetries in percent.



(b) Tagging asymmetry from $D^0 \to K\pi$.



(c) $\mathcal{A}_{\varepsilon}^{\pi\,\mathrm{tag}}$-corrected signal asymmetry.



(d) Folded signal asymmetry, *i.e.* $\mathcal{A}_{CP}$. The uncertainty is statistical and includes the contribution from the control samples.

Figure 7.2: CP asymmetry results on data ($428\,\mathrm{fb}^{-1}$).

fit (0.25%) and the $\mathcal{A}_{\varepsilon}^{\pi\,\mathrm{tag}}$ uncertainty computed as described in Section 6.4.5 (0.12%). The uncertainty is about twice as large as that on MC, as expected (the MC sample is four times larger).

7.3 VALIDATION AND CROSS CHECKS

In this section I describe the checks I perform to validate my results and ensure their internal consistency. All these checks are performed *before* the box opening.

In Section 7.3.1, I verify the assumptions behind the "folding" method for correcting the $\mathcal{A}_{\mathrm{prod}}$. In Section 7.3.2 and Section 7.3.3 I check the effect of my specific choices of $\cos\theta_{\mathrm{CM}}$ binning and fit models by testing alternatives. Finally, in Section 7.3.4, I split my data sample in several parts with criteria that may induce a bias on $\mathcal{A}_{CP}$ and verify the consistency of the results on the single parts.

7.3.1 *Production asymmetry correction*

The assumptions behind the "folding" methods were presented in Section 3.3. In summary, I need the opposite $\mathcal{A}_{\mathrm{prod}}$ contributions in the bins of each symmetric pair. To achieve this, I built $\cos\theta_{\mathrm{CM}}$ bins narrow enough that the reconstruction efficiency is approximately constant within each bin. The reconstruction efficiency on MC as a function of $\cos\theta_{\mathrm{CM}}$ is shown in Figure 5.1b, and reported again in Figure 7.3 in a "folded" plot.

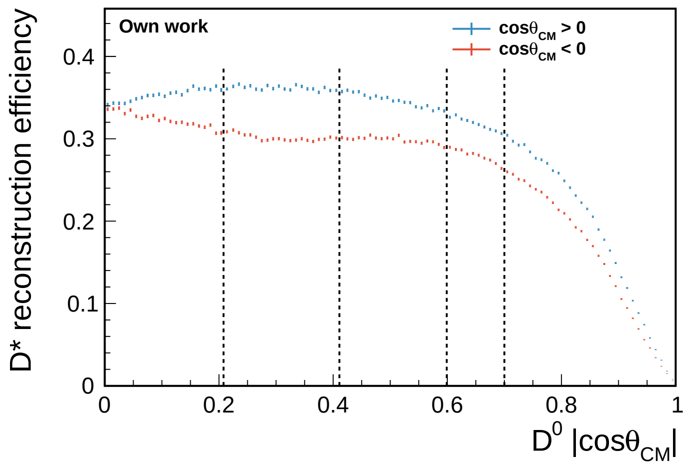


Figure 7.3: Reconstruction efficiency of the signal $D^0 \to \pi^+\pi^-\pi^0$ decays on MC, as a function of $|\cos\theta_{\mathrm{CM}}|$, split by positive (blue) and negative (red) $\cos\theta_{\mathrm{CM}}$.

The efficiency is clearly not constant within all bins. However, the requirement of opposite $\mathcal{A}_{\mathrm{prod}}$ contributions in bins of a pair can be expanded to the next order in the following way:

- the efficiency within each bin is approximately *linear*; and
- the average $|\cos\theta_{\mathrm{CM}}|$ of the signal candidates in the bins of a pair is the same.

I assume that the first condition is verified. For what concerns the second condition, Figure 7.4 shows the difference in average $|\cos\theta_{\mathrm{CM}}|$ between the two bins of each symmetric pair. The largest difference occurs in the central bins (4 and 5, *i.e.* around $\cos\theta_{\mathrm{CM}} = 0$), and amounts to $(2.6 \pm 0.2) \times 10^{-3}$.

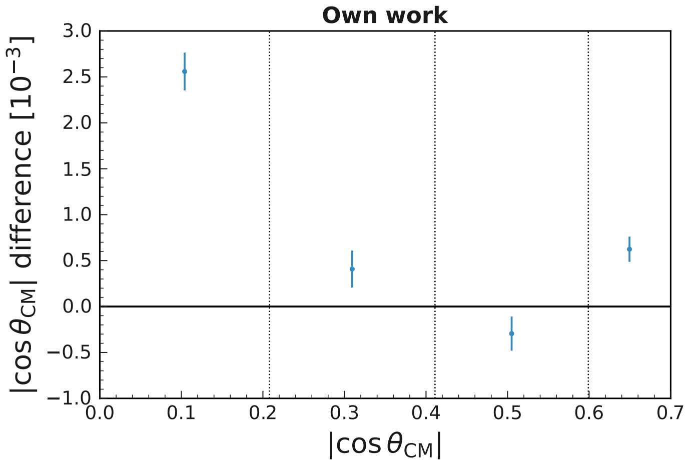


Figure 7.4: Difference between the average $|\cos\theta_{\mathrm{CM}}|$ of the signal candidates in each symmetric pair of bins.

To evaluate the impact of this difference, we take the slope of $\mathcal{A}_{\mathrm{prod}}(\cos\theta_{\mathrm{CM}})$ from Equation 3.7. Knowing that this formula underestimates the production asymmetry on data, we take the maximum slope, which occurs at $\cos\theta_{\mathrm{CM}} = 0$ and amounts to $\sim 2.3\%$ per unit of $\cos\theta_{\mathrm{CM}}$. This results in a bias on $\mathcal{A}_{CP}$ of 0.006%. Being two orders of magnitude below my statistical uncertainty, this effect is completely negligible.

7.3.2 *Choice of* $\cos\theta_{\mathrm{CM}}$ *binning*

To verify that the specific choice of $\cos\theta_{\mathrm{CM}}$ binning does not impact the $\mathcal{A}_{CP}$ measurement, I repeat the whole procedure with five alternative binnings. These are reported in Table 7.1, and graphically shown in Figure 7.5. The first alternative binning (number 0) has fourteen equally-spaced bins instead of eight, while the others have eight bins with different choices of edges.

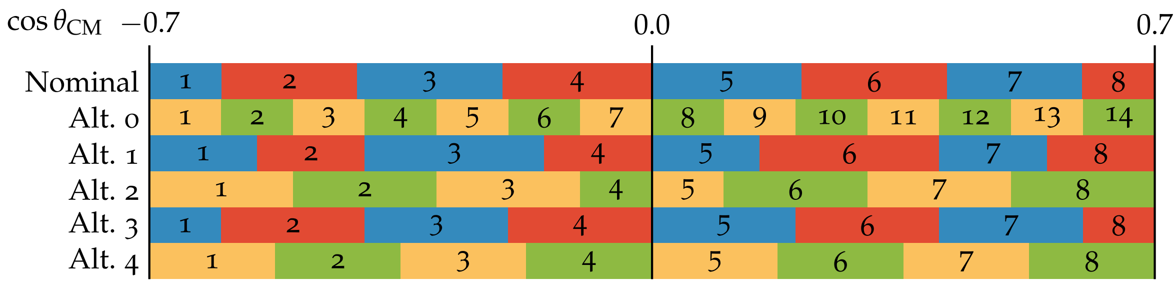


Figure 7.5: The alternative $\cos\theta_{\mathrm{CM}}$ binnings chosen to test the impact of the binning choice on the $\mathcal{A}_{CP}$ measurement, compared to the nominal binning.

Figure 7.6 reports the $\mathcal{A}_{CP}$ measurements resulting for the various alternative binnings, compared to the nominal results. The uncertainties in the figure are computed as the *difference*

Bin num.	Nominal	Alternatives 0	1	2	3	4
Bin 1	-0.7	-0.7	-0.7	-0.7	-0.7	-0.7
Bin 2	-0.599	-0.6	-0.55	-0.5	-0.6	-0.525
Bin 3	-0.411	-0.5	-0.4	-0.4	-0.4	-0.35
Bin 4	-0.208	-0.4	-0.15	-0.1	-0.2	-0.175
Bin 5	0	-0.3	0	0	0	0
Bin 6	0.208	-0.2	0.15	0.1	0.2	0.175
Bin 7	0.411	-0.1	0.4	0.3	0.4	0.35
Bin 8	0.599	0	0.55	0.5	0.6	0.525
Bin 9	0.7	0.1	0.7	0.7	0.7	0.7
Bin 10	–	0.2	–	–	–	–
Bin 11	–	0.3	–	–	–	–
Bin 12	–	0.4	–	–	–	–
Bin 13	–	0.5	–	–	–	–
Bin 14	–	0.6	–	–	–	–
	–	0.7	–	–	–	–

Table 7.1: The alternative $\cos\theta_{\mathrm{CM}}$ binnings chosen to test the impact of the binning choice on the $\mathcal{A}_{CP}$ measurement, compared to the nominal binning.

in quadrature between the uncertainties of the alternative and nominal measurements: this is because the measurements are completely correlated (they are performed on the same sample). The alternative binning results are all compatible with the nominal one.

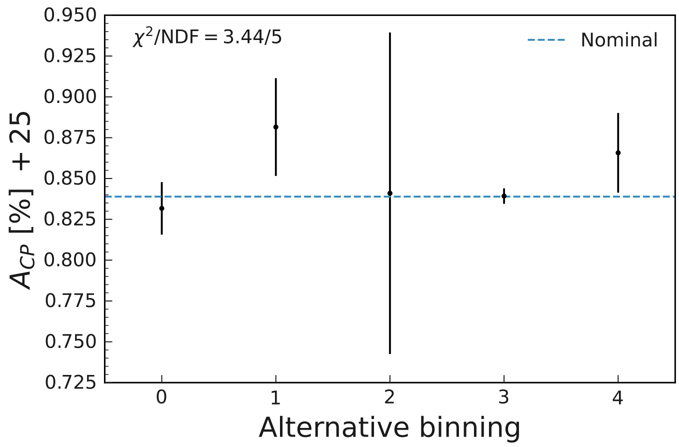


Figure 7.6: $\mathcal{A}_{CP}$ measurements with the alternative binnings (on data, closed-box) compared to the nominal result. The error bars show the difference in quadrature between the uncertainty of the result with alternative binning, and that with the nominal binning.

7.3.3 *Choice of fit models*

To estimate the impact of the specific choice of fit models on the $\mathcal{A}_{CP}$ measurement, I repeat the $D^0 \to \pi^+\pi^-\pi^0$ fit on data with the following modifications.

- Splitting all parameters by $\cos\theta_{\text{CM}}$ bin, instead of using shared parameters; the central value μ of the D^0 invariant mass probability density function (PDF) of the $D^0 \to K\pi\pi^0$ component is an exception, as the component is too small for the fit to determine this value separately in each bin.
- Splitting the following parameters by D^0 flavor, one at a time:
  - δ, γ, λ and μ of the signal ΔM PDF;
  - δ, γ, λ and μ of the signal D^0 invariant mass PDF;
  - p_0 and p_1 of the combinatorial ΔM PDF; and
  - p_1 and p_2 of the combinatorial D^0 invariant mass PDF.
- Using the following alternative PDFs for some components:
  - a Crystal Ball PDF for both the ΔM and D^0 invariant mass distributions of the signal component, with the shape parameters fixed from the fit to the signal component alone on MC;
  - a `RooDstD0Bg` PDF (defined in Section A.3) for the ΔM distribution of the combinatorial component, with x_0 fixed from the fir to the combinatorial component alone on MC; and

– a Johnson S_U PDF with $\lambda = a(M - M_0)^2 + c$ (where M is the D^0 invariant mass) for the ΔM distribution of the $D^0 \to K\pi\pi^0$ component, with γ, δ, a, c and M_0 fixed from the fit to the $D^0 \to K\pi\pi^0$ component alone on MC.

Additionally, I repeat the $\mathcal{A}_\varepsilon^{\pi\,\mathrm{tag}}$ measurement procedure splitting the γ and λ parameters of the signal component's Johnson S_U by D^0 flavor (one at a time, for both tagged and untagged samples).

The results from these tests are reported in Table 7.2. All are compatible with zero within ~ 2 standard deviations; said standard deviations are computed as the difference in quadrature between the alternative and nominal results uncertainties, since the measurements are completely correlated (they are done on the same sample).

Modification	$\Delta\mathcal{A}_{CP}$ [%]
Signal channel fit	
Split all parameters by $\cos\theta_{\mathrm{CM}}$ bin	0.00 ± 0.09
Split δ (signal ΔM) by flavor	-0.44 ± 0.17
Split γ (signal ΔM) by flavor	0.014 ± 0.007
Split λ (signal ΔM) by flavor	-0.24 ± 0.10
Split μ (signal ΔM) by flavor	0.00 ± 0.02
Split δ (signal M) by flavor	0.15 ± 0.20
Split γ (signal M) by flavor	0.03 ± 0.03
Split λ (signal M) by flavor	0.19 ± 0.09
Split μ (signal M) by flavor	0.002 ± 0.003
Split p_0 (combinatorial ΔM) by flavor	0.04 ± 0.02
Split p_1 (combinatorial ΔM) by flavor	-0.002 ± 0.002
Split p_1 (combinatorial M) by flavor	0.00 ± 0.01
Split p_2 (combinatorial M) by flavor	0.16 ± 0.12
Fit with alternative PDFs	-0.02 ± 0.05
Control channel fits	
Split γ by flavor	-0.03 ± 0.03
Split λ by flavor	0.15 ± 0.09

Table 7.2: Change in the measured $\mathcal{A}_{CP}$ due to the use of modified fit models.

7.3.4 *Asymmetry consistency in portions of the dataset that may bias the $\mathcal{A}_{CP}$*

In this section I repeat the whole $\mathcal{A}_{CP}$ measurement independently in subsets of the data sample based on criteria that may induce a bias in the $\mathcal{A}_{CP}$ measurement, and verify that:

- the measurements are compatible with each other; and

- the mean of the measurements, obtained by fitting them with a constant function, is compatible with the nominal measurement.

SPLITTING BY AZIMUTHAL ANGLE Different detector sensors are crossed by particles that travel towards different azimuthal angles ϕ: while detector effects are known to have very small dependence on this, ϕ-dependent biases can not be excluded *a priori*.

I split my sample based on the laboratory-frame ϕ angle of the D^0 meson, placing my edges at -0.57π, -0.07π, 0.43π, and 0.93π. These edges are chosen so that they occur at the overlap between the ladders of the innermost PXD layer, as shown in Figure 7.7.

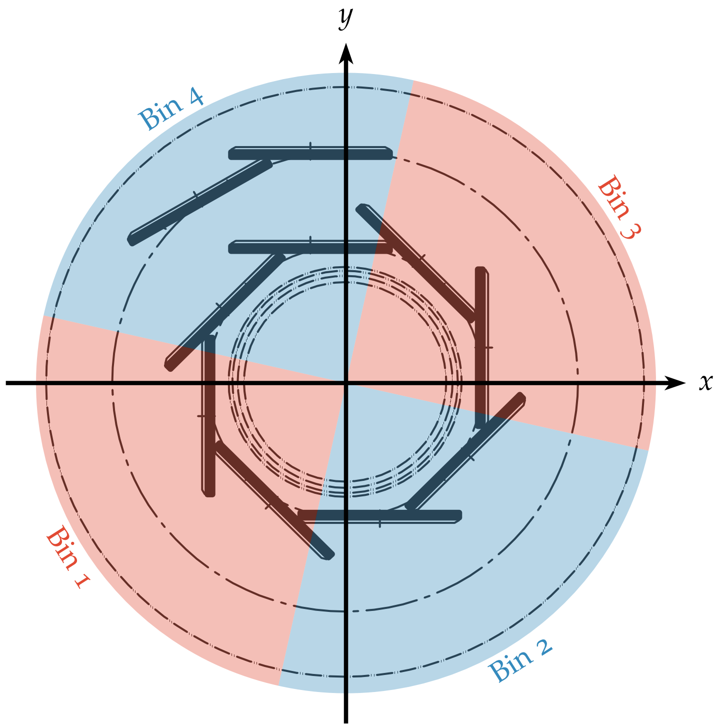


Figure 7.7: The ϕ bins for the consistency check overlaid on a schematic view of PXD.

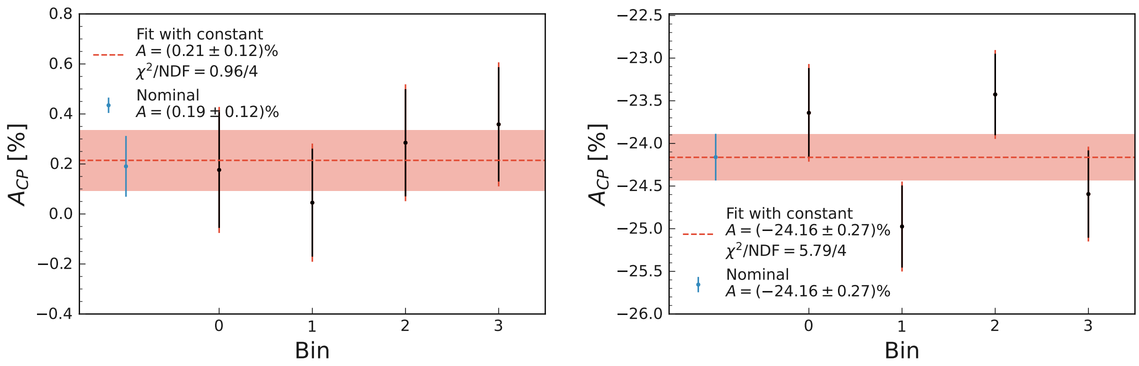
Figure 7.8 compares the resulting $\mathcal{A}_{CP}$ measurements with the nominal one. I find no deviation.

SPLITTING BY DATA-TAKING PERIODS The detector condition changes over time. This could, in principle, bias the $\mathcal{A}_{CP}$ measurement.

I split my sample in five data-taking periods:

0. from 2019 (the beginning of the data-taking) to the end of 2020;
1. from January to June 2021;
2. from October to December 2021;
3. from January to April 2022; and
4. from May to June 2022 (the beginning of long shutdown 1).

These subsamples have roughly the same population.

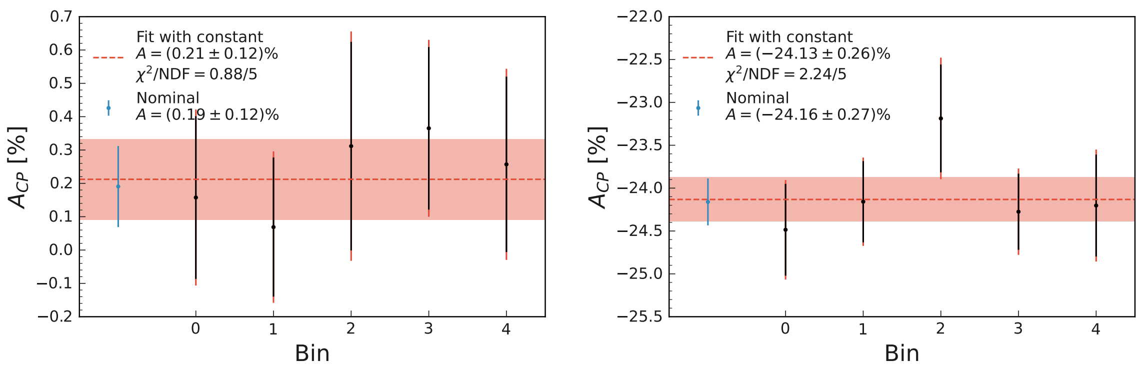


(a) Simulation.

(b) Closed-box data.

Figure 7.8: $\mathcal{A}_{CP}$ results in the four ϕ bins. The values for the four bins are fitted with a constant function (red band and dashed line), which can then be compared to the $\mathcal{A}_{CP}$ from the nominal fit (blue point with error bar).

Figure 7.9 compares the resulting $\mathcal{A}_{CP}$ measurements with the nominal one. I find no deviation.



(a) Simulation.

(b) Closed-box data.

Figure 7.9: $\mathcal{A}_{CP}$ results in the five data-taking periods. The values for the five bins are fitted with a constant function (red band and dashed line), which can then be compared to the $\mathcal{A}_{CP}$ from the nominal fit (blue point with error bar).

8 SYSTEMATIC UNCERTAINTIES EVALUATION

In this chapter I discuss the various sources of systematic uncertainty that I consider. These are summarized in Table 8.1 and described in the following sections. All systematic uncertainties are estimated before the box opening.

Source	Uncertainty
Signal channel	
D^0 reconstruction asymmetry	0.06%
Fit parameters fixed from MC	0.02%
Fit biases	0.06%
Residual mismodeling	0.03%
Control channel	
Fit biases	0.04%
Residual mismodeling	0.07%
Reweighting	0.03%
Wrong-sign $D^0 \to K\pi$ decays	0.01%
Total systematic	**0.13%**
Statistical	
Signal channel	0.25%
Control channel	0.12%
Total statistical	**0.27%**

Table 8.1: Summary of the systematic uncertainties affecting the $\mathcal{A}_{CP}$ measurement. The total systematic uncertainty is the quadrature sum of the individual components, and the statistical uncertainty is reported for reference.

8.1 D^0 RECONSTRUCTION ASYMMETRY

As introduced in Section 3.2.1, the asymmetry induced by the reconstruction of the D^0 daughters is small enough that I can neglect it, and treat it as a systematic uncertainty. I estimate it on Monte Carlo (MC) in the following way.

1. I consider all the *generated* signal decay chains, *i.e.* prompt $D^* \to [D^0 \to \pi^+\pi^-\pi^0]\pi$ decays. I compute their asymmetry, which corresponds to the production asymmetry $\mathcal{A}_{\text{prod}}$.
2. Out of all the generated signal decay chains, I consider only those for which the D^0 daughters have been correctly reconstructed, surviving my selection criteria. Note that

this is a requirement on the single pions: I do *not* require the whole decay chain to be ever reconstructed. I compute also the asymmetries of these generated decay chains, which contains both the $\mathcal{A}_{\text{prod}}$ and the D^0 reconstruction asymmetry.

3. I subtract the former asymmetry ($\mathcal{A}_{\text{prod}}$) from the latter ($\mathcal{A}_{\text{prod}} + \mathcal{A}_{\varepsilon}^{D^0}$), leaving only the D^0 reconstruction asymmetry.

I do this independently in each $\cos\theta_{\text{CM}}$ bin, then compute the average to estimate the impact on $\mathcal{A}_{CP}$. This is shown in Figure 8.1. Numerical results are reported in Table 8.2.

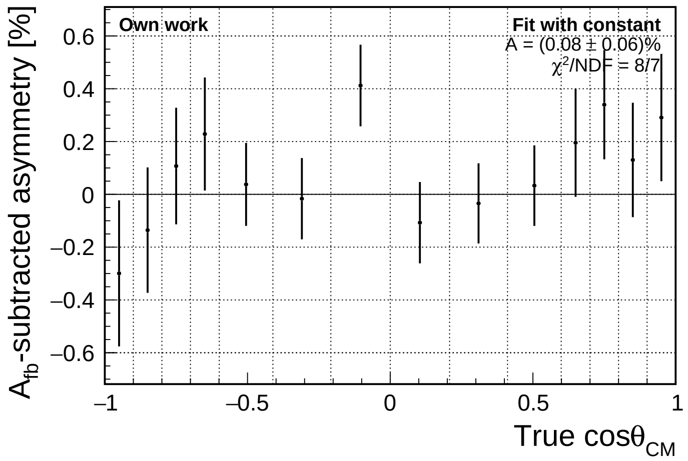


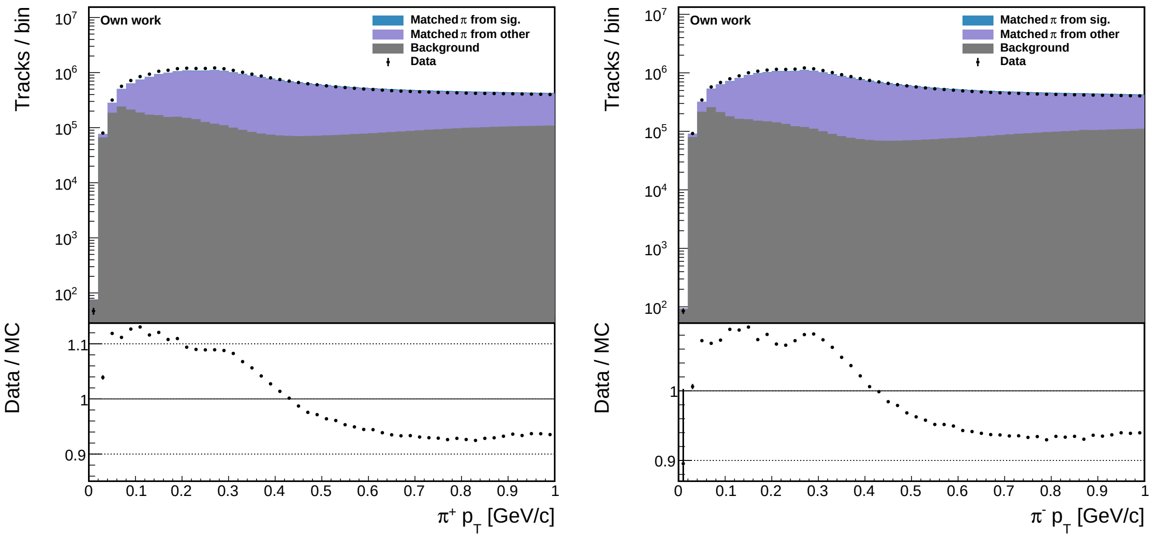
Figure 8.1: D^0 reconstruction asymmetry as a function of the $\cos\theta_{\text{CM}}$ bin on MC (1712 fb^{-1}). The plot extends beyond the signal fit range $|\cos\theta_{\text{CM}}| \leq 0.7$, but only the asymmetries within this range are considered in the fit.

$\cos\theta_{\text{CM}}$ bin	Asymmetries [%] Production	D^0 total	D^0 reconstruction
1	0.97 ± 0.12	1.20 ± 0.21	0.23 ± 0.21
2	0.81 ± 0.09	0.85 ± 0.16	0.04 ± 0.16
3	0.82 ± 0.10	0.81 ± 0.15	-0.02 ± 0.15
4	0.16 ± 0.10	0.57 ± 0.15	0.41 ± 0.15
5	-0.25 ± 0.10	-0.35 ± 0.15	-0.11 ± 0.15
6	-0.71 ± 0.10	-0.75 ± 0.15	-0.03 ± 0.15
7	-0.81 ± 0.09	-0.78 ± 0.15	0.03 ± 0.15
8	-0.94 ± 0.12	-0.74 ± 0.20	0.20 ± 0.20

Table 8.2: Production asymmetry (step 1), total D^0 asymmetry (step 2), and D^0 reconstruction asymmetry (step 3) computed on simulation for each $\cos\theta_{\text{CM}}$ bin.

As the asymmetry is compatible with zero in each bin, I fit them with a constant function to verify this hypothesis. The result, $(0.08 \pm 0.06)\%$, is compatible with zero; also, the values in each bin are compatible with a constant function ($\chi^2 = 8$ with 7 degrees of freedom).

Since this is estimated on MC, as an additional check, I compare the p_T distributions of the charged pions between MC and data (these are the variables the show the largest D^0–$\bar{D}^0$ difference, as shown in Figure 3.1). This is shown in Figure 8.2.



(a) Positive pion. (b) Negative pion.

Figure 8.2: Data/MC comparison of the p_T distributions of the charged D^0 daughters.

Since there are statistically significant data/MC differences, I repeat the asymmetry measurement after reweighting the MC sample to match these distributions. The result is $(0.10 \pm 0.06)\%$, again compatible with zero. Therefore, I assign a systematic uncertainty of 0.06% from this contribution.

8.2 FIT PARAMETERS FIXED FROM SIMULATION

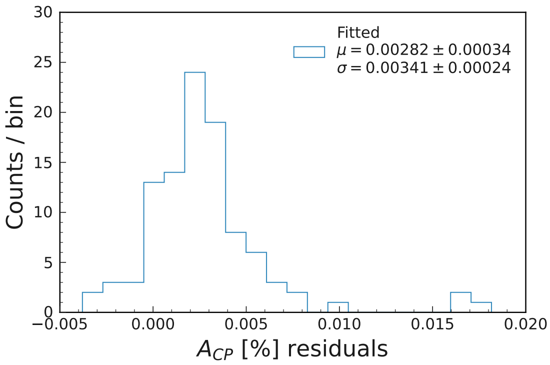
Many parameters of the signal channel fit are fixed to the values found on MC. To evaluate the impact of this choice on the $\mathcal{A}_{CP}$ measurement, I repeat the measurement on data a number of times while varying the parameters by Gaussian shifts. Then, I compute the standard deviation of the $\mathcal{A}_{CP}$ residuals (*i.e.* the differences between measurements with parameter variations and nominal measurement), and assign it as systematic uncertainty.

I do this separately for four groups of parameters.

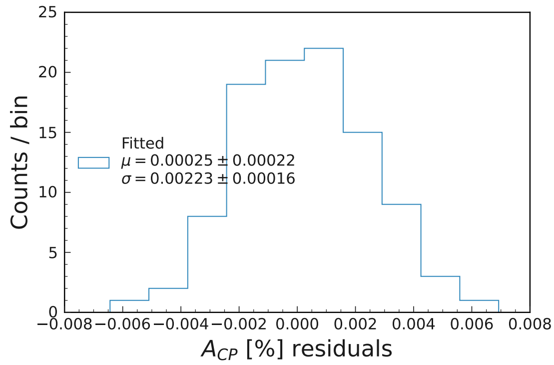
1. The random-pion component fraction: for this, I repeat the measurement 100 times, assume no correlation among the $\cos\theta_{CM}$ bins, and use 25% of the fraction values as standard deviations of the Gaussian shifts. The standard deviation of the residuals distribution is $(0.0034 \pm 0.0002)\%$.
2. The random-pion component raw asymmetry: for this, I repeat the measurement 100 times and assume no correlation among the $\cos\theta_{CM}$ bins, but I use the statistical uncertainties from MC as standard deviations of the Gaussian shifts. The standard deviation of the residuals distribution is $(0.0022 \pm 0.0002)\%$.

3. The mis-reconstructed signal component fraction: for this, I repeat the measurement 100 times, assume no correlation among the $\cos\theta_{\mathrm{CM}}$ bins, and use 25% of the fraction values as standard deviations of the Gaussian shifts. The standard deviation of the residuals distribution is $(0.0122 \pm 0.0009)\%$.
4. The 58 remaining parameters determined by fitting the single components alone on MC: for these, I repeat the measurement 2000 times, use the uncertainties from the fits as standard deviations of the Gaussian shifts, and take the correlations estimated by the fits into account when drawing the shifts. The standard deviation of the residuals distribution is $(0.0139 \pm 0.0002)\%$.

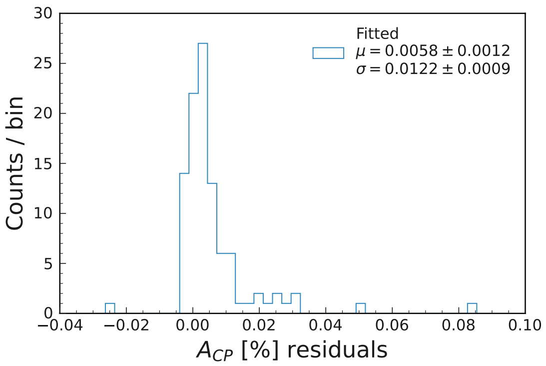
The distributions of the $\mathcal{A}_{CP}$ residuals are shown in Figure 8.3. The choice to vary the random-pion and mis-reconstructed signal fractions by 25% is rather conservative, and is derived from the 20% disagreement between data and MC signal abundance before the reweighting.



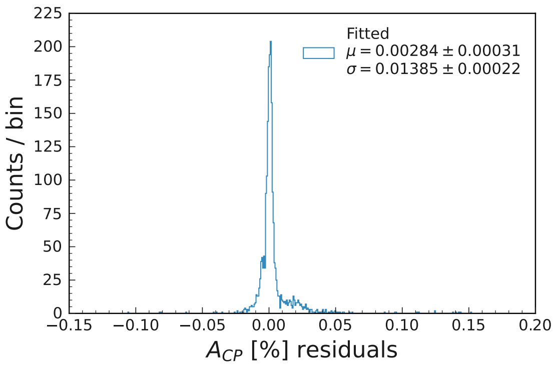
(a) Random-pion fraction.



(b) Random-pion asymmetry.



(c) Mis-reconstructed signal fraction.



(d) All other fixed parameters.

Figure 8.3: $\mathcal{A}_{CP}$ residual distributions for the variations of the fixed signal fit parameters.

The combined systematic uncertainty from all parameters determined on simulation, which I compute as the quadrature sum of the four aforementioned standard deviations, is $(0.0189 \pm 0.0006)\%$.

8.3 FIT BIASES AND NON-LINEARITIES

To check for possible biases in the fits themselves, I perform *linearity tests* based un pseudoexperiments. I treat the signal $D^0 \to \pi^+\pi^-\pi^0$ sample, tagged $D^0 \to K\pi$ sample, and untagged $D^0 \to K\pi$ sample separately and independently from each other. I produce the pseudoexperiments from the probability density function (PDF) used for the fit, with the parameters set to the values obtained when fitting the data. I draw the number of candidates to generate for each component and D^0 flavor from a Poisson distribution whose mean is given by the total yield fitted on data n^x_{tot} and the *true* raw asymmetry for the pseudoexperiment $\mathcal{A}^x_{\text{true}}$.

The true asymmetry of the signal component is always drawn from a uniform distribution between -10% and 10% (different for each pseudoexperiment). The asymmetry in each $\cos\theta_{\text{CM}}$ bin is drawn independently of the others bins.

The true asymmetries for the background components are determined differently in two series of tests.

- In the first series, they are drawn from independent uniform distributions between -10% and 10% as for the signal component: this is to evaluate the impact of different background asymmetries on the $\mathcal{A}_{CP}$ measurement.
- In the second series, they are fixed to the values fitted on data: this is to ensure that the specific values found in the data sample do not induce any particular bias.

The linearity test consists in fitting the measured-*vs.*-true signal asymmetry graph with a line:

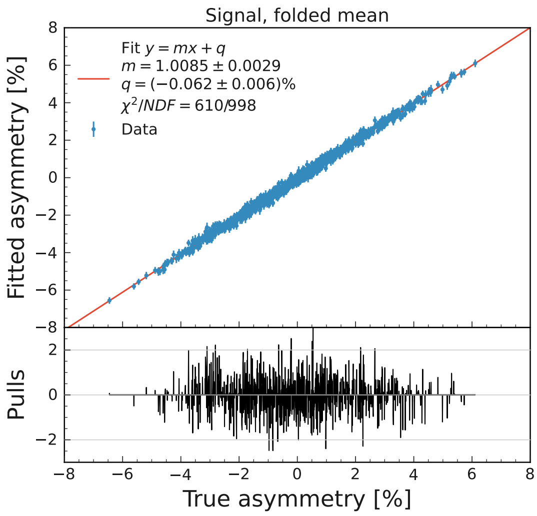
$$\mathcal{A}_{\text{fit}} = m \cdot \mathcal{A}_{\text{true}} + q \qquad \Rightarrow \qquad \text{bias} = \mathcal{A}_{\text{fit}} - \mathcal{A}_{\text{true}} = (m-1)\mathcal{A}_{\text{true}} + q\,. \tag{8.1}$$

If the fit is unbiased, the slope should be unitary, the intercept should be zero, and the graph should be consistent with a line. I test the latter condition by checking the pulls for structures or patterns, and I interpret deviations of the slope and intercept from their expected values as biases, which I use to assign systematic uncertainties.

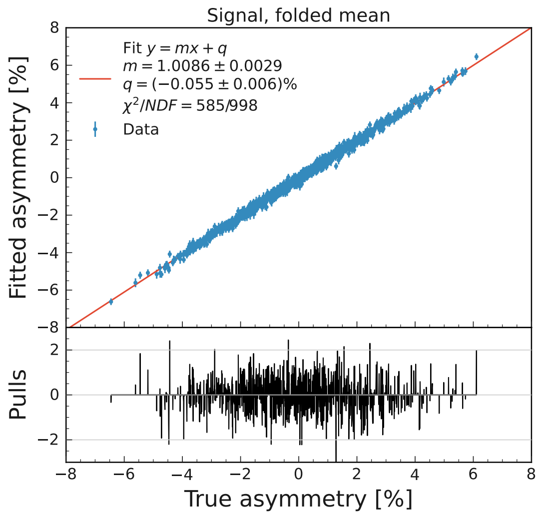
Figure 8.4 shows the results for the $D^0 \to \pi^+\pi^-\pi^0$ fit, using the "folded mean" of the asymmetries to obtain the bias on the $\mathcal{A}_{CP}$ measurement directly. The intercept is $q = (-0.062 \pm 0.006)\%$ for the first test (varying background asymmetries) and $q = (-0.055 \pm 0.006)\%$ for the second test (fixed background asymmetries). Since the values are compatible, I assign the largest as systematic uncertainty. The non-linearity, instead, produces no substantial effect: it is $m - 1 = (0.86 \pm 0.29)\%$, which has to be multiplied by the true asymmetry, which is $\sim 1\%$ on data, thus resulting in a bias $< 0.01\%$. The pulls do not present any pattern.

Figure 8.5 shows the results for the tagged $D^0 \to K\pi$ sample fit. The intercept is $q = (0.022 \pm 0.004)\%$ for the first test and $q = (0.033 \pm 0.004)\%$ for the second test. Since the values are compatible, I assign the largest as systematic uncertainty. The non-linearity is even smaller than that of the signal channel fit, and the pulls do not present any unwanted pattern.

Figure 8.6 shows the results for the untagged $D^0 \to K\pi$ sample fit. The intercept is $q = (-0.032 \pm 0.002)\%$ for the first test and $q = (-0.029 \pm 0.002)\%$ for the second test. Since the values are compatible, I assign the largest as systematic uncertainty. The non-linearity is even smaller than that of the tagged sample fit, and the pulls do not present any particular pattern.

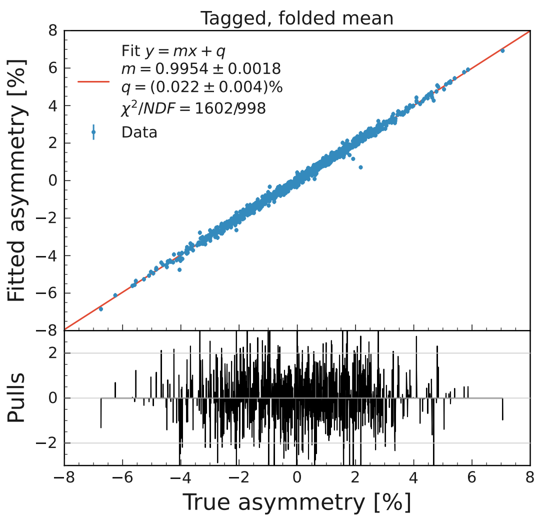


(a) Varying background asymmetries.

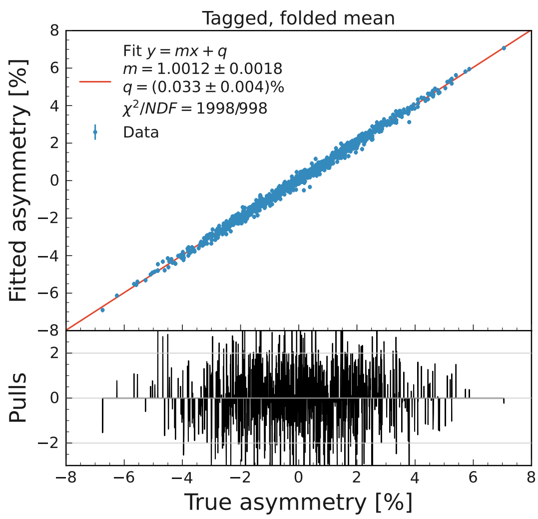


(b) Fixed background asymmetries.

Figure 8.4: Linearity tests on the "folded mean" of the raw asymmetry, $D^0 \to \pi^+\pi^-\pi^0$ fit.

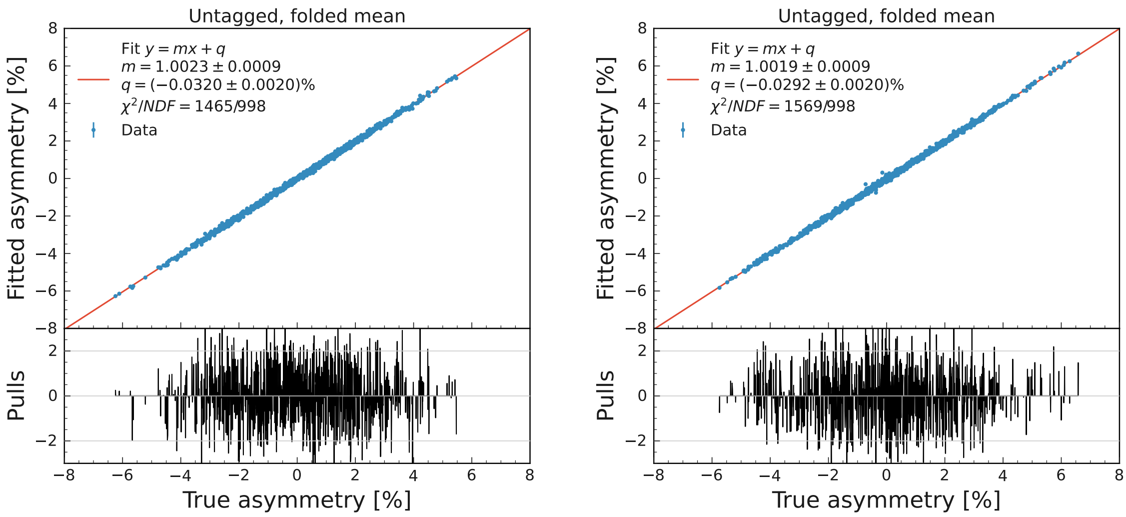


(a) Varying background asymmetries.



(b) Fixed background asymmetries.

Figure 8.5: Linearity tests on the "folded mean" of the raw asymmetry, $D^0 \to K\pi$ tagged sample fit.



(a) Varying background asymmetries.

(b) Fixed background asymmetries.

Figure 8.6: Linearity tests on the "folded mean" of the raw asymmetry, $D^0 \to K\pi$ untagged sample fit.

8.4 RESIDUAL MISMODELING

The PDFs I use do not perfectly model the actual candidate distributions. For the signal $D^0 \to \pi^+\pi^-\pi^0$ sample, the residual mismodeling is mostly due to the assumption that the 2D PDFs can be factorized into the product of two non-correlated 1D PDFs, which neglects the existing correlations between the fit variables, especially for what concerns the $D^0 \to K\pi\pi^0$ background component. For the $D^0 \to K\pi$ samples, instead, it is mostly due to the difficulty of finding an empirical parametrization that fully captures the resolution of the signal peaks for such large samples. Note that the mismodeling I find on data is well reproduced by the MC simulation.

To evaluate the impact of this mismodeling on the $\mathcal{A}_{CP}$ measurement, I repeat the whole procedure 1000 times for different bootstrap samples derived from the MC and data samples. The bootstrap samples are as large as the data, with Poisson-fluctuating number of candidates. To preserve the correlation between tagged and untagged $D^0 \to K\pi$ samples, I produce their bootstraps at the same time and ensure that candidates from selected common events appear in both bootstrap samples.

I measure the raw asymmetries, $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$ ($D^0 \to K\pi$ samples), and $\mathcal{A}_{CP}$ (signal sample) for each bootstrap. Then, I define *residuals* and *pulls* as

$$\text{residual}_i = \mathcal{A}_i - \mathcal{A}_{\text{true}}, \qquad \text{pull}_i = \frac{\mathcal{A}_i - \mathcal{A}_{\text{true}}}{\sigma_{\mathcal{A},i}}, \tag{8.2}$$

where $\mathcal{A}_i$ is the asymmetry measured on the bootstrap sample i, $\sigma_{\mathcal{A},i}$ is the uncertainty estimated for said asymmetry measurement, and $\mathcal{A}_{\text{true}}$ is the true asymmetry computed on the original sample from which the bootstrap sample is derived. When working with data samples, $\mathcal{A}_{\text{true}}$ is replaced by the asymmetry measured on the original sample, of course limiting the checks that can be performed.

The residual distributions should have zero mean, showing that the measurements are unbiased. This, of course, is more meaningful on MC, where the true value is used as reference, but also works as a consistency check on data. The pull distributions should be Gaussian with zero mean and unitary standard deviation, showing that the uncertainties are correctly estimated. Finally, I also check the correlation of the asymmetries among the $\cos\theta_{\mathrm{CM}}$ bins to confirm that they are negligible as expected ($D^0 \to K\pi$ samples) or estimated by the fits ($D^0 \to \pi^+\pi^-\pi^0$ sample).

$D^0 \to K\pi$ SAMPLES In Figure 8.7 I show the residual and pull distributions for the $\mathcal{A}_{\varepsilon}^{\pi\,\mathrm{tag}}$ on MC for one $\cos\theta_{\mathrm{CM}}$ bin. The former peaks around zero with no unexpected structure, and the latter is compatible with a Gaussian distribution.

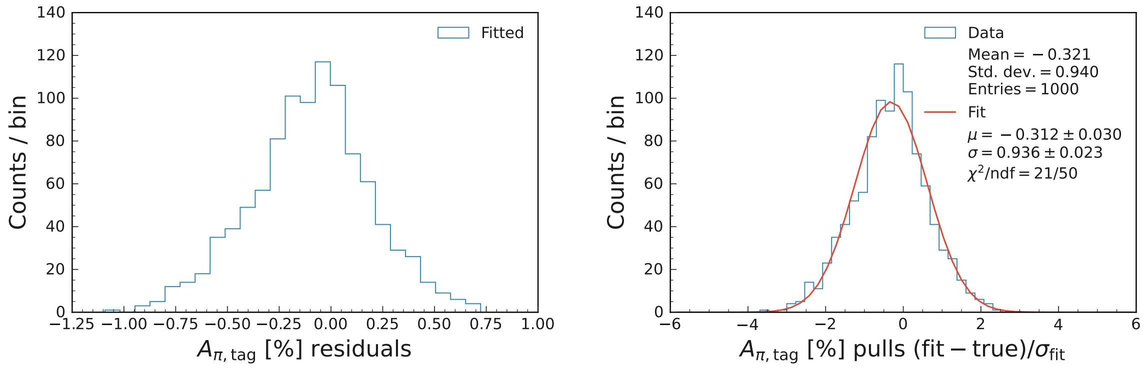


Figure 8.7: $\mathcal{A}_{\varepsilon}^{\pi\,\mathrm{tag}}$ residuals and pulls on MC for bin 4, *i.e.* $\cos\theta_{\mathrm{CM}}$ from -0.208 to 0. The pull distribution is fitted with a Gaussian PDF.

Figure 8.8 reports the means of the residual distributions and the standard deviations of the pull distributions of the various $\cos\theta_{\mathrm{CM}}$ bins on MC. The means of the residuals in the single bins show deviations from zero; however, it is the mean of this biases, computed as prescribed by the “folding” method, that estimates the impact on the $\mathcal{A}_{CP}$ measurement: $(-0.071 \pm 0.004)\%$. Since this is significantly different from zero, I assign the full extent of its central value as systematic uncertainty. The standard deviations of the pulls are all compatible with one, proving that the uncertainties are correctly estimated.

Figure 8.9 reports the standard deviations of the pull distributions of the various $\cos\theta_{\mathrm{CM}}$ bins on data. Again, all are compatible with one, proving that the uncertainties are correctly estimated.

Similar results are found for the $\mathcal{A}_{\mathrm{raw}}^{\mathrm{tagged}}$ and $\mathcal{A}_{\mathrm{raw}}^{\mathrm{untagged}}$. Finally, the correlations among the $\cos\theta_{\mathrm{CM}}$ bins are found to be compatible with zero, as expected.

$D^0 \to \pi^+\pi^-\pi^0$ SAMPLE In Figure 8.10 I show the residual and pull distributions for the $\mathcal{A}_{\varepsilon}^{\pi\,\mathrm{tag}}$-corrected signal $\mathcal{A}_{\mathrm{raw}}$ on MC for one $\cos\theta_{\mathrm{CM}}$ bin. The former peaks around zero with no unexpected structure, and the latter is compatible with a Gaussian distribution.

Figure 8.11 reports the means of the residual distributions and the standard deviations of the pull distributions of the various $\cos\theta_{\mathrm{CM}}$ bins on MC. While the means of the residuals in the single bins show some deviations from zero, the mean of this biases, computed as prescribed by the “folding” method, is compatible with zero: $(-0.003 \pm 0.008)\%$. Note that this mean is

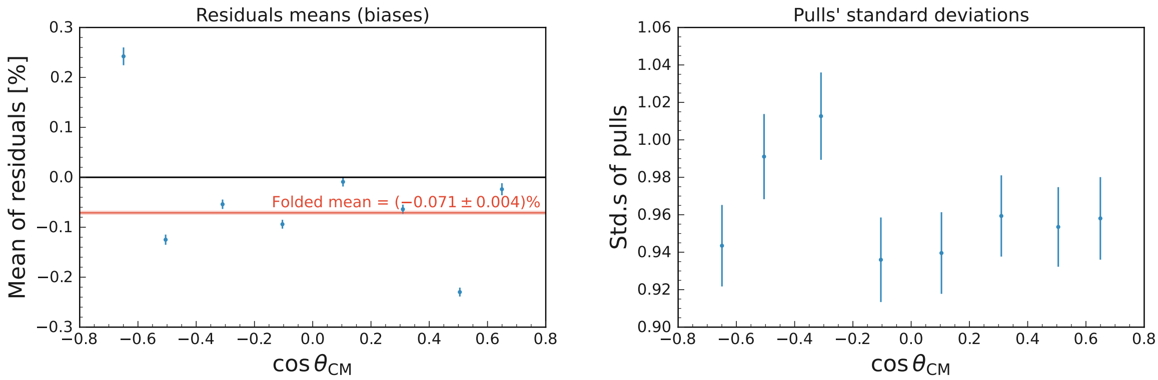


Figure 8.8: $\mathcal{A}_{\varepsilon}^{\pi\,\mathrm{tag}}$ residual means and pull standard deviations on MC as a function of the $\cos\theta_{\mathrm{CM}}$ bin.

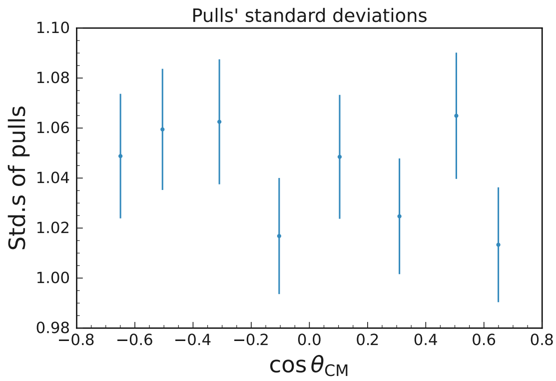


Figure 8.9: $\mathcal{A}_{\varepsilon}^{\pi\,\mathrm{tag}}$ pull standard deviations on data as a function of the $\cos\theta_{\mathrm{CM}}$ bin.

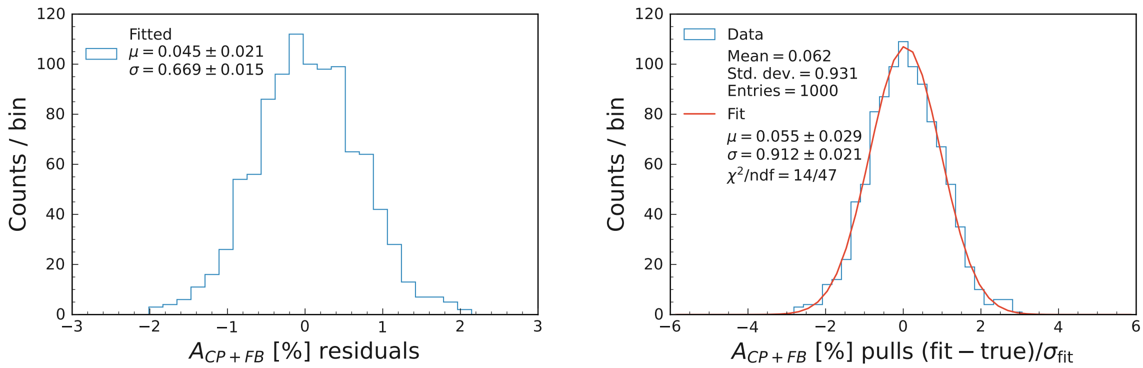


Figure 8.10: $\mathcal{A}_{\varepsilon}^{\pi\,\mathrm{tag}}$-corrected signal $\mathcal{A}_{\mathrm{raw}}$ residuals and pulls on MC for bin 4, *i.e.* $\cos\theta_{\mathrm{CM}}$ from -0.208 to 0. The pull distribution is fitted with a Gaussian PDF.

the bias on the final $\mathcal{A}_{CP}$ measurement. The standard deviations of the pulls are all compatible with one, proving that the uncertainties are correctly estimated.

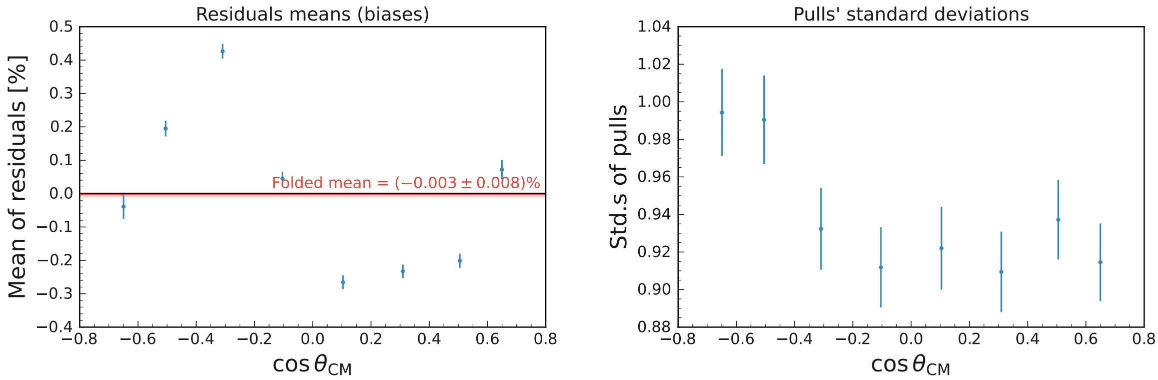


Figure 8.11: $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$-corrected signal $\mathcal{A}_{\text{raw}}$ residual means and pull standard deviations on MC as a function of the $\cos\theta_{\text{CM}}$ bin.

Figure 8.12 reports the standard deviations of the pull distributions of the various $\cos\theta_{\text{CM}}$ bins on data. Only a slight deviation from one, by about 10%, can be seen.

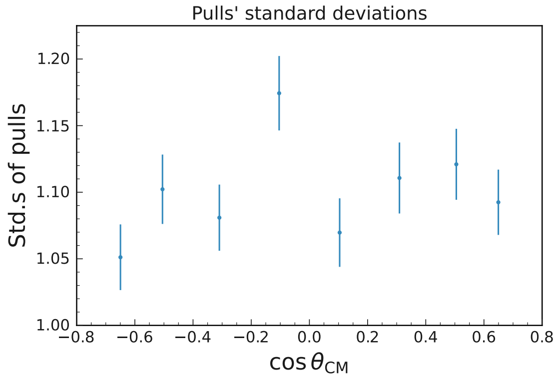
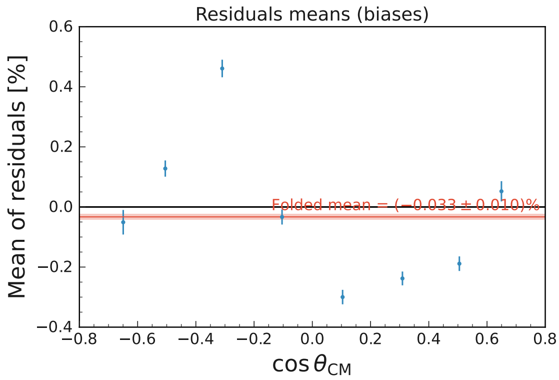


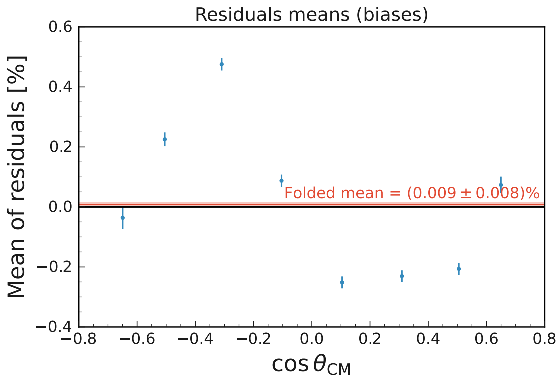
Figure 8.12: $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$-corrected signal $\mathcal{A}_{\text{raw}}$ pull standard deviations on data as a function of the $\cos\theta_{\text{CM}}$ bin.

The same results are found on the $\mathcal{A}_{\text{raw}}$ and folded means ($\mathcal{A}_{CP}$ measurements). Finally, the correlations among the $\cos\theta_{\text{CM}}$ bins are found to be compatible with zero, as expected.

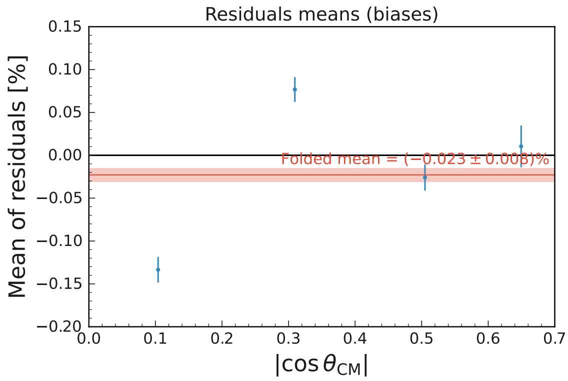
$D^0 \to \pi^+\pi^-\pi^0$ WITH MODIFIED ASYMMETRY As an additional check, I repeat the bootstrap tests for the signal channel on MC, but artificially reweighting the signal component so that its raw asymmetry is twice (Figure 8.13a) or half (Figure 8.13b) as much as the original value in each $\cos\theta_{\text{CM}}$ bin, or so that the true $\mathcal{A}_{CP}$ is $\pm 1\%$ (Figure 8.13c and Figure 8.13d). This test is meant to ensure that the bias does not depend on the specific asymmetry of the MC. The largest bias I get is $(-0.033 \pm 0.010)\%$, and I use the absolute value of its central value as systematic uncertainty.



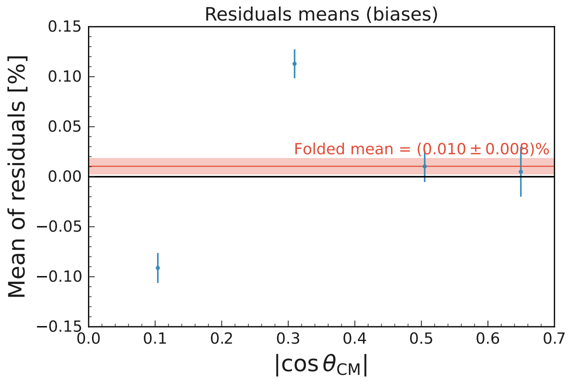
(a) Signal raw asymmetry $\times 2$.



(b) Signal raw asymmetry $\times 0.5$.



(c) True $\mathcal{A}_{CP}$ set to $+1\%$.



(d) True $\mathcal{A}_{CP}$ set to -1%.

Figure 8.13: $\mathcal{A}_{\varepsilon}^{\pi\,\mathrm{tag}}$-corrected signal $\mathcal{A}_{\mathrm{raw}}$ residuals means on MC, with modified signal asymmetry, as a function of the $\cos\theta_{\mathrm{CM}}$ bin.

8.5 CONTROL CHANNEL REWEIGHTING

To evaluate the impact of the reweighting procedure used for the $D^0 \to K\pi$ samples on the final $\mathcal{A}_{CP}$ measurement, I repeat the $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$ measurement without using weights. This produces a difference of $(-0.058 \pm 0.010)\%$ in the final $\mathcal{A}_{CP}$. Since the reweighting procedure produces a significant improvement in the agreement between original and target samples, I conservatively assign half of this variation, *i.e.* 0.03%, for this contribution. This also covers any possible effect from the signal distribution extraction with the $_s\mathcal{P}lot$ technique.

Also, I repeat the measurement of $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$ 1000 times while randomly varying the values of the weights by Gaussian shifts, using the statical uncertainties of the weights as standard deviations. Figure 8.14 shows the distribution of the resulting $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$ measurements, compared to the nominal one. The mean of the distribution is compatible with the nominal value, and the standard deviation is $\sim 0.005\%$. Therefore, this second result does not significantly contribute to the systematic uncertainty.

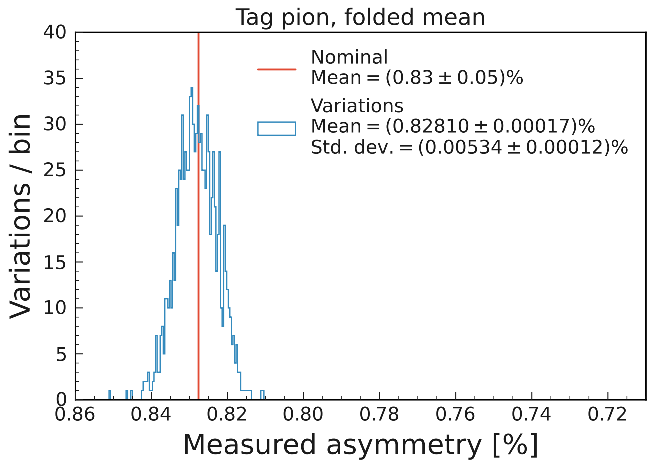


Figure 8.14: Distribution of the $\mathcal{A}_\varepsilon^{\pi\,\text{tag}}$ computed while varying the $D^0 \to K\pi$ samples weights, compared to the nominal value. This test is performed on MC.

8.6 WRONG-SIGN $D^0 \to K\pi$ DECAYS

When fitting the untagged $D^0 \to K\pi$ sample, I neglect the "wrong-sign" doubly-Cabibbo-suppressed component. As this component naturally mimics the signal, it is captured by the signal PDF of the fit. As the incorrect D^0 flavor is naturally deduced, it causes the maximum possible dilution of the signal asymmetry.

Let n_true and $\overline{n}_\text{true}$ be the *actual* numbers of D^0 and $\overline{D}^0$ mesons that decay into $K\pi$ (either right or wrong sign), and let f be the fraction that decays "wrong-sign", which is $(0.379 \pm 0.018)\%$ [15]. Then, the numbers of *observed* D^0 and $\overline{D}^0$ mesons are

$$n_\text{obs} = (1-f) \cdot n_\text{true} + f \cdot \overline{n}_\text{true}, \qquad \overline{n}_\text{obs} = (1-f) \cdot \overline{n}_\text{true} + f \cdot n_\text{true}\,, \tag{8.3}$$

and the observed asymmetry is

$$\mathcal{A}_{\text{obs}} = \frac{(1-f)n_{\text{true}} + f\overline{n}_{\text{true}} - (1-f)\overline{n}_{\text{true}} + fn_{\text{true}}}{n_{\text{true}} + \overline{n}_{\text{true}}} \tag{8.4a}$$

$$= (1-2f)\frac{n_{\text{true}} - \overline{n}_{\text{true}}}{n_{\text{true}} + \overline{n}_{\text{true}}} = (1-2f)\mathcal{A}_{\text{true}}\,. \tag{8.4b}$$

Therefore, the impact of neglecting the wrong-sign decays is a correction factor $1-2f = 0.9924 \pm 0.0004$ on the measured asymmetry. This translates to an absolute bias on $\mathcal{A}_{CP}$ of $(0.007\,56 \pm 0.000\,03)\%$.

In principle I should take into account the possibility that the fraction f is different for the two flavors due to reconstruction effects. However, I measure in MC that the two values are compatible with each other and with the nominal expected value: $(0.382 \pm 0.002)\%$ and $(0.377 \pm 0.002)\%$.

9 CONCLUSIONS AND FUTURE PROSPECTS

In this thesis I presented a measurement of the time-integrated *CP* asymmetry in the $D^0 \to \pi^+\pi^-\pi^0$ decay channel, which I performed using the dataset of e^+e^- collisions collected by the Belle II experiment between 2019 and 2022, corresponding to $428\,\mathrm{fb}^{-1}$. For this channel, the Standard Model (SM) predicts an asymmetry of $\mathcal{O}(10^{-4}$–$10^{-3})$, with large theoretical uncertainties. Beyond-SM contributions may enhance the asymmetry. Therefore, the motivation for this measurement lays both in testing the SM, and in exploring possible new physics contributions, with particular regard to those models in which the *up* quark sector has a special role.

The result I find,

$$\mathcal{A}_{CP}(D^0 \to \pi^+\pi^-\pi^0) = (0.29 \pm 0.27(\mathrm{stat}) \pm 0.13(\mathrm{sys}))\%, \tag{9.1}$$

is compatible with conservation of the *CP* symmetry in this decay, and with the previous measurements [41, 42]. This measurement improves the precision of the current world's best [41] by 34% in terms of the dominant statistical uncertainty, and also by 24% in terms of the systematic uncertainty, despite an increase in the available integrated luminosity of only about 10% ($428\,\mathrm{fb}^{-1}$ *vs.* $385\,\mathrm{fb}^{-1}$). The improvement in precision per unit luminosity was made possible by the superior performance of the Belle II detector, especially for what concerns neutral pion reconstruction, and by the choice to adopt more relaxed selection criteria and use more advanced fitting procedures to measure the raw asymmetries.

At the time of writing this document, this result is considered *preliminary* by the Belle II collaboration. The bulk of the internal approval process has already taken place, proving the robustness of the analysis and allowing me to present this open-box result in this document and at the EPS-HEP 2025 conference. The remaining part of the process involves one more collaboration-wide review step based on the publication draft, which I am actively working on. Barring unexpected issues, I expect this measurement to be published as a journal paper later this year.

This measurement falls within a series of *CP* violation measurements in the charm sector already published by the Belle II experiment, and several others that will be published in the near future. These include the $\mathcal{A}_{CP}$ measurements in $D^0 \to \pi^0\pi^0$ and $D^+ \to \pi^+\pi^0$ decays [8, 9], which help constrain the origin of the *CP* violation observed in $D^0 \to \pi^+\pi^-$ decays [4, 7]; the $\mathcal{A}_{CP}$ measurements in $D^0 \to K^0_S K^0_S$ decays [31, 32], the latter of which uses an innovative flavor tagging technique based on the non-signal particles in the event; the first model-independent measurement of the D^0–$\overline{D}^0$ mixing parameters at a *B*-factory experiments, using $D^0 \to K^0_S\pi^+\pi^-$ decays [33]; and several *CP* asymmetry measurements in charmed hadron decays. Many of these are world-leading results, especially for what concerns channels with neutral particles in the final state, for which the Belle II detector and its collision environment are particularly suited. And, many of these are currently limited by the size of the available dataset, and thus their precision will naturally improve in the future.

Concerning $D^0 \to \pi^+\pi^-\pi^0$ decays, future improvements and extensions of this analysis are possible and envisioned. First, this measurement is limited by its statistical uncertainty, therefore its precision will naturally improve as the Belle II experiment accumulates more integrated luminosity. Second, the systematic uncertainties may be improved once they become limiting:

- I estimated the D^0 reconstruction asymmetry using the simulated sample, therefore its uncertainty will shrink as larger samples are produced;
- I used a very conservative estimate for the systematic uncertainty due the control sample reweighting, mainly because of its small magnitude compared to the statistical uncertainty; and
- the distribution modeling may be improved in the future, if necessary, exploiting the experience accumulated with this work.

Last, more information may be extracted from this three-body channel by exploiting its Dalitz plot distribution [40]: while the timeline was beyond the boundaries of this work, feasibility and preliminary studies for Dalitz-plot dependent $\mathcal{A}_{CP}$ measurements and for amplitude analyses are already taking place.

In conclusion, I measured the time-integrated *CP* asymmetry of the $D^0 \to \pi^+\pi^-\pi^0$ decay using the Belle II dataset collected between 2019 and 2022. I found no evidence of *CP* violation, and improved the current world's best precision by 34% despite having only 10% more integrated luminosity. The preliminary result is yet to be definitively approved by the Belle II collaboration, and is expected to lead to a journal publication later this year. Finally, even better precision can be expected by repeating this measurement in the future, and more information may be gained by exploiting the Dalitz plot distribution of the decay, using this work as a preliminary study towards the optimal analysis strategy.

ACKNOWLEDGMENTS

I would like to thank all the people without whom this work would not have been possible.

First, I would like to thank my supervisors, Francesco Forti and Giulia Casarosa, for guiding and supporting me through this work, for helping me pursue my interests, and for showing me new ones.

I would like to thank all the people in the Belle II Pisa group, and especially Alice, Giuliana, Stefano, Luigi, Foteini, Alessandro, Guglielmo, Surya, and Eugenio, for the help, the chances to learn, and most of all for the time spent together.

I would like to thank the people of the *charm*ing Belle II working group and my reviwers Michel Bertemes, Angelo di Canto, Valerio Bertacchi and Michele Veronesi, for helping me strengthen this work and make it into what it is now.

I would like to thank the many Belle II people for the wonderful time, at KEK and elsewhere, these years: Cecilia, Laura, Simone, Giovanni, Alessandro, Federico, Yannik, Marike, Max, . . .

And finally, I would like to thank my Mom and Dad, for everything.

BIBLIOGRAPHY

[1] 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.

[2] B. Aubert et al. (*BABAR* Collaboration). "Observation of *CP* Violation in the B^0 Meson System." In: *Phys. Rev. Lett.* 87.9 (2001), p. 091801. DOI: 10.1103/PhysRevLett.87.091801.

[3] A. Abulencia et al. (CDF Collaboration). "Observation of $B_s^0 - \overline{B}_s^0$ Oscillations." In: *Phys. Rev. Lett.* 97.24 (2006), p. 242003. DOI: 10.1103/PhysRevLett.97.242003.

[4] R. Aaij et al. (LHCb Collaboration). "Observation of *CP* Violation in Charm Decays." In: *Phys. Rev. Lett.* 122 (May 2019), p. 211803. DOI: 10.1103/PhysRevLett.122.211803. arXiv: 1903.08726.

[5] B. Aubert et al. (*BABAR* Collaboration). "Evidence for $D^0{-}\overline{D}^0$ Mixing." In: *Phys. Rev. Lett.* 98.21 (May 2007), p. 211802. DOI: 10.1103/PhysRevLett.98.211802.

[6] M. Starič et al. (Belle Collaboration). "Evidence for $D^0{-}\overline{D}^0$ Mixing." In: *Phys. Rev. Lett.* 98.21 (May 2007), p. 211803. DOI: 10.1103/PhysRevLett.98.211803.

[7] R. Aaij et al. (LHCb Collaboration). "Measurement of the Time-Integrated *CP* Asymmetry in $D^0 \to K^-K^+$ Decays." In: *Phys. Rev. Lett.* 131.9 (Aug. 2023), p. 091802. DOI: 10.1103/PhysRevLett.131.091802. arXiv: 2209.03179.

[8] I. Adachi et al. (Belle II Collaboration). "Measurement of the time-integrated *CP* asymmetry in $D^0 \to \pi^0\pi^0$ decays at Belle II." In: *Phys. Rev. D* 112.1 (July 2025), p. 012006. DOI: 10.1103/vg9c-xvdc. arXiv: 2505.02912 [hep-ex].

[9] I. Adachi et al. (Belle II Collaboration). "Measurement of the *CP* asymmetry in $D^+ \to \pi^+\pi^0$ decays at Belle II." In: *Phys. Rev. D* 112.3 (Aug. 2025), p. L031101. DOI: 10.1103/3kqk-ldm4. arXiv: 2506.07879 [hep-ex].

[10] M. Thomson. *Modern Particle Physics*. Cambridge University Press, 2013.

[11] *Status of Higgs Boson Physics (rev.)* In: [15]. URL: https://pdg.lbl.gov/2024/reviews/rpp2024-rev-higgs-boson.pdf.

[12] *CKM Quark-Mixing Matrix (rev.)* In: [15]. URL: https://pdg.lbl.gov/2024/reviews/rpp2024-rev-ckm-matrix.pdf.

[13] L. Wolfenstein. "Parametrization of the Kobayashi-Maskawa Matrix." In: *Phys. Rev. Lett.* 51.21 (Nov. 1983), pp. 1945–1947. DOI: 10.1103/PhysRevLett.51.1945.

[14] I. I. Bigi and A. I. Sanda. In: *CP Violation*. 2nd ed. Cambridge University Press, 2009. Chap. 8.4, p. 154. DOI: 10.1017/CBO9780511581014.

[15] S. Navas et al. (Particle Data Group). "Review of particle physics." In: *Phys. Rev. D* 110.3 (2024), p. 030001. DOI: 10.1103/PhysRevD.110.030001.

[16] J. Charles et al. (CKMfitter Group). "*CP* violation and the CKM matrix: assessing the impact of the asymmetric *B* factories." In: *Eur. Phys. J. C* 41 (2005). Updated results and plots available at: http://ckmfitter.in2p3.fr/, pp. 1–131. DOI: 10.1140/epjc/s2005-02169-1. arXiv: hep-ph/0406184.

[17] M. Bona et al. (UTfit Collaboration). "New UTfit Analysis of the Unitarity Triangle in the Cabibbo-Kobayashi-Maskawa scheme." In: *Rend. Fis. Acc. Lincei* 34 (2023). Updated results and plots available at: http://www.utfit.org/, pp. 37–57. DOI: 10.1007/s12210-023-01137-5. arXiv: 2212.03894.

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

[19] R. L. Garwin, L. M. Lederman, and M. Weinrich. "Observations of the Failure of Conservation of Parity and Charge Conjugation in Meson Decays: the Magnetic Moment of the Free Muon." In: *Phys. Rev.* 105.4 (Feb. 1957), pp. 1415–1417. DOI: 10.1103/PhysRev.105.1415.

[20] J. I. Friedman and V. L. Telegdi. "Nuclear Emulsion Evidence for Parity Nonconservation in the Decay Chain $\pi^+ \to \mu^+ \to e^+$." In: *Phys. Rev.* 106.6 (June 1957), pp. 1290–1293. DOI: 10.1103/PhysRev.106.1290.

[21] G. Culligan, S. G. F. Frank, and J. R. Holt. "Longitudinal Polarization of the Electrons from the Decay of Unpolarized Positive and Negative Muons." In: *Proc. Physical Society* 73.2 (Feb. 1959), p. 169. DOI: 10.1088/0370-1328/73/2/303.

[22] L. Landau. "On the conservation laws for weak interactions." In: *Nucl. Phys.* 3.1 (1957), pp. 127–131. DOI: 10.1016/0029-5582(57)90061-5.

[23] A. D. Sakharov. "Violation of *CP* invariance, *C* asymmetry, and baryon asymmetry of the universe." In: *Zh. Eksp. Teor. Fiz. (JETP)* 5 (1967), pp. 32–35. DOI: 10.1070/PU1991v034n05ABEH002497.

[24] *CP Violation in the Quark Sector (rev.)* In: [15]. URL: https://pdg.lbl.gov/2024/reviews/rpp2024-rev-cp-violation.pdf.

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

[26] H.-Y. Cheng and C.-W. Chiang. "Revisiting *CP* violation in $D \to PP$ and VP decays." In: *Phys. Rev. D* 100.9 (2019), p. 093002. DOI: 10.1103/PhysRevD.100.093002.

[27] S. Schacht and A. Soni. "Enhancement of charm *CP* violation due to nearby resonances." In: *Phys. Lett. B* 825 (2022), p. 136855. DOI: 10.1016/j.physletb.2021.136855.

[28] I. Bediaga, T. Frederico, and P. C. Magalhães. "Enhanced Charm *CP* Asymmetries from Final State Interactions." In: *Phys. Rev. Lett.* 131.5 (2023), p. 051802. DOI: 10.1103/PhysRevLett.131.051802.

[29] M. Gavrilova, Y. Grossman, and S. Schacht. "Determination of the $D \to \pi\pi$ ratio of penguin over tree diagrams." In: *Phys. Rev. D* 109.3 (2024), p. 033011. DOI: 10.1103/PhysRevD.109.033011.

[30] Y. Grossman, A. L. Kagan, and J. Zupan. "Testing for new physics in singly Cabibbo suppressed D decays." In: *Phys. Rev. D* 85.11 (June 2012), p. 114036. DOI: 10.1103/PhysRevD.85.114036.

[31] I. Adachi et al. (Belle and Belle II Collaborations). "Measurement of the time-integrated CP asymmetry in $D^0 \to K^0_S K^0_S$ decays using Belle and Belle II data." In: *Phys. Rev. D* 111.1 (2025), p. 012015. DOI: 10.1103/PhysRevD.111.012015.

[32] I. Adachi et al. (Belle and Belle II Collaborations). "Measurement of the time-integrated CP asymmetry in $D^0 \to K^0_S K^0_S$ decays using opposite-side flavor tagging at Belle and Belle II." In: *Phys. Rev. D* 112.1 (July 2025), p. 012017. DOI: 10.1103/8x1h-39dp. arXiv: 2504.15881 [hep-ex].

[33] I. Adachi et al. (Belle and Belle II Collaborations). "Model-independent measurement of $D^0 - \overline{D}^0$ mixing parameters in $D^0 \to K^0_S \pi^+ \pi^-$ decays at Belle and Belle II." In: *Phys. Rev. D* 111.11 (2025), p. 112011. DOI: 10.1103/7p5s-f9qp.

[34] Y. Grossman, A. L. Kagan, and Y. Nir. "New physics and CP violation in singly Cabibbo suppressed D decays." In: *Phys. Rev. D* 75 (Feb. 2007), p. 036008. DOI: 10.1103/PhysRevD.75.036008. arXiv: hep-ph/0609178.

[35] Alexander L. Kagan and Luca Silvestrini. "Dispersive and absorptive CP violation in $D^0 - \overline{D}^0$ mixing." In: *Phys. Rev. D* 103.5 (2021), p. 053008. DOI: 10.1103/PhysRevD.103.053008.

[36] I. I. Bigi and A. I. Sanda. In: *CP Violation*. 2nd ed. Cambridge University Press, 2009. Chap. 14, p. 286. DOI: 10.1017/CBO9780511581014.

[37] I. I. Bigi and A. I. Sanda. In: *CP Violation*. 2nd ed. Cambridge University Press, 2009. Chap. 14.3, p. 296. DOI: 10.1017/CBO9780511581014.

[38] Mikael Chala et al. "$\Delta\mathcal{A}_{CP}$ within the Standard Model and beyond." In: *JHEP* 2019.7 (2019), p. 161. DOI: 10.1007/JHEP07(2019)161.

[39] Aubert, B. and others (*BABAR* Collaboration). "Measurement of CP Violation Parameters with a Dalitz Plot Analysis of $B^\pm \to D_{\pi^+\pi^-\pi^0} K^\pm$." In: *Phys. Rev. Lett.* 99 (Dec. 2007), p. 251801. DOI: 10.1103/PhysRevLett.99.251801. arXiv: hep-ex/0703037.

[40] Avital Dery et al. "Probing the $\Delta U = 0$ rule in three body charm decays." In: *JHEP* 2021.5 (2021), p. 179. DOI: 10.1007/JHEP05(2021)179. arXiv: 2101.02560.

[41] B. Aubert et al. (*BABAR* Collaboration). "Search for CP violation in neutral D meson Cabibbo-suppressed three-body decays." In: *Phys. Rev. D* 78 (Sept. 2008), p. 051102. DOI: 10.1103/PhysRevD.78.051102. arXiv: 0802.4035.

[42] K. Arinstein et al. (*BABAR* Collaboration). "Measurement of the ratio $B(D^0 \to \pi^+\pi^-\pi^0) / B(D^0 \to K^-\pi^+\pi^0)$ and the time-integrated CP asymmetry in $D^0 \to \pi^+\pi^-\pi^0$." In: *Phys. Lett. B* 662.2 (2008), pp. 102–110. DOI: 10.1016/j.physletb.2008.02.054. arXiv: 0801.2439.

[43] R. Aaij et al. (LHCb Collaboration). "Search for CP violation in the phase space of $D^0 \to \pi^+\pi^-\pi^0$ decays with the energy test." In: *JHEP* 2023.9 (2023), p. 129. DOI: 10.1007/JHEP09(2023)129. arXiv: 2306.12746.

[44] R. Aaij et al. (LHCb Collaboration). "Search for Time-Dependent *CP* Violation in $D^0 \to \pi^+\pi^-\pi^0$ Decays." In: *Phys. Rev. Lett.* 133.10 (2024), p. 101803. DOI: 10.1103/PhysRevLett.133.101803. arXiv: 2405.06556.

[45] B. Aslan and G. Zech. "Statistical energy as a tool for binning-free, multivariate goodness-of-fit tests, two-sample comparison and unfolding." In: *Nucl. Instrum. Methods A* 537.3 (2005), pp. 626–636. DOI: 10.1016/j.nima.2004.08.071.

[46] A. Di Canto et al. "Novel method for measuring charm-mixing parameters using multibody decays." In: *Phys. Rev. D* 99.1 (Jan. 2019), p. 012007. DOI: 10.1103/PhysRevD.99.012007.

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

[48] J. E. Augustin et al. "Discovery of a Narrow Resonance in e^+e^- Annihilation." In: *Phys. Rev. Lett.* 33.23 (Dec. 1974), pp. 1406–1408. DOI: 10.1103/PhysRevLett.33.1406.

[49] J. J. Aubert et al. "Experimental Observation of a Heavy Particle *J*." In: *Phys. Rev. Lett.* 33.23 (Dec. 1974), pp. 1404–1406. DOI: 10.1103/PhysRevLett.33.1404.

[50] S. W. Herb, D. C. Hom, L. M. Lederman, et al. "Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions." In: *Phys. Rev. Lett.* 39.5 (Aug. 1977), pp. 252–255. DOI: 10.1103/PhysRevLett.39.252.

[51] A. J. Bevan et al. "The Physics of the B Factories." In: *Eur. Phys. J. C* 74.11 (Nov. 2014), p. 3026. DOI: 10.1140/epjc/s10052-014-3026-9. arXiv: 1406.6311.

[52] D. Andrews et al. "Observation of a Fourth Upsilon State in e^+e^- Annihilations." In: *Phys. Rev. Lett.* 45.4 (Aug. 1980), pp. 219–221. DOI: 10.1103/PhysRevLett.45.219.

[53] G. Finocchiaro et al. "Observation of the Υ''' at the Cornell Electron Storage Ring." In: *Phys. Rev. Lett.* 45.4 (July 1980), pp. 222–225. DOI: 10.1103/PhysRevLett.45.222.

[54] Y. Ohnishi et al. "Accelerator design at SuperKEKB." In: *PTEP* 2013.3 (Mar. 2013), 03A011. DOI: 10.1093/ptep/pts083.

[55] URL: https://www.belle2.org/.

[56] URL: https://www-superkekb.kek.jp/.

[57] S. Baird. *Accelerators for pedestrians*. 2007. URL: https://cds.cern.ch/record/1017689.

[58] P. Raimondi. *Talk given at the 2nd SuperB workshop*. Frascati, 2006. URL: http://www.lnf.infn.it/conference/superb06/talks/raimondi1.ppt.

[59] P. Raimondi, D. N. Shatilov, and M. Zobov. *Beam-Beam Issues for Colliding Schemes with Large Piwinski Angle and Crabbed Waist*. 2007. arXiv: physics/0702033.

[60] M. Aiba. "Injection: Electron Beams." In: *Proceedings of the CAS-CERN Accelerator School on Beam Injection, Extraction and Transfer*. Ed. by Bernhard Holzer. Vol. 5. 2018. DOI: 10.23730/CYRSP-2018-005.

[61] P. M. Lewis et al. "First measurements of beam backgrounds at SuperKEKB." In: *Nucl. Instrum. Methods A* 914 (2019), pp. 69–144. DOI: 10.1016/j.nima.2018.05.071. arXiv: 1802.01366.

[62] T. Abe et al. *Belle II Technical Design Report*. 2010. arXiv: 1011.0352.

[63] K. Adamczyk et al. (Belle II SVD Collaboration). "The design, construction, operation and performance of the Belle II silicon vertex detector." In: *J. Inst.* 17.11 (Nov. 2022), P11042. DOI: 10.1088/1748-0221/17/11/P11042. arXiv: 2201.09824.

[64] T. Kuhr et al. "The Belle II Core Software ." In: *Computing Softw. Big Science* 3.1 (Nov. 2018). DOI: 10.1007/s41781-018-0017-9. arXiv: 1809.04299.

[65] R. Brun and F. Rademakers. "ROOT – An object oriented data analysis framework." In: *Nucl. Instrum. Methods A* 389.1 (1997), pp. 81–86. DOI: 10.1016/S0168-9002(97)00048-X.

[66] D. J. Lange. "The EvtGen particle decay simulation package." In: *Nucl. Instrum. Methods A* 462.1 (2001), pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4. URL: https://evtgen.hepforge.org/.

[67] T. Sjöstrand et al. "An introduction to PYTHIA 8.2." In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024.

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

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

[70] E. Barberio and Z. Wąs. "PHOTOS – a universal Monte Carlo for QED radiative corrections: version 2.0." In: *Comput. Phys. Commun.* 79.2 (1994), pp. 291–308. DOI: 10.1016/0010-4655(94)90074-4.

[71] S. Agostinelli et al. "Geant4 – a simulation toolkit." In: *Nucl. Instrum. Methods A* 506.3 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[72] V. Bertacchi et al. (Belle II Tracking Group). "Track finding at Belle II." In: *Comput. Phys. Commun.* 259 (2021), p. 107610. DOI: 10.1016/j.cpc.2020.107610. arXiv: 2003.12466.

[73] D. Buskulic et al. "The forward-backward asymmetry for charm quarks at the Z pole." In: *Physics Letters B* 352.3 (1995), pp. 479–486. ISSN: 0370-2693. DOI: 10.1016/0370-2693(95)00519-Q.

[74] URL: https://phab.hepforge.org/T219.

[75] P. Avery. *Applied Fitting Theory*. I, IV, VI vols. 1999. URL: https://www.phys.ufl.edu/~avery/fitting.html.

[76] Thomas Keck. "FastBDT: A Speed-Optimized Multivariate Classification Algorithm for the Belle II Experiment." In: *Comput. Softw. Big Sci.* 1.1 (2017). DOI: 10.1007/s41781-017-0002-8.

[77] S. Longo et al. "CsI(Tl) pulse shape discrimination with the Belle II electromagnetic calorimeter as a novel method to improve particle identification at electron-positron colliders." In: *Nucl. Instrum. Meth. A* 982 (2020), p. 164562. DOI: 10.1016/j.nima.2020.164562. arXiv: 2007.09642 [physics.ins-det].

[78] J.-F. Krohn et al. "Global decay chain vertex fitting at Belle II." In: *Nucl. Instrum. Methods A* 976 (2020), p. 164269. DOI: 10.1016/j.nima.2020.164269.

[79] Wouter D. Hulsbergen. "Decay chain fitting with a Kalman filter." In: *Nucl. Instrum. Methods A* 552.3 (2005), pp. 566–575. DOI: 10.1016/j.nima.2005.06.078.

[80] M. Pivk and F. R. Le Diberder. "$_s\mathcal{P}lots$: A statistical tool to unfold data distributions." In: *Nucl. Instrum. Meth. A* 555.1 (2005), pp. 356–369. DOI: 10.1016/j.nima.2005.08.106.

[81] N. L. Johnson. "Systems of frequency curves generated by methods of translation." In: *Biometrika* 36 (1949), pp. 149–176. DOI: 10.1093/biomet/36.1-2.149.

[82] *Approximation for Mean and Variance of a Ratio*. URL: https://www.stat.cmu.edu/~hseltman/files/ratio.pdf.

APPENDICES

A USEFUL FORMULAS

A.1 COMPOSITION OF ASYMMETRIES

When two sources of asymmetry contribute to the total asymmetry of a sample, the total asymmetry can, in general, be written as

$$\mathcal{A} = \frac{X_1 X_2 - \overline{X}_1 \overline{X}_2}{X_1 X_2 + \overline{X}_1 \overline{X}_2} \tag{A.1}$$

where the X_i can be, for example, efficiencies for one flavor, while the $\overline{X}_i$ are the corresponding values for the opposite flavor. The asymmetries due to each source can then be written as

$$\mathcal{A}_i = \frac{X_i - \overline{X}_i}{X_i + \overline{X}_i} \tag{A.2}$$

from which it follows that

$$X_i = \frac{X_i + \overline{X}_i}{2}(1 + \mathcal{A}_i) \qquad \text{and} \qquad \overline{X}_i = \frac{X_i + \overline{X}_i}{2}(1 - \mathcal{A}_i)\,. \tag{A.3}$$

If we then substitute X_2 and $\overline{X}_2$ in Equation A.1 using Equation A.3, we get

$$\mathcal{A} = \frac{\frac{X_2+\overline{X}_2}{2}}{\frac{X_2+\overline{X}_2}{2}} \cdot \frac{X_1(1+\mathcal{A}_2) - \overline{X}_1(1-\mathcal{A}_2)}{X_1(1+\mathcal{A}_2) + \overline{X}_1(1-\mathcal{A}_2)} = \frac{X_1 - \overline{X}_1 + \mathcal{A}_2(X_1 + \overline{X}_1)}{X_1 + \overline{X}_1 + \mathcal{A}_2(X_1 - \overline{X}_1)}\,. \tag{A.4}$$

We can then multiply and divide the numerator by $X_1 - \overline{X}_1$, and the denominator by $X_1 + \overline{X}_1$, getting

$$\mathcal{A} = \frac{X_1 - \overline{X}_1}{X_1 + \overline{X}_1} \cdot \frac{1 + \mathcal{A}_2 \frac{X_1+\overline{X}_1}{X_1-\overline{X}_1}}{1 + \mathcal{A}_2 \frac{X_1-\overline{X}_1}{X_1+\overline{X}_1}} = \mathcal{A}_1 \frac{1 + \mathcal{A}_2/\mathcal{A}_1}{1 + \mathcal{A}_2\mathcal{A}_1} = \frac{\mathcal{A}_1 + \mathcal{A}_2}{1 + \mathcal{A}_1\mathcal{A}_2}\,. \tag{A.5}$$

If we assume the asymmetries to be small, we can approximate

$$\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{O}(\mathcal{A}_i^3)\,. \tag{A.6}$$

Finally, for the case of more than two sources, we can apply the formula recursively. For example, with three sources we get

$$\frac{X_1X_2X_3 - \overline{X}_1\overline{X}_2\overline{X}_3}{X_1X_2X_3 + \overline{X}_1\overline{X}_2\overline{X}_3} = \frac{\mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3 + \mathcal{A}_1\mathcal{A}_2\mathcal{A}_3}{1 + \mathcal{A}_1\mathcal{A}_2 + \mathcal{A}_1\mathcal{A}_3 + \mathcal{A}_2\mathcal{A}_3} = \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3 + \mathcal{O}(\mathcal{A}_i^3)\,. \tag{A.7}$$

A.2 MAXIMUM MOMENTUM OF A D MESON IN A $B\overline{B}$ EVENT

In the collision center of mass (CM) frame, the momentum p_{CM} of a D^* meson produced by the decay of a B meson is kinematically limited to be at most

$$\max p_{\mathrm{CM}}(D^*) = \frac{E_{\mathrm{CM}}(B) \cdot p(D^*) + p_{\mathrm{CM}}(B) \cdot E(D^*)}{M(B)}\,, \tag{A.8}$$

where I used the energy and momentum of the B meson in the collision frame as a function of the collision energy $\sqrt{s}$, assuming an $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ event. These are given by

$$E_{\mathrm{CM}}(B) = \frac{\sqrt{s}}{2}, \qquad p_{\mathrm{CM}}(B) = \sqrt{E^2_{\mathrm{CM}}(B) - M^2(B)}, \tag{A.9}$$

and the maximum energy and momentum of the D^* in the B meson rest frame, assuming $B^0 \to D^{*-}e^+\nu$ to be the B decay with minimum mass recoiling against the D^*, is

$$\max E(D^*) = \frac{M^2(B) - M^2(e) + M^2(D^*)}{2M(B)}, \qquad p(D^*) = \sqrt{E^2(D^*) - M^2(D^*)}. \tag{A.10}$$

When the collision energy $\sqrt{s}$ is at the $\Upsilon(4S)$ resonance peak, $\max p_{\mathrm{CM}}(D^*) \simeq 2.45\,\mathrm{GeV}/c$.

The same computation can be repeated for a D^0 produced by the decay of a B meson, using $B^+ \to \bar{D}^0 e^+\nu$ as the B decay with minimum mass recoiling against the D^0. In this case, when the collision energy $\sqrt{s}$ is at the $\Upsilon(4S)$ resonance peak, $\max p_{\mathrm{CM}}(D^0) \simeq 2.5\,\mathrm{GeV}/c$.

A.3 DISTRIBUTIONS AND PDFS

In this section I report the definitions of some probability density functions (PDFs) that I use for the fits.

JOHNSON'S S_U-DISTRIBUTION This is defined as in [81]:

$$p(x|\mu,\lambda,\gamma,\delta) = \frac{\delta}{\lambda\sqrt{2\pi}} \frac{1}{\sqrt{1+\left(\frac{x-\mu}{\lambda}\right)^2}} \exp\left(-\frac{1}{2}\left(\gamma + \delta \sinh^{-1}\left(\frac{x-\mu}{\lambda}\right)\right)\right). \tag{A.11}$$

The most probable value (MPV) of the distribution is given by

$$m = \mu - \lambda \sinh\left(\frac{\gamma}{\delta}\right). \tag{A.12}$$

To use the MPV as a fit parameter, I define the following function

$$J_\mu(m,\lambda,\gamma,\delta) = m + \lambda \sinh\left(\frac{\gamma}{\delta}\right), \tag{A.13}$$

which I use as the μ parameter of the Johnson's distribution where appropriate.

CRYSTAL BALL DISTRIBUTION The generalized asymmetric double-sided Crystal Ball PDF has an asymmetric Gaussian core and asymmetric power-law tails. It is defined as

$$p(x|x_0,\sigma_L,\sigma_R,\alpha_L,n_L,\alpha_R,n_R) \propto \begin{cases} A_L \cdot (B_L - \frac{x-x_0}{\sigma_L})^{-n_L} & \frac{x-x_0}{\sigma_L} < -\alpha_L \\ \exp\left(-\frac{1}{2}\left(\frac{x-x_0}{\sigma_L}\right)^2\right) & x \le x_0 \\ \exp\left(-\frac{1}{2}\left(\frac{x-x_0}{\sigma_R}\right)^2\right) & \frac{x-x_0}{\sigma_R} \le \alpha_R \\ A_R \cdot (B_R + \frac{x-x_0}{\sigma_R})^{-n_R} & \text{otherwise} \end{cases}, \tag{A.14}$$

where x_0 is the MPV of the distribution, σ_L and σ_R are the left and right Gaussian standard deviations, n_L and n_R are the left and right power-law exponents, and α_L and α_R are the left and right transition points between Gaussian and power-law tail expressed in units of $\sigma_{L,R}$. The transitions between core and tails preserve the continuity of the PDF and its first derivative. This is ensured by the A and B factors, which are defined as

$$A_i = \left(\frac{n_i}{|\alpha_i|}\right)^{n_i} \cdot \exp\left(-\frac{|\alpha_i|^2}{2}\right), \qquad B_i = \frac{n_i}{|\alpha_i|} - |\alpha_i| \,. \tag{A.15}$$

RooDstD0Bg This is an empirical model for the ΔM combinatorial background, defined in the RooFit package as

$$p(x|x_0, A, B, C) \propto \left(1 - \exp\left(-\frac{x - x_0}{C}\right)\right) \cdot \left(\frac{x}{x_0}\right)^A + B \cdot \left(\frac{x}{x_0} - 1\right) . \tag{A.16}$$

A.4 INDEPENDENCE OF ASYMMETRY AND TOTAL NUMBER OF CANDIDATES

Let us consider a sample of D^0 and $\overline{D}^0$ mesons. Let N be the number of D^0 mesons and $\overline{N}$ be the number of $\overline{D}^0$ mesons. Let also $T = N + \overline{N}$ be the total number of mesons of both flavors, and $\mathcal{A} = (N - \overline{N})/(N + \overline{N}) = (N - \overline{N})/T$ be the asymmetry.

I compute the covariance of $\mathcal{A}$ and T, assuming that N and $\overline{N}$ follow independent Poisson distributions with means $\mathrm{E}[N]$ and $\mathrm{E}[\overline{N}]$ respectively:

$$\mathrm{Cov}(\mathcal{A}, T) = \mathrm{E}[\mathcal{A}T] - \mathrm{E}[\mathcal{A}]\,\mathrm{E}[T] = \mathrm{E}[N - \overline{N}] - \mathrm{E}\left[\frac{N - \overline{N}}{N + \overline{N}}\right] \mathrm{E}[N + \overline{N}] \,. \tag{A.17}$$

For the expectation value of the ratio we use the second-order approximation from [82]:

$$\mathrm{E}\left[\frac{N - \overline{N}}{N + \overline{N}}\right] = \frac{\mathrm{E}[N - \overline{N}]}{\mathrm{E}[N + \overline{N}]} - \frac{\mathrm{Cov}(N - \overline{N}, N + \overline{N})}{\mathrm{E}^2[N + \overline{N}]} + \frac{\mathrm{Var}(N + \overline{N})\,\mathrm{E}[N - \overline{N}]}{\mathrm{E}^3[N + \overline{N}]}, \tag{A.18a}$$

$$\mathrm{Cov}(N - \overline{N}, N + \overline{N}) = \mathrm{Cov}(N, N) + \mathrm{Cov}(N, \overline{N}) - \mathrm{Cov}(\overline{N}, N) - \mathrm{Cov}(\overline{N}, \overline{N}) \tag{A.18b}$$

$$= \mathrm{Var}(N) - \mathrm{Var}(\overline{N}) = \mathrm{E}[N] - \mathrm{E}[\overline{N}] = \mathrm{E}[N - \overline{N}], \tag{A.18c}$$

$$\mathrm{Var}(N + \overline{N}) = \mathrm{Var}(N) + \mathrm{Var}(\overline{N}) = \mathrm{E}[N] + \mathrm{E}[\overline{N}] = \mathrm{E}[N + \overline{N}], \tag{A.18d}$$

$$\Rightarrow \mathrm{E}\left[\frac{N - \overline{N}}{N + \overline{N}}\right] = \frac{\mathrm{E}[N - \overline{N}]}{\mathrm{E}[N + \overline{N}]} - \frac{\mathrm{E}[N - \overline{N}]}{\mathrm{E}^2[N + \overline{N}]} + \frac{\mathrm{E}[N + \overline{N}]\,\mathrm{E}[N - \overline{N}]}{\mathrm{E}^3[N + \overline{N}]} \tag{A.18e}$$

$$= \frac{\mathrm{E}[N - \overline{N}]}{\mathrm{E}[N + \overline{N}]}, \tag{A.18f}$$

$$\Rightarrow \mathrm{Cov}(\mathcal{A}, T) = \mathrm{E}[N - \overline{N}] - \frac{\mathrm{E}[N - \overline{N}]}{\mathrm{E}[N + \overline{N}]}\,\mathrm{E}[N + \overline{N}] = 0\,. \tag{A.18g}$$

q.e.d.

B PRELIMINARY SELECTIONS

In order to limit the number of particle combinations and the size of intermediate files, I apply a preliminary selection before and during the creation of signal candidates. This preliminary selection is relatively loose, so that optimization studies and adjustments can still be performed. Nevertheless, this preliminary selection is always implicitly applied to all distributions shown in this document. The criteria are summarized in Table B.1.

Variable	Selection
Photons	
`clusterNHits`	> 1.5
θ	within CDC acceptance ($17°$ to $150°$)
E	$> 25\,\mathrm{MeV}$ (forward endcap and barrel ECL sections)
E	$> 40\,\mathrm{MeV}$ (backward endcap ECL section)
Neutral pion	
M	between $100\,\mathrm{MeV}/c^2$ and $170\,\mathrm{MeV}/c^2$
Charged particles	
d_r	$< 2\,\mathrm{cm}$
$\lvert d_z \rvert$	$< 4\,\mathrm{cm}$
θ	within CDC acceptance ($17°$ to $150°$)
D^0 **meson**	
M	between $1.75\,\mathrm{GeV}/c^2$ and $1.95\,\mathrm{GeV}/c^2$
D^* **meson**	
ΔM	$< 160\,\mathrm{MeV}/c^2$
p_{CM}	$> 2\,\mathrm{GeV}/c^2$

Table B.1: The preliminary selection criteria used for signal candidates.

C DATA RESULTS FOR EACH $\cos\theta_{\mathrm{CM}}$ BIN

Table C.1 and Table C.2 report the asymmetry results for the various data samples in each $\cos\theta_{\mathrm{CM}}$ bin. The following sections contain plots of the fits for each $\cos\theta_{\mathrm{CM}}$.

$\cos\theta_{\mathrm{CM}}$ bin	Asymmetries [%]				
	$\mathcal{A}_{\mathrm{raw}}^{\pi\pi\pi^0}$	$\mathcal{A}_{\mathrm{raw}}^{\mathrm{tagged}}$	$\mathcal{A}_{\mathrm{raw}}^{\mathrm{untagged}}$	$\mathcal{A}_{\varepsilon}^{\pi\,\mathrm{tag}}$	Corrected $\mathcal{A}^{\pi\pi\pi^0}$
1	5.2 ± 1.0	2.6 ± 0.6	1.9 ± 0.4	0.8 ± 0.6	4.4 ± 1.2
2	5.1 ± 0.7	4.31 ± 0.34	1.76 ± 0.22	2.54 ± 0.33	2.6 ± 0.8
3	5.7 ± 0.7	5.77 ± 0.32	1.39 ± 0.19	4.38 ± 0.31	1.4 ± 0.8
4	5.2 ± 0.7	4.92 ± 0.33	1.24 ± 0.20	3.68 ± 0.32	1.5 ± 0.7
5	0.8 ± 0.6	1.19 ± 0.32	-0.08 ± 0.21	1.27 ± 0.31	-0.5 ± 0.7
6	-1.3 ± 0.6	0.08 ± 0.31	-0.19 ± 0.20	0.28 ± 0.30	-1.6 ± 0.7
7	-2.0 ± 0.6	-0.73 ± 0.29	-1.07 ± 0.19	0.34 ± 0.29	-2.3 ± 0.7
8	-2.8 ± 0.9	-2.0 ± 0.4	-1.61 ± 0.26	-0.4 ± 0.4	-2.4 ± 1.0

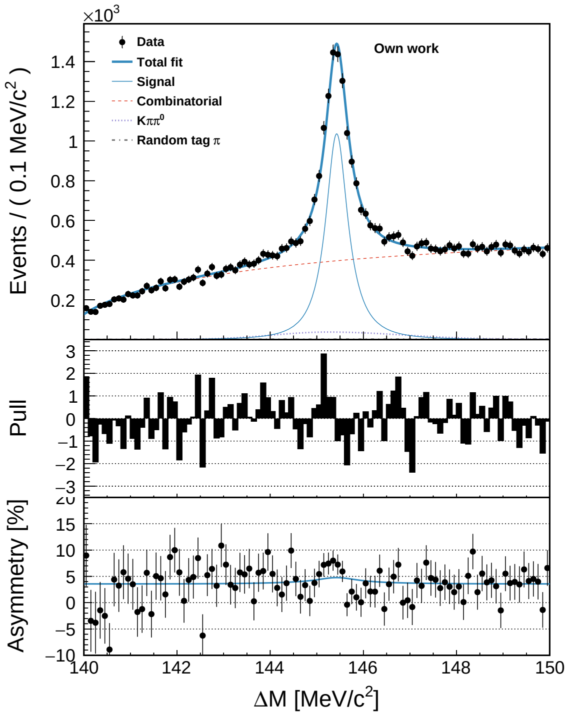
Table C.1: For the data samples and in each $\cos\theta_{\mathrm{CM}}$ bin: $D^0 \to \pi^+\pi^-\pi^0$, tagged $D^0 \to K^-\pi^+$, and untagged $D^0 \to K^-\pi^+$ raw asymmetries; tag pion detection asymmetry; and $\mathcal{A}_{\varepsilon}^{\pi\,\mathrm{tag}}$-corrected $D^0 \to \pi^+\pi^-\pi^0$ asymmetry.

$\cos\theta_{\mathrm{CM}}$ bin pair	Asymmetries [%]		
	Backward bin	Forward bin	Average
1 and 8	4.4 ± 1.2	-2.4 ± 1.0	1.0 ± 0.8
2 and 7	2.6 ± 0.8	-2.3 ± 0.7	0.2 ± 0.5
3 and 6	1.4 ± 0.8	-1.6 ± 0.7	-0.1 ± 0.5
4 and 5	1.5 ± 0.7	-0.5 ± 0.7	0.5 ± 0.5

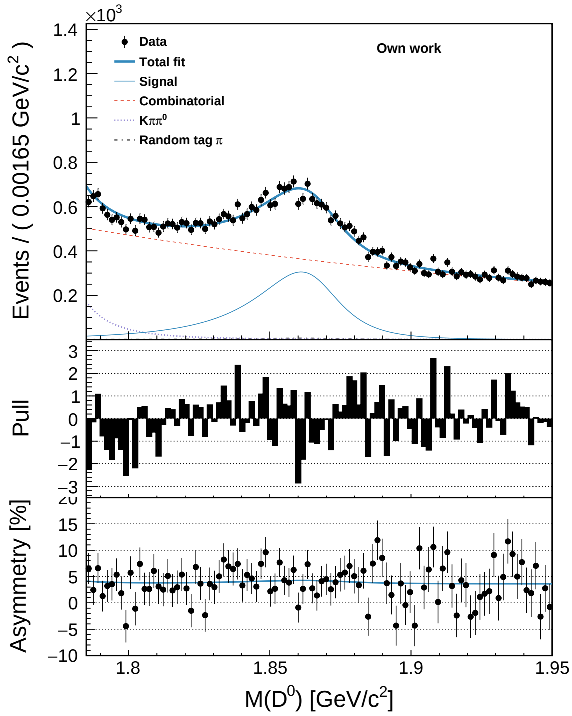
Table C.2: $\mathcal{A}_{\varepsilon}^{\pi\,\mathrm{tag}}$-corrected $D^0 \to \pi^+\pi^-\pi^0$ asymmetry averages for each $\cos\theta_{\mathrm{CM}}$ bin pair.

C.1 SIGNAL CHANNEL FIT

Figures C.1 through C.8 report the $D^0 \to \pi^+\pi^-\pi^0$ channel fit results for each $\cos\theta_{\mathrm{CM}}$ bin separately.

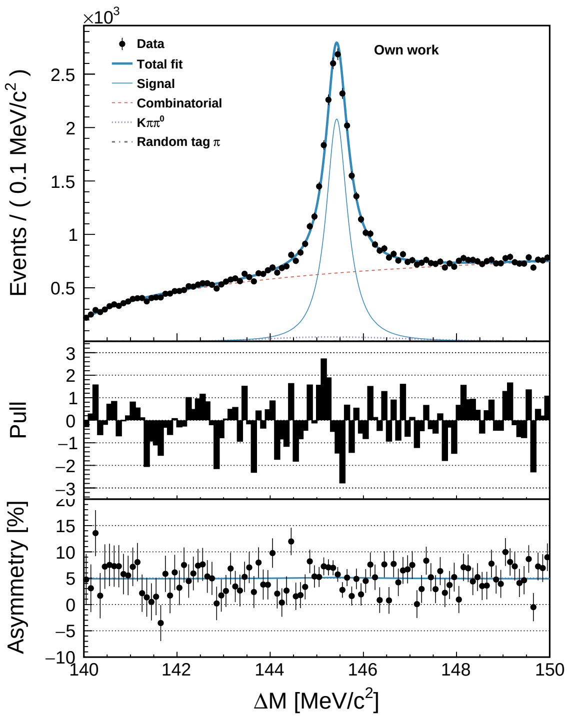


(a) ΔM distribution.

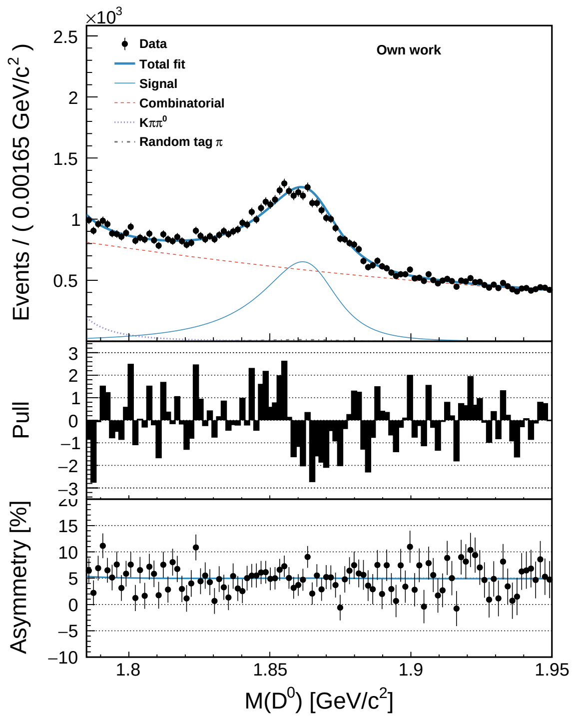


(b) D^0 invariant mass distribution.

Figure C.1: Final fit of the data sample for $\cos\theta_{\mathrm{CM}}$ bin 1, *i.e.* $-0.7 < \cos\theta_{\mathrm{CM}} < -0.599$.

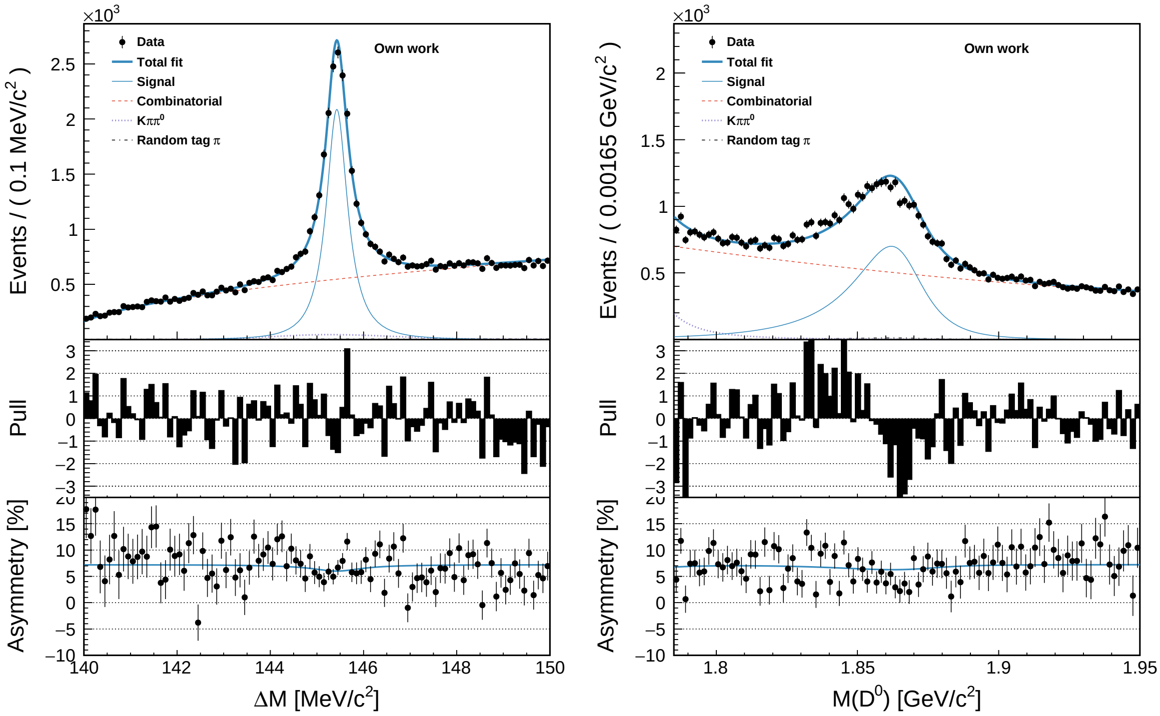


(a) ΔM distribution.



(b) D^0 invariant mass distribution.

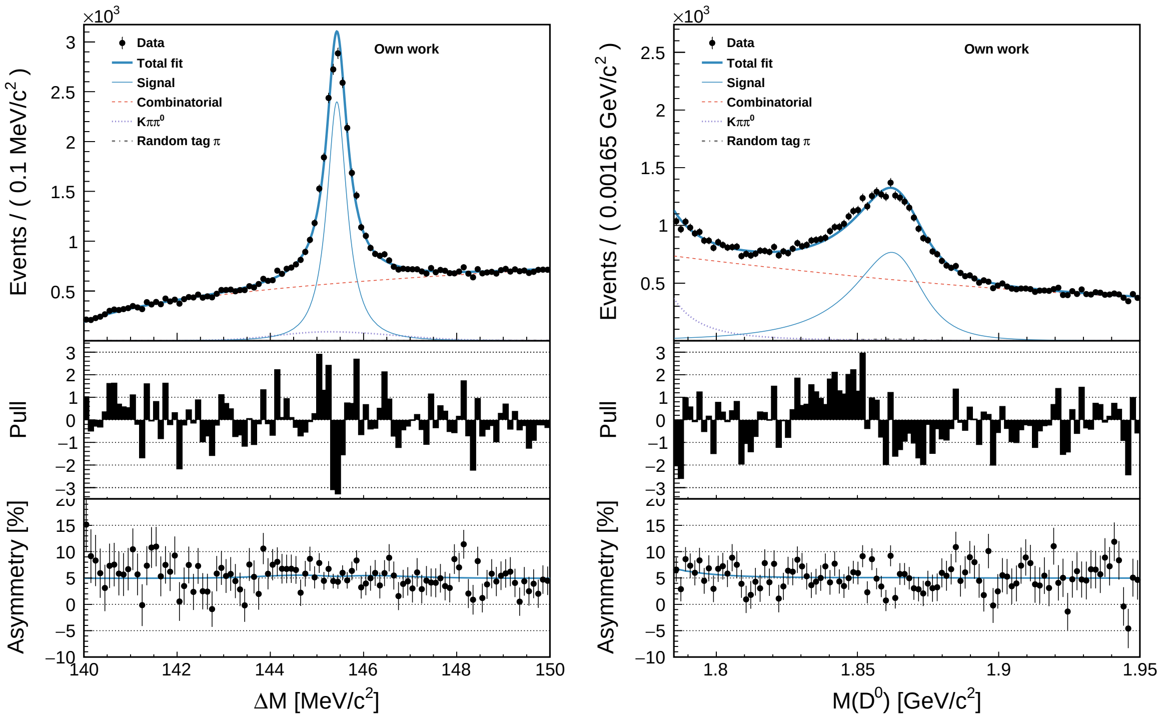
Figure C.2: Final fit of the data sample for $\cos\theta_{\mathrm{CM}}$ bin 2, *i.e.* $-0.599 < \cos\theta_{\mathrm{CM}} < -0.411$.



(a) ΔM distribution.

(b) D^0 invariant mass distribution.

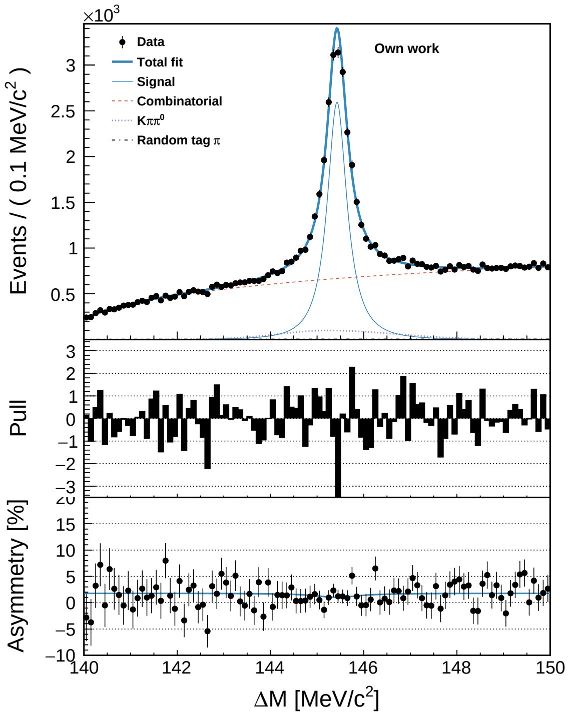
Figure C.3: Final fit of the data sample for $\cos\theta_{\mathrm{CM}}$ bin 3, *i.e.* $-0.411 < \cos\theta_{\mathrm{CM}} < -0.208$.



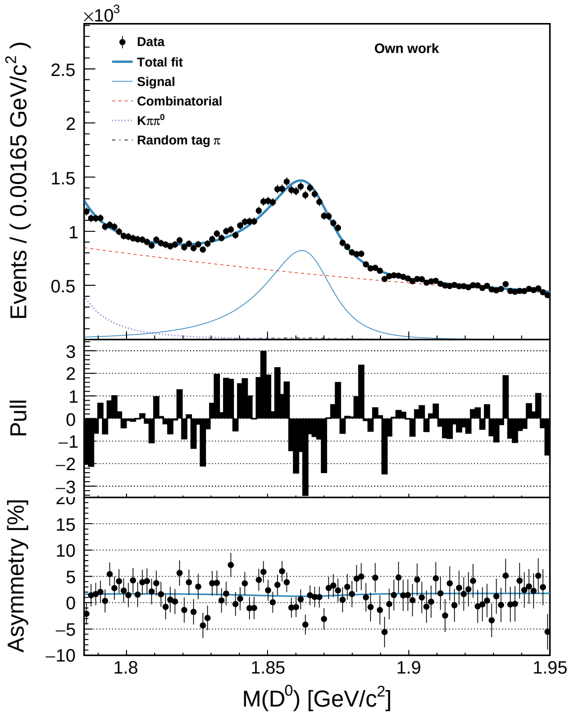
(a) ΔM distribution.

(b) D^0 invariant mass distribution.

Figure C.4: Final fit of the data sample for $\cos\theta_{\mathrm{CM}}$ bin 4, *i.e.* $-0.208 < \cos\theta_{\mathrm{CM}} < 0$.

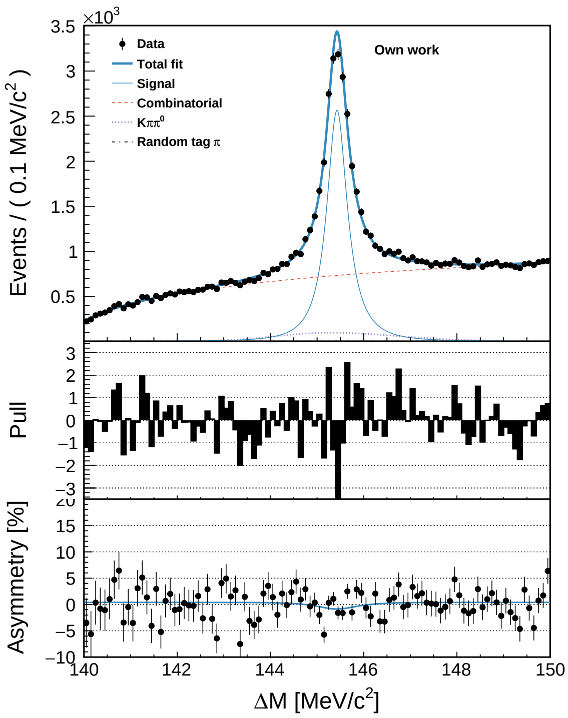


(a) ΔM distribution.

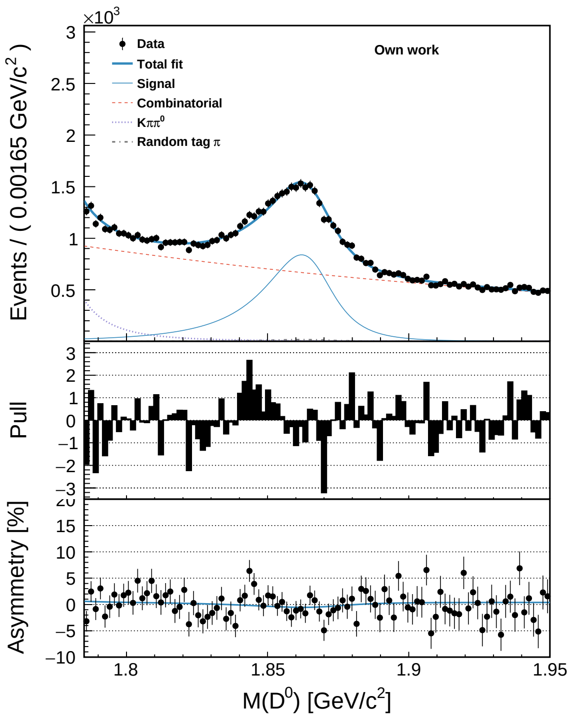


(b) D^0 invariant mass distribution.

Figure C.5: Final fit of the data sample for $\cos\theta_{\mathrm{CM}}$ bin 5, *i.e.* $0 < \cos\theta_{\mathrm{CM}} < 0.208$.

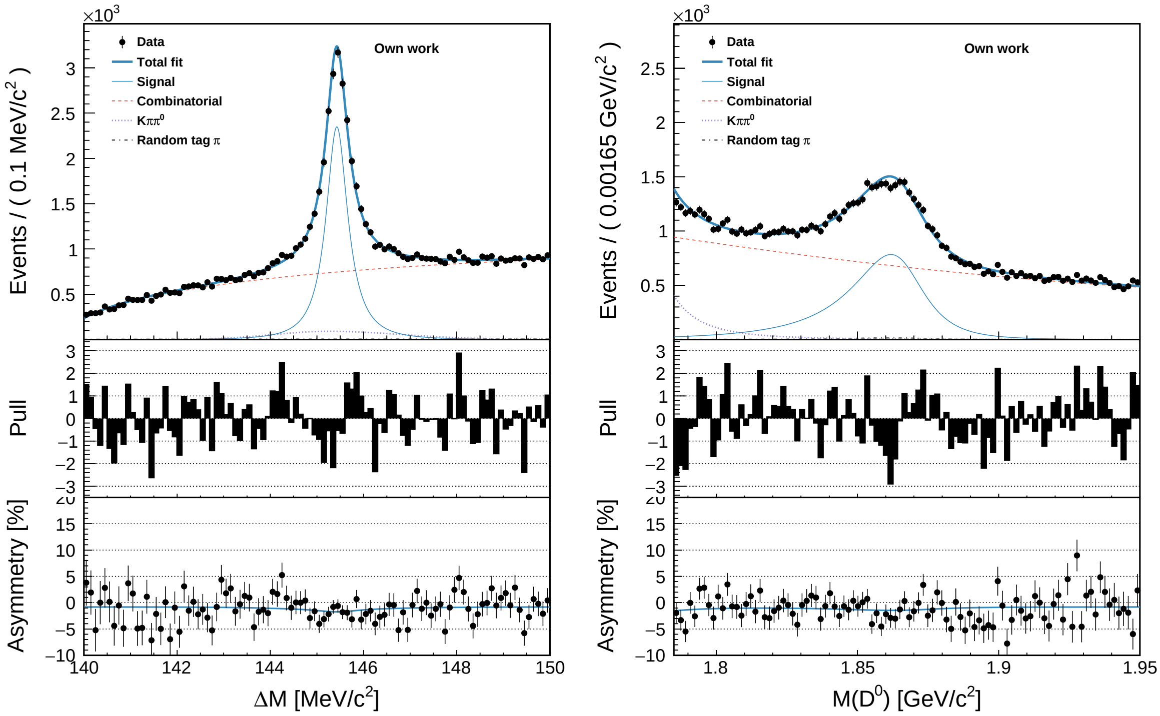


(a) ΔM distribution.



(b) D^0 invariant mass distribution.

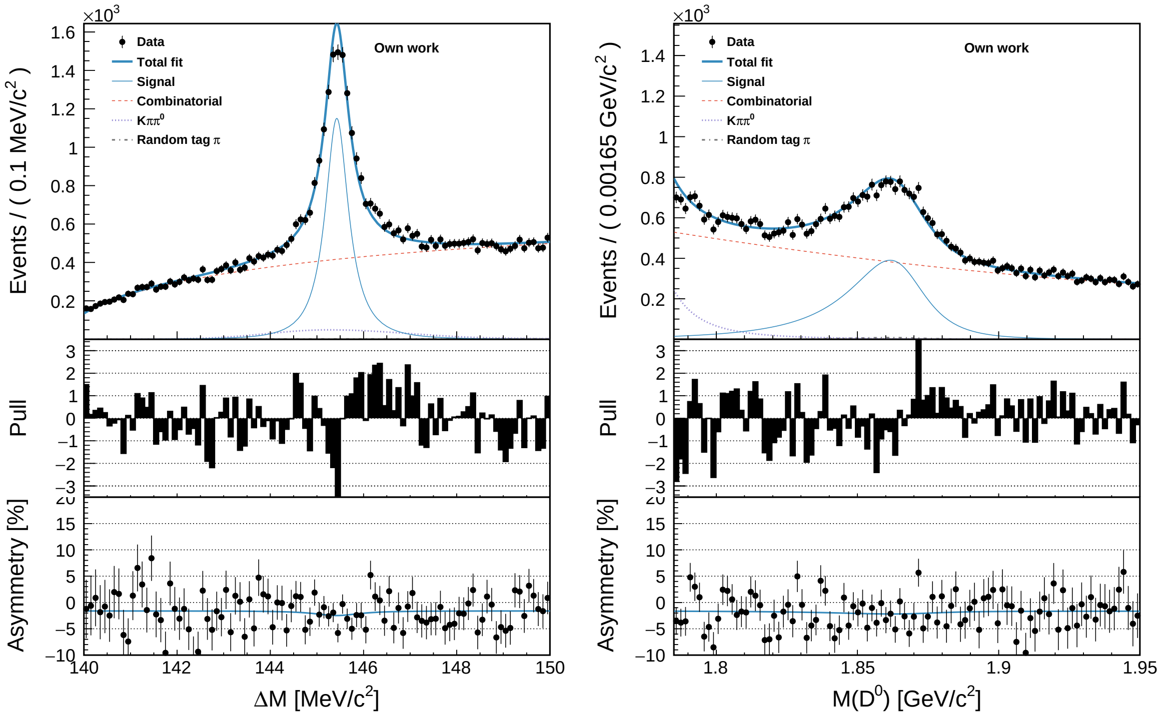
Figure C.6: Final fit of the data sample for $\cos\theta_{\mathrm{CM}}$ bin 6, *i.e.* $0.208 < \cos\theta_{\mathrm{CM}} < 0.411$.



(a) ΔM distribution.

(b) D^0 invariant mass distribution.

Figure C.7: Final fit of the data sample for $\cos\theta_{\text{CM}}$ bin 7, *i.e.* $0.411 < \cos\theta_{\text{CM}} < 0.599$.



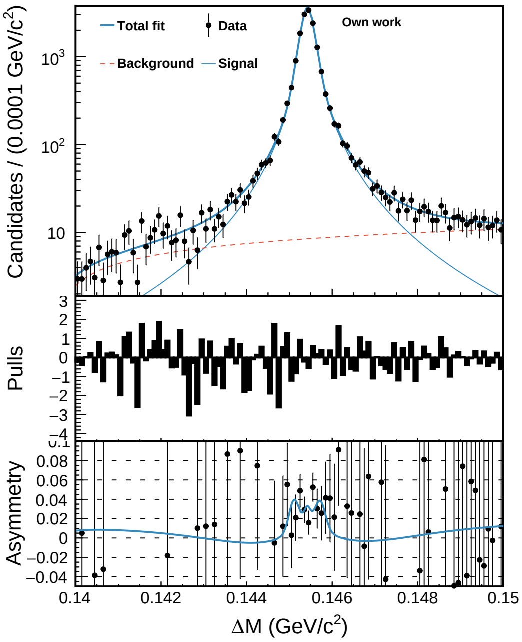
(a) ΔM distribution.

(b) D^0 invariant mass distribution.

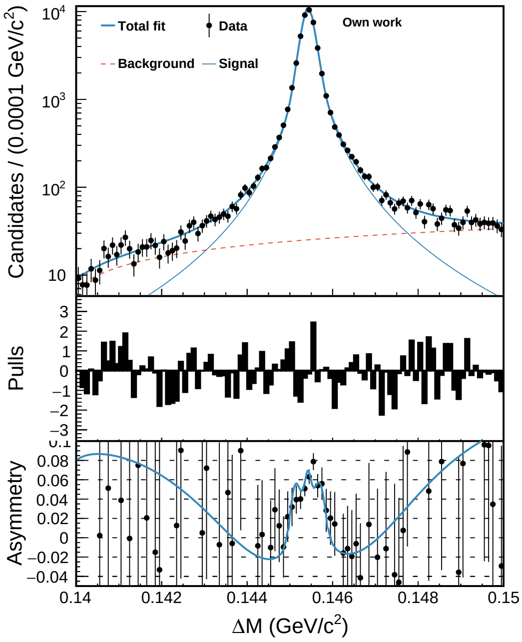
Figure C.8: Final fit of the data sample for $\cos\theta_{\text{CM}}$ bin 8, *i.e.* $0.599 < \cos\theta_{\text{CM}} < 0.7$.

C.2 TAGGED $D^0 \to K\pi$ FIT

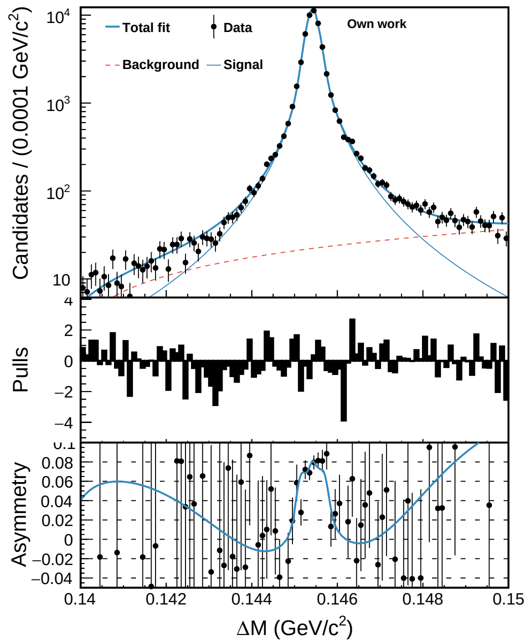
Figures C.9 and C.10 report the tagged $D^0 \to K\pi$ sample fit results for each $\cos\theta_{\mathrm{CM}}$ bin separately.



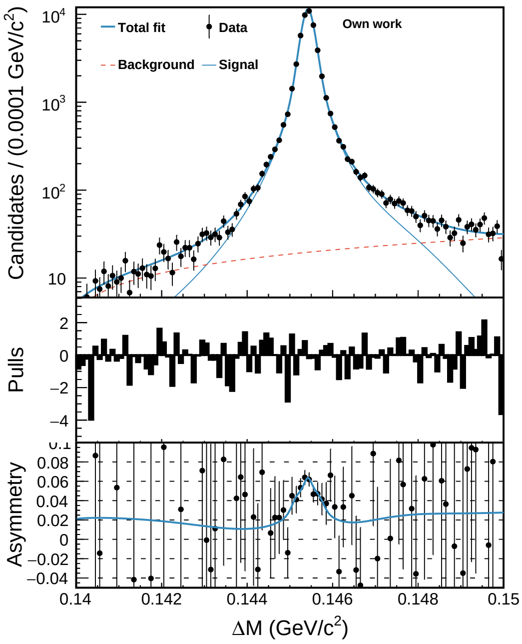
(a) $\cos\theta_{\mathrm{CM}}$ bin 1, *i.e.* $-0.7 < \cos\theta_{\mathrm{CM}} < -0.599$.



(b) $\cos\theta_{\mathrm{CM}}$ bin 2, *i.e.* $-0.599 < \cos\theta_{\mathrm{CM}} < -0.411$.

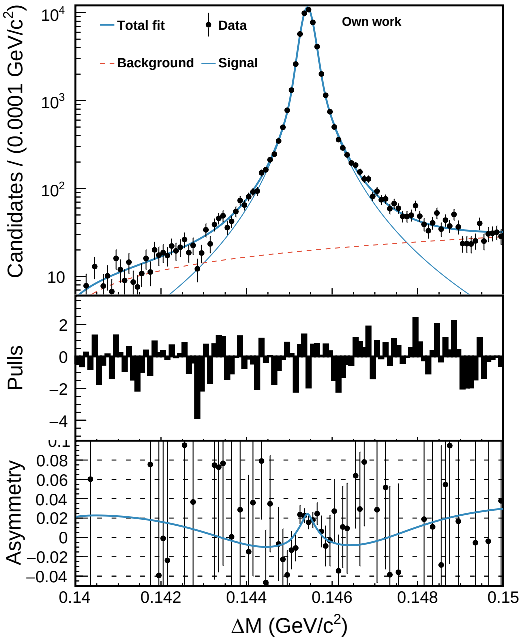


(c) $\cos\theta_{\mathrm{CM}}$ bin 3, *i.e.* $-0.411 < \cos\theta_{\mathrm{CM}} < -0.208$.

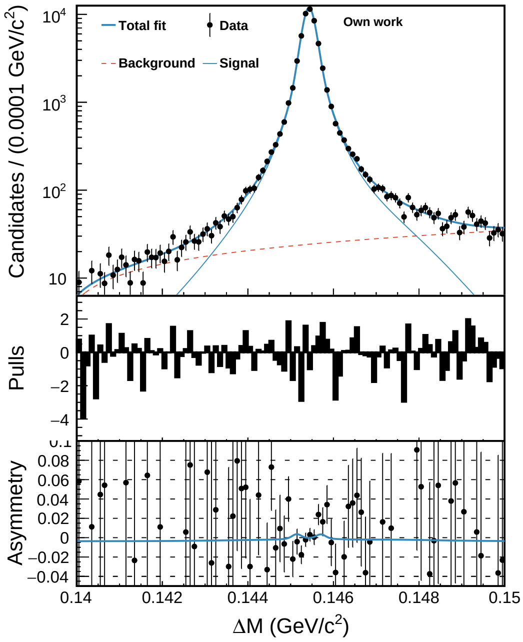


(d) $\cos\theta_{\mathrm{CM}}$ bin 4, *i.e.* $-0.208 < \cos\theta_{\mathrm{CM}} < 0$.

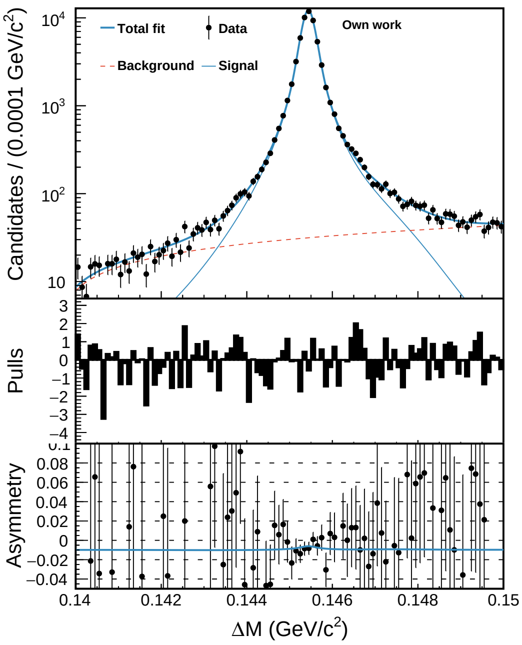
Figure C.9: Asymmetry fits of the tagged $D^0 \to K\pi$ sample on data.



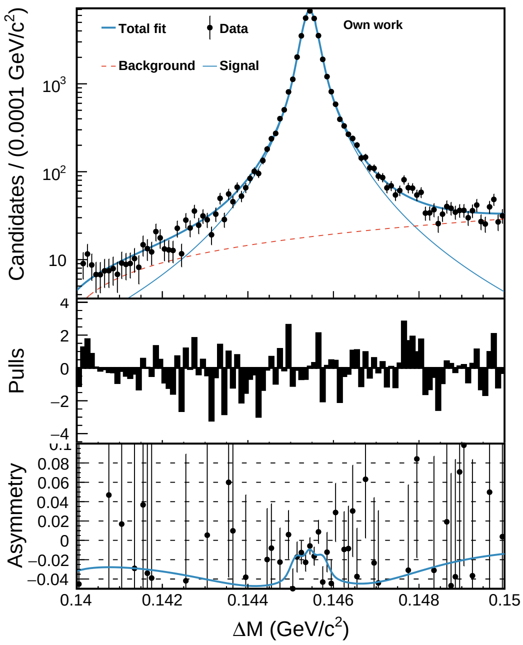
(a) $\cos\theta_{\text{CM}}$ bin 5, *i.e.* $0 < \cos\theta_{\text{CM}} < 0.208$.



(b) $\cos\theta_{\text{CM}}$ bin 6, *i.e.* $0.208 < \cos\theta_{\text{CM}} < 0.411$.



(c) $\cos\theta_{\text{CM}}$ bin 7, *i.e.* $0.411 < \cos\theta_{\text{CM}} < 0.599$.

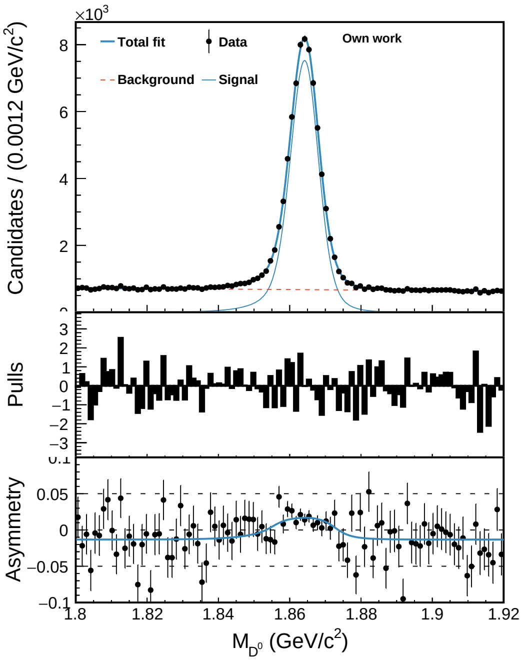


(d) $\cos\theta_{\text{CM}}$ bin 8, *i.e.* $0.599 < \cos\theta_{\text{CM}} < 0.7$.

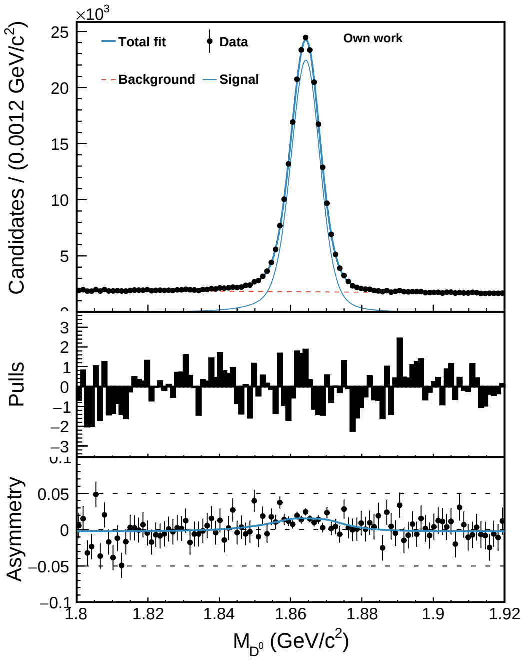
Figure C.10: Asymmetry fits of the tagged $D^0 \to K\pi$ sample on data.

C.3 UNTAGGED $D^0 \to K\pi$ FIT

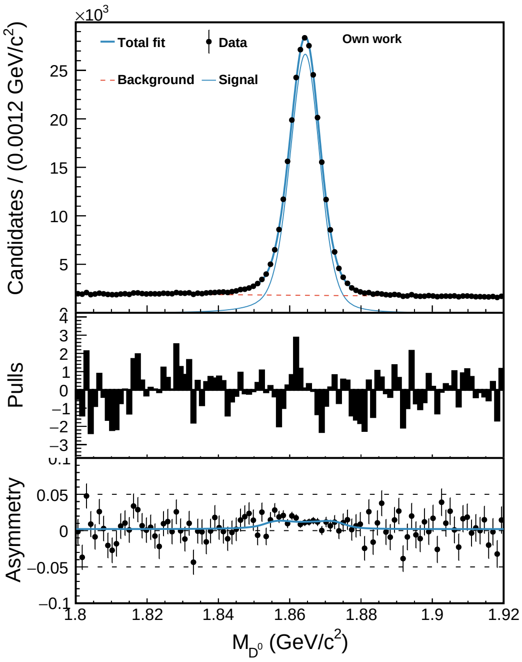
Figures C.11 and C.12 report the untagged $D^0 \to K\pi$ sample fit results for each $\cos\theta_{CM}$ bin separately.



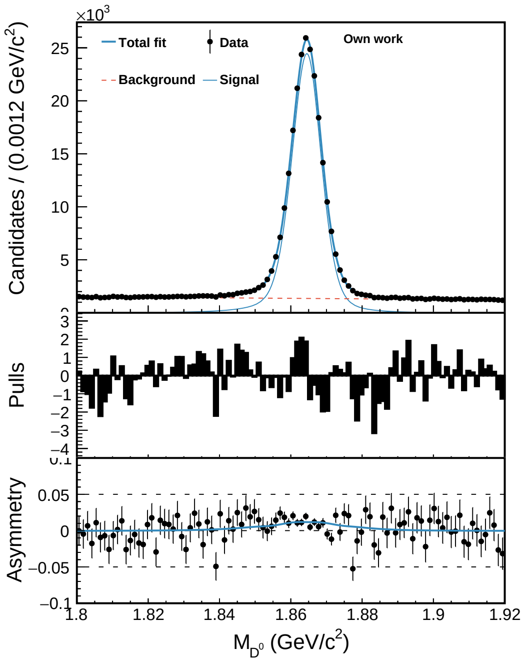
(a) $\cos\theta_{CM}$ bin 1, *i.e.* $-0.7 < \cos\theta_{CM} < -0.599$.



(b) $\cos\theta_{CM}$ bin 2, *i.e.* $-0.599 < \cos\theta_{CM} < -0.411$.

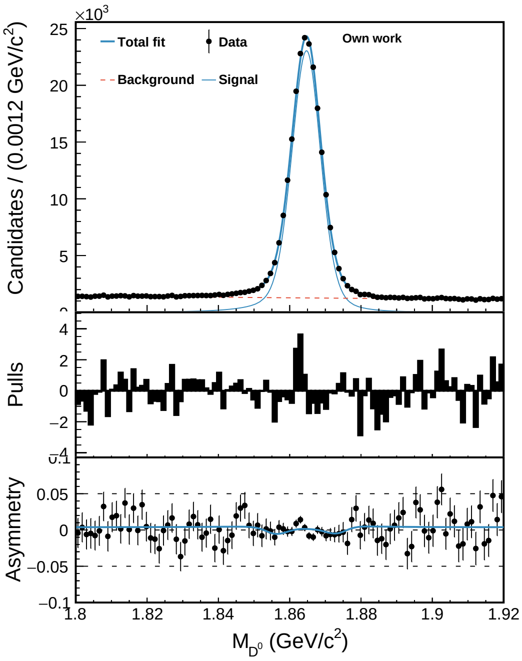


(c) $\cos\theta_{CM}$ bin 3, *i.e.* $-0.411 < \cos\theta_{CM} < -0.208$.

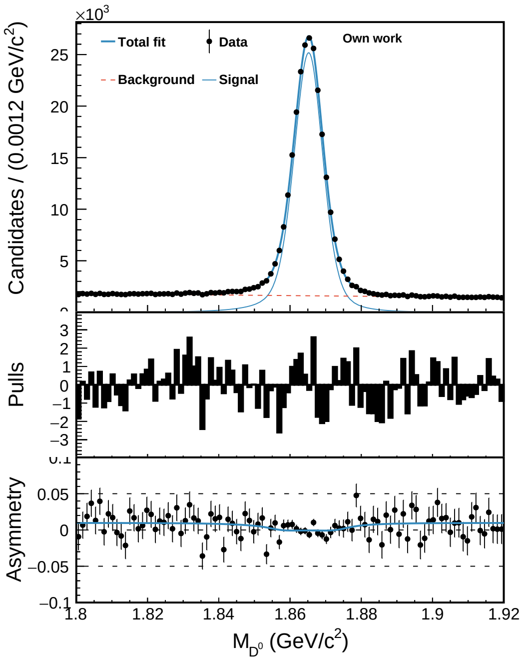


(d) $\cos\theta_{CM}$ bin 4, *i.e.* $-0.208 < \cos\theta_{CM} < 0$.

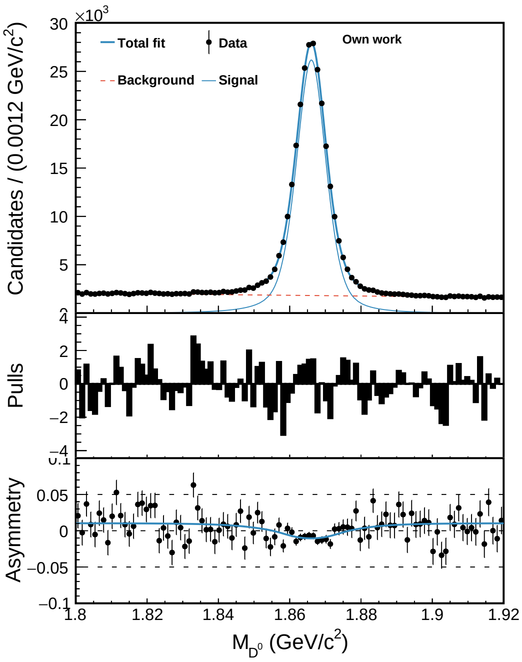
Figure C.11: Asymmetry fits of the untagged $D^0 \to K\pi$ sample on data.



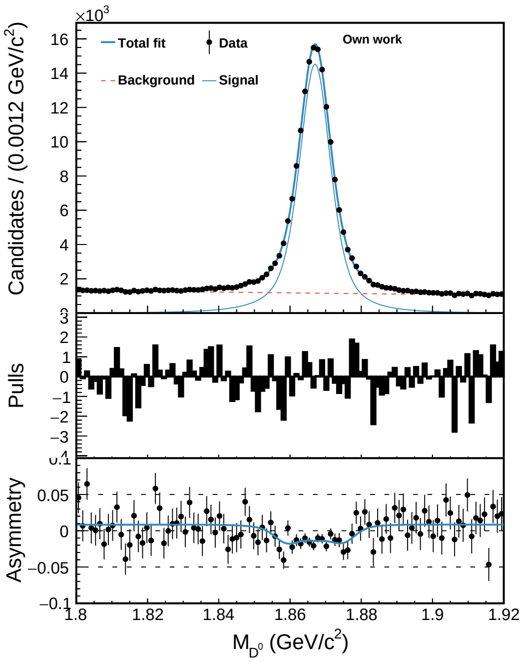
(a) $\cos\theta_{\mathrm{CM}}$ bin 5, *i.e.* $0 < \cos\theta_{\mathrm{CM}} < 0.208$.



(b) $\cos\theta_{\mathrm{CM}}$ bin 6, *i.e.* $0.208 < \cos\theta_{\mathrm{CM}} < 0.411$.



(c) $\cos\theta_{\mathrm{CM}}$ bin 7, *i.e.* $0.411 < \cos\theta_{\mathrm{CM}} < 0.599$.



(d) $\cos\theta_{\mathrm{CM}}$ bin 8, *i.e.* $0.599 < \cos\theta_{\mathrm{CM}} < 0.7$.

Figure C.12: Asymmetry fits of the untagged $D^0 \to K\pi$ sample on data.