Pontifícia Universidade Católica
do Rio de Janeiro

Victoria Ramos de Oliveira

Search for CP violation in the $D^+ \to \pi^-\pi^+K^+$ phase space in the LHCb Experiment

Dissertação de Mestrado

Dissertation presented to the Programa de Pós–graduação em Física of PUC Rio in partial fulfillment of the requirements for the degree of Mestre em Física.

Advisor: Profa Carla Göbel Burlamaqui de Mello

Rio de Janeiro
October 2023

Pontifícia Universidade Católica
do Rio de Janeiro

Victoria Ramos de Oliveira

Search for CP violation in the $D^+ \to \pi^-\pi^+K^+$ phase space in the LHCb Experiment

Dissertation presented to the Programa de Pós–graduação em Física of PUC Rio in partial fulfillment of the requirements for the degree of Mestre em Física. Approved by the Examination Committee:

Profa Carla Göbel Burlamaqui de Mello
Advisor
Departamento de Física – PUC Rio

Profa. Melissa Maria Cruz Torres
Universidad Nacional Autonoma de Honduras

Profa. Patrícia Camargo Magalhães
Universidade de Bristol

Rio de Janeiro, October 6th, 2023

Victoria Ramos de Oliveira

Graduated in Physics by the Universidade Federal de Viçosa.

Ficha Catalográfica

Oliveira, Victoria Ramos de

Search for CP violation in the $D^+ \to \pi^-\pi^+K^+$ phase space in the LHCb Experiment / Victoria Ramos de Oliveira; advisor: Carla Göbel Burlamaqui de Mello. – 2023.

92 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Física, 2023.

Inclui bibliografia

1. Física – Teses. 2. Física de Charme. 3. Violação de CP. 4. Decaimento de três corpos. I. Mello, Carla Göbel Burlamaqui de. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Física. III. Título.

CDD: 530

To my parents, my brother and my boyfriend for their support
and encouragement.

Acknowledgments

First, I would like to thank God for giving me the strength to overcome all the difficult paths and for carrying me until here. Without Him none of this would have been possible. Every time I was scared to go on, I always found comfort in His words and His presence with me, specially in the little things that would ended up making my day, like someone singing in the metro exactly what I needed to hear. God is good all the time, all the time God is good.

I would like to thank my parents, Victor and Rosemeire, for all the unconditional love and support. I know how hard they worked so I could study and follow my dreams. This work presented here is yours too. It is not easy to stay away from home but you never let me felt alone and were always just a phone call away. Everything I am today is because of you, thank you so much! I love you guys with all my heart.

I would also like to thank my brother, Lucas. When I was just a little girl all I asked for was a brother and I've loved you since the moment I found out you were coming to this world. Thank you for being my brother, my friend and for being there every time I needed, I love you very deeply.

To my boyfriend, Guilherme, thank you for have always believed in me and pushed me forward, cheering and celebrating every achievement. Your love and support carried me through all this process and made it so much lighter even 200 km away. It was not easy being in different cities but you always find a way to be present and cheer me up. I love you so much.

To my supervisor Carla Göbel, thank you for the opportunity of working with you and for guiding me through all this project. Being your student was a pleasure and truly enriched my development as a scientist. Thank you for every kind word (and chocolate) when I was desperate and also for the hard words that I sometimes needed to hear, it made all the difference. And thank you for always take the time to go one step back and explain the things I don't know. You're an inspiration for me as a scientist and woman in physics.

To the LHCb Rio Charm Group professors, Alberto Reis, Sandra Amato and Erica Polycarpo, I would like to thank for all the physics discussions and many good suggestions to improve my analysis, not only on the weekly group meetings but everytime I asked for help. I would also like to thank my group colleagues "charmosinhos", Renata, Gustavo, Alexandre, Ivanildo, Caio and Leonardo, for the informal meetings, helping sessions, the drinks on conferences and Munchkin matchs that made this work so much easier. To Juan and Lucas Falcão who welcomed me and helped me a lot during my stay at CERN. And finally, an special thank to Beatriz, Felipe and Lucas Romão for the friendship we build and that made this journey more lighter and fun, whether it was debugging a code, drinking Pisco or making inside jokes.

I would also like to thank professor Oswaldo Del Cima, who was my supervisor as an undergrad student and first introduced me to the world of particle physics. Thank your for encourage me to pursue this path and for lending me your apartment at Rio these two years during my Masters, you're a life saver.

To my friends at PUC, in special the people from Cavernão. It was not easy to start this Master during pandemic times and especially changing to a new Institution, but you guys made me feel welcome since the begging. Thank you for all the dart matchs, early lunches and for make my days more lively and joyful. I would also like to thank all the administrative staff at PUC-Rio, Gizelda, Eduardo, Marcia and Julinho, and the excellent professors I had.

Finally, I also want to express my gratitude to all my friends that even from far away have supported me, in particular Ana and Joyce who are the kindest persons I know and for whom I'm grateful to have been friends for more than 10 years. Love you both!

To CNPq, FAPERJ and PUC-Rio, I am grateful for the aids granted, without which this work would not have been accomplished.

"This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001".

Abstract

Oliveira, Victoria Ramos de; Mello, Carla Göbel Burlamaqui de (Advisor). **Search for CP violation in the $D^+ \to \pi^- \pi^+ K^+$ phase space in the LHCb Experiment**. Rio de Janeiro, 2023. 92p. Dissertação de Mestrado – Departamento de Física, Pontifícia Universidade Católica do Rio de Janeiro.

The Standard Model (SM) of particle physics is currently the best theory to describe the interactions between elementary particles and their properties. The observation of CP violation in a variety of weak processes in the quark sector — allowing an absolute distinction of particles and antiparticles — is well explained by the SM (through so-called Cabibbo-Kobayashi-Maskawa ansatz). CP violation is one of the necessary conditions for baryogenesis and might be the key to explain the matter-antimatter asymmetry in the universe.

This dissertation presents the search for CP violation in the doubly Cabibbo suppressed (DCS) $D^+ \to \pi^- \pi^+ K^+$ decay, using data collected by LHCb from 2016-2018 of *pp* collisions with a centre of mass energy of 13 TeV, corresponding to an integrated luminosity of 5.6 fb^{1}. The goal of this analysis is the implementation of a model-independent technique to perform a statistical comparison between the phase space, called Dalitz Plot (DP), of particle and antiparticle decay channel, searching for local differences in the distribution of events between the two DPs. First, a selection process was executed to remove background contributions from other charm decays, as well as a multivariate analysis using machine learning algorithms to reduce combinatorial background levels. After the selection, a final sample of ¨6M signal candidates was obtained, which is nowadays the largest sample ever obtained for a DCS D^+ decay channel, allowing an outstanding sensitivity for CPV search. This is the first CPV search in the studied decay channel. This analysis is performed blinded, meaning that there is no actual result for the signal region at this first stage and in order to guarantee that there are no nuisance asymmetries, from production and detection effects, tests were performed using the background region, the control channel $D^+ \to K^- \pi^+ \pi^+$ and the Monte Carlo simulated sample.

Keywords

Charm physics; CP violation; Three-body decay.

Resumo

Oliveira, Victoria Ramos de; Mello, Carla Göbel Burlamaqui de. **Busca de violação de CP no espaço de fase do decaimento $D^+ \to \pi^-\pi^+K^+$ no Experimento LHCb**. Rio de Janeiro, 2023. 92p. Dissertação de Mestrado – Departamento de Física, Pontifícia Universidade Católica do Rio de Janeiro.

O Modelo Padrão (SM) da física de partículas é atualmente a melhor teoria para descrever as interações entre partículas elementares e suas propriedades. A observação da violação do CP em diversos processos fracos no setor de quarks - permitindo uma distinção absoluta entre partículas e antipartículas - é bem explicada pelo SM (através do chamado Cabibbo-Kobayashi-Maskawa *ansatz*). A violação do CP é uma das condições necessárias para a bariogênese e pode ser a chave para explicar a assimetria matéria-antimatéria no universo.

Esta dissertação apresenta a busca por violação de CP no decaimento $D^+ \to \pi^-\pi^+K^+$ duplamente suprimido por Cabibbo (DCS), usando dados coletados pelo LHCb de 2016-2018 de colisões *pp* com uma energia de centro de massa de 13 TeV, correspondendo a uma luminosidade integrada de 5,6 fb^{1}. O objetivo desta análise é a implementação de uma técnica independente de modelo para realizar uma comparação estatística entre o espaço de fase, denominado Dalitz Plot (DP), de partícula e antipartícula no canal de decaimento, buscando diferenças locais na distribuição de eventos entre os dois DPs. Primeiro, foi feito um processo de seleção para remover contribuições de *background* de outros decaimentos de charme, bem como uma análise multivariada utilizando algoritmos de aprendizado de máquina para reduzir os níveis de *background* combinatoriais. Após a seleção, foi obtida uma amostra final 6M de candidatos a sinal, que é hoje a maior amostra já obtida para um canal de decaimento D^+ DCS, permitindo uma excelente sensibilidade para busca de VCP. Essa é a primeira busca de VCP no canal de decaimento estudado. Esta análise é realizada de forma cega, o que significa que não há resultado para a região do sinal nesta primeira etapa e, para garantir que não haja assimetrias espúrias, como efeitos de produção e detecção, foram realizados testes para a região de *background*, para o canal de controle $D^+ \to K^-\pi^+\pi^+$ e para a amostra simulada de Monte Carlo.

Palavras-chave

Física de Charme; Violação de CP; Decaimento de três corpos.

Table of contents

List of figures

List of tables

A estrada vai longe. É preciso perder pra ganhar. Entender é o escombro.

Cícero Rosa Lins, *Banzo.*

1
Introduction

The observed asymmetry between matter and antimatter in the universe is, nowadays, one of the biggest unanswered problems in physics. One of the necessary conditions to explain this asymmetry is the violation of the combined symmetries of charge conjugation (C) and parity (P) [1]. C and P are two important discrete transformations in particle physics and the product CP was, for a long time, thought to be an exact symmetry of nature, which is indeed true for strong and electromagnetic interactions, but not for weak processes [2]. Since the first observation of CP violation in neutral kaons system in 1964 [3], it was disproved the general belief that particle and antiparticle would behave in the same way and it brought to light the decay rate asymmetry between a process and its associated conjugate [4].

The Standard Model (SM) of particle physics is currently the best theory to describe the interactions between elementary particles, and their properties. CP violation arises within the SM through a complex phase in the quark mixing matrix of the weak interaction. Even though this description helps to explain the baryon asymmetry [5], it is not complete and new sources of CP violation beyond the SM are necessary [6].

The observation of CP violation in weak interactions is firmly established for K and B mesons, but it was only recently that it was first observed for neutral D mesons [7]. The predictions within the SM for charm decays are very small [4].

For charm decays, according to the SM, CPV occurring directly from decays (i.e. not associated to D^0- $\overline{D}^0$ mixing) is only predicted for so-called Cabibbo suppressed processes, observed through the different interference pattern for $D^0(D^+)$ and $\overline{D}^0(D^-)$ between amplitudes leading to the same final state. Other than that, for three-body decays, the two-dimensional phase space, called Dalitz plot (DP), may present an ample pattern of interfering resonant states that might produce local CP asymmetries larger than phase-space integrated ones. Due to this, one possible approach to search for CP violation effects around the phase space is by a statistical comparison between the DP for particle and antiparticle, looking for local differences in the distribution of events in the two DPs.

For this dissertation, the decay $D^+ \to \pi^-\pi^+K^+$ is studied. This decay is doubly Cabibbo suppressed (DCS) and, as such has no expectation of CPV effects within the SM, so any observation of this phenomena would be an indication of new physics (NP). Furthermore, besides the possibility of being a source of NP effects, DCS decays may play an important role in helping to understand the weak decay mechanism of charm hadrons [8]. This work uses data samples obtained in pp collisions with a centre-of-mass energy of 13 TeV collected by the LHCb detector during 2016-2018, that corresponds to an integrated luminosity of 5.6 fb^{-1}. The $D^+ \to \pi^-\pi^+K^+$ sample analysed here is nowadays the largest sample ever obtained for a DCS D^+ decay channel, allowing correspondingly and unprecedented sensitivity for CPV searches.

The remainder of this dissertation is structured as follows. In Chapter 2, the theoretical fundamentals of particle physics and the SM are briefly discussed, followed by a discussion of CP violation, the charm sector and 3-body decays.

In Chapter 3, an overview of the LHCb experiment is presented, with the description of the technical parts of the detectors and of the LHC accelerator. Chapter 4 reports the selection processes towards our final samples of $D^+ \to \pi^-\pi^+K^+$ decays, with also a description of simulation samples, and the description of the full requirements applied to reduce the background level on the samples.

Chapter 5 describes the technique to search for CP violation, the *Mirandizing*, in the $D^+ \to \pi^-\pi^+K^+$ sample. The results of the studies performed to search for asymmetries in the background region, on the control channel $D^+ \to K^-\pi^+\pi^+$ and on simulated samples are also presented, with the prospects of this analysis for the *unblind* procedure, that is, the implementation of the actual method to the signal sample.

Finally, Chapter 6 presents the conclusions of this dissertation. Additional information concerning this analysis are included as an Appendix.

2
Theoretical Fundamentals

2.1
The Standard Model

The Standard Model (SM) of Particle Physics is currently the theory best describing the properties of the Universe's fundamental constituents, the elementary particles, and their interactions. The SM model particles are divided into two groups: the bosons and the fermions. The spin-1 (vector) bosons are called force-carrier particles [9]:

– photon (γ): is the gauge boson of the QED (Quantum Electrodynamics) and has zero mass and zero electric charge. It is responsible to mediate the electromagnetic interaction.

– gluon (g): is the gauge boson of QCD (Quantum Chromodynamcs), is massless just like the photon and mediates the strong interaction. The charge of QCD is the *colour*.

– W^+, W^- and Z^0: are the gauge bosons that mediate the weak interaction described through the electroweak theory. Differently from the photon and the gluon, these bosons have mass and are in fact much more massive than a proton. The W^+ and W^- mediate the weak charged-current interaction and are responsible for the nuclear β-decay, among others, while the Z boson is the carrier of the weak neutral-current.

The other group of the SM particles are the fundamental fermions. They have half-integer spin and are divided into two groups within the SM, the quarks and the leptons. As can be seen on Fig. 2.1, each group has six particles gathered in doublets, giving rise to three families of particles (or generations). The lightest and most stable particles are those from the first generation. As the generation increases, there is also an increase in the mass. Due to the charged weak force, the second and third generation represent less stable particles, that quickly decay to more stable ones. The six known leptons are denoted by $e, \nu_e, \mu, \nu_\mu, \tau,$, and ν_τ, and by definition do not take part in strong interactions. The particles that do take part in strong interactions are the six quarks u, d, s, c, b, and t. The different species of quarks and leptons are called *flavours*. The quarks are defined collectively as q and do not propagate freely due to a phenomenon called colour confinement and so, they are found only in compound states, the hadrons. The hadrons are typically baryons (bound

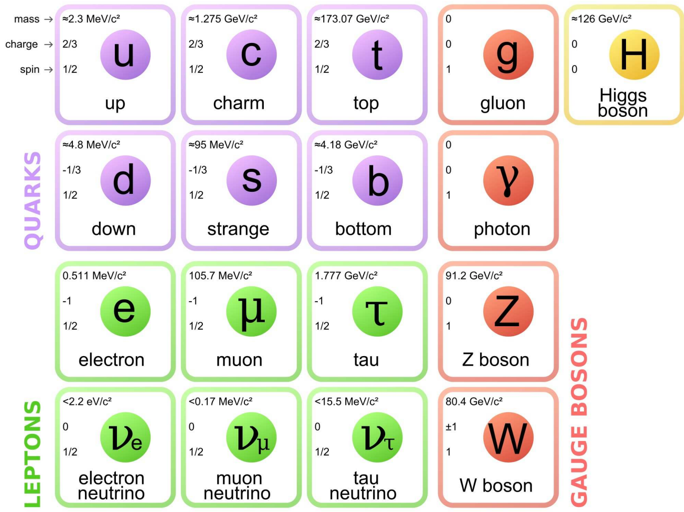


Figure 2.1: Standard Model of Elementary Particles representation [10].

states of three quarks or antiquarks, like protons - uud - and neutrons - udd) or mesons (bound state of a quark and an antiquark - $q\bar{q}$), although recently new states like tetraquarks ($q\bar{q}q\bar{q}$) and pentaquarks ($qqqq\bar{q}$) have been observed.

The last known element of the SM is the Higgs boson, that is different from all the other particles. It is a spin-0 particle and is the only scalar boson in the theory. The Higgs boson has a mass of $m_H \approx 125$GeV and was discovered in 2012 by the ATLAS and CMS experiments at the Large Hadron Collider (LHC) [11, 12]. It is responsible for the mechanism through which all other particles acquire mass, known as the Brout-Englert-Higgs (BEH) Mechanism.

The SM is a non-abelian gauge theory with spontaneous symmetry breaking and a gauge symmetry group given by $SU(3)_C \times SU(2)_L \times U(1)_Y$ before broken. The interactions can be summarised by the transformation properties with respect to this gauge group and the algebra generators of the group with whom the gauge bosons are associated with [13]:

$$
\begin{array}{ccccc}
SU(3)_C & \times & SU(2)_L & \times & U(1)_Y \\
\downarrow & & \downarrow & & \downarrow \\
8G^\alpha_\mu & & 3W^\alpha_\mu & & B_\mu \\
\alpha = 1, ..., 8 & & \alpha = 1, 2, 3 & &
\end{array}
\tag{2-1}
$$

The first part corresponds to QCD and the particles that transform with respect to this are said to be colour-charged and to interact "strongly" with the eight gluons represented by the fields $G^\alpha_\mu(x)$. The factor $SU(2)_L$ is associated to the three spin-1 fields $W^\alpha_\mu(x)$ and the subscript "L" indicates that this quantum number is only carried by left-hand fermions[1]. And finally, the $U(1)_Y$ term is

[1]The chirality of a Dirac fermion ψ is defined through the operator γ^5, that has

associated with one $B_\mu(x)$ and the subscript "Y" is to represent the weak *hypercharge*. The $SU(2)_L \times U(1)_Y$ part of the group relates to the Glashow-Salam-Weinberg model, which takes into account the fact that the SM is chiral, i.e. distinguishes between left-hand and right-hand chirality of the particles. These four bosons are related to the three gauge bosons of the weak interaction $W^{\pm}, Z^0$ that gain mass after the spontaneous symmetry breaking due the interaction with the Higgs field, and the photon that remain massless. All the other fermions also acquire mass through the BEH mechanism.

2.2 CP Violation in the SM

2.2.1 CPT Symmetry

The beauty of a theory is constantly associated with the presence of symmetries in it and, in fact, the last two centuries of physics research can be summarised in the search for symmetries in nature and in the laws by which it is governed. In the study of particle physics two important transformations are the charge conjugation (C) and the parity conjugation (P). Both are discrete operations and the combined product of CP plays an important role in the understanding of not only particle physics but the evolution of the universe, being CP symmetry violation stated as one of the necessary conditions for baryogenesis [1]. Withal, other important discrete operation is the time reversal (T) and with that the studies of CP, CPT and T may provide answers to fundamental questions such as the observed dominance of matter over antimatter in the universe [2].

The charge conjugation C can be described as an operation that conjugates all additive quantum numbers, like the electric charge, the baryonic number and the leptonic number. For a generic state $|\psi\rangle$ with angular momentum J and electric charge Q: $C|\psi(J,Q)\rangle = |\psi(J,-Q)\rangle$. The parity conjugation P, when applied, inverts all spatial coordinates with respect to the origin, $\vec{x} \to -\vec{x}$. However, this transformation do not affect pseudo-vectors neither spin. The conjugation of the combined product CP turns particles into antiparticles and vice-versa. And, at last, the time reversal operation T is the transformation $T : t \to -t$. So, the application of the T operator would have the effect of changing the velocity sign without changing the position: $t \to -t, \quad \vec{x} \to \vec{x}, \quad \vec{p} \to -\vec{p}$.

eigenvalues +1 and -1. The sign of the eigenvalues is equal to the particle's chirality, being +1 for right-handed particles and -1 for left-handed. By the action of the projection operators $\frac{1}{2}(1 \pm \gamma^5)$ on ψ, any Dirac field can be project into its left- or right-handed component.

If CP were an exact symmetry the laws of nature for particle and antiparticle would be the same, and that is indeed true for gravitational, electromagnetic and strong interactions [14]. But in spite of the general belief that CP would also be a good symmetry in weak interactions, the observation of CP violation (CPV) for the first time in processes involving neutral kaons in 1964 [3] brought to light the asymmetry between a process and its associated conjugate at fundamental level, meaning that particle and antiparticle behave differently [4] and can be unambiguously distinguished.

On the other hand, the product CPT is believed to be a valid symmetry and indeed so far all observations indicate that CPT is a fundamental symmetry of nature. The CPT theorem states that *"any quantum field theory based on a Hermitian, local, normal-ordered Lagrangian which is invariant under Lorentz transformations, and for which the usual field commutation and anti-commutation rules hold, is also invariant under the transformation corresponding to the product of C, P, and T, taken in any order, irrespectively of its symmetry under the three inversions separately"* [15]. The CPT symmetry establishes that the lifetime of a particle and is antiparticle are the same, and so thus are the total decay widths.

2.2.2 The CKM Matrix Elements

The SM can be divided into three parts [16]:

$$\mathcal{L}_{SM} = \mathcal{L}_{kinetic} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa} \tag{2-2}$$

For the first one, considering the quark doublets Q_L and the lepton doublets L_L^I,

$$\mathcal{L}_{kinetic} = \begin{cases} i\overline{Q_{Li}}\gamma_\mu \left(\partial^\mu + \frac{i}{2}g_s G_a^\mu \lambda_a + \frac{i}{2}gW_b^\mu \tau_b + \frac{i}{6}g' B^\mu\right)\delta_{ij}Q_{Lj} & \text{, for } Q_L \\ i\overline{L_{Li}}\gamma_\mu \left(\partial^\mu + \frac{i}{2}gW_b^\mu \tau_b - \frac{i}{6}g' B^\mu\right)\delta_{ij}L_{Lj} & \text{, for } L_L^I \end{cases} \tag{2-3}$$

and here G_a^μ, W_b^μ and B^μ are the same ones presented on Eq. 2-1, g and g' are the coupling constants and γ_μ are the Dirac gamma matrices. Besides that, there are also the generators, that play a crucial role in defining the behaviour and interactions of particles and help defining the structure of the Lagrangian:

– L_a for the $SU(3)_C$, with the 3×3 Gell-Mann matrices $\frac{1}{2}\lambda_a$ for triplets and 0 for singlets;

– T_b for the $SU(2)_L$, with the 2×2 Pauli matrices $\frac{1}{2}\tau_b$ for doublets and 0 for singlets;

– Y's for $U(1)_Y$ charges.

This first part of the Lagrangian is flavour universal due to the unit matrix in flavour space δ_{ij} and it also conserves CP.

The second term on Eq. 2-2 is associated to the Higgs scalar self-interactions,

$$\mathcal{L}_{Higgs} = \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \tag{2-4}$$

where ϕ is the Higgs field. This part also conserves CP since the SM scalar sector has only a single doublet. Last but not least, the Yukawa interaction terms in $\mathcal{L}_{SM}$ are

$$-\mathcal{L}_Y^\eta = \begin{cases} Y_{ij}^d\overline{Q_{Li}}\phi D_{Rj} + Y_{ij}^u\overline{Q_{Li}}(i\tau_2\phi^\dagger)U_{Rj} + \text{h.c.} & \text{, for } \eta = \text{q, quark interaction} \\ Y_{ij}^e\overline{L_{Li}}\phi E_{Rj} + \text{h.c.} & \text{, for } \eta = \text{l, lepton interaction} \end{cases} \tag{2-5}$$

where Y_{ij}^f are the 3×3 complex matrices and i, j are the generation labels. And, in despite of the $\mathcal{L}_{kinetic}$ and $\mathcal{L}_{Higgs}$ terms, this term of the Lagrangian is CP violating and usually flavour dependent ($Y^f \not\propto 1$) [16].

In the SM, the quarks masses and mixings arrives from the Yukawa interactions with the Higgs condensate, present on Eq. 2-5, when ϕ acquires a vacuum expectation value ($\langle\phi\rangle = (0, v/\sqrt{2})$) and yields mass terms for the quarks. Thus, the Yukawa interactions give rise to the mass matrices,

$$M_q = \frac{v}{\sqrt{2}}Y^q \tag{2-6}$$

The mass basis corresponds, by definition, to diagonal mass matrices and one can always find unitary matrices V_{qL} and V_{qR} such that

$$V_{qL}M_qV_{qR}^\dagger = M_q^{diag} \equiv \frac{v}{\sqrt{2}}\lambda_q \tag{2-7}$$

The four unitary matrices V_{dL}, V_{dR}, V_{uL} and V_{uR} are then the ones required to transform to the mass basis. The charged-current $W^\pm$ interactions for quarks, that is the interactions of the charged $SU(2)_L$ gauge bosons $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2)$, couple to the physical quarks with a more complicated form in the mass basis,

$$-\frac{g}{\sqrt{2}}\overline{U_{Li}}\gamma^{\mu}V_{ij}D_{Lj}W_{\mu}^{+} + \text{h.c.} \tag{2-8}$$

where U_L and D_L are the left-handed quark mass eigenstates of *up*-type (u, c, t) and *down*-type (d, s, b), respectively. The V_{ij} above represent the elements of a 3×3 unitary matrix

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{2-9}$$

and this is the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix for quarks, which measures the mis-match between the matrices that diagonalize the U and D quark mass terms. As a result of V_{CKM} not being diagonal, the $W^{\pm}$ gauge bosons can couple to quark mass eingenstates of different generations, with a strength given by the relative element of the matrix [9]. Although the structure of the weak interactions of quarks and leptons is the same, the leptonic sector presents an almost perfect universality in the weak charged-current, that is, the different leptons, electron (e^-), muon (μ^-) and tau (τ^-) have the same interaction strengths. This picture is different in the quark sector since the weak eigenstates and the mass eigenstates are different.

The CKM matrix is an extension of the 2×2 Cabibbo mixing matrix [17]. In order for the mixing of quarks flavour to restore the universality of weak interactions, Cabibbo introduced a mixing angle that is now known as the Cabibbo angle θ_C. With this, the weak eigenstates d', s' are related to the mass eigenstates by,

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} \tag{2-10}$$

However, as 2×2, this matrix is real and no CP violation can occur since CPV requires a complex phase difference between the quark and anti-quark transitions, which is not possible with only two quark families [18].

A matrix with complex elements would open the possibility for CPV. The solution to this was proposed in 1973 by Kobayashi and Maskawa [19], who have included a third family of quarks and extended the 2×2 matrix to a 3×3 one that is the nowadays known quark mixing matrix. It has a complex phase, which introduces a mismatch between the quark and anti-quark transitions.

The CKM matrix is a mandatory ingredient for CPV in the SM in the quark sector.

The CKM matrix can be parameterised in many equivalent ways by three mixing angles and a CP-violating phase. A standard choice is [20]

$$V_{CKM} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{2-11}$$

where $s_{ij} = \sin\theta_{ij}$, $c_{ij} = \cos\theta_{ij}$ and δ is the phase that allows CPV to arise in flavour changing processes in the SM. All the angles can be chosen to lie in the first quadrant, so $s_{ij}, c_{ij} \geq 0$, and each angle is labelled with the indexes corresponding to the mixing of two families.

From experiments, it is known that $s_{13} \ll s_{23} \ll s_{12} \ll 1$. This hierarchy can be evidenced by a parametrization where the four mixing parameters are (λ, A, ρ, η), instead of the three angles and one complex phase, and $\lambda = \sin\theta_C \approx |V_{us}| \approx 0.23$ plays the role of the expansion parameter whiles η accounts for the CPV phase. This is known as the Wolfenstein parametrization [21],

$$V_{CKM} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda + \frac{1}{2}A^2\lambda^5[1 - 2(\rho + i\eta)] & 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4[1 + 4A^2] & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 + \frac{1}{2}A\lambda^4[1 - 2(\rho + i\eta)] & 1 - \frac{1}{2}A^2\lambda^4 \end{pmatrix} + \mathcal{O}(\lambda^6) \tag{2-12}$$

It is easy to see on matrix 2-12 how different the magnitude of one matrix element is from the other, and also that the matrix is almost diagonal and almost real. Likewise, a closer look reveals that the diagonal elements tend to approach to one while the off diagonal terms tend to get smaller as farther they get from diagonal. Since the CKM matrix elements are fundamental parameters of the SM, it is important to determine them precisely. From the unitarity of the CKM matrix ($V^\dagger_{CKM}V_{CKM} = I$) some conditions are implied

$$\begin{array}{lcl} \sum_i V_{ij}V_{ik}^* = \delta_{jk} & , & i = u, c, t \text{ and } j, k = d, s, b \\ \sum_j V_{ij}V_{kj}^* = \delta_{ik} & , & j = d, s, b \text{ and } i, k = u, c, t \end{array} \tag{2-13}$$

and from the six vanishing combinations present in Eq. 2-13, the so-called Unitary Triangles can be constructed and are defined in the complex plane, all of them with the same area, but most of them very squashed. The most important one has sides of similar size:

$$\begin{array}{c} V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \\ \downarrow \\ \dfrac{V_{ud}V_{ub}^*}{V_{td}V_{tb}^*} + 1 + \dfrac{V_{td}V_{tb}^*}{V_{td}V_{tb}^*} = 0 \end{array} \tag{2-14}$$

for which the internal angles can be defined as

$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad \gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) \tag{2-15}$$

as can be seen in Fig. 2.2

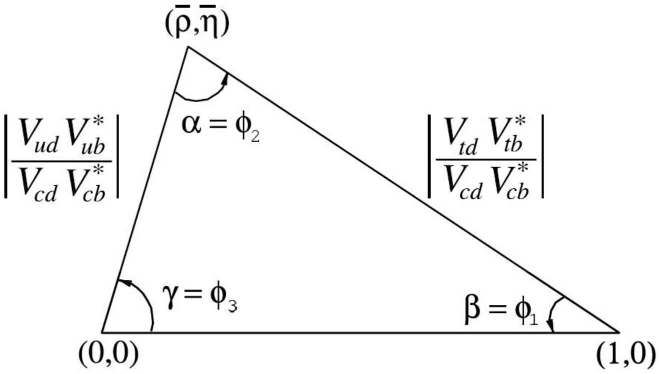


Figure 2.2: Sketch of the unitary triangle [22].

Since the CKM matrix is almost diagonal, quarks have a tendency to change flavour inside the same family, as illustrated in Fig. 2.3.

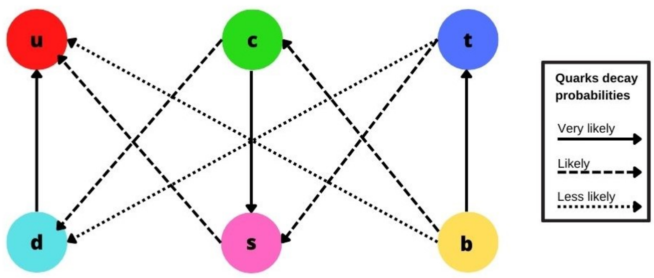
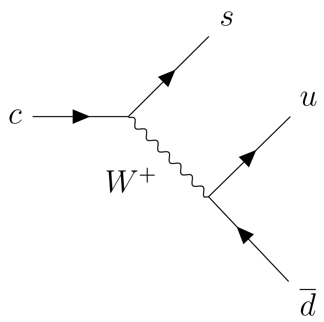


Figure 2.3: Probabilities of quarks decay proportional to $|V_{ij}|^2$.

For the case of non-leptonic weak quark transitions, the decay amplitude is proportional to the product of the two matrix elements present in the processes. Furthermore, for charm decays, only the 2×2 Cabibbo submatrix is relevant up to $\mathcal{O}(\lambda^4)$ and the flavour changes are classified based on their λ suppression as:

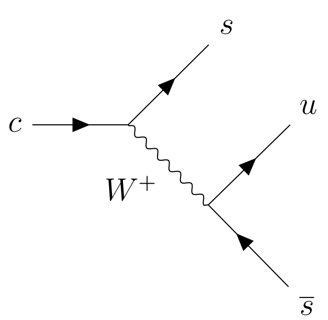
– Cabibbo Favoured (CF): both are diagonal elements, $c \to s(u\bar{d})$

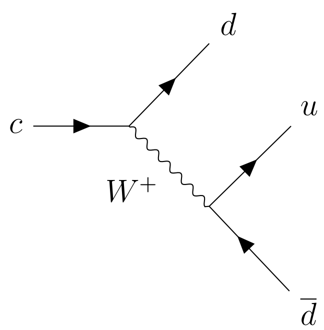
$$V_{cs}V_{ud}^* \sim \mathcal{O}(1)$$



– Cabibbo Suppressed (CS): one element of the diagonal and one element off-diagonal, $c \to s(u\bar{s})$ or $c \to d(u\bar{d})$

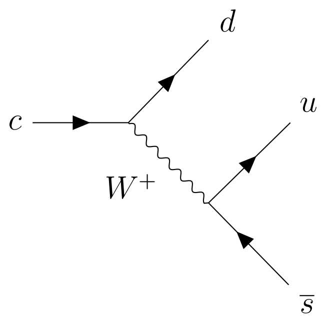
$$V_{cs}V_{us}^* \sim \mathcal{O}(\lambda)$$





– Doubly Cabibbo Suppressed (DCS): both elements are off-diagonal, $c \to d(u\bar{s})$

$$V_{cd}V_{us}^{*} \sim \mathcal{O}(\lambda^2)$$



2.3 CP violation in decays

As it has been discussed, CPV arises in the SM owing to an irreducible phase in the CKM matrix. For this phase to be observable, interference is needed between processes with different weak phase leading to the same final state. In the context of hadrons decay, there is a rich phenomenology of CP violation that can be classified in terms of how this complex phase manifests [15]. Suppose a hadron P decaying into a final state f with decaying amplitude $A(P \to f)$. The associated charge-conjugated process is $\overline{P} \to \overline{f}$ with amplitude $A(\overline{P} \to \overline{f})$. Three special cases are [2, 18]:

(A) the particle decay rate of $P \to f$ differs from the conjugate process $\overline{P} \to \overline{f}$. $\Longrightarrow$ **Direct CP violation**

(B) when both P and $\overline{P}$ decay to the same final states $P \to f \leftarrow \overline{P}$. The CP asymmetry is a result of these two decay routs interference. This process is only possible for neutron hadrons. $\Longrightarrow$ **Indirect CP violation in mixing**

(C) the decay $\overline{D}^0 \to K^+\pi^+$ can occur directly or through the intermediate process $\overline{D}^0 \to D^0 \to K^+\pi^+$, where $\overline{D}^0$ oscillates to D^0 introducing a phase difference. Once again, this process only occurs with neutron hadrons, where particle and antiparticle can oscillate between their identities . $\Longrightarrow$ **Indirect CP violation in the interference between mixing and decay**

For charged charm decays (as the $D^+ \to \pi^-\pi^+K^+$ discussed in this dissertation), the only possible way this effect could be observed is by direct CP violation. However, this can only take place if the decay amplitude can be broken into at least two different paths. This interference produces complex phases that can be of two types: CP-even and CP-odd, regarding to the change of sign under a CP transformation. Within the SM, the CP-odd phases come strictly from some weak transitions whereas the CP-even phases appear due to strong or electromagnetic interactions. For this matter, the former is usually referred to as weak phase and the second as strong phase [4].

The simplest possible case is to consider two decay amplitudes resulting in the same final state f,

$$\begin{aligned} A(P \to f) &= |A_1|e^{i(\delta_1+\phi_1)} + |A_2|e^{i(\delta_2+\phi_2)} \\ A(\overline{P} \to \overline{f}) &= |A_1|e^{i(\delta_1-\phi_1)} + |A_2|e^{i(\delta_2-\phi_2)} \end{aligned} \tag{2-16}$$

where $\delta_{1,2}$ are the strong phases, $\phi_{1,2}$ are the weak phases and $A_{1,2}$ are the amplitudes of the intermediate processes. A CP violation observable that can be defined is the asymmetry A_{CP},

$$A_{CP} = \frac{|A(P \to f)|^2 - |A(\overline{P} \to \overline{f})|^2}{|A(P \to f)|^2 + |A(\overline{P} \to \overline{f})|^2} \tag{2-17}$$

which, for the example of Eq. 2-16 becomes,

$$A_{CP} = \frac{2A_{21}\sin(\delta_1-\delta_2)\sin(\phi_1-\phi_2)}{1 + A_{21}^2 + 2A_{21}\cos(\delta_1-\delta_2)\cos(\phi_1-\phi_2)} \tag{2-18}$$

where $A_{21} = |A_2/A_1|$. So, from Eq. 2-18 it is clear that in order for direct CPV to be observed it must exists a difference between the weak phases of the two processes and also a difference between the strong phases, resulting in $A_{CP} \neq 0$. Wherefore, both strong and weak phases are necessary for CP violation. The weak phase has its origins, within the SM, in the CKM matrix and it is thus possible to predict which decays could potentially manifest CPV effects. On the other hand, the strong phase origin could be attributed to two sources: short-distance effects, like due to penguin diagrams (one-loop processes in which a quark changes flavour, via a W loop) contribution, or long-distance effects from final-state interactions.

2.3.1
CP violation in charm decays

The study for CP violation is an important tool for the search of New Physics (NP) beyond the SM. The future leaves room for a variety of experiments and for additional sources of CP violation [23]. This phenomenon is already established experimentally for the mesons K and B, involving the quarks *strange* and *beauty*, respectively. Still, it was not until 2019 that it was observed the direct CPV on the charm sector in D decays, with the result $\Delta A_{CP}(KK - \pi\pi) = A_{CP}(D^0 \to K^-K^+) - A_{CP}(D^0 \to \pi^-\pi^+) = (-15.4 \pm 2.9) \times 10^{-4}$ by the LHCb Collaboration [7].

Direct CPV effects in the charm sector are predicted to be very small within the SM, with expectation size of asymmetries in the decay rates of D and $\overline{D}$ of order of $\mathcal{O}(10^{-4})$ or less [14] and are only possible in singly Cabibbo suppressed decays, where different amplitudes can produce final states with the same flavour content [4]. For these decays, the CPV effects may occur through the interference between the $c \to d(u\overline{d})$ tree and highly suppressed $c \to u(\overline{q}q)$ penguin processes, flavour SU(3) symmetry breaking or rescattering effects. Final state interactions, like rescattering effects, have been recently demonstrated in theoretical studies to play an important role in CPV effects [24], producing the interference necessary to magnify the CPV in the $D^0 \to \pi^-\pi^+$ and $D^0 \to K^-K^+$ amplitude decays [25].

Therefore, the analysis of $D^+ \to \pi^-\pi^+K^+$,[2] that is a doubly Cabibbo suppressed decay, is a great laboratory to search for new physics effects. Besides that, the search for direct CPV in three-body charm decays has also the advantage of potentially being more sensitive to localised effects in the phase space, since these asymmetries can be larger than the integrated ones [16].

2.3.2 $D^+ \to \pi^-\pi^+K^+$

This decay is a doubly-Cabibbo suppressed mode with a quark level transition $c \to d(u\overline{s})$ that can have mainly three types of topology contributing: annihilation and tree-level emission of W (with and without colour suppression), as can be seen in Fig. 2.4.

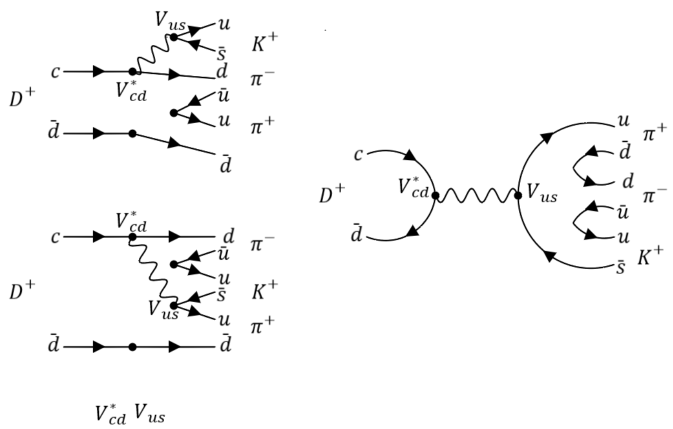


Figure 2.4: $D^+ \to \pi^-\pi^+K^+$ topologies: tree-level emission of W (top left), tree-level emission of W colour suppressed (bottom left) and annihilation diagram (Right).

[2]Charge conjugated decays are implicit, otherwise mentioned explicitly.

Also, the $d\bar{d}$ (or $u\bar{u}$) and $d\bar{s}$ pairs may form different resonances as intermediate states which decay to $\pi^-\pi^+$ and $K^+\pi^-$, respectively, creating a pattern in the phase space. These decay combinations are explicit in Fig. 2.5. From a previous analysis [26], the resonant structures are dominated by $\rho(770)^0$ and $K^{*0}(892)$ and these vector resonances account for about 90% of the decay fraction, although there is also a presence of the tensor $K_2^*(1430)$ in the $m^2_{\pi K}$ mass combination, and of the scalar $f_0(980)$ in $m^2_{\pi\pi}$. In the SM, there are no processes with different weak phases leading to the same final state and, hence, no prediction of CPV effects for this decay. Therewith, this study opens the possibility to search for NP beyond the SM.

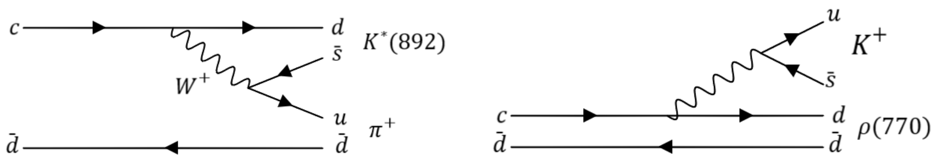


Figure 2.5: $D^+ \to \pi^-\pi^+K^+$ with main resonances highlighted.

2.4 Tree-body decays

Differently from the two-body decays, where the modulus of the momenta of the final state particles are determined by the energy-momentum conservation relations, the three-body final state has a much richer dynamics [27]. For a mother particle with mass M, spin-0 and 4-momentum P decaying in three daughter particles with masses m_1, m_2 and m_3, 4-momentum p_1, p_2 and p_3 and also being spin-0, the energy-momentum conservation constraints are:

$$\begin{aligned} E_P &= \sum_{i=1}^{n} E_i \\ E_i^2 &= m_i^2 + \vec{p}_i^2 \\ \vec{p}_P &= \sum_{i=1}^{n} \vec{p}_i \end{aligned} \tag{2-19}$$

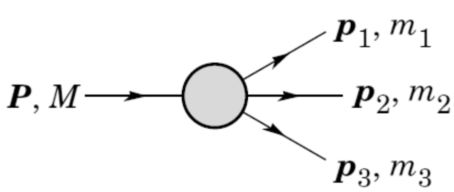


Figure 2.6: Schematic of a three-body decay [14].

This system has 12 degrees of freedom (d.o.f.) due to the three 4-vector for momenta of the final state particles. However, these 12 variables are constrained by the four energy-momentum conservation equations and mass relations constraints, which reduces this number to 5 d.o.f. and, beyond that, since the initial state is isotropic in the mother's rest frame, the final state cannot depend on the three Euler angles describing its orientation as a whole. For this, the daughters must be produced in an isotropic way [27] and other three d.o.f. are removed. In this way, only two independent variables are left, as can be seen on Table 2.1, and the kinematic quantities can be written in terms of these.

Constraints	**Degree of freedom**
3 four-vectors	12
Energy-Momentum conservation relations	-4
3 Masses relations	-3
3 Euler angles	-3
Remaining d.o.f	**2**

Table 2.1: Three-body decay degrees of freedom.

For these three particles, the distribution of momenta can populate a 2-dimensional representation called phase space. The phase space defined in terms of invariant 2-body combination masses is called the Dalitz plot (DP) and includes all possible momenta configuration that may happen in the decay in the mother's centre of mass. These invariant masses are defined as

$$\begin{aligned} s_{12} &= (p_1 + p_2)^2 = (P - p_3)^2 \\ s_{13} &= (p_1 + p_3)^2 = (P - p_2)^2 \\ s_{23} &= (p_2 + p_3)^2 = (P - p_2)^2 \end{aligned} \tag{2-20}$$

and from this, the following relation can be obtained,

$$s_{12} + s_{13} + s_{23} = s + m_1^2 + m_2^2 + m_3^2 \tag{2-21}$$

where $\sqrt{s} = M$ is the mother's mass.

The boundaries of the DP are defined uniquely by M, m_1, m_2 and m_3. Since energy and momentum are conserved, there are maximal and minimal possible values defining these boundaries,

$$(m_i + m_j)^2 \quad \leq \quad s_{ij} \quad \leq \quad (M - m_k)^2 \qquad , \quad i, j, k = 1, 2, 3 \tag{2-22}$$

that is, when s_{ij} is minimal, particles i and j are produced at rest, in the mother's centre of mass referential, so p_k is maximal. On the other hand, when

s_{ij} is maximal, particle k is produced at rest and p_i and p_j are maximal and moving in opposite directions, so $s_{ij} = M - m_k$ and all the kinetic energy is transferred to the pair ij. Choosing s_{12} and s_{13} variables, the DP kinematics are determined by two functions, s_{12}^{+} and s_{12}^{-} [27], the superior and inferior limits, respectively,

$$s_{12}^{\pm} = m_1^2+m_2^2-\frac{1}{2s_{23}}[\,(s_{23}-s+m_1^2)(s_{23}+m_2^2-m_3^2)\mp\lambda^{\frac{1}{2}}(s_{23},s,m_1^2)\lambda^{\frac{1}{2}}(s_{23},m_2^2,m_3^2)] \tag{2-23}$$

where $\lambda(x,y,z) = (x-y-z)^2$ is the Källén function. A generic DP can be seen on Fig. 2.7.

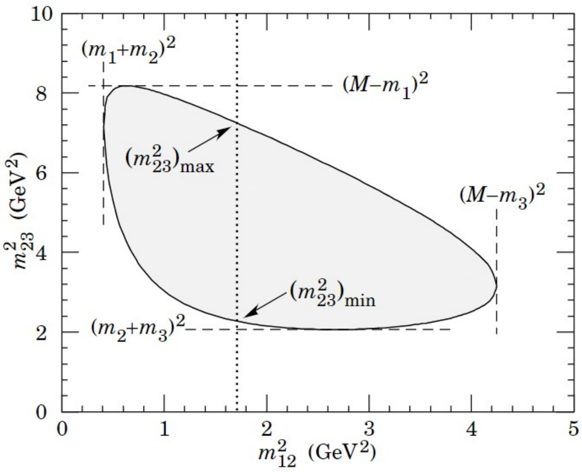


Figure 2.7: Generic Dalitz plot for a three-body final state [14].

Three-body decays in principle could occur through two different topologies: the three particles of the final state are produced directly from the mother particle, the so-called non-resonant decay; or by the formation of resonances as intermediate states that can decay strongly to the final state particles, the resonant decay. A sketch of these processes is represented on Fig. 2.8, where m_3 in the resonant decay is referred to as the companion hadron and is assumed to not interact with the $m_{1,2}$ system.

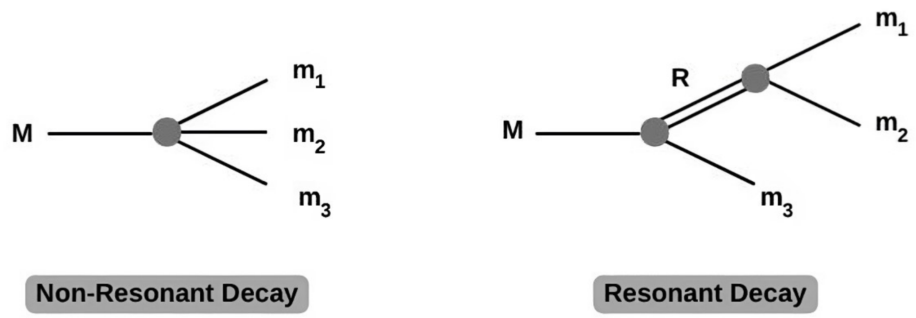


Figure 2.8: Representation of a non-resonant and a resonant tree-body decay.

All the information about the dynamics of the decay is given by the differential decay rate, that is defined by [14]

$$d\Gamma = \frac{1}{(2\pi)^3}\frac{1}{32M^3}|A|^2 ds_{12}ds_{13}. \tag{2-24}$$

In Eq. 2-24, $|A|^2$ is the square modulus of the decay amplitude and contains all the decay dynamics information. If this quantity is constant, the events will uniformly populate the allowed region of the DP. In this sense, a non-uniformity in the plot provides visually apparent information of the dynamics, for example in the case of $D^+ \to \pi^-\pi^+K^+$, bands are seen when $m_{K\pi} = m_{K^*(892)}$ and $m_{\pi\pi} = m_{\rho(770)}$, reflecting the appearance of the decay chains $D^+ \to K^*(892)\pi^+$ and $D^+ \to \rho(770)K^+$, respectively.

The simplest and most common approach to characterize the contribution of these structures is the Isobar model, in which the total amplitude is described as a coherent sum of all resonant amplitudes [4], as well as a contribution of non-resonant amplitudes (NR),

$$A_{Total} = a_{NR}e^{i\delta_{NR}} + \sum_i a_i e^{i\delta_i} A_i \tag{2-25}$$

where here $a_{NR}e^{i\delta_{NR}}$ describes the NR contribution, and this amplitude is assumed to be constant for D decays. The summation part i accounts for all possible resonant amplitudes that could be produced and the a_i and δ_i are, respectively, the magnitude and the complex phase associated with each resonant contribution, and where the complex phases incorporate both strong and weak phases effects. This model assumes a quasi-two-body approximation (2+1) where the companion hadron is not a product of the resonance and its interaction with the other two mesons is neglected.

Within the Isobar model, the general form of the individual amplitudes is,

$$A_R = {}^J F_P {}^J F_R \times {}^J M_R \times BW_R \tag{2-26}$$

where J is the spin of the resonance; ${}^{J}F_{M,R}$ are the Blatt-Weisskopf damping factors [28] and account for the P and R finite sizes; ${}^{J}M_R = (-2|p_k||p_i|)^J P_L(\cos\theta_{ki}^{R_{ij}})$ is the angular factor to ensure angular momentum conservation, where P_L is the Legendre polynomial, described by the Zemach formalism [29]. And finally, BW_R is usually a relativistic Breit-Wigner propagator given by,

$$BW_R = \frac{1}{m_R^2 - s_{ij} - i m_R \Gamma_R(s_{ij})}. \tag{2-27}$$

These 2-body resonances appear on the DP as a band corresponding to the pair of particles that forms the resonance and, from the angular factor, resonances with different spins have different signatures across the DP. Some examples can be seen on Fig. 2.9

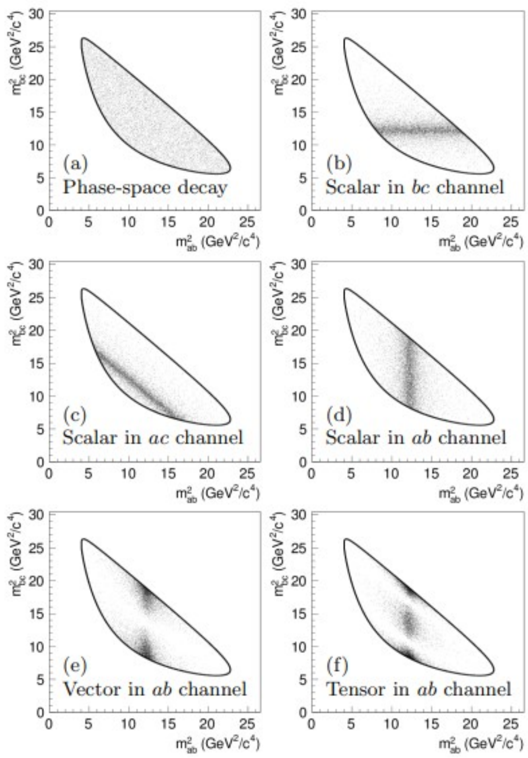


Figure 2.9: Example Dalitz plots for a decay $M \to m_a m_b m_c$ with (a) phase-space decay, (b-d) one scalar resonance appearing in various decay channels, (e, f) vector and tensor resonances. Adapted from [30].

CP violation may arise from the interference between decays through different resonances, potentially introducing differences in the sign and magnitude of the CPV across the DP. Experimentally, this can be measured by comparing the yields of the particle and antiparticle decays through the so-called raw asymmetry,

$$A_{raw} = \frac{N(P \to f) - N(\overline{P} \to \overline{f})}{N(P \to f) + N(\overline{P} \to \overline{f})}. \tag{2-28}$$

which also induces effects caused by production mechanisms for particle and antiparticle, and charge detection asymmetries. These effects, if there, need to be quantified or subtracted in order to have acces to A_{CP}.

As already mentioned, 3-body decays are usually dominated by resonant structures and the distribution of events across the DP is the result of the superposition of the different amplitudes, and the interference patterns depend directly on the strong and weak phases involved. Such a rich dynamics allows for the strong phase to be originated from different sources, enabling localised CP asymmetries to be stronger than phase space integrated ones, and even to change sign [4]. In contrast, a two-body decay does not have a phase space since the daughters momentum is well known and the only possible observable to search for CP violation is the comparison between the total number of events for particle and antiparticle decays.

One possible approach for the search for localised CP effects is to separate the DP for particle (D^+) and antiparticle (D^-) and divide these DPs into small regions, called bins. Then, a direct comparison is performed between the pairs of DPs bins to search for significant differences in the distribution of events for particle and antiparticle. This strategy makes it also possible to pin down the regions of the phase space where the CP violation is manifested. Although the SM predicted effect is small for Cabibbo suppressed charm decays, with enough data it is possible to increase the sensitivity to it. In the cases of existing NP effects, the sensitivity for it could also be enhanced in higher statistics. This can be the case of the DCS decay channel $D^+ \to \pi^-\pi^+K^+$, purpose of this dissertation, for which no search for CPV has yet been carried out.

3
The LHCb Experiment

3.1
The Large Hadron Collider (LHC)

The Large Hadron Collider (LHC) is a two-ring-superconducting-hadron accelerator and collider installed in a 27 km tunnel between 45m and 170m below the surface, localised at CERN (the European Organization for Nuclear Research) in the border of Switzerland and France. The tunnel was originally built for LEP (Large Electron-Positron Collider), between 1984 and 1989, and the existing civil engineering structures were fully used, in addition to some required modifications. The LHC is linked to the CERN accelerator complex, that acts as an injector, by two transfer channels of approximately 2.5 km length each [31]. The accelerator complex is shown in Fig. 3.1. One of the main goals of this huge machine is to answer fundamental questions of Physics, regarding the interactions and structure of the elementary particles, within an attempt to recreate the primordial Universe conditions. Also, it is proposed to validate the SM of Particle Physics as well as search for NP beyond it.

The accelerator was originally designed for proton-proton (pp) collisions with a centre-of-mass energy of 14 TeV and a luminosity of 10^{34} cm$^{-2}s^{-1}$, and could also be used to collide heavy (Pb) ions with an energy of 2.8 TeV

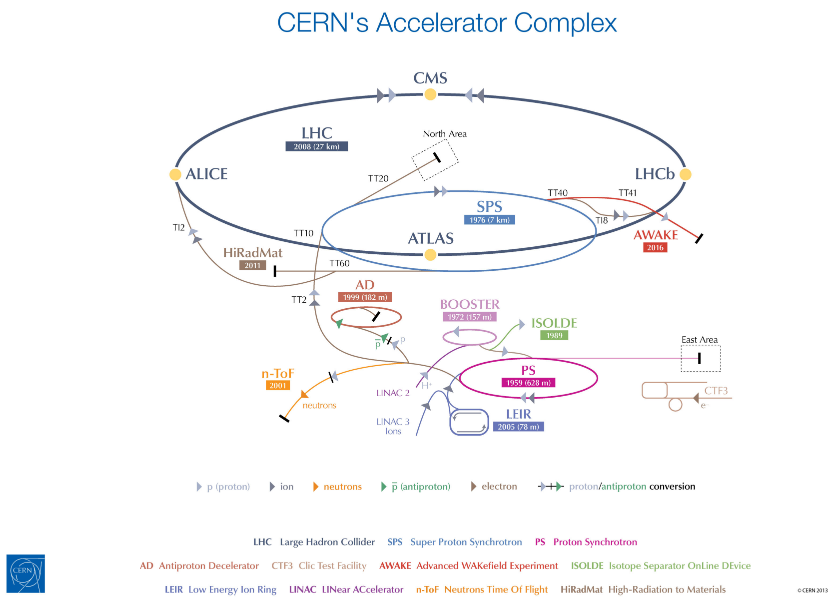


Figure 3.1: CERN accelerator complex [32]

per nucleon and a peak luminosity of 10^{27} cm$^{-2}s^{-1}$ [31]. The luminosity L relates number of events per second generated in the accelerator, N, and the cross-section for the event under study, σ, by $N = L\sigma$. The protons are extracted from a hydrogen gas by ionisation with an electric field and led to the accelerator chain using electromagnetic fields. First, these protons are taken to 50 MeV of energy in a linear accelerator, the Linac2, and then the beam proceeds to a series of synchrotron accelerators. The first is the Proton Synchrotron Booster (PSB) where the protons are accelerated to 1.4 GeV, then on the Proton Synchrotron (PS) they go up to 25 GeV and finally the Super Proton Synchrotron (SPS) where the beam reaches 450 GeV. After this, the proton beam is injected in the LHC and separated in two beam pipes that go on different directions and are accelerated to their maximum energy, currently 6.8 TeV. The two beams are brought into collision in four distinct places, that are the four main experiments of LHC: ATLAS, CMS, ALICE and LHCb.

ATLAS (A Toroidal LHC ApparatuS) is a general-purpose detector designed to exploit a large range of physics phenomena pushing the full discovery potential of the LHC, and it is the biggest of the four main experiments. Together with the CMS, ATLAS is responsible for the measurement of the Higgs boson in 2012. The detector tracks and identifies particles to investigate a large range of physics, from the study of the Higgs boson and top quark to the search for extra dimensions and particles that could make up dark matter, as well as precision measurements of the SM [33].

CMS (Compact Muon Solenoid) is also a general purpose experiment, but more compact. Although the main goals of the detector are essentially the same of ATLAS, it uses different techniques and technologies, and also a different type of magnet-system design [34].

ALICE (A Large Ion Collider Experiment) is an experiment dedicated to study heavy ion physics, focused on the strong interactions of matter on a quark-gluon-plasma state, that forms at extreme energy densities. And also, to better understand the phenomena of confinement and chiral-symmetry restoration [35].

The LHCb (Large Hadron Collider beauty) experiment is dedicated to study the asymmetry between matter and antimatter, searching for CPV effects on decays involving quarks b and c, as well as studying rare decays and spectroscopy. It is an arm forward detector, differently from ATLAS and CMS, that uses a series of sub-detectors to track mainly forward particles [36].

The LHC has finalised two accelerator runs, the Run 1 in 2011-2012, and the Run 2 from 2015-2018. The LHC Run 3 has started in 2022 breaking a new energy world record of 13.6 TeV in its first stable-beam collisions and is scheduled to last until the end of 2025. A long shutdown period is experienced between the runs in order to do the maintenance and upgrade of the detectors.

3.2 The LHCb detector

The LHCb detector, presented in Fig 3.2, is a single-arm forward spectrometer that covers the pseudorapidity range $2 < \eta < 5$ (where $\eta = -\ln[\tan \frac{\theta}{2}]$ and θ is the angle with respect to the beam axis), in the aim to optimise the quantity of particles reconstructed in this angular acceptance and based on the $b\bar{b}$ production angles at the LHC beam energy, as schematised in Fig 3.3.

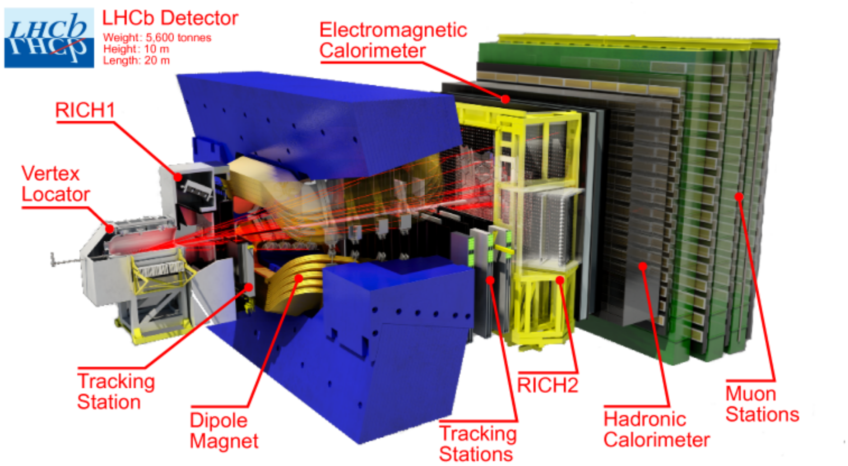


Figure 4.2: The LHCb detector 3D layout, including all its subdetectors.

Figure 3.2: LHCb detector 3D illustration [37].

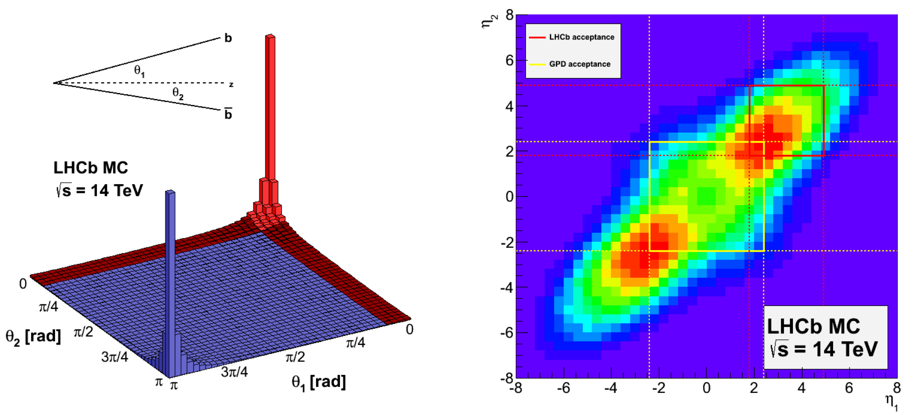


Figure 3.3: Distribution of $b\bar{b}$ production as a function of the angle with respect to the beam axis (left) and as a function of the pseudorapidity (right) at $\sqrt{s} = 14$ TeV, created using PYTHIA8 and CTEQ6 NLO considering all the five flavours (u, d, s, c, b) in the PDFs and including the processess, weighted according to their cross-sections, $q\bar{q} \to b\bar{b}$, $gg \to b\bar{b}$, $q\bar{q} \to b\bar{b}g$ (where $q \neq b$), $b\bar{b} \to b\bar{b}g$ and $gg \to b\bar{b}g$ [38].

The LHCb is specially designed to study heavy flavour physics, i.e. physics of the b (and c) quark and perform various precision measurements, CP violation searches, angular observable of the CKM matrix and the study rare decays. Furthermore, it is also adequate to searches for NP beyond SM and other exotic effects.

To study heavy hadrons decays, it is important to measure with high precision the point where the particle is created (the so-called primary vertex, PV) and the point where it decays (the so-called secondary vertex, SV) in order to calculate the lifetime of the particle and other topological quantities, as well as obtaining a first identification of the decays based on the lifetimes. At the same time, it is crucial to have a good particle identification system to properly identify and separate the decays.

To perform this measurements, the machine is composed of many subdetectors that include a high-precision tracking system comprising a silicon-strip vertex detector circling the collision point, the Vertex-Locator (VELO); a large-area silicon-strip detector, the Trigger Tracker (TT), that is located upstream a dipole magnet (with bending power of approximately 4 Tm and responsible for deflecting the trajectory of charged particles), and three stations of silicon-strip detectors and straw drift tubes, T1, T2, T3 located downstream the magnet. The ring-imaging Cherenkov detectors, RICH1 and RICH2, provide information on particle identification to distinguish charged hadrons and the calorimeters (ECAL, HCAL, SPD/PS) are responsible to measure the energy and position of the particles, also assisting on the identification process of electrons, hadrons and photons. Finally, the muon stations (M1-M5) are responsible for identifying the muons. The detector components are illustrated in Fig. 3.4.

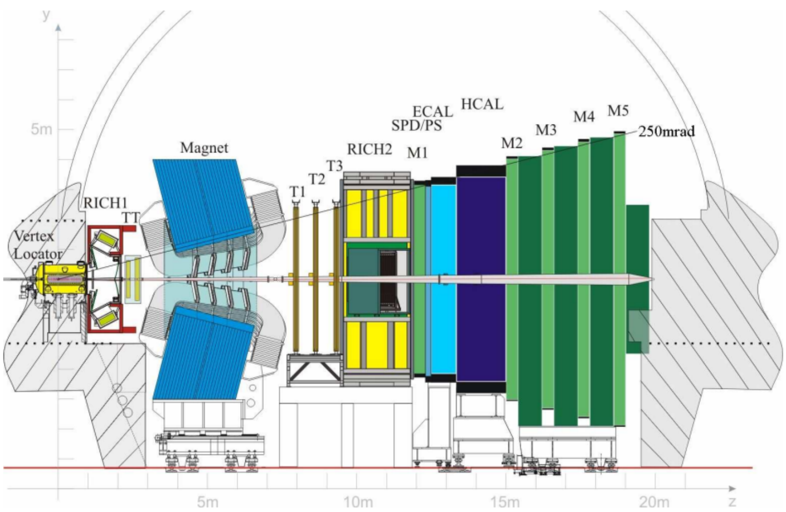


Figure 3.4: Illustration of LHCb detector components [36].

3.2.1 The tracking system

3.2.1.1 VELO

The VELO [39] is a silicon microstrip detector positioned around the pp interaction point and provides precise measurements of track coordinates in order to identify the PV and SV, which are distinctive features of b- and c-hadron decays. It consists of 42 silicon modules arranged along the beam axis, each providing a measure of the r and ϕ coordinates, being 21 modules in each side of the semi-circumference that surrounds the beam. An schematic representation can be seen in Fig. 3.5. The sensors are positioned only 7mm from the LHC beams, with a few centimetres space between each module in the z axis to ensure that each track produced within the 300 mrad LHCb acceptance interacts with at least four VELO stations. The VELO sensors are retractable and remain open while the LHC beams are circulating but closes very near the interaction point for the collisions.

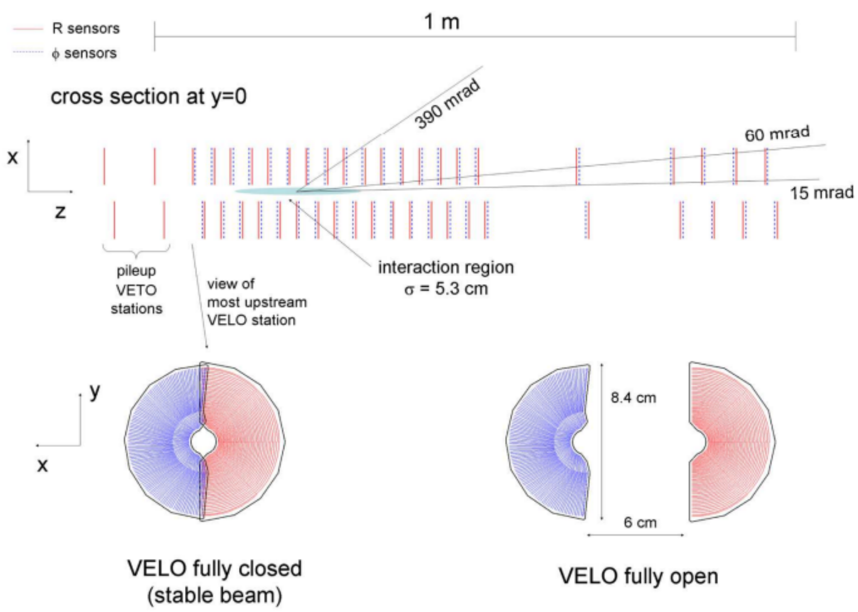


Figure 3.5: Schematic representation of the VELO detector and cross section. The z direction is defined along the beam axis into the detector, y is vertical and x is horizontal [36].

3.2.1.2 The Magnet

The magnet is responsible for bending the trajectory of charged particles to extract information about their momenta [40]. It was designed with saddleshaped coils in a window-frame yoke with sloping poles in order to match the required detector acceptance and generates a magnetic field of 4 Tm for tracks of 10m length in the y direction [36]. To reduce asymmetries originating from overall detector efficiencies, systematic errors and nuisance charge asymmetries, the LHCb inverts the polarity of the magnetic field. In this way, the data is taking with the direction of the field pointing either up or down, and this configurations receive the names *MagUp* and *MagDown*. An illustration of the magnet is presented in Fig. 3.6.

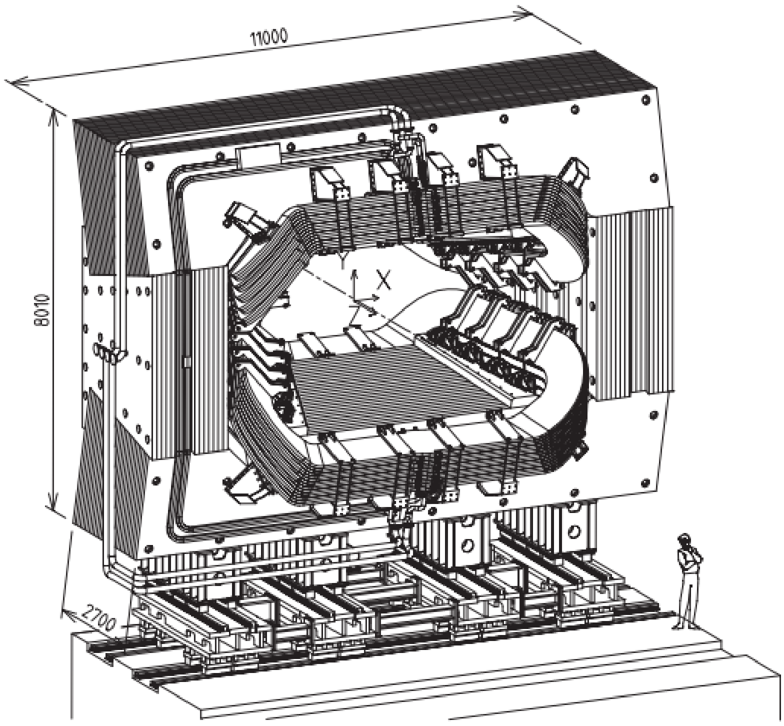


Figure 3.6: Illustration of LHCb magnet dipole [36]

.

3.2.1.3
Tracking stations

The TT is located upstream of the magnets and provides information about tracks with low momentum. It is composed of four layers of silicon microstrip sensors, with the two inner layers having rotations of $-5°$ and $+5°$ relative to the first and last vertically oriented layers to achieve better resolution. This arrangement is shown in Fig. 3.7. The TT covers the whole angular acceptance of LHCb and is localised between RICH1 and the Magneto. Each silicon strip has a resolution of $200\mu m$.

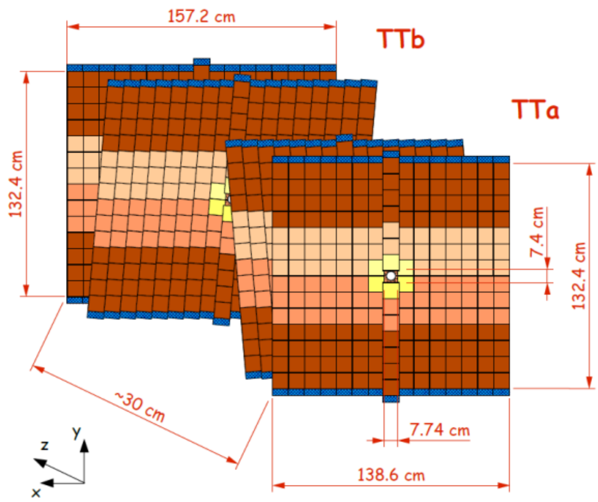


Figure 3.7: Schematic view of the four layers of the Trigger Tracker stations [41].

The other tracking stations, T1-T3, are divided into two regions: the Inner Tracker (IT) and the Outer Tracker (OT). The IT's are located near the beam pipe, positioned in the centre of the tracking stations and consist of four boxes, having each box the same four silicon strip layers configuration as the TT [42]. On the other hand, the OT is located in the outer region and are arranged into two staggered layers of straw tubes drift chambers [43]. This is illustrated in Fig. 3.8.

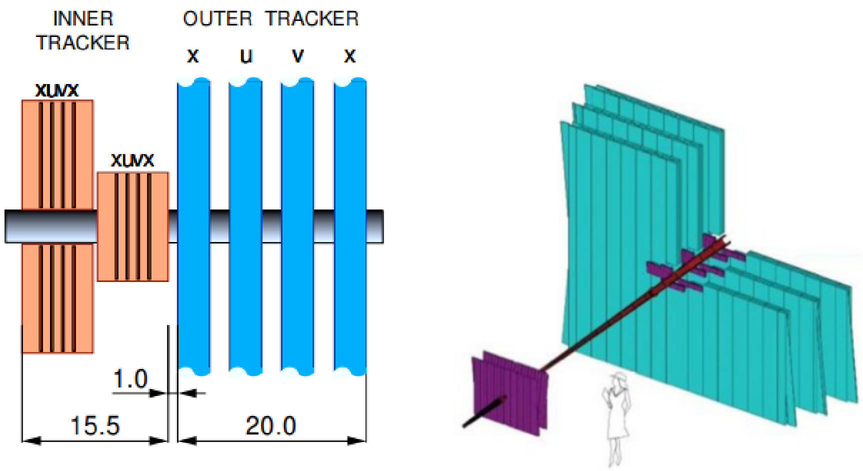


Figure 3.8: Schematic representation of the top view of a tracking station (left) with dimensions along the beam axis given in cm [42] and illustration of the trajectory stations (right) with the IT (purple) and the OT (blue) [41].

3.2.2 Ring-Imaging Cherenkov System - RICH

As already emphasised, particle identification (PID) has a crucial role on LHCb operations, whether it is to distinguish between decays with similar topology, reduce background levels coming from random combination of tracks, separate hadronic decays from leptonic ones and mainly to distinguish between pions and kaons which is fundamental to study beauty and charm hadrons. For this purpose, the RICH detectors are used. At large polar angles the momentum spectrum is softer while at small polar angles the momentum spectrum is harder, thus, in order to cover the full momentum range the PID system consists of two RICH detectors, RICH1 and RICH2.

These detectors utilise the Cherenkov effect, that occurs when a charged particle cross a dielectric medium with a velocity greater than the phase velocity of light in that medium causing a momentary polarisation and, as the medium relaxes back to the ground state, it emits photons (radiation). These photons form a cone with an angle θ_c with respect to the trajectory given by:

$$\cos\theta_c = \frac{1}{n\beta} = \frac{1}{n}\frac{\sqrt{p^2c^2 + m^2c^4}}{pc} \tag{3-1}$$

where n is the refractive index of the material, β is the ratio between the particle's velocity and the speed of light (c), m is the particle's rest mass and p is the particle's momentum.

The combination of the momentum reconstructed by the tracking system with the Cherenkov angle, which provides information on the particle velocity, can then be used to discriminate particles of different masses and relate a probability for PID. In both RICH detectors the focusing of the Cherenkov light is obtained by a combination of spherical and flat mirrors to reflect the image out of the spectrometer acceptance [36]. The relation between the Cherenkov angle and the particle momentum can be seen in Fig. 3.9.

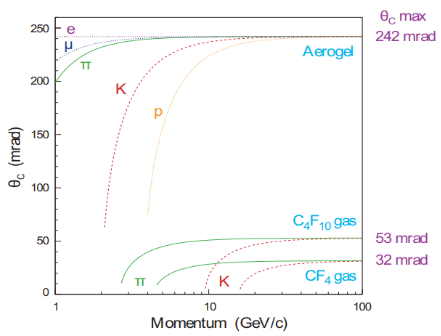


Figure 3.9: Cherenkov angle *versus* particle momentum for the RICH detectors [36].

The RICH1 is the upstream detector, located between the VELO and the TT, and is responsible for the particles in the low momentum range ($\sim 1-60$ GeV) using aerogel C_4F_{10} radiators. It has a wide acceptance covering the full LHCb acceptance from ± 25 mrad to ± 300 mrad (horizontal) and ± 250 mrad (vertical). The RICH2 uses a CF_4 radiator and is located downstream, between the T3 and the muon stations. It covers a limited angular acceptance of $\sim \pm 15$ mrad to ± 120 mrad (horizontal) and ± 100 mrad (vertical) to focus on the region where the high momentum particles (from ~ 15 GeV up to beyond 100 GeV) are produced. An schematic representation of the RICH detectors can be seen in Fig. 3.10.

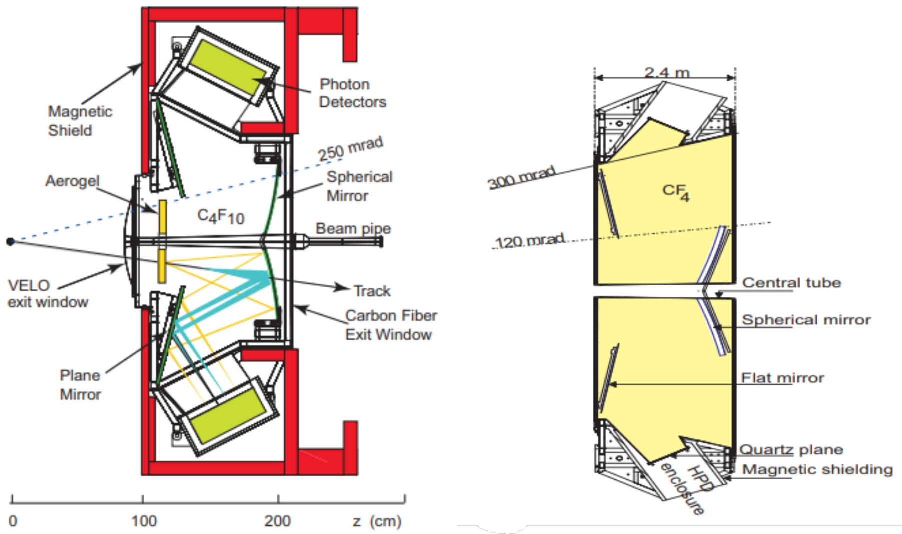


Figure 3.10: Side view of the schematic layout of RICH1 (left) and top view of the schematic of RICH2 [36].

3.2.3 Calorimeters

The LHCb calorimeter system is composed by the electromagnetic (ECAL) and hadronic (HCAL) calorimeters, the Scintilator Pad Detector (SPD) and the PreShower (PS) and is responsible to measure the transverse energy E_T and position of the particles that produce electromagnetic or hadronic showers. This information is the basis of the Level 0 trigger and, therefore, has to be provided with sufficient selectivity and in a very short time. This system helps on the identification of electrons, photons and hadrons although the most demanding identification is of the electrons [44].

The ECAL and HCAL are made with plates of 2 mm lead and 16 mm iron absorber material, respectively, alternating with 4mm scintillating plates. An incident particle produces new particles with lower energy creating a shower after travelling a certain distance inside a dense material. These new particles pass then through the scintillators producing photons that are collected by the Photo-Multiplier Tubes (PMT) and the number of photons detected is proportional to the energy of the incident particle. An illustration of the signal deposited in different parts of the detector is shown in Fig. 3.11.

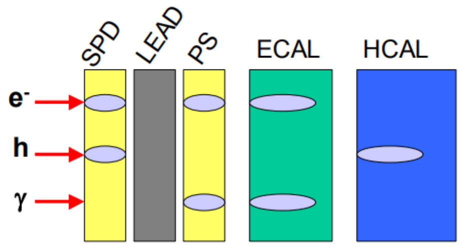


Figure 3.11: Signal deposited on all the parts of the calorimeter by an electron, a hadron, and a photon [45].

The ECAL purpose is to measure the energy of electrons and photons and to reconstruct π^0. It has an energy resolution of $\sigma_E/E = 10\%\sqrt{E}\otimes 1\%$. On the other hand, the HCAL measures the energy of protons, neutrons, pions and kaons and has an energy resolution of $\sigma_E/E = 80\%\sqrt{E} \otimes 10\%$. The SPD and PS participate in this process by contributing in the rejection of particles. The SPD rejects electrons with high transverse momentum in neutral pion decay and discriminates electrons and photons showers while the PS is responsible for the rejection of the background of charged pions. The detectors are illustrated in Fig. 3.12.

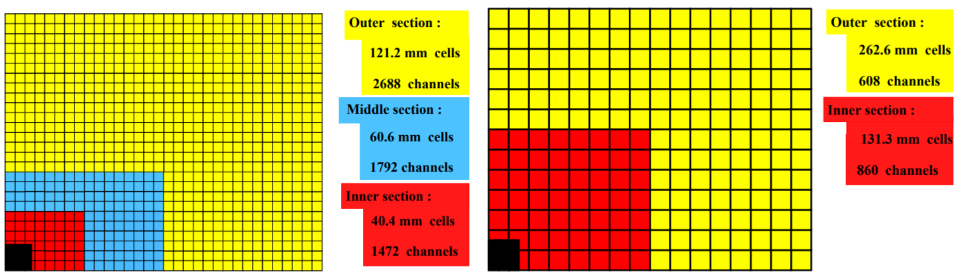


Figure 3.12: Illustration of the SPD/PS and ECAL detectors (left) and HCAL (rigth). One quarter of the detector front face is shown [44].

3.2.4 Muon stations

Muons are present in the final state on interesting b- and c-hadron decays at the LHCb detector and thus, muon triggering and offline muon identification is fundamental for the LHCb analysis. The muons from semi-leptonic b decays play an important role in CP asymmetry and oscillation measurements by tagging the initial state flavour of the accompanying neutral B mesons and these decays can also be sensitivity to NP beyond the SM [36].

The LHCb muon system is composed of 5 stations (M1-M5) covering an area of 435 m^2, the layout is shown in Fig. 3.13. Each Muon station is divided into four regions, R1-R4, that increase distance from the beam axis. M1 is used to improve the p_T measurement in the trigger and for this it is placed in front of the calorimeters. The other stations (M2-M5) are located downstream the calorimeters and are intercalated with 80 cm iron filter to select penetrating muons and avoid background from hadrons. The stations M1-M3 have high spatial resolution on the bending plane and are used to define the track direction and also to calculate the p_T of the muon candidate with a resolution of 20%. Stations M4 and M5, on the other hand, have limited spatial resolution and are mainly used for identification of penetrating particles. This is the last piece of the LHCb detector and should, ideally, be crossed only by muons, which are long-lived quasi-stable particles.

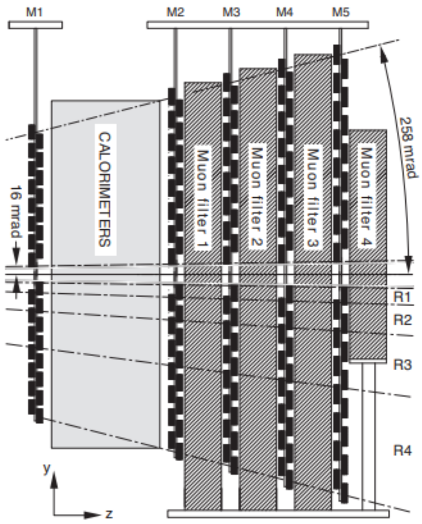


Figure 3.13: Schematic representation of the muon system side-view [36].

With exception of only M1, that uses Gas Electron Multiplier (GEM) detectors, all the other stations are constructed with Multi Wire Proportional Chamber (MWPC) technology. The MWPCs use a gas mixture of Ar, CO_2 and CF_4 and the chambers produce an electron shower when a muon passes. These electrons are directed to the anode while the ions are taken to the cathode and produce an electrical signal. The GEMs are made of a mixture of Ar, CO_2 and CF_4 and three metal layers intercalated between the anode and the cathode plates and with a high density of holes in order to collect the ionising electrons. In this way, the trajectory of the muon trespassing the system can be determined.

3.2.5
The trigger system

The LHCb experiment operates at an average luminosity of 2×10^{32} cm^{-2}s^{-1}, which is much lower than the maximum design luminosity of LHC (1×10^{34} cm^{-2}s^{-1}), in order to limit the number of interactions per beam crossing, allowing then the reconstruction of the PV and SV and also reducing the radiation damage to the detectors. Due to the LHC bunch structure and low luminosity, the crossing frequency with interactions visible by the spectrometer is about 40 MHz and since this rate cannot be all storage, the trigger system is responsible for reducing it to a few kHz. This reduction is performed in two trigger levels: Level-0 (L0) and High Level Trigger (HLT) [34].

3.2.5.1
Level-0 trigger

The first level of trigger is responsible for reducing the LHC beam crossing rate of 40 MHz to the rate of 1 MHz, which can then be read out by the entire detector. It is implemented using custom made electronics and operates synchronously with the 40 MHz bunch crossing frequency. By combining information of hadrons, electrons and photons with high E_T in the calorimeters, high p_T muons in the muon chambers and number of primary pp interactions from the VELO in each bunch crossing, the L0 is able to select the events that are going to be triggered for posterior analysis. This decision is processed by the L0 decision unit (L0DU).

The events that fire the trigger can be classified according to the presence or not of a signal candidate. If the trigger is fired by the signal candidate track itself, it is classified as Triggered on Signal (TOS), otherwise if it was triggered by some other effect in this event, it is classified as Triggered Independent of Signal (TIS).

3.2.5.2
High level trigger

Differently from the L0, the HLT trigger is executed asynchronously on a processor farm by commercially available equipment. After passing through the L0, to reduce the event rate from 1 MHz down to 2 kHz the HLT uses the full event data to make the selection. This trigger level is divided in two stages, HLT1 and HLT2. The TOS and TIS classification are also applied in this level.

The purpose of HLT1 is to reduce the data rate to $\sim$ 30 KHz by applying some requirements. It performs the partial reconstruction of the event by using information from VELO, the T-stations and the muon chambers. This reduction is necessary for the HLT2 to perform the full patter recognition on the remaining events. At the HLT2 stage, the full event reconstruction is done and the data rate is reduced to only few kHz. Here, two types of trigger lines are introduced, exclusive and inclusive. The exclusive lines are optimised for each decay while for the inclusive lines only generic topological requirements are imposed. The trigger process is illustrated in Fig.

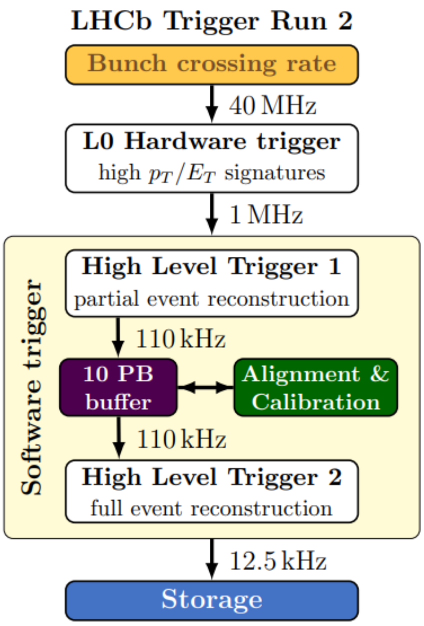


Figure 3.14: Overview of the LHCb trigger system [46].

3.3 The LHCb Upgrade I

The LHCb was primarily designed for precision measurements in heavy-flavour physics and to search for new physics through studies of CP-violation. However, the experiment has demonstrated excellent capabilities in handling other domains, like electroweak measurements, heavy ion and fixed target physics, and is moving towards becoming a general purpose experiment covering the forward region. To allow for this wider physics program, during CERN's Long Shutdown 2 (LS2) the experiment infrastructure passed through its major upgrade in both hardware and software level [47]. The layout of the upgraded LHCb detector is presented in Fig. 3.15.

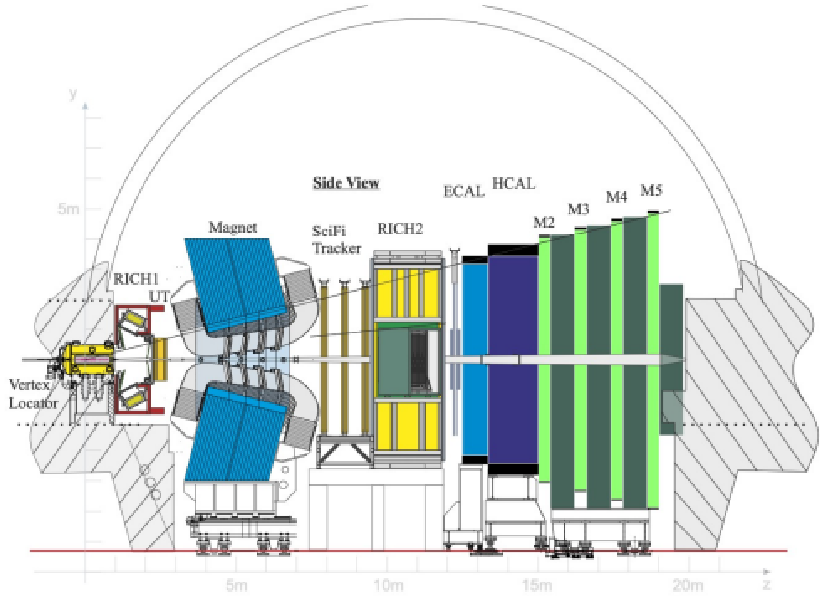


Figure 3.15: Side view of the layout of the upgraded LHCb detector [47].

The VELO detector was almost completely replaced by the VELOPIX chips. This upgraded detector is a hybrid silicon pixel detector capable of collecting signal hits from 256×256 pixels. It is also closer to the beam axis, 5.1 mm as opposed to the previous 8.4 mm, and offers an improved hit resolution and simpler tracker reconstruction. The detector is arranged into 52 modules that brings together the silicon detectors, their cooling, powering, readout and mechanical supporting into a single repeating unit.

The RICH1 and RICH2 have both been refurbished to cover up the more challenging data-taking of the LHC Run 3. The photon detection system has been redesigned with two types of 64-channel multi-anode photomultiplier tubes being used to select and delete single photons at the same time they provide excellent spatial resolution and low background noise. Also, the optical system of RICH1 was redesigned to spread the Cherenkov rings over a large surface, reducing then the number of photons in the hottest region.

The TT was replaced by a new upstream tracker (UT) that uses innovative silicon-microstrip sensors. It is composed of four planes of silicon-microstrip detectors mounted on both sides of a vertical structure called stave. As for the T-stations (T1-T3), they were replaced by a new type of station based on scintillating fibres (SciFi) with silicon photomultiplier (SiPM) read out.

At the software level, the LHCb trigger system has experienced a radical change targeting a software-only trigger system. The whole detector will now read at the full rate of 40 MHz and the new trigger system was designed to perform real-time analysis, allowing the event selection to be done in a more precise and flexible way by the software. This new two stage system is composed on a first stage by a inclusive high level trigger 1, HLT1, and on a second stage by HLT2. A summary of the LHCb Upgrade I is presented in Fig. 3.16.

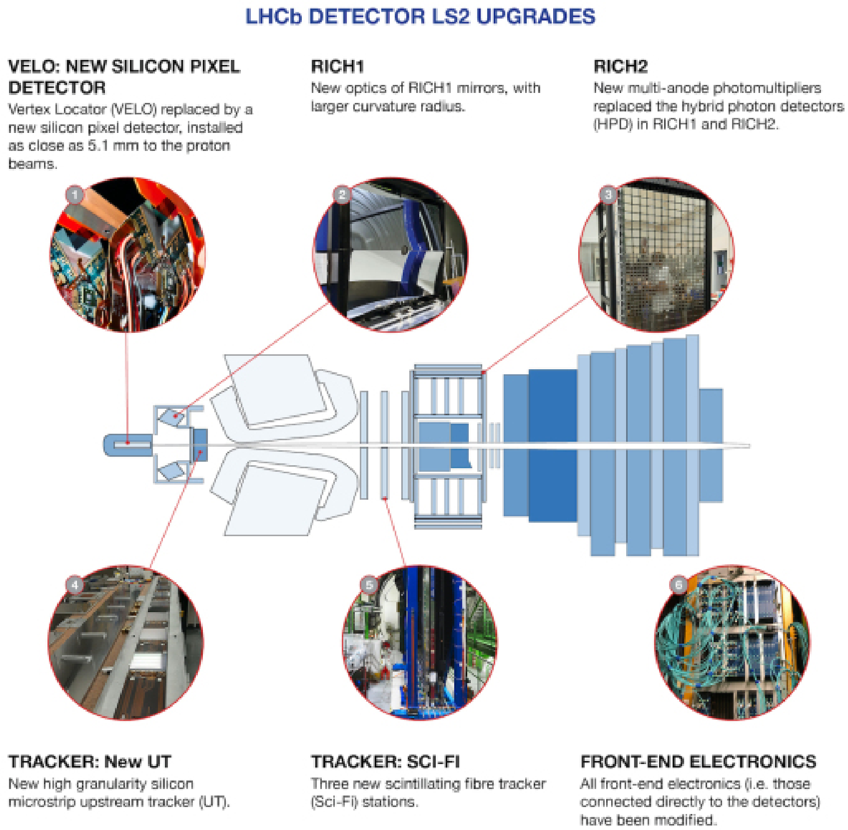


Figure 3.16: LHCb Upgrade I summary [48].

4
Data Sample

This chapter describes the data selection procedure performed to obtain the final data sample of $D^+ \to \pi^-\pi^+K^+$ used in this work. From now on, the particles ordering will be defined as $D^+ \to \pi^-(1)\pi^+(2)K^+(3)$. The selection process can be divided into two steps: online and offline. The former is performed by the trigger systems (L0, HLT1 and HLT2) during the data taking at hardware and software levels. After the trigger, the data sample is ready to be analysed and, then, an offline selection is performed in order to remove events that are not of our interest, the so-called background events.

4.1
Selection variables

In order to select the decays of interest, reconstructed information is used to set up the selection criteria. The variables used are mainly based on topological characteristics of the decay and on particle identification, and are associated to physical quantities that are good discriminators such that it is possible to distinguish between signal and background. In this section, the relevant variables for this analysis are defined.

The topology of a 3-body decay is shown in Fig. 4.1, with: the primary vertex (PV), where the D meson is produced; the secondary vertex (SV), where it decays; the distance between these two points, the so-called flight distance (FD); and the impact parameter (IP), that is the minimum distance of a particle's trajectory to the PV. Along with these quantities, other important discriminatory information can be obtained from: the χ^2_{IP}, defined as the difference in the vertex-fit χ^2 of a given PV reconstructed with and without the particle being considered; the χ^2_{FD}, defined as the square of FD over the square of its uncertainty; the momentum (p) and the transverse momentum (p_T), that is the direction perpendicular to the beam axis; the angle (DIRA) between the reconstructed D candidate momentum of the particle and the flight direction; the lifetime (τ) of the particle; the pseudorapidity $\eta = -\ln[\tan\frac{\theta}{2}]$, that describes the angle of the particle relative to the beam axis; the Pointing $= \frac{p\sin\theta}{p\sin\theta + \sum_i^3 p_{T_i}}$, that is a comparison of D^+ momentum perpendicular to the flight direction to reconstructed final state particle's transverse momenta, where θ is the angle between D^+ momentum and the flight direction; and the $\pi^-\pi^+K^+$ invariant mass combination (M).

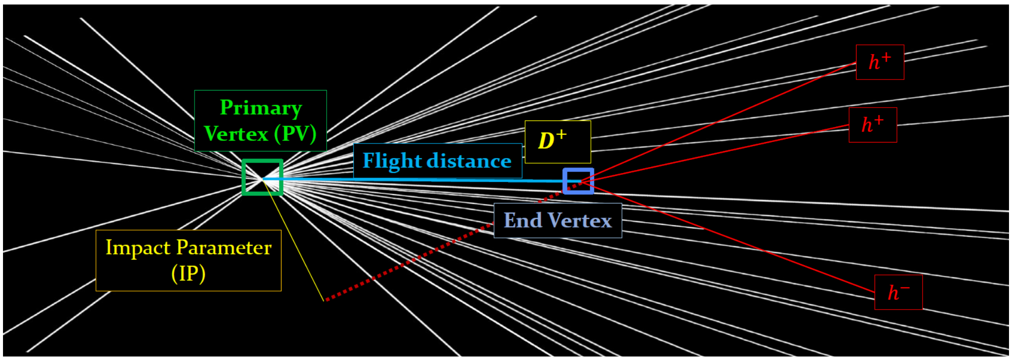


Figure 4.1: Illustration of a $D^+ \to h^- h^+ h^+$ decay at LHCb. Courtesy of my dear friend Felipe Almeida.

The particle identification (PID) criteria comes from the RICH detectors and are used to distinguish between protons, kaons and pions. These variables can be either PIDK, that is the delta-log-likelihood of being a kaon with respect to the pion hypothesis; or can be ProbNNk/ProbNNpi, which are outputs of multivariate techniques that result in a single probability for the given particle hypothesis of being a kaon/pion, which is created combining tracking and PID information. On the other hand, photons, electrons and other hadrons are identified by a calorimeter system consisting of scintillating-pad (SPD), being the number of SPD hits (nSPDHits) another selection criteria.

4.2
Data Set

The analysis is performed using the data samples from proton-proton (pp) collisions with a centre-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 5.6 fb^{-1} collected by the LHCb detector from 2016 to 2018.

These samples come directly from the exclusive HLT2 Turbo line `Hlt2CharmHadDpToKpPimPip`, that selects the $D^+ \to \pi^- \pi^+ K^+$ candidates which require them to be TOS on the HLT1 Track Lines `Hlt1TrackMVA` and `Hlt1TwoTrackMVA`. The selection criteria are shown in Table 4.1[1]. The candidates are required to be of TIS at L0. This requirements are necessary to ensure a good control of the online selection and in order for the events to represent inelastic collisions.

[1]Natural units, $\hbar = c = 1$, are used throughout this work.

Variables	cuts
Daughter requirements	
Track χ^2/ndf	<3.0
p_T [MeV]	>250
χ^2_{IP} (PV)	>4.0
PIDK (pions)	<5
PIDK (kaons)	>5
Combination cuts	
Mass [MeV]	1779-1959
$\sum p_T$ [MeV]	>3000
p_T (2 of 3 tracks) [MeV]	>400
p_T (1 of 3 tracks) [MeV]	>1000
χ^2_{IP} (2 of 3 tracks)	>10
χ^2_{IP} (1 of 3 tracks)	>50
Mother cuts	
Track vertex χ^2/DOF	<6
lifetime [ps]	>0.4
acos(DIRA) [mrad]	<10
Mass [MeV]	1789-1949
TisTosSpec	HLT1.*Track.*Decision%TOS

Table 4.1: HLT2 selection criteria.

4.3 Monte Carlo simulated samples

Full LHCb simulations of the signal channel $D^+ \to \pi^-\pi^+K^+$ were used as a proxy for signal events in the multivariate analysis, and were also used to determine the parameters of the signal PDFs for the data mass spectrum fit, as it will be discussed in Section 4.4.3 and Section 4.5.2, respectively. Besides that, these samples are also used in the CP violation studies to test for nuisance charge asymmetries. This data is simulated with the Gauss framework [49] using specialised programs. The first step is the generation of events using Pythia 8 [50] to reproduce generic pp collisions at 13 TeV with the same configuration and operating conditions of the LHCb detectors. These simulations start from the hard process using parton distribution functions that describes the relative composition of the protons as a function of the momentum of incoming protons to the outgoing partons that will produce showers due to QCD confinement. Then, using EvtGen [51], the hadron decay is introduced and these generated events pass to the GEANT4 [52, 53] simulating detector to reproduce the propagation and interaction of the particles with the detector material. To reconstruct the events, the *Moore* package is used to simulate the trigger stages (L0 and HLT).

In order to generate large MC samples, configuration files with generator level cuts, presented in Table 4.2, are provided and HLT2 filter are use, with the trigger lines of the signal channel `Hlt2CharmHadDpToKpPimPip`. These MC samples were generated with resonant structures based on results of previous analysis [26], which included $K^{*0}(892)$, $\rho^0(770)$, $K^*(1410)$, $K^{*0}(1430)$ and $f_0(980)$.

Variable	Cut
each daughter $p >$	2.0 GeV
each daughter $p_T >$	0.25 GeV
D^+ $p >$	14.0 GeV
D^+ $p_T >$	2.1 GeV

Table 4.2: Cuts applied at generator level.

4.4 Offline selection of $D^+ \to \pi^-\pi^+K^+$

An offline selection with loose requirements was applied on top of the trigger ones to produce the samples for the analysis. Such requirements can be seen in Table 4.3.

nSPDHits <1000
1.5 $<\eta_{\text{daughters}}$ <5
$p_{\text{daughters}}$ <100 GeV
χ^2_{IP} <12
D^+: 1805 <M <1935 MeV

Table 4.3: Additional pre-selection requirements.

The mass distribution of $\pi^-\pi^+K^+$ the whole data sample after these pre-selection requirements is shown in Fig. 4.2, and the total number of events is given in Table 4.4, separated by year and magnet polarity.

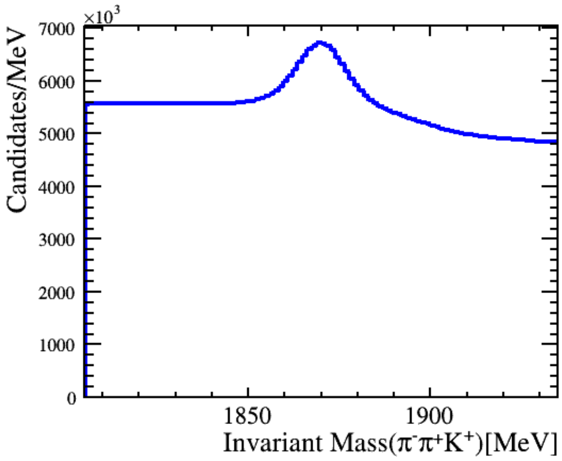


Figure 4.2: $D^+ \to \pi^-\pi^+K^+$ mass distribution.

Year	Polarity	N. of candidates ($\times 10^7$)
2016	MagUp	9.4
	MagDown	9.2
2017	MagUp	8.3
	MagDown	8.6
2018	MagUp	10.1
	MagDown	9.3
Total		55.2

Table 4.4: Number of $D^+ \to \pi^-\pi^+K^+$ candidates after the selection.

As can been seen from the mass distribution, this data sample has a significant level of background contamination that needs to be further reduced before the analysis, since it can either introduce nuisance charge asymmetries or dilute potential CP violation signals. The strategy used to treat the background include requirements applied on some discriminatory variables to reduce specific backgrounds, such as cross-feed from other channels. After these requirements were applied, in order to reduce some remaining contamination and to reduce the combinatorial background level, another study is performed using maximization of signal significance to define the selection requirement. Finally, to improve the signal significance and the purity of the data sample, a multi-variate analysis was applied using machine learning techniques to reduce even more the combinatorial background.

4.4.1 Charm background

The specific background contamination comes from the cross-feed from other decay channels due to a mis-identification (mis-ID) of a daughter particle. The main sources of background for the $D^+ \to \pi^-\pi^+K^+$ decay are the fully reconstructed decays $D_s^+ \to K^-K^+\pi^+$, $D^+ \to K^-\pi^+\pi^+$, $\Lambda_c^+ \to \pi^-\pi^+p$ and $D^+ \to K_S^0K^+$, where the K_S^0 decays to π^- π^+. The $D_s^+ \to K^-K^+\pi^+$ decay is the dominant contamination and appears from the K^+ being mis-identified as a pion. This contamination is visible at the left mass sideband in Fig. 4.2, the background level on this side being more accentuated than that on the right side. Now, for the $D^+ \to K^-\pi^+\pi^+$ cross-fed there is a double $K-\pi$ mis-ID, and although this contamination is slightly present on the right sideband, it spreads over the whole mass spectrum. These two backgrounds can be significantly suppressed by specific PID requirements.

To reduce the $D_s^+ \to K^-K^+\pi^+$ and $D^+ \to K^-\pi^+\pi^+$ contamination, both p1_PIDK and p1_ProbNNk variables were tested for particle 1 with different cut values applied in the invariant mass spectrum, where the events were reconstructed assigning a kaon or a pion mass to a daughter in order to match the the decay being studied. It was also compared the effectiveness between the p1_PIDK and the p1_ProbNNk requirements. The results for both requirements can be seen in Figs. 4.3 and 4.4 and the efficiency associated to each requirement can be seen in Fig. 4.5. The efficiency is calculated as the ratio between the number of signal candidates after the requirement was applied and the number of signal candidates before any requirement. From the invariant mass plots it can be seen that both contamination are reduced to a small level with p1_ProbNNk < 0.02 and this requirement shows a better efficiency if compared with the p1_PIDK.

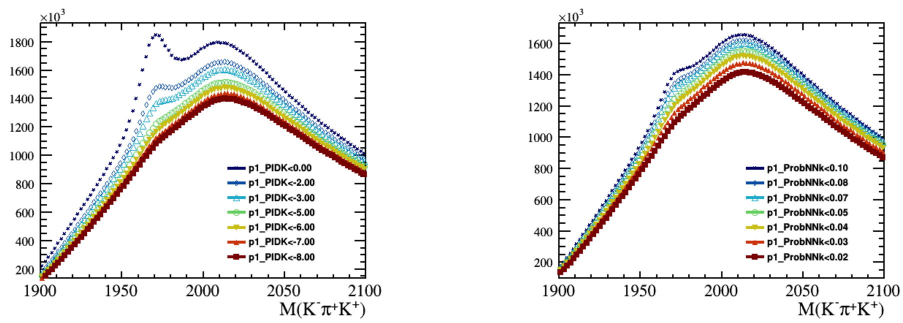


Figure 4.3: Three-body invariant mass reconstructed by assigning a kaon mass to particle 1 (π^-), for multiple values of p1_PIDK and p1_ProbNNk requirements.

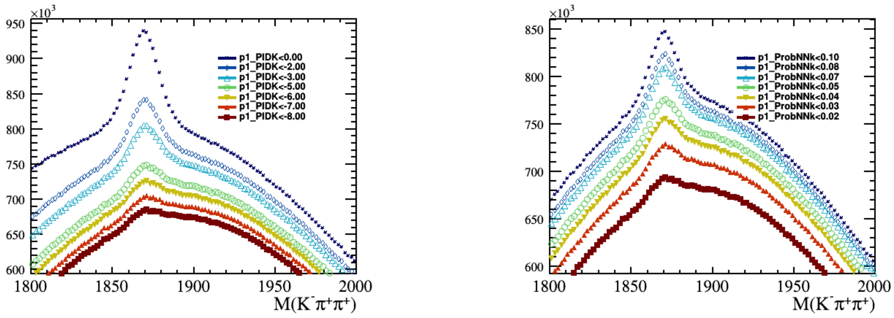


Figure 4.4: Three-body invariant mass reconstructed by assigning a kaon mass to the particle 1 (π^-) and a pion mass to particle 3 (K^+), for multiple values of p1_PIDK and p1_ProbNNk requirements.

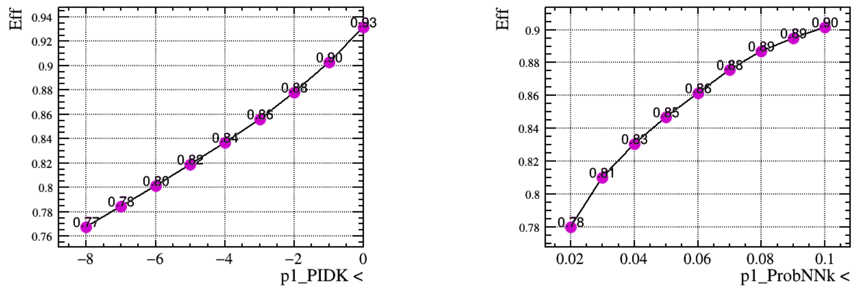


Figure 4.5: Signal efficiency associated to each value of p1_PIDK and p1_ProbNNk requirements.

After analysing particle 1 and applying the requirement p1_ProbNNk $<$ 0.02, the approach was to maximise the significance for particle 3 in order to reduce some remaining contamination from other charm decays, again comparing PIDK *versus* ProbNNk, where

$$\text{Significance} = \frac{S}{\sqrt{S+B}}, \tag{4-1}$$

and S and B are the estimated number of signal and background candidates with a 40 MeV mass window for the signal region and 20 MeV for each side band as shown in Table 4.5. The value of S is obtained by sideband subtraction, assuming a linear background distribution over the mass spectrum.

Signal region	1850 - 1890 MeV
Left wing	1810 - 1830 MeV
Right wing	1910 - 1930 MeV

Table 4.5: $D^+ \to \pi^-\pi^+K^+$ signal and background mass windows.

The maximisation of significance shows no peak for p3_PIDK, while for for p3_ProbNNk it is on p3_ProbNNk $>$ 0.64 as can be seen in Fig. 4.6.

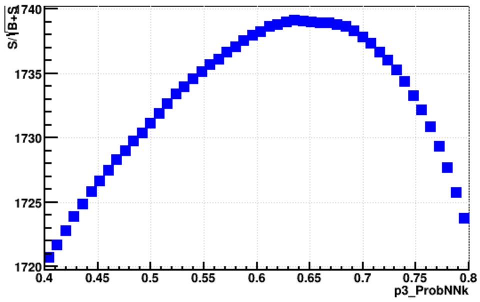


Figure 4.6: Significance calculated for different requirements of p3_ProbNNk.

For the $\Lambda_c^+ \to \pi^-\pi^+p$ contamination, although it is not easily visible on the data mass distribution on a first look, the invariant mass reconstruction with a proton assigned to particle 3 shows a prominent peak. For this contamination, the ProbNNk requirements already applied reduce considerably this cross-fed, being quite small in the signal region. For the $D^+ \to K_S^0 K^+$ contamination it has the same final state as the main decay channel, not constituting a mis-ID, and on the mass reconstruction for the $\pi^-\pi^+$ invariant mass a peak for the K_S^0 contribution can be seen. All the pre-selection requirements already applied seems to reduce considerably this unwanted K_S^0. Both contamination contributions before and after the selection requirements can be seen in Fig. 4.7.

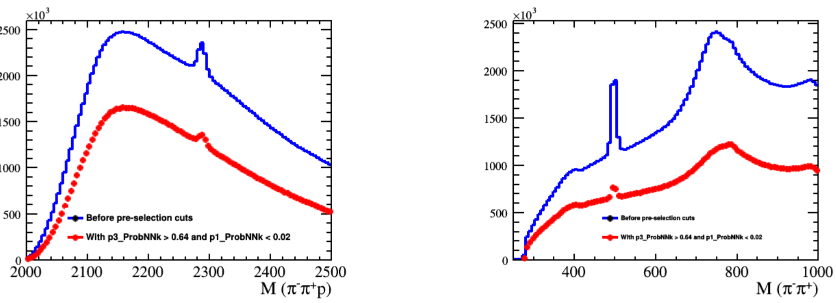


Figure 4.7: Invariant mass reconstruction assigning a proton mass to particle 3 (K^+) for the $\Lambda_c^+ \to \pi^-\pi^+p$ contamination (left) and the invariant mass for the pair $\pi^-\pi^+$ for the $D^+ \to K_S^0 K^+$ contamination (right).

Besides these important contaminations, another relevant contribution for the background comes from the $\overline{D}^0 \to \pi^-\pi^0K^+$, where the π^0 is replaced by a random π^+ that comes from PV. As can be seen in Fig. 4.8, a tight requirement $\chi^2_{\rm IP,2} > 30$, to ensure this particle comes from a SV, is sufficient to remove most of this contamination.

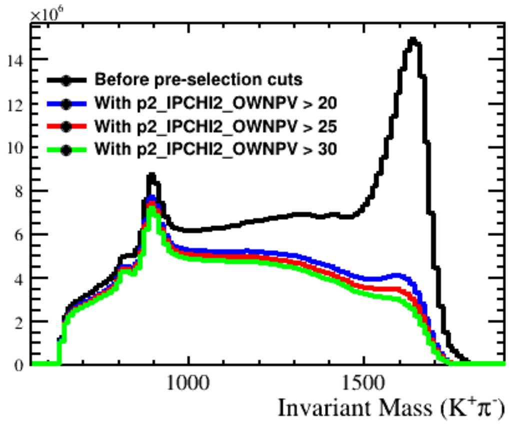


Figure 4.8: Two-body invariant mass reconstructed for the pair $m_{3,1}$ with different values of $\chi^2_{\text{IP},2}$ requirements applied.

4.4.2
Monte Carlo simulated samples

All the selection criteria applied on the data sample were applied on the MC samples, except the PID requirements. The PID criteria need to be emulated using a data-driven method since the response of the RICH detector is not well modelled and this could lead to inaccurate results. To obtain this, the PIDCalib package [54], which is a set of tools used to calculate the efficiency on the PID requirements, is used and PID selection criteria is derived from calibration samples. The mass distribution after the data selection criteria with the PID weight applied and the DP for the MC sample are presented in Fig. 4.9.

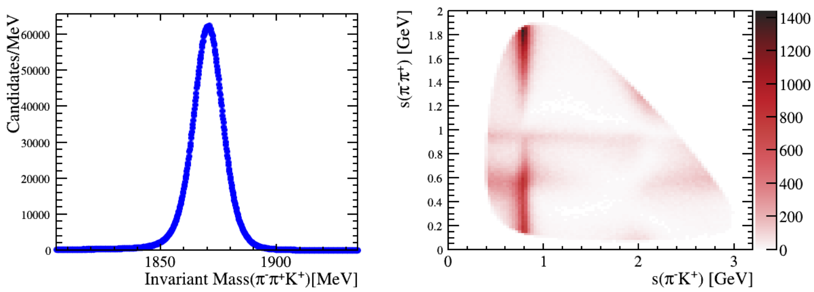


Figure 4.9: Mass distribution and Dalitz plot for the MC sample after the selection criteria.

4.4.2.1
Reweight Procedure

The simulated sample is compared to the data to check for differences in kinematics, as can be seen in Fig. 4.10 for the variables $\chi^2_{\text{IP},D}$ and $\log \text{IP}$. In order to overcome these differences, a kinematic weight, the "reweight", needs to be performed and applied to match these distributions. This procedure is carried out by using the GBReweighter, from `hep_ml` library [55], on the same variables used on the MVA. The training is performed using approximately 1M events from the MC samples and 1M events from the sPlotted data to select only signal candidates. The sPlot technique [56] is used to attribute weights to the data sample in order to distinguish between signal and background events. The result after the reweight process can be seen in Fig. 4.11.

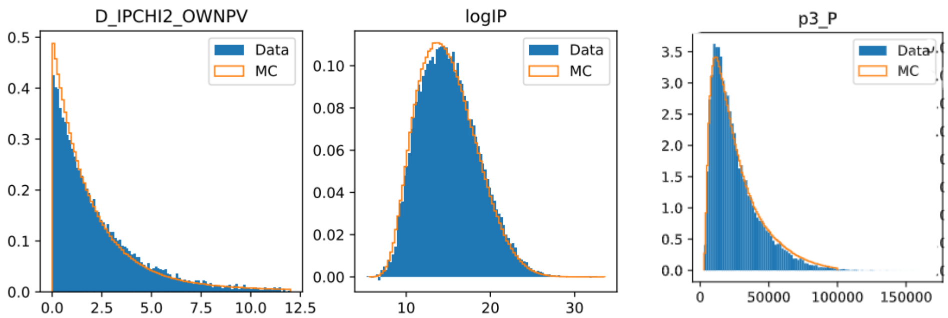


Figure 4.10: $\chi^2_{\mathrm{IP},D}$, $\log \mathrm{IP}$ and $p3\_P$ distributions for $D^+ \to \pi^-\pi^+K^+$ signal decays before the reweight process.

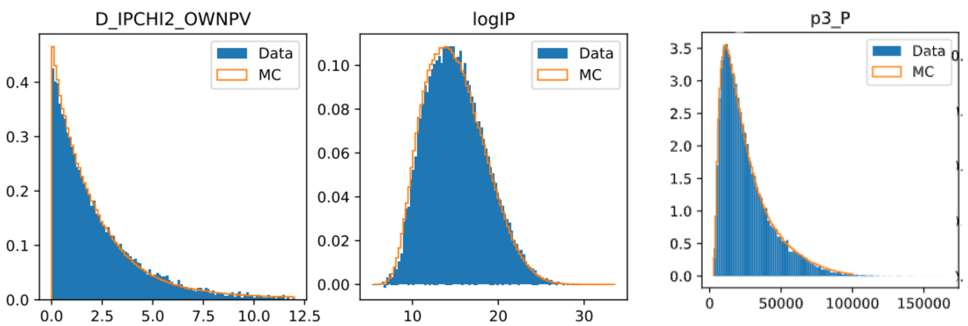


Figure 4.11: $\chi^2_{\mathrm{IP},D}$, $\log \mathrm{IP}$ and $p3\_P$ distributions for $D^+ \to \pi^-\pi^+K^+$ signal decays after the reweight process.

4.4.3 Multi-variate analysis - MVA

After the selection process described previously, a multi-variate analysis (MVA) that uses a neural network is applied to further reduce the combinatorial background and to improve the signal significance and purity of the samples. The MVA analysis is based on classification learning between what is signal and what is background and, during this stage, samples are provided so the algorithm can train, evaluate and apply more precise and efficient requirements to remove the background. The method chosen in this analysis is the Boosted Decision Tree (BDT) [57] technique. This technique consists of a binary tree that takes a set of input features and splits input data recursively based on those features and the boosting means that each tree is dependent on prior trees and these trees are combined into a strong classifier. This binary tree makes decisions one variable at a time in order to classify the input data.

Before the training and testing phase, the samples per each year were divided into so-called A and B samples, being these samples filled with 50% of the MagDown samples and 50% of the MagUp samples, as illustrated in the diagram on Fig. 4.12. This process was executed in order to overcome kinematic differences between the different years/polarities and it is executed on both data sample and Monte Carlo simulated samples.

Figure 4.12: Schematic representation of the divisions of samples A and B.

For the training and testing phase, two sub-samples of 1 million candidates from MC events were used as a proxy for signal and two sub-samples of 1 million candidates from the sidebands were used as proxy for the background. Of these two 1M samples, one is used on the training stage and the other on the test. In order to avoid bias, if the training is performed on an A sample, the test needs to be done on a B sample and vice versa. Nine discriminatory variables were used by the algorithm: IP_{χ^2}, $\log \mathrm{IP}$, FD, FD_{χ^2}, POINTING, Vertex_{χ^2}, DIRA, p and p_T. The BDT response for the training and testing samples of 2016_A can be seen in Fig. 4.13.

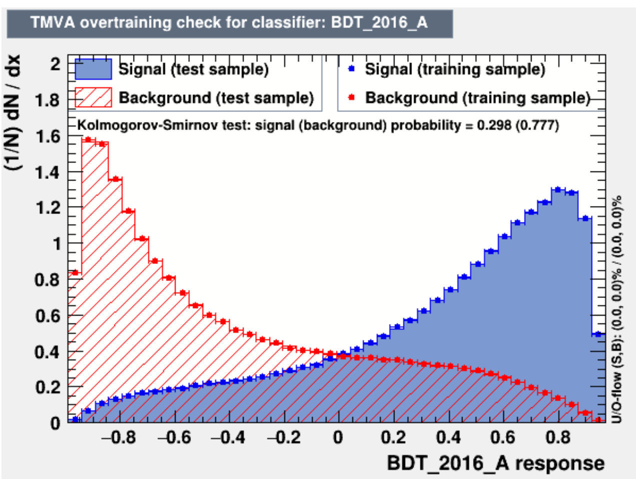


Figure 4.13: BDT response for training and testing samples of 2016_A.

The responses were then applied in the 2016-2018 A and B samples using the 6-folded approach to avoid overtraining. By this method, for each of the six samples it is evaluated the response of the other five and each process of this produces a BDT classifier, creating a new selection variable called valBDT. The final classifier, the valBDT_mean, is then the average of these five classifiers produced. This technique is illustrated in Fig. 4.14.

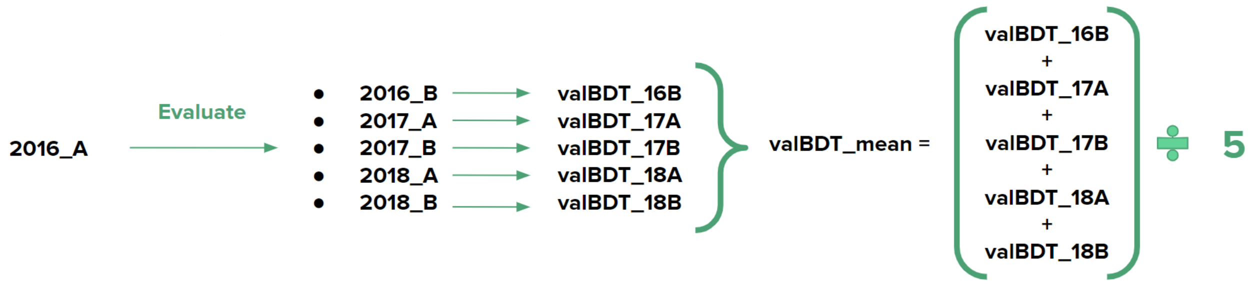


Figure 4.14: Schematic representation of the construction of the BDT classifier.

The requirement on the value of the valBDT_mean classifier was determined by looking at the significance, in Fig. 4.15, and was applied at valBDT_mean > 0.08.

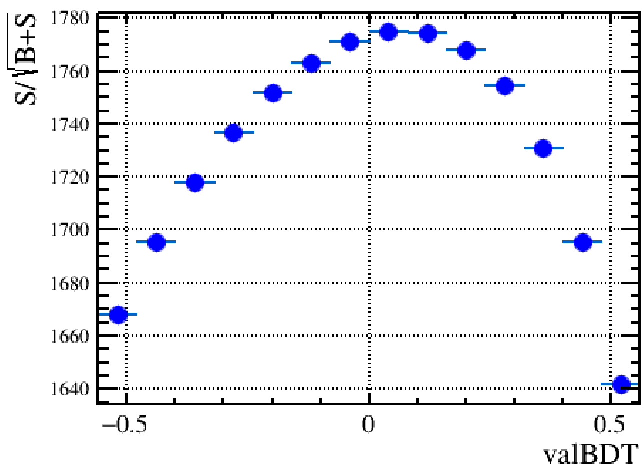


Figure 4.15: Significance for the valBDT_mean requirement for all samples.

This concludes the offline selection applied on the data sample and the requirements are summarised on Table 4.6.

Requirements	$D^+ \to \pi^-\pi^+K^+$
Pre-selection	$1805 < M < 1905$ [MeV]
	χ^2_{IP} <12
	nSPDHits <1000
	1.5 <$\eta_{\mathrm{daughters}}$ <5
	$p_{\mathrm{daughters}}$ <100 GeV
PID	$\mathrm{ProbNNk}_1$ <0.02
	$\mathrm{ProbNNk}_3$ >0.64
Specific background	$\mathrm{IP}_{\chi^2,2}$ >30
BDT	valBDT_mean >0.08

Table 4.6: Offline selection requirements summary.

The mass distribution after all the selection requirements is shown in Fig. 4.16. After the selection process, a mass fit needs to be performed to determine the number of signal events in the data sample. Some parameters used on the fit to model the signal events were fixed to those from the MC samples.

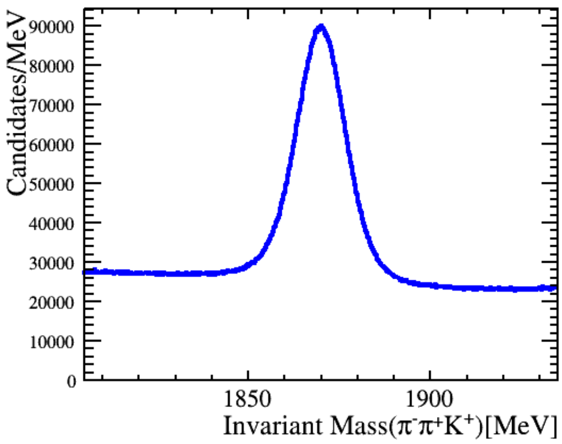


Figure 4.16: $D^+ \to \pi^-\pi^+K^+$ mass distribution after all the selection requirements.

4.5 Mass spectrum fit and final samples

4.5.1 Monte Carlo simulated samples

After applying the selection criteria, the PIDCalib and reweight weights, and the MVA, a mass fit was performed on the MC samples to obtain some parameters for the data sample mass fit. For the mass fit, the Roofit [58] package was used and for the signal probability density function (PDF) it was used a Gaussian and two Crystal Balls (CB) [59],

$$\begin{aligned} \mathcal{P}_{signal}(m) = f_G \times G(\mu, \sigma_G) + (1 - f_G)\times \\ [f_{CB} \times CB_1(\mu, R_1\sigma_G, \alpha_1, N_1) + (1 - f_{CB}) \times CB_2(\mu, R_2\sigma_G, \alpha_2, N_2)] \end{aligned} \tag{4-2}$$

The CB consist of a Gaussian core portion and a power-law low-end tail, below a certain threshold, and is used to account for the fact that the mass distribution is not a perfect Gaussian. In that way, the CBs are used one for each "tail" of the mass distribution. The results are presented in Table. 4.7 and in Fig. 4.17.

Parameters		Results
	N_{sig}	7 258 030 ± 2 694
Signal components fractions'	f_G	0.5433 ± 0.0057
	f_{CB}	0.2080 ± 0.0094
Gaussian parameters	μ	1870.41 ± 0.0036
	σ_G	5.5197 ± 0.0165
Crystal Ball parameters	R_1	1.6788 ± 0.0202
	R_2	1.6600 ± 0.0045
	α_1	0.6215 ± 0.0469
	α_2	-2.1692 ± 0.0111
	N_1	3.3250 ± 0.4600
	N_2	2.9181 ± 0.0602

Table 4.7: MC fit results.

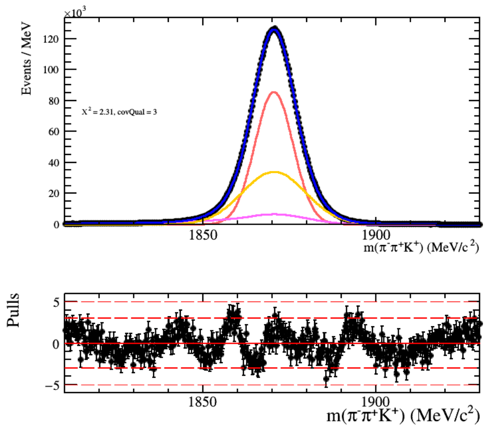


Figure 4.17: Monte Carlo fit using the PDF in Eq. 4-2.

4.5.2 Mass fit and final sample statistics

After all the selection requirements, the final samples were fitted to obtain their statistics and purity. As in the MC sample fit, the Roofit package was used. For the signal PDF it was used a Gaussian and two CB, same expression as presented in Eq. 4-2, with the CB parameters fixed from the MC sample. The parameters were presented in Table 4.7. To model the background, a third-order Bernstein polynomial was used as PDF,

$$\mathcal{P}_{bkg}(m) = \sum_{i=0}^{3} a_i \binom{3}{i} m^i \cdot (1-m)^{3-i}. \tag{4-3}$$

The final PDF used for the fit, combining Eqs. 4-2 and 4-3, is then given by Eq. 4-4,

$$\mathcal{P} = \mathcal{N}_{sig}\mathcal{P}_{sig} + \mathcal{N}_{bkg}\mathcal{P}_{bkg} \tag{4-4}$$

where $\mathcal{N}_{sig}$ and $\mathcal{N}_{bkg}$ represent the number of signal events candidates and background events candidates, respectively. The fit can be seen in Fig. 4.18 and the results in Table 4.8.

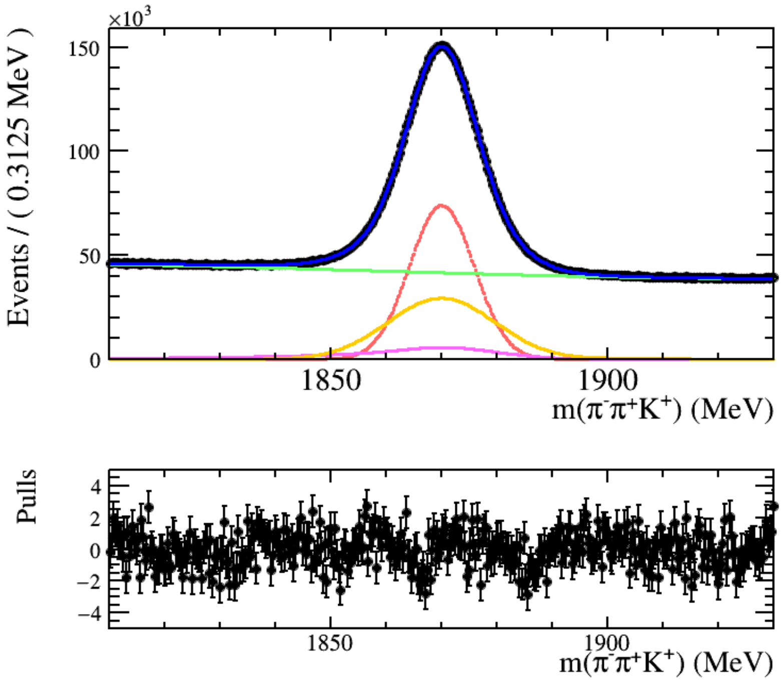


Figure 4.18: Mass fit of $D^+ \to \pi^-\pi^+K^+$ final sample..

Parameters	Result
μ	1869.95 ± 0.0047
σ_G	5.8195 ± 0.0048
a_0	1.1782 ± 0.0028
a_1	1.1388 ± 0.0050
a_3	1.0041 ± 0.0035

Table 4.8: Data sample fit results.

An effective sigma σ_{eff} can be calculated from the fit standard deviation σ_G,

$$\sigma_{eff} = \sqrt{f_G\sigma_G^2 + (1-f_G)f_{CB}\sigma_{CB_1}^2 + (1-f_G)(1-f_{CB})\sigma_{CB_2}^2} \tag{4-5}$$

and thus, from Eq. 4-5, the signal yields can obtained within a region of $2\sigma_{eff}$ showed in Table 4.9

σ_{eff}	7.825 ± 0.006 [MeV]
Signal region	1854.3 - 1885.6 [MeV]
Purity	58.48 ± 0.02 (%)

Table 4.9: Signal region obtained from the σ_{eff}.

and the yields are presented in Table 4.10,

Full spectrum yields ($\times 10^6$)		Signal region yields ($\times 10^6$)	
Background	Signal	Background	Signal
16.062 ± 0.006	6.351 ± 0.005	4.170 ± 0.002	5.875 ± 0.005

Table 4.10: Signal and background yields obtained from the mass fit for $D^+ \to \pi^-\pi^+K^+$.

4.5.3 Dalitz plot

In the $D^+ \to \pi^-\pi^+K^+$ Dalitz Plot, as mentioned in Chapter 2, there is a clear sign of the resonant contributions $K^*(892)$, on the mass $m^2_{\pi K}$, and $\rho(770)$, on $m^2_{\pi\pi}$, as can be seen in Fig. 4.19.

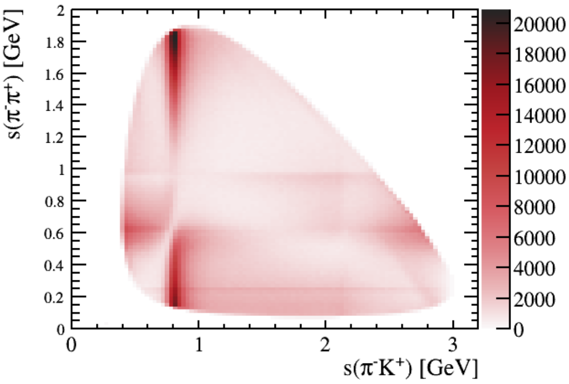


Figure 4.19: Dalitz plot of the final sample full mass spectrum (left) and $2\sigma_{eff}$ signal region (right) of $D^+ \to \pi^-\pi^+K^+$.

The observation of the DP shows a beautiful pattern of interference between these two resonant structures, which creates a region with a higher density of events around $s(\pi^-K^+) \approx 0.8$ GeV^2 and $s(\pi^-\pi^+) \approx 0.6$ GeV^2. In addition, a closer analysis reveals other two regions where a slight linear concentrations of events can be seen, indicating the presence of another resonant structures. On the mass $s(\pi^-K^+) \approx 2.1$ GeV, it may indicate an interference between the decay chains $K^*(1410)$ and $K^{*0}(1430)$, and for $s(\pi^-\pi^+) \approx 0.98$ GeV^2 the contribution of $f_0(980)$.

4.6 Control channel: $D^+ \to K^-\pi^+\pi^+$ selection

The decay $D^+ \to K^-\pi^+\pi^+$ is Cabibbo-favoured and, thus, has no CPV effects predicted within the SM. Therefore, since it can not introduce asymmetries resulting from CP violation, this decay was chosen as a control channel for this analysis in order to check for production and detection asymmetry effects, as discussed later in Chapter 5. The origin of the production asymmetry comes from the fact the pp collisions are not charged symmetric which may result in a particle production being favoured compared to its antiparticle, for example the $D^-(\bar{c}d)$ is slightly more produced than the $D^+(c\bar{d})$. On the other hand, detection asymmetry is related to charge asymmetries introduced by the detector.

The selection criteria used on the control channel are the same ones adopted for the main channel. The $D^+ \to K^-\pi^+\pi^+$ has its exclusive Turbo line, `Hlt2CharmHadDpToKmPipPip`, with the same HLT2 requirements as the main channel, showed on Table 4.1. Due to large number of events in this channel, the samples were divided into smaller samples, as showed in Table 4.11, with trigger level requirements and the additional pre-selection requirements.

Year		Number of samples	Number of events per sample
2016	Down	19	$\approx$ 13M
	Up	19	$\approx$ 13M
2017	Down	19	$\approx$ 13M
	Up	19	$\approx$ 13M
2018	Down	21	$\approx$ 13M
	Up	21	$\approx$ 13M

Table 4.11: $D^+ \to K^-\pi^+\pi^+$ subsamples created for each year/polarity.

These subsamples also went trough the specific requirements of $D^+ \to \pi^-\pi^+K^+$ the offline selection, applying the PID requirement for the kaon daughter of the main channel on the kaon daughter of the control channel, and the PID requirement for particle 1 (π^-) of the main channel on particle 3 (π^+) of the control channel. The requirement to remove the $\overline{D}^0 \to \pi^-\pi^0K^+$contribution was applied on the second daughter, the π^+, as it was on the main channel. These requirements are summarised on Table 4.12.

Requirements	$D^+ \to \pi^-(1)\pi^+(2)K^+(3)$ to $D^+ \to K^-(1)\pi^+(2)\pi^+(2)$
PID	$\mathrm{ProbNNk}_3$ <0.02
	$\mathrm{ProbNNk}_1$ >0.64
Specific background	$\chi^2_{\mathrm{IP},2}$ >30

Table 4.12: Offline selection requirements applied on $D^+ \to K^-\pi^+\pi^+$ summary.

To be consistent with the main channel analysis, the subsamples were also divided into A and B samples and the MVA response obtained on the main channel was evaluated on them, again using the 6-folded method. After the MVA application, the BDT requirement valBDT_mean > 0.08 was also applied on the control channel samples.

After this selection process, the yields for the control channel are presented in Table 4.13 for the full mass spectra and for the signal region, defined on the same mass window as in Table 4.5. The yields are approximately five times higher the main channel yields. The invariant mass distribution after the selection process can be seen in Fig. 4.20.

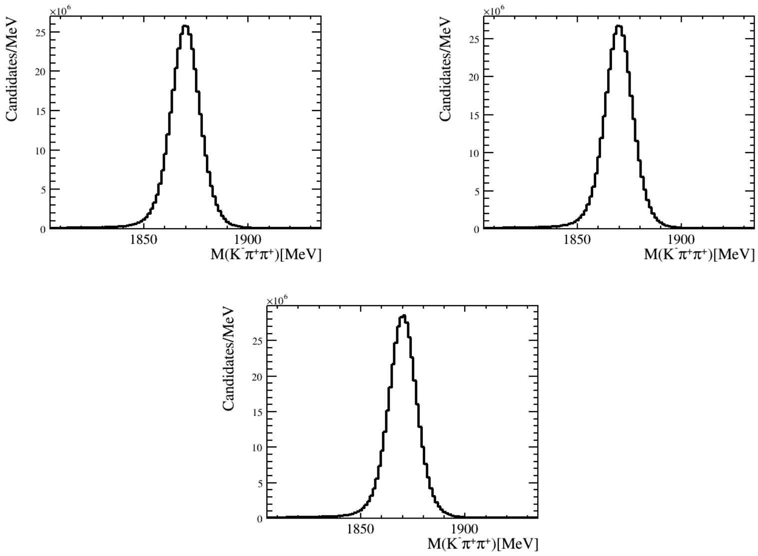


Figure 4.20: Invariant mass distribution of the samples: 2016 (top left), 2017 (top right) and 2018 (bottom) for the decay $D^+ \to \pi^-\pi^+K^+$.

Year		Full spectrum ($\times 10^7$)	Signal region ($\times 10^7$)
2016	A	1.798	1.732
	B	1.798	1.732
	Total	3.596	3.464
2017	A	1.807	1.743
	B	1.807	1.743
	Total	3.614	3.486
2018	A	1.925	1.858
	B	1.925	1.858
	Total	3.850	3.716
Total		11.6	10.6

Table 4.13: $D^+ \to K^-\pi^+\pi^+$ final statistics for the full mass spectrum and the signal region after the selection process.

5
Search for local CPV in the $D^+ \to \pi^-\pi^+K^+$ phase space

5.1
Analysis strategy

One approach to search for CP violation effects across the $D^+ \to \pi^-\pi^+K^+$ phase space is the *Mirandizing* method [60, 61], which consists of a statistical comparison between the DP of particle and antiparticle. This technique does not depend on the amplitude parameterisation of the decay's DP, i.e. it is model independent, and it is a simple and fast tool to search for local CP violation effects.

The method consists in dividing the DP into sections called bins and a significance of the difference between D^+ and D^- candidates is computed bin per bin,

$$\mathcal{S}^i_{CP} = \frac{N^i(D^+) - \alpha N^i(D^-)}{\sqrt{\alpha\left(\delta^2_{N^i(D^+)} + \delta^2_{N^i(D^-)}\right)}}, \quad \alpha = \frac{N_{\text{tot}}(D^+)}{N_{\text{tot}}(D^-)} \tag{5-1}$$

where $N^i(D^\pm)$ is the observed number of decays in the i^{th} bin of the DP, $N_{\text{tot}(D^\pm)}$ is the sum over all bins and $\delta^2_{N^i(D^\pm)}$ is are uncertainties. The factor α is introduced to account for global asymmetry effects, like production asymmetry originated from the parent meson, that can lead to an overall charge asymmetry, or charge detection asymmetries. This effect most of the times is expected to be constant over the DP. Unfortunately, at the same time the α may also removes global CP asymmetries. Nevertheless, the technique is confirmed to be sensitive to the presence of local effects. The other terms in Eq. 5-1,

For the case where there is no statistically significant local charge asymmetries, the only difference between the compared phase spaces is be due to statistical fluctuations. To test this hypothesis, a χ^2 test, which can be obtained from the S^i_{CP} distribution, is performed,

$$\chi^2 = \sum_{i=1}^{N_{\text{bins}}} \left(S^i_{CP}\right)^2 \tag{5-2}$$

and, by using the χ^2 and the number of degrees of freedom (NDOF = N_{bins} - 1, that is equal to the number of bins minus one subtracted from the global α constraint), a p-value can be obtained. The p-value that results from this test represents the probability of obtaining, for a given number of d.o.f, and under the assumption of no CP violation (the null-hypothesis in this case), a

χ^2 as high as the value observed [62]. This p-value quantifies the confidence level on the measurement that the D^+ and D^- Dalitz Plots are statistically compatible, i.e, that they differ only by statistical fluctuations and not by a significant difference that could indicate CPV.

A statistical significance in terms of the standard deviation σ between samples can be obtained from the p-value using Gaussian statistics. With this, a difference of 3σ would correspond to a p-value ~ 0.03. In order to claim an effect of CP violation is observed, a p-value $< 3 \times 10^{-7}$ is required, corresponding to the 5σ threshold. However, this method only allows to verify whether if CP violation is observed or not but does not quantify the corresponding asymmetry, for this result it would require the calculation of the A_{CP} mentioned in Chapter 2.

In order to reject the null-hypothesis, it is essential to ensure that the asymmetry being probed is indeed from CP violation and do not come from eventual nuisance asymmetries. The control channel $D^+ \to K^-\pi^+\pi^+$, that has the same final state particles but with different charge, and has a similar topology to $D^+ \to \pi^-\pi^+K^+$, was used to study nuisance local effects, such as differences due to the reconstruction, selection or trigger efficiencies, left-right detector asymmetries and different production mechanisms for D^+ and D^-. Besides the control channel, the Monte Carlo samples of $D^+ \to \pi^-\pi^+K^+$ are also used to test for detection and production asymmetries (mentioned in Section 4.6). In addition, another test is performed in the background regions of the decay of interest to ensure that no effects of charge-asymmetry coming from these events is introduced to data. If the study probes that no local asymmetries are observed for the control channel, the MC samples and on the signal channel's background, the test can then be applied to the main decay channel.

As demonstrated in a recent analysis [63], the presence of background events in the calculation can introduce biases in the S_{CP}, for large samples and low purity after the selection. Due to this, a new and more accurate implementation of the original method was used to perform these studies in this work, the *fit-per-bin method.* By this approach, the number of events per bin of the phase space is obtained by fitting the corresponding invariant-mass spectra of $D^+ \to \pi^-\pi^+K^+$ in each bin and collecting the signal yields $N^i(D^\pm)$ and their uncertainties $\delta^2_{N^i(D^\pm)}$. This technique allows to remove the background effects and also suits to the different signal and background shapes around the DP.

The definition of the binning scheme is important for increasing the method sensitivity. Binning schemes with $\mathcal{O}(20)$ have been shown to be ideal since by increasing the number of bins may additional bin dilute the statistical precision in the phase space. For these studies, three types of binning schemes were applied: uniform, adaptive and physics motivated. The uniform binning is constructed by dividing the phase space into equal-sized bins. Adaptive binning, on the other hand, has different-sized bins, but all bins have the same number of events. And finally, the physics motivated binning is designed by hand using information about the main resonant structures of the decay in order to be more sensitive to local effects that may be introduced by them.

The configuration of binning schemes chosen was: two pairs of uniform binning, being one of 5×5 grid with 19 effectively occupied bins and one of 8×8 grid with 39 occupied bins; one adaptive binning with 25 bins and one physics motivated with 24 bins. These binning schemes can be seen in Fig. 5.1 where the Dalitz plot for the signal region is displayed after all selection requirements.

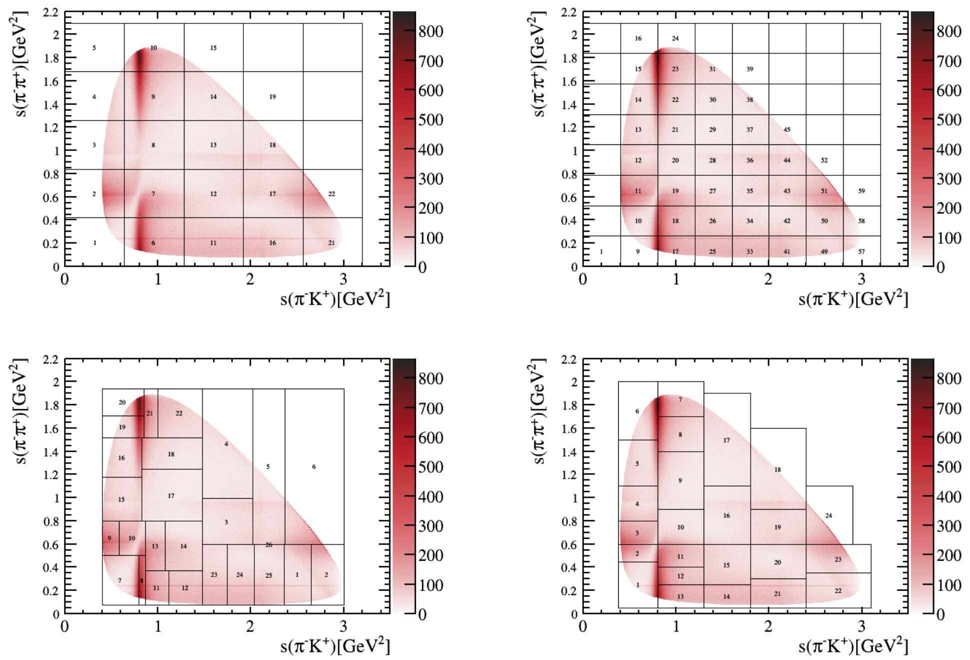


Figure 5.1: DP binning schemes plotted for the events of the $D^+ \to \pi^-\pi^+K^+$ final sample signal region. Uniform 5×5 grid (top left), uniform 8×8 grid (top right), adaptive with 25 bins (bottom left) and physics motivated with 24 bins (bottom right). The binning configuration number is shown.

As it have been adopted by many analysis in the LHCb Collaboration, in order to ensure that the final result is not driven by *experimenter's bias* [64], this analysis is being carried out blinded. A blinded analysis means that all necessary steps and checks need to be performed without actually looking the actual main observable, in this case the S_{CP} values, at the main channel. Results for the CPV search for $D^+ \to \pi^-\pi^+K^+$ will only be accessed after it is demonstrated far confidence in the method's response and any other possible source of asymmetry is identified, such that it does not spoil the method.

5.2 Null-test

To validate mainly the method of fit-per-bin, a null-test was performed using samples that were generated dividing the final data sample of $D^+ \to \pi^-\pi^+K^+$ randomly into two parts. The size of the samples 1 and 2 is proportional to the number of events of D^+ and D^-, respectively, but no charge distinction was made, characterising a null-test. The statistics of these samples are presented in Table 5.1. To obtain the $N^i(D^\pm)$ values, an integrated invariant mass fit, i.e. with no distinction of charge, was performed for each bin to extract the signal yields and its uncertainties, the fit-per-bin method. An small asymmetry of the order of 0.5% between the two samples is introduced. This asymmetry introduces a fake α factor, typical of D^+ and D^- decays. For this test, the physics binning scheme was used. Additionally, as adopted on previous analysis [65–67], the test was also performed using the total number of events within the signal region (defined in Table 4.5) in each bin for $N^i(D^\pm)$ and the uncertainties were simply the square root of the number of candidates in each bin. This step was performed also as a comparison to the fit-per-bin results and all three binning schemes were used here.

This test was performed mainly to verify the quality of fit model for the different bins of the Dalitz plot. For the fit-per-bin method, the fit for some bins is shown in Fig. 5.2 and on Appendix A. For all bins, the results demonstrate that good fits were obtained in all regions of the Dalitz plot. The result for the null-test of both studies is presented in Table 5.2 and the $\mathcal{S}_{CP}$ distributions can be seen in Fig. 5.3 for the regular event counting and in Fig. 5.4 for the fit-per-bin. As it can be seen, the results for both tests show the expected result of *p-values* compatible with null hypothesis.

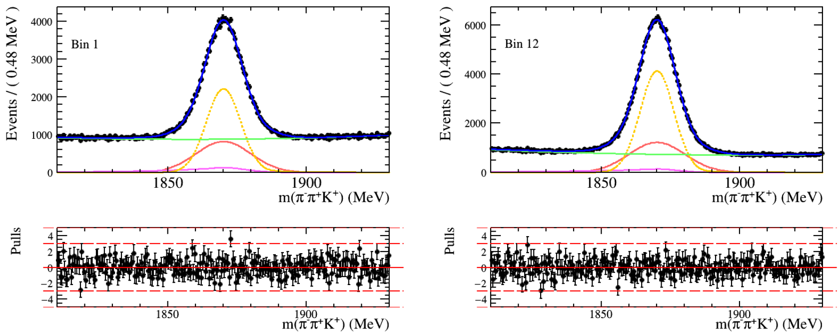


Figure 5.2: Invariant mass fit of bin 1 of sample 1 (left) and bin 12 of sample 2 (right), used on the fit-per-bin method.

Sample	Total events	Signal Yield	Events in $2\sigma_{eff}$	Purity in $2\sigma_{eff}$
1	11 206 920	3 197 590 ± 3808	5 037 490	(58.71 ± 0.03)%
2	10 963 930	3 124 410 ± 3764	4 924 790	(58.68 ± 0.03)%

Table 5.1: Statistics of the $D^+ \to \pi^-\pi^+K^+$ null-test samples.

Binning		Standard event count		fit-per-bin	
		$\chi^2/ndof$	p-value (%)	$\chi^2/ndof$	p-value (%)
Uniform	5 × 5	31.2 / 20	5.2	-	-
	8 × 8	56.5 / 41	5.3	-	-
Adaptive	25	30.4 / 24	13.8	-	-
Physics	24	43.3 / 23	0.6	30.2 / 23	14.3

Table 5.2: Null-test results for $D^+ \to \pi^-\pi^+K^+$.

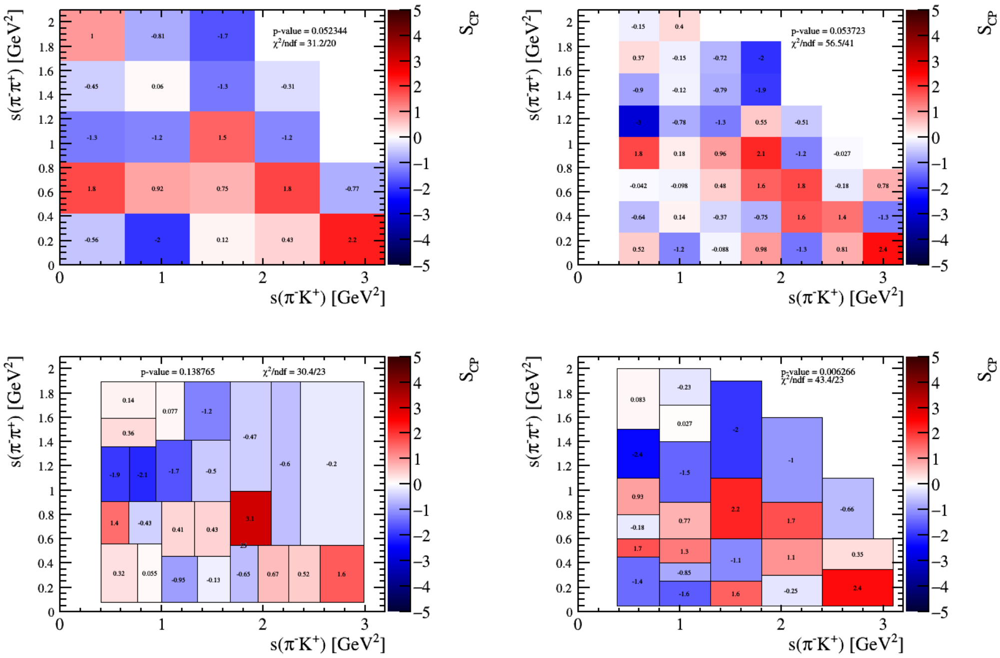


Figure 5.3: $D^+ \to \pi^-\pi^+K^+$ null test counting all entries in the bin (standard Mirandizing test), for uniform 5 × 5 grid (top left), uniform 8 × 8 grid (top right), adaptive with 25 bins (bottom left) and physics motivated with 24 bins (bottom right).

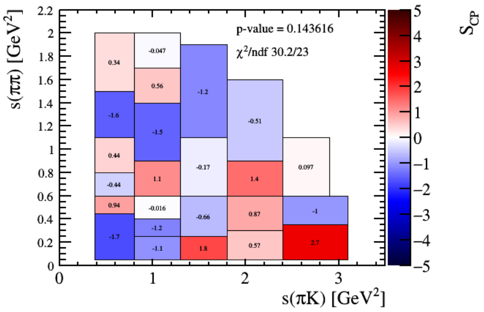


Figure 5.4: $D^+ \to \pi^-\pi^+K^+$ null test using the fit-per-bin method for the physics motivated binning with 24 bins.

5.3 Study of nuisance asymmetries

As part of the blind analysis requirements, tests for the presence nuisance asymmetries are performed. For these tests, the background region of the main channel and the signal region of the control channel were analysed, as well as the MC samples to guarantee the method is not sensitive to nuisance asymmetries.

5.3.1 Background

The background events of the main channel were also tested to look for any possible charge-asymmetries in these regions that could cause a fake CP violation signal in the unblided analysis for main channel.

For this test, the events within the sidebands of the invariant mass spectra were considered to obtain the $N^i(D^{\pm})$ for each bin. The mass windows used were the same defined in Table 4.5, and the number of events in this region is 5.3×10^6 in total. The fit-per-bin method was also performed to obtain the background yields (while maintaining the signal yields blinded) and their uncertainties to be used on the calculus of the $\mathcal{S}_{CP}$.

The results for this background study are shown in Table 5.3 and on Fig. 5.5 and Fig. 5.6. All the results presented demonstrate that the method is not sensitive for local nuisance asymmetries in the background region.

Binning		Standard event count		fit-per-bin	
		$\chi^2/ndof$	p-value (%)	$\chi^2/ndof$	p-value (%)
Uniform	5×5	15.8 / 18	59.9	-	-
	8×8	50.7 / 38	8.2	-	-
Adaptive	25	23.2 / 23	44.6	-	-
Physics	24	22.0 / 23	51.8	29.9 / 23	15.3

Table 5.3: Background check for $D^+ \to \pi^-\pi^+K^+$.

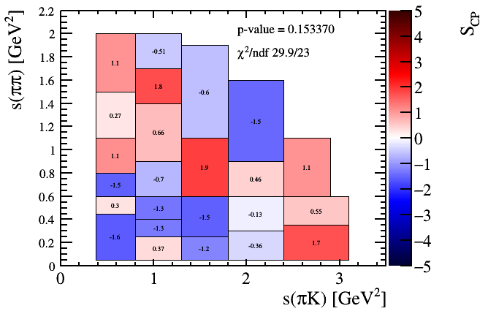


Figure 5.5: $D^+ \to \pi^-\pi^+K^+$ background test using the fit-per-bin method for the physics motivated binning with 24 bins.

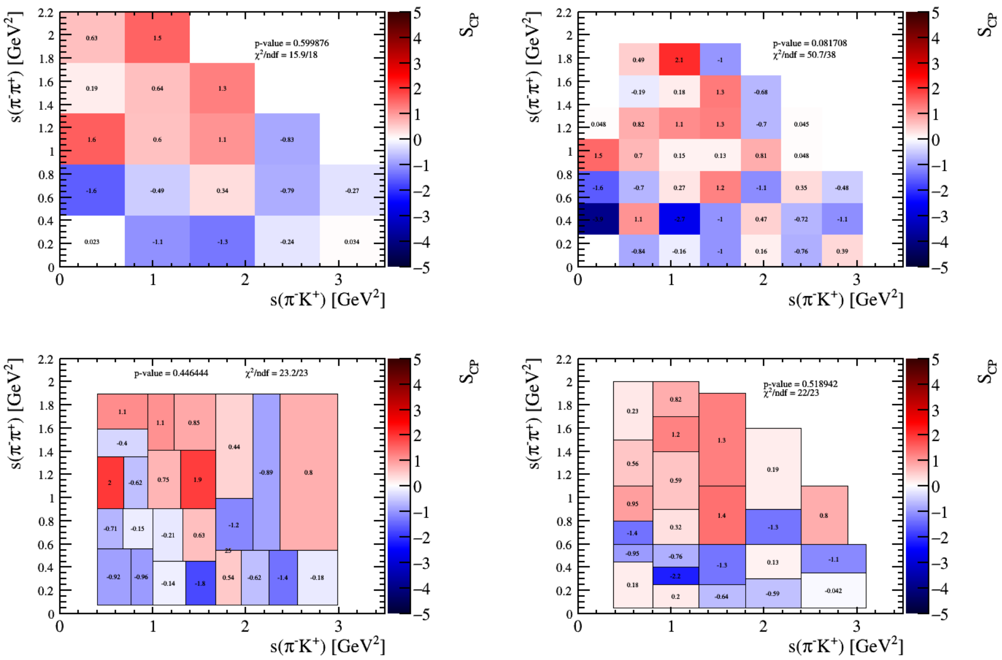


Figure 5.6: $D^+ \to \pi^-\pi^+K^+$ background test using standard event counting in each bin for uniform 5×5 grid (top left), uniform 8×8 grid (top right), adaptive with 25 bins (bottom left) and physics motivated with 24 bins (bottom right).

5.3.2 Monte Carlo

A test was performed on the Monte Carlo samples to look for local charge asymmetries effects introduced by the detector and asymmetries effects originated by the parent meson production.

For the MC simulated samples the fit-per-bin was not applied here once there is no need to remove the background contribution. To determine the quantities $N^i(D^\pm)$, the total number of entries in the D^+ and D^- Dalitz plots are consider for the calculation and the uncertainties are the square root of the number of candidates in each bin.

For this test, uniform binning scheme with 5×5 and 8×8 grids, adaptive binning with 25 bins and physics motivated binning with 24 bins were used. The number of events is 7.7×10^6 in total after all selection criteria, including weights for PIDCalib.

The results for the test are shown in Table 5.4 and on Fig. 5.7, and the test results are compatible with no evidence of local asymmetries present in the MC samples.

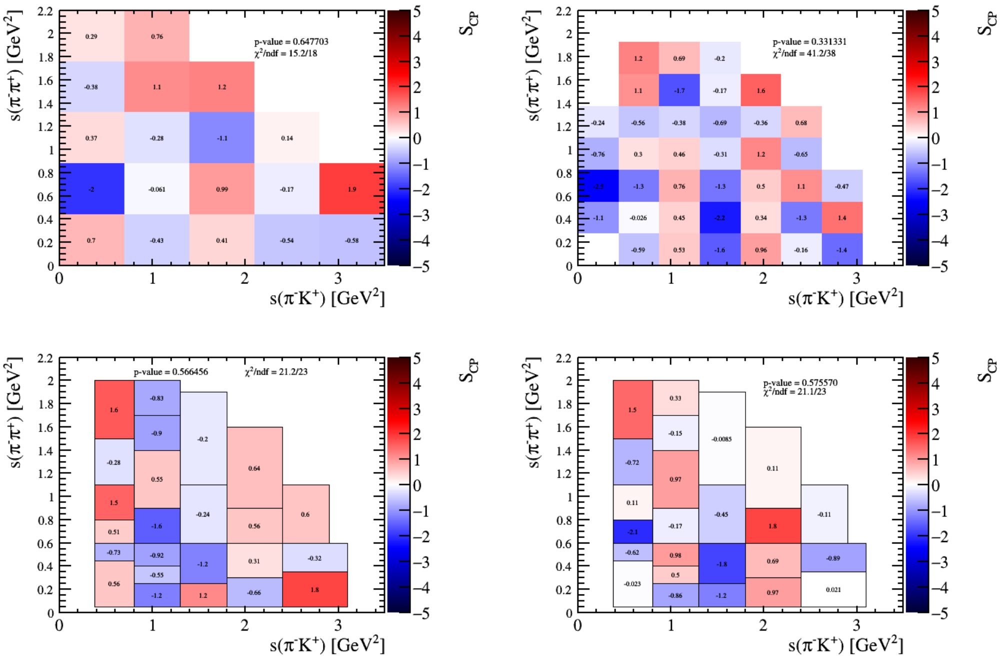


Figure 5.7: Distribution of p-value responses over the search in the MC samples for uniform 5×5 (top left) and 8×8 (top right) grids, adaptive with 25 bins (bottom left) and physics motivated with 24 bins (bottom right).

Binning		$\chi^2/ndof$	p-value (%)
Uniform	5×5	15.2 / 19	64.7
	8×8	41.2 / 39	33.1
Adaptive	25	21.2 / 24	56.6
Physics motivated	24	21.1 / 23	57.5

Table 5.4: p-value responses over the search for charge asymmetries in the MC samples after the selection criteria.

5.3.3 Control channel

Last but no least, a test was performed in the control channel $D^+ \to K^-\pi^+\pi^+$ to look for local nuisance asymmetries, like detection and production asymmetries, since it has the same final state as the main channel excepting for different charge signs. These samples present a high purity and, thus, just like on the MC study, there was no need to perform the fit-per-bin to obtain the $N^i(D^\pm)$ values. All the samples presented in Table 4.11 were selecting events within the signal region defined in Table 4.5. After the selection requirements the statistics for these samples are approximately 5 times larger than the signal channel.

To perform the search, the uniform and adaptive binning schemes were applied to the DP, presented in Fig. 5.8, with 5×5 and 8×8 grids for the uniform and 25 bins for the adaptive. The total number of entries in the D^+ and D^- Dalitz plots are considered to obtain the $N^i(D^\pm)$ and the uncertainties are the square root of the number of candidates in each bin. The results for one sample are presented in Fig. 5.9. The p-value distribution for all the samples can be seen in Fig. 5.10. By the results obtained in this analysis for $D^+ \to K^-\pi^+\pi^+$, it can be concluded that there is no evidence of charge asymmetry present in the control channel.

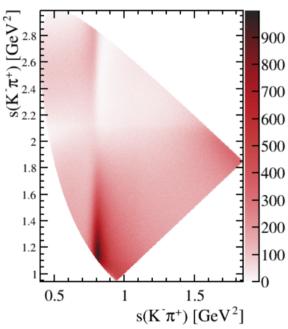


Figure 5.8: Dalitz plot for one sample of the control channel $D^+ \to K^-\pi^+\pi^+$.

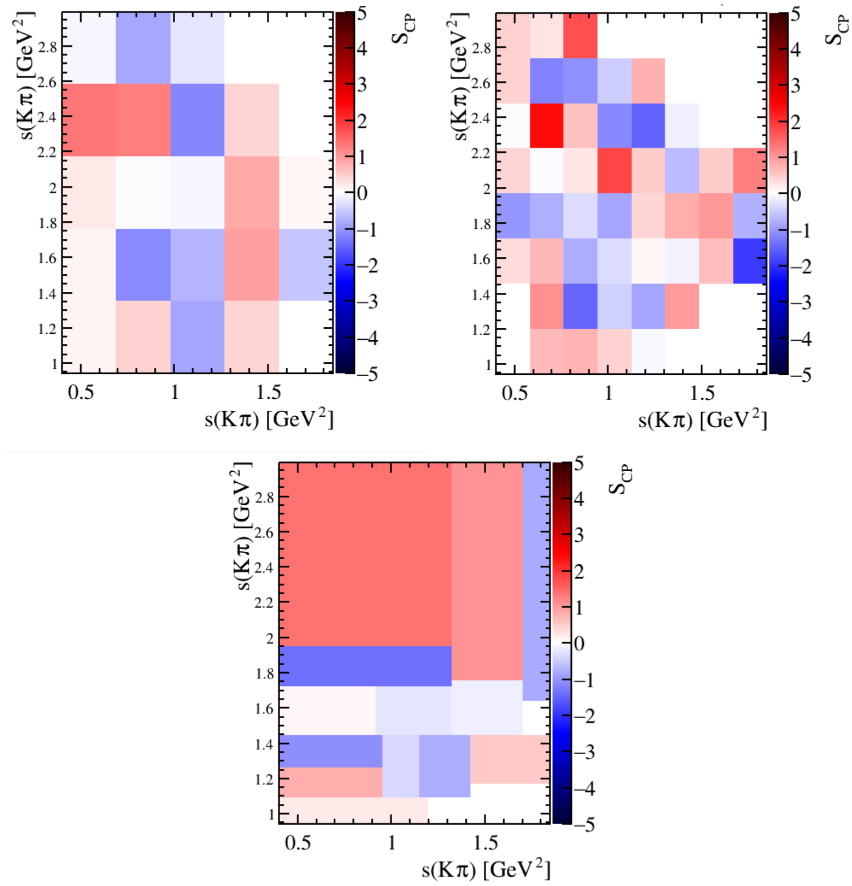


Figure 5.9: $\mathcal{S}_{CP}$ distribution for the control channel $D^+ \to K^-\pi^+\pi^+$ using 5×5 grid (top left) and 8×8 grid (top right) for the uniform and 25 bins (bottom) for the adaptive binning scheme.

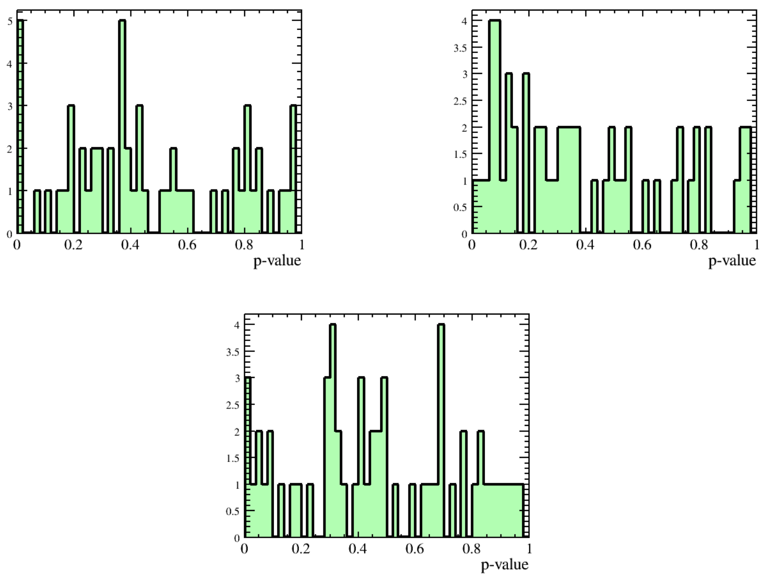


Figure 5.10: Distribution of p-value responses over the search in all samples of the control channel for uniform 5×5 grid (top left) and 8×8 grid (top right) and adaptive with 25 bins (bottom).

5.4 Blind analysis: procedure and prospects

Since this is a blind analysis, as already mentioned previously, the only available results until this point are from the nuisance asymmetries tests presented in this Chapter. This analysis, as part of a LHCb analysis, follows the Collaboration procedure. The first stage is the blind analysis, where the method and all the test results are reviewed by the correspondent Working Group (WG), which in this analysis is the Charm WG. After this step, the analysis is authorised to enter the internal Review Committee (RC) process, where the full analysis is reviewed and is, then, approved to be unblinded. After this, the analysis is eligible to receive the "approval to go to paper" and the final paper is reviewed widely by the whole Collaboration. The present analysis is on the way to start the WG review process.

After receiving the authorisation to unblind this analysis, the method chosen to take the p-value for the signal region is the fit-per-bin method.

As mentioned, for this analysis to be sensitive to CP violation, the p-value obtained should be lower than the 5σ threshold, i.e., p-value $< 3 \times 10^{-7}$. For a value above this threshold, it indicates that the studied sample is not statistically sensitive to CP violation in the $D^+ \to \pi^-\pi^+K^+$ channel, as predicted by the SM for doubly-Cabibbo suppressed decays and no effects of NP are originated from this decay. On the other hand, a value below this threshold is an evidence for CPV in the studied channel and this channel is a possible source for NP beyond the SM.

6
Conclusions

This dissertation reports the analysis of the doubly Cabibbo suppressed decay $D^+ \to \pi^-\pi^+K^+$ for CP violation studies. According to the Standard Model of particle physics, there is no prediction for CPV effects in DCS charm decays and in that way this analysis represents a direct search for new physics.

The phase space of a three-body decay, the Dalitz Plot, often presents a rich dynamics with the presence of resonant structures, and the interference between these resonant structures may favours the observation of mensurable local CPV effects higher then the phase space integrated ones. That way, to perform this search the DPs for D^+ and D^- can be divided into bins and the yields can be compared bin per bin (via the $\mathcal{S}_{CP}$ observable) to measure how significantly different they are; this is called the *Mirandizing* method.

For this analysis, data collected by the LHCb from 2016-2018 was used and a selection process was performed in order to reduce contributions from cross-fed of other charm decays, together with a multi-variate analysis to reduce combinatorial background contributions. This step is necessary in order to minimise spurious asymmetries that may come from these contamination and dilute potential CP violation signals, and also increase the statistical significance of the sample. After the selection process, a final sample of about 6M decays was obtained, with a purity of 58%. This is nowadays the largest sample ever obtained for a DCS D^+ decay channel.

The model-independent Mirandizing method chosen for this analysis is the fit-per-bin method, where the full mass spectra of each bin of the DP is fitted to obtain the signal yields and their uncertainties and to remove background contributions in the $\mathcal{S}_{CP}$ calculation. The fit-per-bin method is used since the presence of large background can introduce biases in the calculations [63]. For this binned search, three types of binning schemes were used: uniform, adaptive and physics motivated binning, in order increase the method's sensitivity.

A null-test was performed in data divided randomly into two samples with no distinction of charge to verify the quality of the fit model for the different bins of the DP and test the internal consistency of the method.

The method was then applied to the background candidates to verify if any possible charge-asymmetries may come from this region, both using counting and fit-per-bin method. By using the background yields while not looking at the signal yields guarantee that the analysis remains blinded. For both tests the results are compatible with no asymmetries in the background region.

A test was also performed on Monte Carlo simulated samples to search for nuisance asymmetries, like detection and production effects. Since these samples are composed of only signal candidates, there is no need to remove the background contribution: to determine the quantities used in the the $\mathcal{S}_{CP}$ only the total number of entries in the D^+ and D^- Dalitz plots were consider for the calculation. The results for this test present no sensitivity for spurious asymmetries effects.

A final test for nuisance asymmetries was performed using the Cabibbo favoured decay $D^+ \to K^-\pi^+\pi^+$ as a control channel. The control channel sample was divided into 59 samples with much higher statistics than the signal channel. For this test, since the samples had a high purity and it was not necessary to overcome the background contribution, only the total entries in each bin were considered for the calculation, just like for the MC test. The results for the control channel present no sensitivity for nuisance effects.

The final step of this analysis is the unblinding procedure of the results for the signal channel $D^+ \to \pi^-\pi^+K^+$, which has not been performed yet. Currently, this analysis is on the way to start the Charm WG review process and follow all the stages until the unblinding approval. After the unblind, the final results will be reported on the analysis paper submitted for publication.

A
Fits per bin for the null-test

Here are presented the fits to the invariant mass distributions per bin of the Dalitz Plot of Sample 1, used on the null-test performed on the data (5.2). The physics motivated binning scheme with 24 total bins is shown in Fig. A.1 and the fit per bin in Figs. A.2 - A.5.

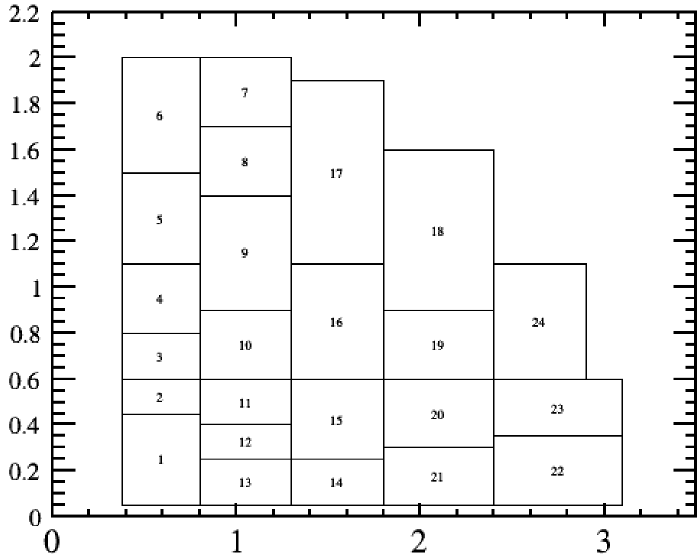


Figure A.1: Physics motivated binning scheme for the $D^+ \to \pi^-\pi^+K^+$ Dalitz Plot.

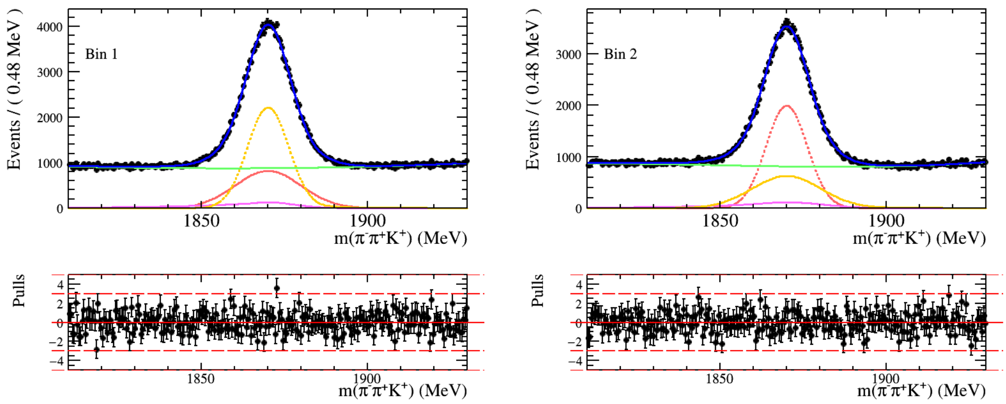


Figure A.2: $D^+ \to \pi^-\pi^+K^+$ invariant mass fit per bin of the physics motivated binning scheme with 24 total bins in sample 1 (bins 1 and 2).

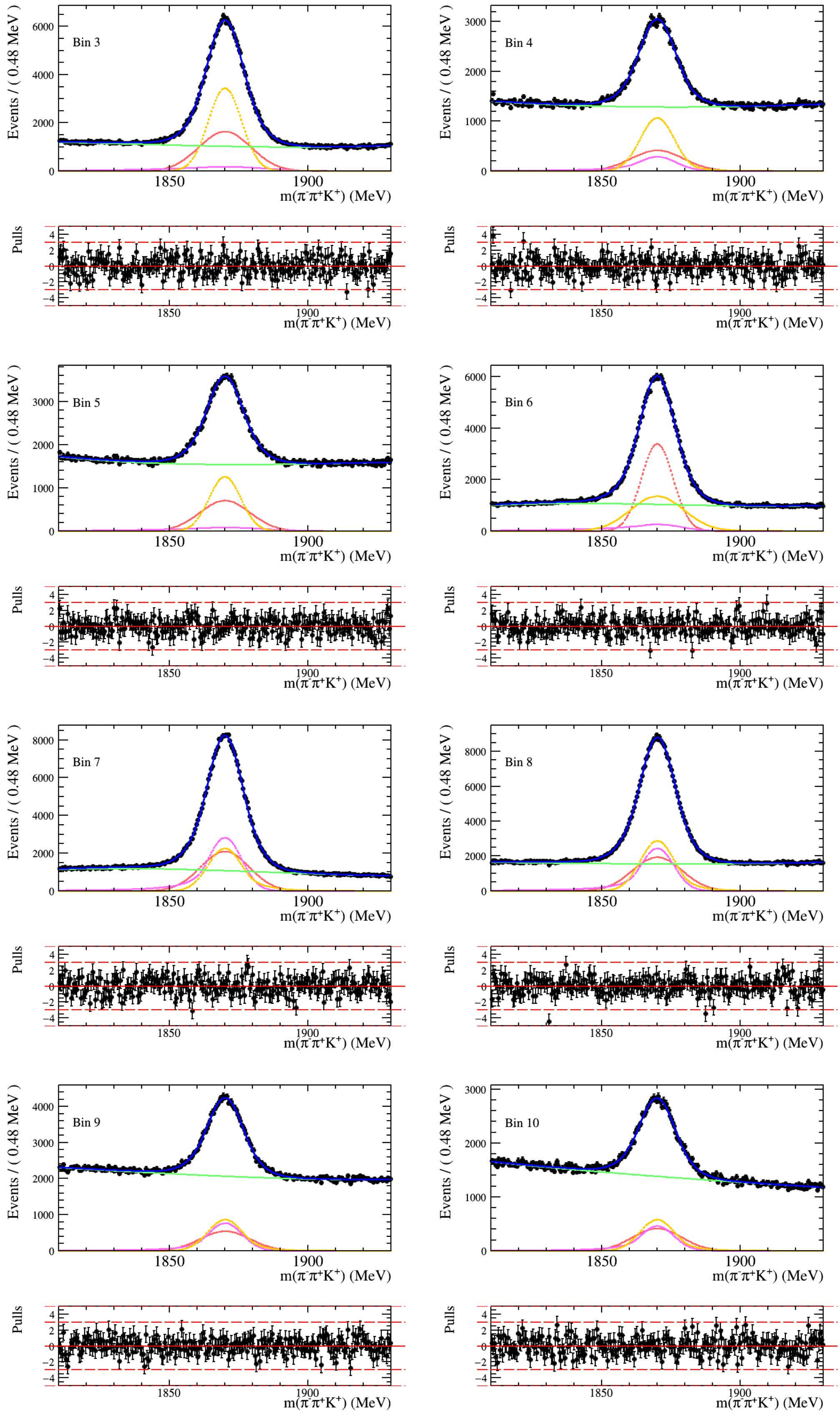


Figure A.3: $D^+ \to \pi^-\pi^+K^+$ invariant mass fit per bin of the physics motivated binning scheme with 24 total bins in sample 1 (bins 3-10).

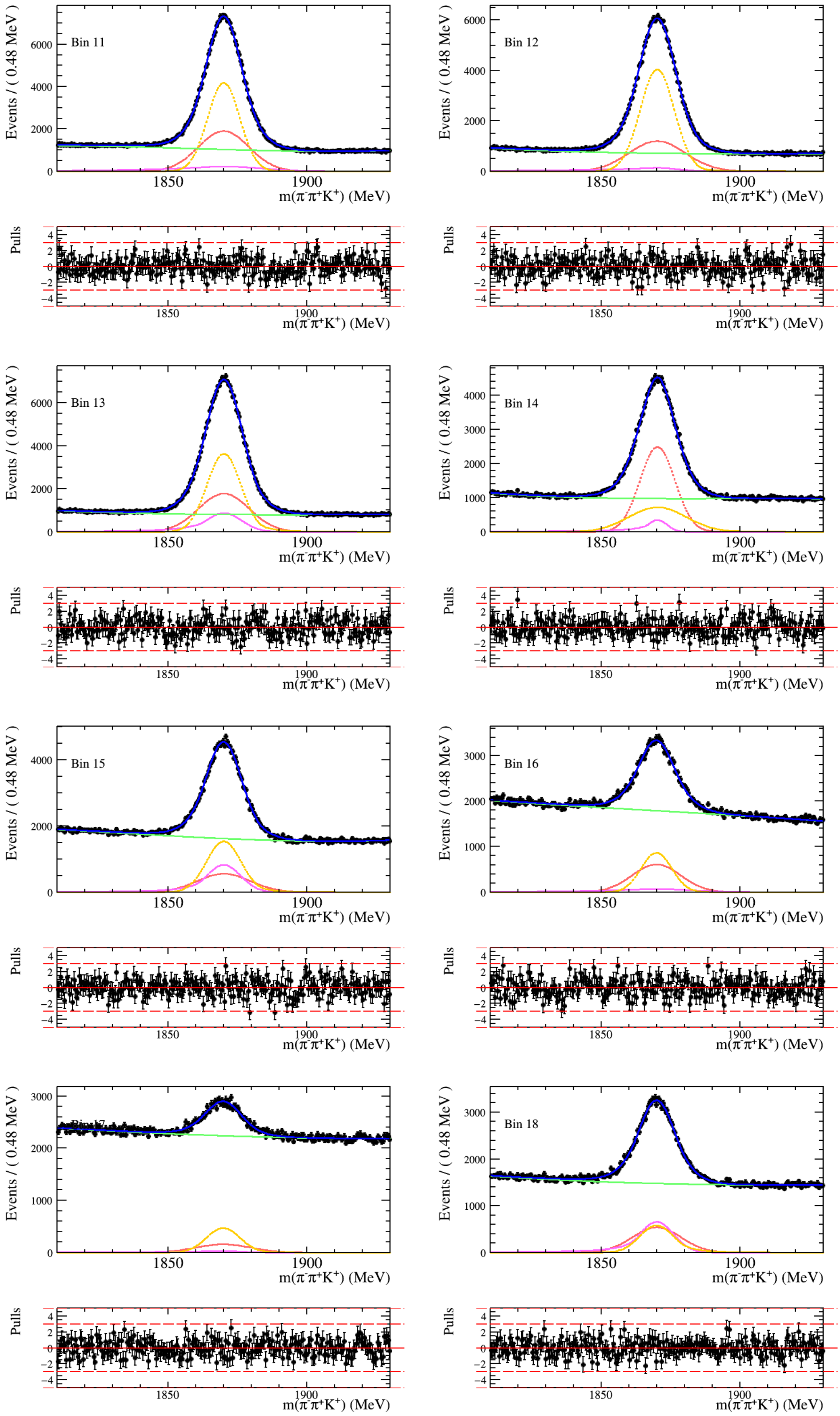


Figure A.4: $D^+ \to \pi^-\pi^+K^+$ invariant mass fit per bin of the physics motivated binning scheme with 24 total bins in sample 1 (bins 11-18).

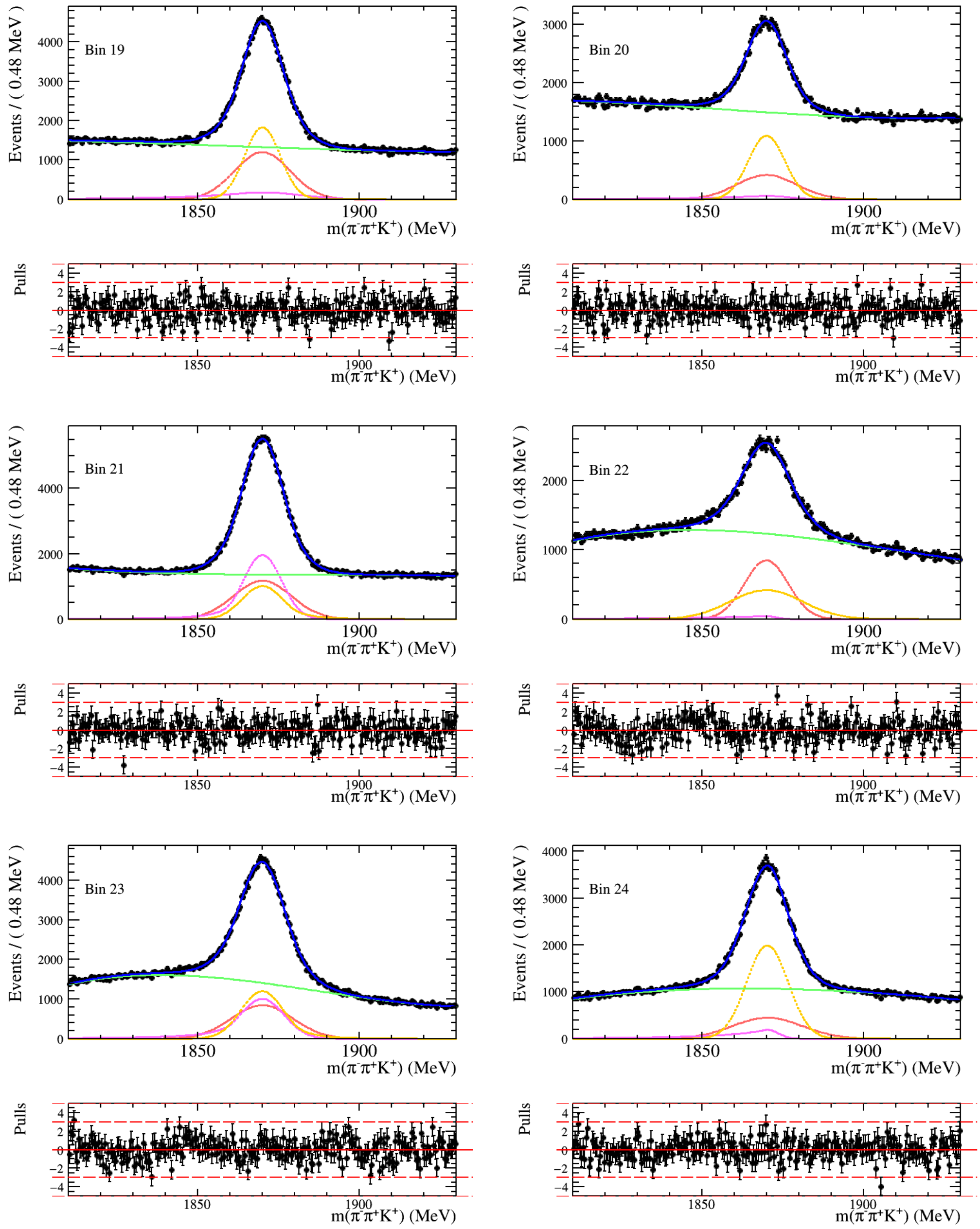


Figure A.5: $D^+ \to \pi^-\pi^+K^+$ invariant mass fit per bin of the physics motivated binning scheme with 24 total bins in sample 1 (bins 19-24).

B
Bibliography

[1] A. D. Sakharov. Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe. *Pisma Zh. Eksp. Teor. Fiz.*, 5:32–35, 1967. doi: 10.1070/PU1991v034n05ABEH002497.

[2] Ikaros I. Bigi and A. I. Sanda. *CP violation*, volume 9. Cambridge University Press, 9 2009. ISBN 978-0-511-57728-4, 978-0-521-84794-0, 978-1-107-42430-2. doi: 10.1017/CBO9780511581014.

[3] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π decay of the k_2^0 meson. *Phys. Rev. Lett.*, 13:138–140, Jul 1964. doi: 10.1103/PhysRevLett.13.138. URL `https://link.aps.org/doi/10.1103/PhysRevLett.13.138`.

[4] I. Bediaga and C. Göbel. Direct cP violation in beauty and charm hadron decays. *Progress in Particle and Nuclear Physics*, 114:103808, 2020. ISSN 0146-6410. doi: https://doi.org/10.1016/j.ppnp.2020.103808. URL `https://www.sciencedirect.com/science/article/pii/S0146641020300557`.

[5] L. Canetti, M. Drewes, and M. Shaposhnikov. Matter and antimatter in the universe. *New Journal of Physics*, 14(9):095012, Sep 2012. doi: 10.1088/1367-2630/14/9/095012. URL `https://doi.org/10.1088%2F1367-2630%2F14%2F9%2F095012`.

[6] Y. Grossman, Y. Nir, and R. Rattazzi. CP violation beyond the standard model. In *Heavy Flavours II*, pages 755–794. WORLD SCIENTIFIC, Jun 1998. doi: 10.1142/9789812812667_0011. URL `https://doi.org/10.1142%2F9789812812667_0011`.

[7] R. Aaij et al. Observation of CP violation in charm decays. *Phys. Rev. Lett.*, 122:211803, May 2019. doi: 10.1103/PhysRevLett.122.211803. URL `https://link.aps.org/doi/10.1103/PhysRevLett.122.211803`.

[8] X. Pan. Doubly cabibbo-suppressed D decays at BESIII. *Suplemento de la Revista Mexicana de Física*, 3(3):0308048 1–6, May 2022. doi: 10.31349/SuplRevMexFis.3.0308048. URL `https://rmf.smf.mx/ojs/index.php/rmf-s/article/view/6102`.

[9] M. Thomson. *Modern Particle Physics*. Cambridge University Press, 2013. doi: 10.1017/CBO9781139525367.

[10] A. B. Arbuzov. Quantum field theory and the electroweak standard model. pages 1–34, 2017. doi: 10.23730/CYRSP-2017-004.1.

[11] The ATLAS Collaboration and G. Aad et al. Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the lhc. *Physics Letters B*, 716(1):1–29, 2012. ISSN 0370-2693. doi: https://doi.org/10.1016/j.physletb.2012.08.020. URL `https://www.sciencedirect.com/science/article/pii/S037026931200857X`.

[12] The CMS Collaboration and S. Chatrchyan et al. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30–61, 2012. ISSN 0370-2693. doi: https://doi.org/10.1016/j.physletb.2012.08.021. URL `https://www.sciencedirect.com/science/article/pii/S0370269312008581`.

[13] C. Burgess and G. Moore. *The Standard Model: A Primer*. Cambridge University Press, 2006. doi: 10.1017/CBO9780511819698.

[14] P. A. Zyla et al. Review of particle physics. *Progress of Theoretical and Experimental Physics*, 2020(8), 2020. doi: 10.1093/ptep/ptaa104.

[15] M. S. Sozzi. *Discrete symmetries and CP violation: From experiment to theory*. Oxford University Press, 2012.

[16] Y. Nir. Flavour physics and CP violation. 2015. doi: 10.5170/CERN-2015-001.123. URL `https://cds.cern.ch/record/2019737`. 34 pages, contribution to the 2013 CERN - Latin-American School of High-Energy Physics, Arequipa, Peru, 6-19 Mar 2013. arXiv admin note: text overlap with arXiv:1010.2666, arXiv:0708.1872.

[17] N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–533, Jun 1963. doi: 10.1103/PhysRevLett.10.531. URL `https://link.aps.org/doi/10.1103/PhysRevLett.10.531`.

[18] A. A. Petrov. CP-violation in charm. 2011. doi: 10.48550/ARXIV.1101.3822. URL `https://arxiv.org/abs/1101.3822`.

[19] M. Kobayashi and T. Maskawa. CP-violation in the renormalizable theory of weak interaction. *Progress of Theoretical Physics*, 49(2):652–657, 02 1973. ISSN 0033-068X. doi: 10.1143/PTP.49.652. URL `https://doi.org/10.1143/PTP.49.652`.

[20] Yao et al. Review of particle physics. *Journal of Physics G*, 33:1+, 2006. URL `http://pdg.lbl.gov`.

[21] L. Wolfenstein. Parametrization of the Kobayashi-Maskawa matrix. *Phys. Rev. Lett.*, 51:1945–1947, Nov 1983. doi: 10.1103/PhysRevLett.51.1945. URL https://link.aps.org/doi/10.1103/PhysRevLett.51.1945.

[22] A. M. Sanchez. Measurement of the gamma angle from tree decays at the LHCb experiment. *EPJ Web Conf.*, 28:12030, 2012. URL https://cds.cern.ch/record/1418719.

[23] J. L. Rosner. CP violation: Past, present, and future. *Brazilian Journal of Physics*, 2001. ISSN 0103-9733. URL https://www.redalyc.org/articulo.oa?id=46413501004.

[24] I. Bediaga, T. Frederico, and P. C. Magalhães. Enhanced charm cp asymmetries from final state interactions. *Phys. Rev. Lett.*, 131:051802, Aug 2023. doi: 10.1103/PhysRevLett.131.051802. URL https://link.aps.org/doi/10.1103/PhysRevLett.131.051802.

[25] R. Alvarez Garrote, J. Cuervo, P. C. Magalhães, and J. R. Peláez. Dispersive $\pi\pi \to k\overline{K}$ amplitude and giant cp violation in b to three light-meson decays at lhcb. *Phys. Rev. Lett.*, 130:201901, May 2023. doi: 10.1103/PhysRevLett.130.201901. URL https://link.aps.org/doi/10.1103/PhysRevLett.130.201901.

[26] J.M. Link et al. Study of the doubly and singly cabibbo suppressed decays $D^+ \to K^+\pi^-\pi^+$ and $D_s^+ \to K^+\pi^-\pi^+$. *Physics Letters B*, 601(1-2):10–19, Nov 2004. doi: 10.1016/j.physletb.2004.09.022. URL https://doi.org/10.1016%2Fj.physletb.2004.09.022.

[27] E. Byckling and K. Kajantie. *Particle Kinematics*. A Wiley-Interscience publication. Wiley, 1973. ISBN 9780471128854. URL https://books.google.com.br/books?id=d_cNAQAAIAAJ.

[28] J. M. Blatt and V. F. Weisskopf. *Theoretical nuclear physics*. Springer, New York, 1952. ISBN 978-0-471-08019-0. doi: 10.1007/978-1-4612-9959-2.

[29] C.s Zemach. Three-pion decays of unstable particles. *Phys. Rev.*, 133:B1201–B1220, Mar 1964. doi: 10.1103/PhysRev.133.B1201. URL https://link.aps.org/doi/10.1103/PhysRev.133.B1201.

[30] A. J. Bevan et al. The physics of the b factories. *The European Physical Journal C*, 74(11), Nov 2014. doi: 10.1140/epjc/s10052-014-3026-9. URL https://doi.org/10.1140/epjc/s10052-014-3026-9.

[31] L. Evans and P. Bryant. LHC machine. *Journal of Instrumentation*, 3(08): S08001, aug 2008. doi: 10.1088/1748-0221/3/08/S08001. URL https://dx.doi.org/10.1088/1748-0221/3/08/S08001.

[32] F. Marcastel. CERN's accelerator complex. la chaîne des accélérateurs du CERN. 2013. URL https://cds.cern.ch/record/1621583. General Photo.

[33] The ATLAS Collaboration and G Aad et al. The ATLAS experiment at the CERN large hadron collider. *Journal of Instrumentation*, 3(08):S08003, aug 2008. doi: 10.1088/1748-0221/3/08/S08003. URL https://dx.doi.org/10.1088/1748-0221/3/08/S08003.

[34] The CMS Collaboration and S Chatrchyan et al. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004, aug 2008. doi: 10.1088/1748-0221/3/08/S08004. URL https://dx.doi.org/10.1088/1748-0221/3/08/S08004.

[35] The ALICE Collaboration and K. Aamodt et al. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002, aug 2008. doi: 10.1088/1748-0221/3/08/S08002. URL https://dx.doi.org/10.1088/1748-0221/3/08/S08002.

[36] The LHCb Collaboration and A. A. Alves Jr et al. The LHCb detector at the LHC. *Journal of Instrumentation*, 3(08):S08005, aug 2008. doi: 10.1088/1748-0221/3/08/S08005. URL https://dx.doi.org/10.1088/1748-0221/3/08/S08005.

[37] P. Koppenburg and V. Vagnoni. Precision physics with heavy-flavoured hadrons. *Adv. Ser. Dir. High Energy Phys.*, 23:31–59, 2015. doi: 10.1142/9789814644150_0002. URL http://cds.cern.ch/record/2059944.

[38] C. Elsasser. $b\bar{b}$ production angle plots. URL https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html.

[39] R. Aaij and A. Affolder et al. Performance of the LHCb Vertex Locator. *Journal of Instrumentation*, 9(09):P09007, sep 2014. doi: 10.1088/1748-0221/9/09/P09007. URL https://dx.doi.org/10.1088/1748-0221/9/09/P09007.

[40] S. Amato, D. Carvalho, and et al. *LHCb magnet: Technical Design Report*. Technical design report. LHCb. CERN, Geneva, 2000. URL https://cds.cern.ch/record/424338.

[41] Large Hadron Collider beauty experiment. Tracking system. URL https://lhcb-outreach.web.cern.ch/detector/tracking-system/.

[42] P. R. Barbosa-Marinho, I. Bediaga, and et al. *LHCb inner tracker: Technical Design Report*. Technical design report. LHCb. CERN, Geneva, 2002. URL http://cds.cern.ch/record/582793.

[43] P. R. Barbosa-Marinho, I. Bediaga, and et al. *LHCb outer tracker: Technical Design Report*. Technical design report. LHCb. CERN, Geneva, 2001. URL https://cds.cern.ch/record/519146.

[44] S. Amato, D. Carvalho, and et al. *LHCb calorimeters: Technical Design Report*. Technical design report. LHCb. CERN, Geneva, 2000. URL http://cds.cern.ch/record/494264.

[45] E. P. Olloqui (on behalf ofthe LHCb Collaboration). LHCb Preshower(PS) and Scintillating Pad Detector (SPD): Commissioning, calibration, and monitoring. *Journal of Physics: Conference Series*, 160(1):012046, apr 2009. doi: 10.1088/1742-6596/160/1/012046. URL https://dx.doi.org/10.1088/1742-6596/160/1/012046.

[46] R. Aaij, S. Akar, and et al. Design and performance of the LHCb trigger and full real-time reconstruction in run 2 of the lhc. *Journal of Instrumentation*, 14(04):P04013, apr 2019. doi: 10.1088/1748-0221/14/04/P04013. URL https://dx.doi.org/10.1088/1748-0221/14/04/P04013.

[47] LHCb collaboration, R. Aaij, A. S. W. Abdelmotteleb, and et al. The LHCb upgrade I, 2023.

[48] LHCb upgrades during LS2. URL https://home.cern/press/2022/LHCb-upgrades-LS2.

[49] M. Clemencic, G. Corti, S. Easo, C. R. Jones, S. Miglioranzi, M. Pappagallo, and P. Robbe (on behalf ofthe LHCb Collaboration). The LHCb Simulation Application, Gauss: Design, evolution and experience. *Journal of Physics: Conference Series*, 331(3):032023, dec 2011. doi: 10.1088/1742-6596/331/3/032023. URL https://dx.doi.org/10.1088/1742-6596/331/3/032023.

[50] T. Sjöstrand, S. Mrenna, and P. Skands. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11):852–867, 2008. ISSN 0010-4655. doi: https://doi.org/10.1016/j.cpc.2008.01.036. URL https://www.sciencedirect.com/science/article/pii/S0010465508000441.

[51] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth. A*, 462:152–155, 2001. doi: 10.1016/S0168-9002(01)00089-4.

[52] S. Agostinelli and J. Allison et al. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003. ISSN 0168-9002. doi: https://doi.org/10.1016/S0168-9002(03)01368-8. URL https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[53] J. Allison, K. Amako, and et al. Geant4 developments and applications. *IEEE Transactions on Nuclear Science*, 53(1):270–278, 2006. doi: 10.1109/TNS.2006.869826.

[54] L. Anderlini, A. Contu, and et al. The PIDCalib package. Technical report, CERN, Geneva, 2016. URL https://cds.cern.ch/record/2202412.

[55] A. Rogozhnikov et al. hep_ml. *GitHub repository*, 2021. doi: 10.5281/zenodo.5500612. URL https://github.com/arogozhnikov/hep_ml/tree/v0.7.0.

[56] M. Pivk and F.R. Le Diberder. sPlots: A statistical tool to unfold data distributions. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 555(1):356–369, 2005. ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2005.08.106. URL https://www.sciencedirect.com/science/article/pii/S0168900205018024.

[57] L Breiman. *Classification and regression trees*. Routledge, 1st edition edition, 1984. doi: https://doi.org/10.1201/9781315139470.

[58] W. Verkerke and D. P. Kirkby. The RooFit toolkit for data modeling. *eConf*, C0303241:MOLT007, 2003.

[59] Tomasz Skwarnicki. *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*. PhD thesis, Cracow, INP, 1986.

[60] I. Bediaga, I. I. Bigi, A. Gomes, G. Guerrer, J. Miranda, and A. C. dos Reis. On a CP anisotropy measurement in the Dalitz plot. *Phys. Rev. D*, 80:096006, Nov 2009. doi: 10.1103/PhysRevD.80.096006. URL https://link.aps.org/doi/10.1103/PhysRevD.80.096006.

[61] BABAR Collaboration and B. Aubert et al. Search for CP violation in neutral D meson cabibbo-suppressed three-body decays. *Phys. Rev. D*,

78:051102, Sep 2008. doi: 10.1103/PhysRevD.78.051102. URL https://link.aps.org/doi/10.1103/PhysRevD.78.051102.

[62] Louis Lyons. *Statistics for Nuclear and Particle Physicists*. Cambridge University Press, 1986. doi: 10.1017/CBO9781139167710.

[63] R. Aaij, A.S.W. Abdelmotteleb, and et al. Search for CP violation in $D^+ \rightarrow K^-K^+K^+$ decays. *J. High Energ. Phys.*, 2023. doi: https://doi.org/10.1007/JHEP07(2023)067.

[64] J. R Klein and A. Roodman. Blind analysis in nuclear and particle physics. *Annual Review of Nuclear and Particle Science*, 55(1):141–163, 2005. doi: 10.1146/annurev.nucl.55.090704.151521. URL https://doi.org/10.1146/annurev.nucl.55.090704.151521.

[65] R. The LHCb Collaboration, Aaij and et al. Search for CP violation in $\Xi_c^+ \rightarrow pK^-\pi^+$decays using model-independent techniques. *The European Physical Journal C*, 2020. doi: 10.1140/epjc/s10052-020-8365-0. URL https://doi.org/10.1140/epjc/s10052-020-8365-0.

[66] R. et al Aaij. Search for time-dependent cp violation in $D^0 \rightarrow K^+K^-$ and $D^0 \rightarrow \pi^+\pi^-$ decays. *Phys. Rev. D*, 104:072010, Oct 2021. doi: 10.1103/PhysRevD.104.072010. URL https://link.aps.org/doi/10.1103/PhysRevD.104.072010.

[67] T. Aaltonen and et al. Measurement of CP-violation asymmetries in $D^0 \rightarrow K_S^0\pi^+\pi^-$. *Phys. Rev. D*, 86:032007, Aug 2012. doi: 10.1103/PhysRevD.86.032007. URL https://link.aps.org/doi/10.1103/PhysRevD.86.032007.