Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea Magistrale in Fisica

Search for CP violation in $D_s^{\pm} \to K_S^0 \pi^{\pm}$, $D^{\pm} \to K_S^0 K^{\pm}$ and $D^{\pm} \to \phi \pi^{\pm}$ decays

Relatore:
Prof. Angelo Carbone

Presentata da:
Serena Maccolini

Correlatori:
Dott. Angelo Di Canto
Dott. Fabio Ferrari

Anno Accademico 2016/2017

Sommario

Il lavoro qui presentato riporta una ricerca preliminare di violazione *CP* nei decadimenti singolarmente-Cabibbo-soppresi $D_s^{\pm} \to K_S^0\pi^{\pm}$, $D^{\pm} \to K_S^0 K^{\pm}$ e $D^{\pm} \to \phi\pi^{\pm}$, ricostruiti in un campione di dati corrispontene a 2 fb^{-1} di luminosità integrata raccolta da LHCb durante il Run-2 (2015-2016). Le asimmetrie di violazione *CP* sono misurate separatemente come

$$\begin{aligned} \mathcal{A}_{CP}(D_s^{\pm} \to K_S^0\pi^{\pm}) &= (0.03 \pm 0.40\,(\text{stat.}))\%, \\ \mathcal{A}_{CP}(D^{\pm} \to K_S^0 K^{\pm}) &= (0.04 \pm 0.15\,(\text{stat.}))\%, \\ \mathcal{A}_{CP}(D^{\pm} \to \phi\pi^{\pm}) &= (-0.124 \pm 0.045\,(\text{stat.}))\%, \end{aligned}$$

assumendo che la violazione *CP* nei decadimenti Cabibbo-favoriti sia trascurabile. I risultati sono consistenti con quelli delle analisi precedenti. Dato che i risultati sono compatibili con zero, non è stata trovata alcuna evidenza di violazione *CP* con la statistica attualmente analizzata. Tuttavia, i risultati sono preliminari poiché ulteriori studi, come quelli di valutazione delle principali sistematiche, sono ancora in corso.

Abstract

The work here presented reports a preliminary search for *CP* violation in the singly-Cabibbo-suppressed $D_s^{\pm} \to K_S^0\pi^{\pm}$, $D^{\pm} \to K_S^0 K^{\pm}$ and $D^{\pm} \to \phi\pi^{\pm}$ decays, reconstructed in a data sample corresponding to 2 fb^{-1} of integrated luminosity collected by LHCb during Run-2 (2015-2016). The individual *CP*-violating asymmetries are measured to be

$$\begin{aligned} \mathcal{A}_{CP}(D_s^{\pm} \to K_S^0\pi^{\pm}) &= (0.03 \pm 0.40\,(\text{stat.}))\%, \\ \mathcal{A}_{CP}(D^{\pm} \to K_S^0 K^{\pm}) &= (0.04 \pm 0.15\,(\text{stat.}))\%, \\ \mathcal{A}_{CP}(D^{\pm} \to \phi\pi^{\pm}) &= (-0.124 \pm 0.045\,(\text{stat.}))\%, \end{aligned}$$

assuming that *CP* violation in the Cabibbo-favoured decays is negligible. The measurements are consistent with those of the previous analyses. As the results are compatible with zero, no evidence of *CP* violation is found with the current analyzed statistics. However, the results are preliminary since additional studies, such as the evaluation of the systematic uncertainties, are still ongoing.

Contents

Introduction

The violation of the combined charge conjugation and parity symmetry, *i.e.* *CP* violation, is one of the necessary conditions to explain the matter-antimatter asymmetry observed in the Universe. In the Standard Model (SM), *CP* violation arises due to an irreducible complex phase in the Cabibbo-Kabayashi-Maskawa (CKM) matrix. However, this mechanism is not enough to explain the observed baryogenesis. Therefore, it is important to search for new sources of *CP* violation.

The study of *CP* violation in charm transitions is a powerful probe of physics beyond the Standard Model, in particular when this couples to the up-type quarks. Thanks to the large samples of charm hadrons that are being collected, LHCb is approaching an unprecedented precision that, for decays of D^0 mesons into two charged hadrons, is becoming close to theoretical expectations. However, no signs of *CP* violation are yet observed. Further investigations in other decay modes are therefore becoming more and more important. The singly Cabibbo-suppressed (CS) decays of the D^+ and D_s^+ mesons, such as the ones into a neutral kaon and a charged hadron and the ones into the ϕ meson and a charged pion, are one of the most sensitive probes for direct *CP* violation in charm decays.

The goal of this thesis is to perform a preliminary measurement of the *CP*-violating asymmetries in the $D_s^\pm \to K_S^0\pi^\pm$, $D^\pm \to K_S^0 K^\pm$ and $D^\pm \to \phi\pi^\pm$ decays, using data collected during Run-2 (2015-2016) by the LHCb detector. The analysis presented in this document uses control channels for the study of instrumental and production asymmetries, namely the Cabibbo-favoured $D^\pm \to K_S^0\pi^\pm$, $D_s^\pm \to K_S^0 K^\pm$ and $D_s^\pm \to \phi\pi^\pm$ decays, where no *CP* violation is expected.

The thesis is structured as follows. In Chapter 1, there is a brief description of the theoretical aspects related to this measurement. In Chapter 2, the LHCb detector, experimental apparatus which provided the data used for the experimental result, is described. The analysis strategy and data selection are explained in Chapter 3. Here preliminary results are presented. A description of the potential systematic effects to the measurement is also reported. Conclusions are discussed in the last Chapter. Complementary material can be found in the Appendices.

Chapter 1

CP violation in charm decays

1.1 Historical background

The concept of symmetry holds a special place in particle theory as the recognition of a symmetry gives insight into the fundamental physics. The most famous summary of this is Noether's theorem which states that symmetries are equivalent to conservation laws. Three discrete symmetries are considered in particle physics: parity, charge and time reversal. Parity, or space inversion, is the reflection in the origin of the space coordinates of a particle (*i.e.* the three space dimensions x, y and z become, respectively, $-x$, $-y$ and $-z$). Charge-conjugation (C) is a mathematical operation that transforms a particle into an antiparticle, inverting the sign of all charges (such as electric charge) while leaving unaffected all other quantities (and in particular space-time related ones such as position, momentum and spin). It implies that every charged particle has an oppositely charged antimatter counterpart, while the antiparticle of an electrically neutral particle may be identical to the particle, as in the case of the neutral π meson. The time reversal (T) represents the inversion of time coordinate t into $-t$. In this thesis, the role of CP violation, which is the violation of the combined conservation laws associated with charge conjugation and parity by the weak nuclear force, will be underlined.

Before 1950s, it was assumed that the P, C and T operators were symmetric for the electromagnetic, strong and weak interactions. However, a series of discoveries bring the physicists to alter significantly their assumption. In particular, the theoretical physicists C. N. Yang and T.-D. Lee, in 1956, they showed that there was no evidence supporting parity invariance in weak interactions [1]. Experiments conducted in the late 1950s in both nuclear and pion's beta decay [2, 3] verified that parity was violated. Moreover, they revealed that charge conjugation symmetry was also broken for these decay processes. A subsequent experiment made by Goldhaber et al. [4], in 1958, showed that the neutrino is left-handed, *i.e.* its spin is anti-parallel with respect to its momentum. It was soon pointed out that the independent application of P or C operators to the left-handed

neutrino (ν_L) led to physical states not observed in nature (right-handed neutrino (ν_R) or left-handed anti-neutrino ($\bar{\nu}_L$), respectively), but that the application of the *CP* operator to the ν_L led to the observed $\bar{\nu}_R$. For this reason it was thought that the *CP* symmetry was indeed conserved. This point was supported by the local T invariance derived from Mach's principle and from the CPT invariance that is natural in Lorentz–invariant quantum field theories [5, 6], *i.e.* no *CP* violation is then allowed if T violation is not found.

Subsequently, in 1964, Cronin and Fitch observed long lived neutral K meson (which is a *CP* eigenstate with negative eigenvalue) decaying into two pions, a *CP* eigenstate with positive eigenvalue, demonstrating that also the *CP* symmetry was not conserved by that weak interactions [7]. At that time, the C and P symmetry violations were incorporated in the basic structure of the unified electroweak theory by representing the left–handed and the right–handed fermions as a doublet and a singlet of the symmetry group $SU(2)_L \times U(1)_Y$. The theory was originally proposed by S. Glashow, S. Weinberg and A. Salam [8–10] to describe the leptons and then extended to the quarks by Nicola Cabibbo, in 1963, by introducing a mixing angle θ_C (the so-called Cabibbo angle). The hypothesis that the object that couples to the up quark via charged-current weak interaction is a superposition of down-type quarks (*i.e.* $d' = d\cos\theta_C + s\sin\theta_C$), arised in order to preserve the universality of the weak interaction [11]. Few years later, in 1970, Glashow, J. Iliopoulos and L. Maiani [12] proposed to explain the non-observation of flavour-changing neutral current (FCNC) processes with the existence of a fourth quark, the charm. Before 1973, the model foresaw *CP* and T invariance with no explanation of the observed neutral K mesons phenomenology. This was an open problem until M. Kobayashi and T. Maskawa [13] noticed that *CP* violation could be allowed in the electroweak model, by the presence of a single unremovable complex phase in the charged current interactions of quarks, if there were at least six quarks. The existence of the charm and bottom quarks were established few years later, respectively in 1974 [14, 15] and in 1977 [16], while the sixth quark, the top, was discovered later in 1995 [17, 18].

Cabibbo-Kobayashi-Maskawa formalism has proven to be very successful in explaining and predicting *CP* violation in different decays. Since its discovery in the K meson system in 1964, a systematic study of the *CP* violation phenomenon has been carried out by a number of experiments in beauty and charm decays. Historically, *CP* violation has been discovered in a indirect way through the study of flavour oscillations of neutral mesons. The existence of *CP* violation in the decays of B^0 mesons was demonstrated by the BaBar and Belle experiments [19, 20] while the first observation of *CP* violation in B_s^0 decays was reported by LHCb Collaboration [21]. *CP* violation has not yet been observed in the charm sector, and only recently the slow mixing rate of the $D^0 - \overline{D}^0$ flavour oscillations has been established [22] opening the venue for the measurement of *CP* violation.

The phenomenology of *CP* violation is particularly interesting since it could reveal

the presence of New Physics beyond the Standard Model. In fact, it is unlikely that the Standard Model provides a complete description of *CP* violation in nature. In addition, *CP* violation is connected with cosmology. One of the main mysteries of the Universe is the cosmological baryon asymmetry[1]. As was pointed out by A. D. Sakharov [23], the necessary conditions for the generation of such an asymmetry is that elementary interactions are *CP*-violating. Model calculations of the baryon asymmetry indicate, however, that the amount of *CP* violation present in the Standard Model seems to be too small to generate the observed baryon asymmetry [24–26]. Several extension of the Standard Model provides new sources of *CP* violation, thus measuring *CP* violation observables that significantly differ from the Standard Model expectation would be a clear sign of New Physics.

It is therefore essential to understand first the picture of *CP* violation arising in the framework of the Standard Model, where the Kobayashi–Maskawa mechanism plays a key role.

1.2 *CP* Violation in and beyond the Standard Model

The Standard Model (SM) describes the interaction between elementary particles, that form the fundamental building blocks of matter, and the fundamental forces of the nature. The theory describes these elementary particles interacting via the strong, electromagnetic and weak forces. The gravitational force is not included in the Standard Model theory.

Within the mathematical framework of the SM, particles and forces are the observable manifestations of scalar, vector or spinor fields which interact with each other according to a well-defined set of rules. The quanta of the scalar and vector fields are bosons, which are defined by their integer spin quantum number and which "carry" forces from one place to another. The bosons include the scalar Higgs particle, the strong vectors, called gluons, and the electroweak vectors $W^{\pm}$, Z^0, and γ. The quanta of the spinor fields are fermions, which have half-integer spins and are categorised as quarks or leptons. The leptons can be observed directly, and comprise the electron (e), the muon (μ), the tau (τ), and their neutrinos ($\nu_{e,\mu,\tau}$). The quarks do not exist apart[2], but make up all other observable massive particles, which are called hadrons. Quarks come in six flavours: up (u), down (d), strange (s), charm (c), bottom (b), and top (t). Quarks and leptons (except neutrinos) are charged, and each particle has a partner of opposite charge, which is referred to as its antiparticle. Most fundamental particles acquire their masses via the Higgs mechanism. The exceptions are the photon and gluon, which are massless, and

[1]This asymmetry is defined as the difference between the number density of baryon and that of anti-baryon, normalized to the entropy of the Universe.

[2]The only exception is the top quark, which may decay via electroweak interaction before it hadronizes.

perhaps the neutrinos. Most of the mass of the more common composite hadrons such as the proton and neutron results from the strong forces binding their constituent quarks together, rather than from their masses.

Mathematically, the Standard Model is defined as a model of elementary particles and their interactions by the local invariance of the Lagrangian with respect to the transformation of the symmetry group

$$G_{SM} = SU(3)_C \times SU(2)_L \times U(1)_Y \tag{1.1}$$

where the $SU(3)_C$ is the non-abelian algebra of the strong forces and the electroweak part is described by $SU(2)_L \times U(1)_Y$ through the Glashow–Weinberg–Salam theory [8–10]. Fermions are divided in three generations, each consisting of five representations:

$$Q^I_{Li}(3,2)_{+1/6}, \quad u^I_{Ri}(3,1)_{+2/3}, \quad d^I_{Ri}(3,1)_{-1/3}, \quad L^I_{Li}(1,2)_{-1/2}, \quad \ell^I_{Ri}(1,1)_{-1}. \tag{1.2}$$

The notations mean that, for example, the left-handed quarks, Q^I_L, are in a triplet (3) of the $SU(3)_C$ group, a doublet (2) of $SU(2)_L$ and carry hypercharge $Y = Q_{EM} - T_3 = +1/6$. The super-index I denotes interaction eigenstates. The sub-index $i = 1, 2, 3$ is the flavour (or generation) index. The Higgs boson is represented as the scalar multiplet

$$\phi(1,2)_{+1/2}, \tag{1.3}$$

defined as

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{1.4}$$

which assumes the vacuum expectation value (VEV)

$$< \phi >= \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}. \tag{1.5}$$

Thus, is often parametrized as

$$\phi = e^{i\frac{\sigma_i}{2}\theta_i} \begin{pmatrix} 0 \\ \frac{v+H^0}{\sqrt{2}} \end{pmatrix}, \tag{1.6}$$

where σ_i are the Pauli's matrices, θ_i are three real fields and H^0 is the Higgs boson field. The non-zero vacuum expectation generates a spontaneous breaking of the gauge group:

$$G_{SM} \rightarrow SU(3)_C \times U(1)_{EM}. \tag{1.7}$$

The Standard Model Lagrangian, $\mathcal{L}_{SM}$, is the most general renormalizable Lagrangian that is consistent with the gauge symmetry G_{SM} of eq. 1.1. It can be divided to four parts:

$$\mathcal{L}_{SM} = \mathcal{L}_{kinetic} + \mathcal{L}_{gauge} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}. \tag{1.8}$$

The kinetic term has the form $i\bar{\psi}\gamma_\mu D^\mu\psi$ where γ_μ are the Dirac matrices, ψ and $\bar{\psi} = \psi^\dagger\gamma^0$ are a Dirac spinor and its adjoint. D^μ is the covariant derivative which has replaced ∂^μ in order to maintain gauge invariance and it is defined as

$$D^\mu = \partial^\mu + ig_s G_a^\mu L_a + igW_b^\mu T_b + ig'B^\mu Y. \tag{1.9}$$

Here G_a^μ are the eight gluon fields, W_b^μ the three weak interaction bosons and B^μ the single hypercharge boson. The L_a's are $SU(3)_C$ generators (the 3×3 Gell-Mann matrices $1/2\,\lambda_a$ for triplets, 0 for singlets), the T_b's are $SU(2)_L$ generators (the 2×2 Pauli matrices $1/2\,\tau_b$ for doublets, 0 for singlets), and Y are the $U(1)_Y$ charges. For example, for the left-handed quarks Q_L^I, we have

$$\begin{aligned}\mathcal{L}_{kinetic}(Q_L) &= i\overline{Q_{Li}^I}\gamma_\mu D^\mu Q_{Li}^I, \\ &= i\overline{Q_{Li}^I}\gamma_\mu\left(\partial^\mu + \frac{i}{2}g_s G_a^\mu\lambda_a + \frac{i}{2}gW_b^\mu\tau_b + \frac{i}{6}g'B^\mu\right)Q_{Li}^I\end{aligned} \tag{1.10}$$

that represents the free propagation of the fermion and the interactions with a generic field A^μ (*e.g.* B^μ) in the form of a charged current $-J_\mu$ (*e.g.* $-\frac{1}{6}g'\overline{Q_{Li}^I}\gamma_\mu Q_{Li}^I$). This part of the interaction Lagrangian is always *CP* conserving.

The second term describes the self-interactions of the gauge fields and is written as

$$-\mathcal{L}_{gauge} = \frac{1}{4}\left(G_a^{\mu\nu}G_{a\,\mu\nu} + W_b^{\mu\nu}W_{b\,\mu\nu} + B^{\mu\nu}B_{\mu\nu}\right) \tag{1.11}$$

where $G_a^{\mu\nu}$, $W_b^{\mu\nu}$ and $B_{\mu\nu}$ are the Yang-Mills tensors, defined as

$$\begin{aligned}G_a^{\mu\nu} &= \partial_\mu G_a^\nu - \partial_\nu G_a^\mu + g_s f^{abc}G_b^\mu G_c^\nu, \\ W_b^{\mu\nu} &= \partial_\mu W_b^\nu - \partial_\nu W_b^\mu + g\varepsilon^{bcd}W_c^\mu W_d^\nu, \\ B^{\mu\nu} &= \partial_\mu B^\nu - \partial_\nu B^\mu.\end{aligned} \tag{1.12}$$

Although there exists an additional term (*i.e.* a gauge-invariant and renormalizable operator) that could introduce *CP* violation, $\mathcal{L}_{gauge}$ is assumed to be *CP* invariant [27].

The Higgs potential, which describes the scalar self-interactions, is given by:

$$\mathcal{L}_{Higgs} = \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \tag{1.13}$$

where λ is the Higgs self-coupling strength and $\mu = v\sqrt{\lambda}$. For the Standard Model scalar sector, where there is a single doublet, this part of the Lagrangian is also *CP* conserving. For extended scalar sector, such as that of a two Higgs doublet model, $\mathcal{L}_{Higgs}$ can be *CP* violating. Even in case that it is *CP* symmetric, it may lead to spontaneous *CP* violation.

The Yukawa interactions between the fermions and the scalar field are given by

$$-\mathcal{L}_{Yukawa} = +Y^d_{ij}\overline{Q^I_{Li}}\phi d^I_{Rj} + Y^u_{ij}\overline{Q^I_{Li}}\tilde{\phi}u^I_{Rj} + Y^\ell_{ij}\overline{L^I_{Li}}\phi\ell^I_{Rj} + h.c., \tag{1.14}$$

where $\tilde{\phi} = i\sigma_2\phi^\dagger$, Y^f are 3×3 complex matrices and h.c. stands for the hermitian conjugate terms. This part of the Lagrangian is, in general, *CP* violating as further described.

After the electroweak symmetry is broken by a non–zero vacuum expectation value v of the Higgs field (see Eq. 1.5)), the Yukawa terms in Eq. 1.14 give rise (upon the replacement $\mathfrak{Re}(\phi^0) \to (v+H^0)/\sqrt{2}$) to the mass matrices of quarks and charged leptons:

$$-\mathcal{L}_M = (M_d)_{ij}\overline{d^I_{Li}}d^I_{Rj} + (M_u)_{ij}\overline{u^I_{Li}}u^I_{Rj} + (M_\ell)_{ij}\overline{\ell^I_{Li}}\ell^I_{Rj} + h.c., \tag{1.15}$$

where $M_f = \frac{v}{\sqrt{2}}Y^f$ and the $SU(2)_L$ doublets are decomposed into their components:

$$Q^I_{Li} = \begin{pmatrix} u^I_{Li} \\ d^I_{Li} \end{pmatrix}, \quad L^I_{Li} = \begin{pmatrix} \nu^I_{Li} \\ \ell^I_{Li} \end{pmatrix}. \tag{1.16}$$

Since the Standard Model neutrinos have no Yukawa interactions, they are predicted to be massless[3].

The mass basis corresponds, by definition, to diagonal mass matrices. We can always find unitary matrices V_{fL} and V_{fR} such that

$$V_{fL}M_fV^\dagger_{fR} = M^{diag}_f, \tag{1.17}$$

with M^{diag}_f diagonal and real. The mass eigenstates are then identified as

$$\begin{aligned} d_{Li} &= (V_{dL})_{ij}d^I_{Lj}, \quad & d_{Ri} &= (V_{dR})_{ij}d^I_{Rj}, \\ u_{Li} &= (V_{uL})_{ij}u^I_{Lj}, \quad & u_{Ri} &= (V_{uR})_{ij}u^I_{Rj}, \\ \ell_{Li} &= (V_{\ell L})_{ij}\ell^I_{Lj}, \quad & \ell_{Ri} &= (V_{\ell R})_{ij}\ell^I_{Rj}, \\ \nu_{Li} &= (V_{\nu L})_{ij}\nu^I_{Lj}. \end{aligned} \tag{1.18}$$

Since the Standard Model neutrinos are massless, $V_{\nu L}$ is arbitrary.

The charged current interactions (that are the interactions of the charged $SU(2)_L$ gauge bosons $W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu)$) for quarks, which in the interaction basis are described by Eq. 1.10, have a complicated form in the mass basis:

$$-\mathcal{L}_{W^\pm} = \frac{g}{\sqrt{2}}\overline{u_{Li}}\gamma^\mu(V_{uL}V^\dagger_{dL})_{ij}d_{Lj}W^+_\mu + h.c.. \tag{1.19}$$

[3]This is mere a consequence of the fact that in the Standard Model there are no "sterile" right–handed neutrinos, $\nu^I_{Ri}(1,1)_0$, thus it is impossible to produce Dirac mass terms of the form $\bar{\nu}^I_{Li}\nu^I_{Ri}$. Therefore, the established observations of neutrino masses [28] necessarily call for an extension of the Standard Model.

The unitary 3×3 matrix,

$$V_{CKM} = V_{uL} V_{dL}^{\dagger}, \quad (V_{CKM} V_{CKM}^{\dagger} = 1), \tag{1.20}$$

is the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix for quarks [11, 13]. By convention, the elements of V_{CKM} are written as follows:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{1.21}$$

A unitary $n \times n$ matrix depends on $(n-1)^2$ physical parameters[4]: $n(n-1)/2$ real angles and $(n-1)(n-2)/2$ complex phases. For $n = 2$, *i.e.* two families, there is just one mixing angle that yields the Cabibbo matrix:

$$V_C = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix}. \tag{1.22}$$

For $n = 3$, there are instead four physical parameters, namely three Euler angles and one phase δ. It is the latter that provides the gateway for *CP* violation (*i.e.* $\mathcal{L}_W \neq \mathrm{CP}\mathcal{L}_W(CP)^{\dagger}$. In the standard parametrization, the elements of the CKM matrix are written as follows:

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{1.23}$$

where $c_{ij} \equiv \cos\theta_{ij}$, $s_{ij} \equiv \sin\theta_{ij}$ and θ_{ij} are the three real mixing parameters.

As a result of the fact that V_{CKM} is not diagonal, the $W^{\pm}$ gauge boson couple to quark (mass eigenstates) of different generations. Within the Standard Model, this is the only source of *flavour-changing* interactions. In principle, there could be additional sources of flavour mixing (and *CP* violation) in the lepton sector and in Z^0 interactions, but within the Standard Model context this does not happen.

1.3 The CKM matrix and the Unitary Triangles

The current knowledge of the CKM matrix elements, as obtained by Particle Data Group [29] using the prescription of Refs. [30, 31], is the following:

$$V_{CKM} = \begin{pmatrix} 0.97434^{+0.00011}_{-0.00012} & 0.22506 \pm 0.00050 & 0.00357 \pm 0.00015 \\ 0.22492 \pm 0.00050 & 0.97351 \pm 0.00013 & 0.0411 \pm 0.0013 \\ 0.00875^{+0.00032}_{-0.00033} & 0.0403 \pm 0.0013 & 0.99915 \pm 0.00005 \end{pmatrix} \tag{1.24}$$

[4] $2n-1$ parameters out of n^2 can be eliminated through the rephrasing of the n up-type and n down-type fermion fields (*i.e.* changing all fermions by the same phase obviously does not affect V_{CKM}).

The observed hierarchy $|V_{ub}| << |V_{cb}| << |V_{us}|, |V_{cd}| << 1$ allows to expand V_{CKM} in powers of $\lambda = |V_{us}| = \sin\theta_C$, as firstly realized by Wolfenstein [32]. The expansion up to $\mathcal{O}(\lambda^3)$ is given by:

$$V_{CKM} = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + (\lambda^4). \tag{1.25}$$

where $\lambda \approx 0.23$, $A \approx 0.82$, $\rho \approx 0.12$, $\eta \approx 0.35$ [29] are the four mixing parameters. In such representation, we clearly see that the CKM matrix is a very special unitary matrix: it is almost diagonal, it is almost symmetric and the matrix elements get smaller the more one moves away from the diagonal.

Various parametrizations differ in the way that the freedom of phase rotation is used to leave a single phase in V_{CKM}. One can define, however, a *CP* violating quantity in V_{CKM} that is independent of the parametrization [33]. This quantity is the Jarlskog invariant J_{CP}, defined through

$$\mathfrak{Im}[V_{ij}V_{kl}V_{il}^*V_{kj}^*] = J_{CP}\sum_{m,n=1}^{3}\epsilon_{ikm}\epsilon_{jln}, \quad (i,j,k,l = 1,2,3). \tag{1.26}$$

CP is violated in the Standard Model only if $J_{CP} \neq 0$ and in fact current measurements tell us [29]:

$$J_{CP} = (3.04^{+0.21}_{-0.20}) \cdot 10^{-5}. \tag{1.27}$$

The usefulness of J_{CP} may not be clear from its formal definition in 1.26, but does give useful insights once the *unitarity triangles* are introduced. The unitarity of the CKM matrix, ($V_{CKM}V_{CKM}^\dagger = 1$), leads to a set of nine equations among the matrix elements, six of which require the sum of three complex quantities to vanish:

$$\underbrace{V_{ud}V_{us}^*}_{\mathcal{O}(\lambda)} + \underbrace{V_{cd}V_{cs}^*}_{\mathcal{O}(\lambda)} + \underbrace{V_{td}V_{ts}^*}_{\mathcal{O}(\lambda)^5} = 0, \tag{1.28}$$

$$\underbrace{V_{ud}^*V_{cd}}_{\mathcal{O}(\lambda)} + \underbrace{V_{us}^*V_{cs}}_{\mathcal{O}(\lambda)} + \underbrace{V_{ub}^*V_{cb}}_{\mathcal{O}(\lambda^5)} = 0, \tag{1.29}$$

$$\underbrace{V_{us}V_{ub}^*}_{\mathcal{O}(\lambda^4)} + \underbrace{V_{cs}V_{cb}^*}_{\mathcal{O}(\lambda)^2} + \underbrace{V_{ts}V_{tb}^*}_{\mathcal{O}(\lambda)^2} = 0, \tag{1.30}$$

$$\underbrace{V_{cd}^*V_{td}}_{\mathcal{O}(\lambda)^4} + \underbrace{V_{cs}^*V_{ts}}_{\mathcal{O}(\lambda)^2} + \underbrace{V_{cb}^*V_{tb}}_{\mathcal{O}(\lambda)^2} = 0, \tag{1.31}$$

$$\underbrace{V_{ud}V_{ub}^*}_{\mathcal{O}(\lambda)^3} + \underbrace{V_{cd}V_{cb}^*}_{\mathcal{O}(\lambda)^3} + \underbrace{V_{td}V_{tb}^*}_{\mathcal{O}(\lambda)^3} = 0, \tag{1.32}$$

$$\underbrace{V_{ud}^*V_{td}}_{\mathcal{O}(\lambda)^3} + \underbrace{V_{us}^*V_{ts}}_{\mathcal{O}(\lambda)^3} + \underbrace{V_{ub}^*V_{tb}}_{\mathcal{O}(\lambda)^3} = 0. \tag{1.33}$$

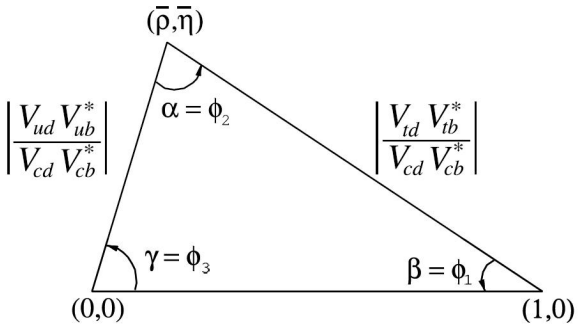


Figure 1.1: Rescaled unitary triangle corresponding to Eq. 1.32.

Each relation can be geometrically represented in the complex plane as a triangle. These are "the unitarity triangles", though the term "unitarity triangle" is usually reserved for the relation 1.32 only.

It can be easily shown that all unitarity triangles have the same area, equal to $|J_{CP}|/2$. The relation between Jarlskog's measure of *CP* violation J_{CP} and the Wolfenstein parameters is given by $J_{CP} \simeq \lambda^6 A^2 \eta$.

The rescaled unitarity triangle is derived from 1.32 by choosing a phase convention such that $(V_{cd}V_{cb}^*)$ is real and dividing the lengths of all sides by $|V_{cd}V_{cb}^*|$. Two vertices of the rescaled unitarity triangle are thus fixed at (0,0) and (1,0). The coordinates of the remaining vertex correspond to the Wolfenstein parameters (ρ, η). The area of the rescaled unitarity triangle is $|\eta|/2$.

Depicting the rescaled unitarity triangle in the (ρ, η) plane, as shown in Fig 1.1, the lengths of the two complex sides are

$$R_u \equiv \sqrt{\rho^2+\eta^2} = \frac{1}{\lambda}\left|\frac{V_{ub}}{V_{cb}}\right|, \quad R_t \equiv \sqrt{(1-\rho)^2+\eta^2} = \frac{1}{\lambda}\left|\frac{V_{td}}{V_{cb}}\right|. \tag{1.34}$$

The three angles of the unitarity triangle are denoted by α, β and γ [34]:

$$\alpha \equiv \arg\left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right], \quad \beta \equiv \arg\left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right], \quad \gamma \equiv \arg\left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right]. \tag{1.35}$$

1.4 *CP* Violation in meson decays

In the previous section, we understood how *CP* violation arises in the Standard Model. Now, we would like to understand the implications of this theory for the phenomenology of *CP* violation in meson decays. Our main focus will be on D-meson decays. To do so, we first present a model independent analysis of *CP* violation in meson decays.

We define decay amplitudes of a D meson (which could be charged or neutral) and its CP conjugate $\bar{D}$ to a multi-particle finale state f and its CP conjugate $\bar{f}$ as

$$A_f =< f|\mathcal{H}|D >, \quad \bar{A}_f =< f|\mathcal{H}|\bar{D} >, \quad A_{\bar{f}} =< \bar{f}|\mathcal{H}|D >, \quad \bar{A}_{\bar{f}} =< \bar{f}|\mathcal{H}|\bar{D} > \quad (1.36)$$

whew $\mathcal{H}$ is the decay Hamiltonian.

CP relates A_f and $\bar{A}_{\bar{f}}$. There are two types of phases that may appear in A_f and $\bar{A}_{\bar{f}}$. Complex parameters in any Lagrangian term that contributes to the amplitude will appear in complex conjugate form in the CP-conjugate amplitude. Thus their phases appear in A_f and $\bar{A}_{\bar{f}}$ with opposite signs. In the Standard Model these phases occur only in the CKM matrix which is part of the electroweak sector of the theory, hence these are often called "weak phases". The weak phase of any single term is convention dependent. However the difference between the weak phases in two different terms in A_f is convention independent because the phase rotations of the initial and final states are the same for every term. A second type of phase can appear in scattering or decay amplitudes even when the Lagrangian is real. Such phases do not violate CP and they appear in A_f and $\bar{A}_{\bar{f}}$ with the same sign. Their origin is the possible contribution from intermediate on-shell states in the decay process, that is an absorptive part of an amplitude that has contributions from coupled channels. Usually the dominant rescattering is due to strong interactions and hence the designation "strong phases" for the phase shifts so induced. Again only the relative strong phases of different terms in a scattering amplitude have physical content, an overall phase rotation of the entire amplitude has no physical consequences.

CP asymmetry in the decay is defined as

$$\mathcal{A}_{CP} = \frac{|A_f|^2 - |\bar{A}_{\bar{f}}|^2}{|A_f|^2 + |\bar{A}_{\bar{f}}|^2} \quad (1.37)$$

and appears as a result of interference among various terms in the decay amplitude. As an example, let us consider a decay process which can proceed through several amplitudes:

$$A_f = \sum_j |A_j| e^{i(\delta_j + \phi_j)}, \qquad \bar{A}_{\bar{f}} = \sum_j |A_j| e^{i(\delta_j - \phi_j)}, \quad (1.38)$$

where δ_j and ϕ_j are strong (CP conserving) and weak (CP violating) phases, respectively. Then, to observe CP violation there must be a contribution from at least two processes with different weak and strong phases in order to have a non vanishing interference term

$$|A_f|^2 - |\bar{A}_{\bar{f}}|^2 = -2 \sum_{i,j} |A_i||A_j| \sin(\delta_i - \delta_j) \sin(\phi_i - \phi_j). \quad (1.39)$$

The phenomenology of CP violation in neutral flavoured meson decays is enriched by the possibility that, besides the decay, it is also possible to have particle-antiparticle

transitions, also known as flavour mixing or oscillations. This process is a manifestation of flavour-changing neutral currents that occurs because of the charm flavour-defined states, D^0 and $\bar{D}^0$ differ from the physical mass eigenstates, D_1 and D_2, through the relation

$$|D_{1,2}\rangle = p\left|D^0(t)\right\rangle \pm q\left|\bar{D}^0(t)\right\rangle. \tag{1.40}$$

The time evolution of the meson can be described by:

$$\left|D^0(t)\right\rangle = g_+(t)\left|D^0(0)\right\rangle + \frac{q}{p}g_-(t)\left|\bar{D}^0(0)\right\rangle, \tag{1.41}$$

$$\left|\bar{D}^0(t)\right\rangle = g_+(t)\left|\bar{D}^0(0)\right\rangle + \frac{q}{p}g_-(t)\left|D^0(0)\right\rangle \tag{1.42}$$

with

$$g_+(t) = e^{-\imath Mt-\Gamma t/2}\cos\left(\frac{x}{2}\Gamma t - \frac{\imath y}{2}\Gamma t\right), \tag{1.43}$$

$$g_-(t) = e^{-\imath Mt-\Gamma t/2}\imath\sin\left(\frac{x}{2}\Gamma t - \frac{\imath y}{2}\Gamma t\right). \tag{1.44}$$

where M and Γ are the average mass and decay width, while x and y are the normalized mass and decay width differences between the two mass eigenstates

$$x = \frac{m_1 - m_2}{\Gamma} \quad \text{and} \quad y = \frac{\Gamma_1 - \Gamma_2}{2\Gamma}. \tag{1.45}$$

Often, experiments publish results for the parameters x' and y' which are related to x and y by a rotation by the strong phase difference δ between Cabibbo-favoured and doubly Cabibbo-suppressed amplitudes.

While *CP* violation in charged meson decays depends only on A_f and $\bar{A}_{\bar{f}}$, in the case of neutral mesons, because of the possibility of flavour oscillations, *CP* violating effects have additional dependences. We then distinguish three types of *CP* violating effects in meson decays:

(i) *CP* violation in the decay is defined by

$$|\bar{A}_{\bar{f}}/A_f| \neq 1.$$

(ii) *CP* violation in mixing is defined by

$$|q/p| \neq 1.$$

(iii) *CP* violation in interference between a decay without mixing, $D^0 \to f$, and a decay with mixing, $D^0 \to \bar{D}^0 \to f$, is defined by

$$\mathfrak{Im}\left(\frac{q}{p}\frac{\bar{A}_f}{A_f}\right) \neq 0.$$

Usually type *(i)* is also known as direct *CP* violation, while type *(ii)* and *(iii)* are referred as indirect *CP* violation.

1.5 Charm and *CP* violation

The charm quark was predicted in the late 1960s to complete the GIM mechanism [12]. The first experimental evidence was the discovery in 1974 of the J/Ψ particle, a $c\bar{c}$ resonance, while the first open charm mesons (D^0, D^+) where found by the Mark I experiment in 1976 [35]. At leading order, the effective weak Lagrangian responsible for Cabibbo-favoured hadronic charm decays, from the Feynman rules, is

$$\mathcal{L} = 4G_F\sqrt{2}V_{cs}V_{ud}^*(\bar{s}_L\gamma_\mu c_L)(\bar{u}_L\gamma^\mu d_L) + h.c. \tag{1.46}$$

where the Dirac spinor fields for the up, down, charm and strange quarks are denoted by the appropriate letter and G_F is the Fermi constant $g^2\sqrt{s}/8m_W^2$. However, the mass of the charm quark m_c of about 1 GeV/c^2 makes perturbative expansions difficult as neither m_c nor $1/m_c$ are small. In addition, the many final state interactions and rescattering possibilities make it challenging to match theoretical predictions to experimental data.

In 2007, the first evidence for the mixing of neutral charm mesons was found at the B-factory experiments by the BaBar and Belle collaborations [36, 37]. This is now well established, and the most recent study at LHCb [38] measures $y' = (0.52 \pm 0.08)\%$ and $x'^2 = (0.036 \pm 0.043) \cdot 10^{-3}$. The x and y parameters are significantly smaller for the charm than for either the K or the B systems, and their small size means that most of neutral D mesons will decay before appreciable mixing can take place.

As discussed in § 1.4, direct *CP* violation in charm decays requires the interference of two amplitudes to the same final state with different strong and weak phases. The two amplitudes must have similar size in order to create appreciable interference. In the charm system, such amplitudes can be found in singly Cabibbo-suppressed decays, where tree and loop (penguin) diagrams with comparable matrix element are possible.

This thesis is concerned with charged D decays, specifically the singly Cabibbo-suppressed channels $D_s^\pm \to K_S^0\pi^\pm$ and $D^\pm \to K_S^0K^\pm$ (collectively referred to as $D_{(s)}^\pm \to K_S^0h^\pm$ in the following) and $D^\pm \to \phi\pi^\pm$. Examples of possible Feynman graphs for these decays are given in Fig. 1.2. Other (not shown) possible Feynman graphs, *e.g.* with different levels of colour- and OZI-suppression[5], exist for the $D_{(s)}$ decays which contribute to the overall matrix element. Unlike in neutral D meson decays, there is no possibility for W exchange diagrams. The tree amplitudes are practically *CP* conserving, since the decays involve $V_{cd}^*V_{ud}$ or $V_{cs}^*V_{us}$ which are real in Wolfenstein parametrization up to $\mathcal{O}(\lambda^4)$ or $\mathcal{O}(\lambda^5)$. *CP* violation can arise from the penguin amplitudes induced by heavy virtual particles such as a the b quark. However, as shown in the Feynman diagrams of Figs. 1.2(b,d,f), the contribution of these amplitudes is strongly suppressed by the small value of the product $V_{cb}^*V_{ub}$. *CP* asymmetry can be roughly estimated as $\arg(V_{ub})$ and is expected to be $\sim \mathcal{O}(0.1\%)$ or below [35]. Hence, for decays which contain

[5]The OZI rule states that any strongly occurring process with a Feynman diagram that can be split in two by cutting only internal gluon lines will be suppressed.

	$\mathcal{A}_{CP}$	Experiment (dataset)
$D_s^\pm \to K_S^0\pi^\pm$		
	$(16.3 \pm 7.3\,(\text{stat.}) \pm 0.3\,(\text{syst.}))\%$	CLEO [40]
	$(0.60 \pm 2.00\,(\text{stat.}) \pm 0.30\,(\text{syst.}))\%$	BaBar [41]
	$(+5.45 \pm 2.50\,(\text{stat.}) \pm 0.33\,(\text{syst.}))\%$	Belle [43]
	$(+0.61 \pm 0.83\,(\text{stat.}) \pm 0.14\,(\text{syst.}))\%$	LHCb (2011) [46]
	$(+0.38 \pm 0.46\,(\text{stat.}) \pm 0.17\,(\text{syst.}))\%$	LHCb (2011-2012) [47]
$D^\pm \to K_S^0 K^\pm$		
	$(4.7 \pm 1.8\,(\text{stat.}) \pm 0.9\,(\text{syst.}))\%$	CLEO [40]
	$(+0.13 \pm 0.36\,(\text{stat.}) \pm 0.25\,(\text{syst.}))\%$	BaBar [41]
	$(+0.12 \pm 0.36\,(\text{stat.}) \pm 0.22\,(\text{syst.}))\%$	Belle [43]
	$(-0.25 \pm 0.28\,(\text{stat.}) \pm 0.14\,(\text{syst.}))\%$	Belle [44]
	$(+0.03 \pm 0.17\,(\text{stat.}) \pm 0.14\,(\text{syst.}))\%$	LHCb (2011-2012) [47]
$D^\pm \to \phi\pi^\pm$		
	$(0.35 \pm 0.30\,(\text{stat.}) \pm 0.15\,(\text{syst.}))\%$	BaBar [42]
	$(+0.51 \pm 0.28\,(\text{stat.}) \pm 0.05\,(\text{syst.}))\%$	Belle [45]
	$(-0.04 \pm 0.14\,(\text{stat.}) \pm 0.14\,(\text{syst.}))\%$	LHCb (2011) [46]

Table 1.1: Summary of previous measurements of *CP*-violating asymmetries in $D_{(s)}^\pm \to K_S^0 h^\pm$ and $D^\pm \to \phi\pi^\pm$ decays.

a K^0, to first order one would expect to observe an asymmetry consistent only with the neutral kaon *CP* violation, with no charm decay phase contribution.

Ref. [39] reports the calculus for *CP* asymmetry in the SM in $D_{(s)}^\pm \to K_S^0 h^\pm$ decays, in addition to predictions for branching ratios. No equivalent prediction is available for the $D^\pm \to \phi\pi^\pm$ decay.

Several studies on *CP* violation in $D_{(s)}^\pm \to K_S^0 h^\pm$ and $D^\pm \to \phi\pi^\pm$ decays have been performed previously by CLEO [40], BaBar [41, 42], Belle [43–45] and LHCb using Run-1 data [46, 47]. Previous results indicate no evidence for *CP* violation and are reported in Tab. 1.1.

As widely described in Chapter 3, the Cabibbo-favoured (CF) $D^\pm \to K_S^0\pi^\pm$, $D_s^\pm \to K_S^0 K^\pm$ and $D_s^\pm \to \phi\pi^\pm$ channels are used as control sample in the measurement of *CP* asymmetries. For completeness, Feynman diagrams of these decays are shown in Fig. 1.3.

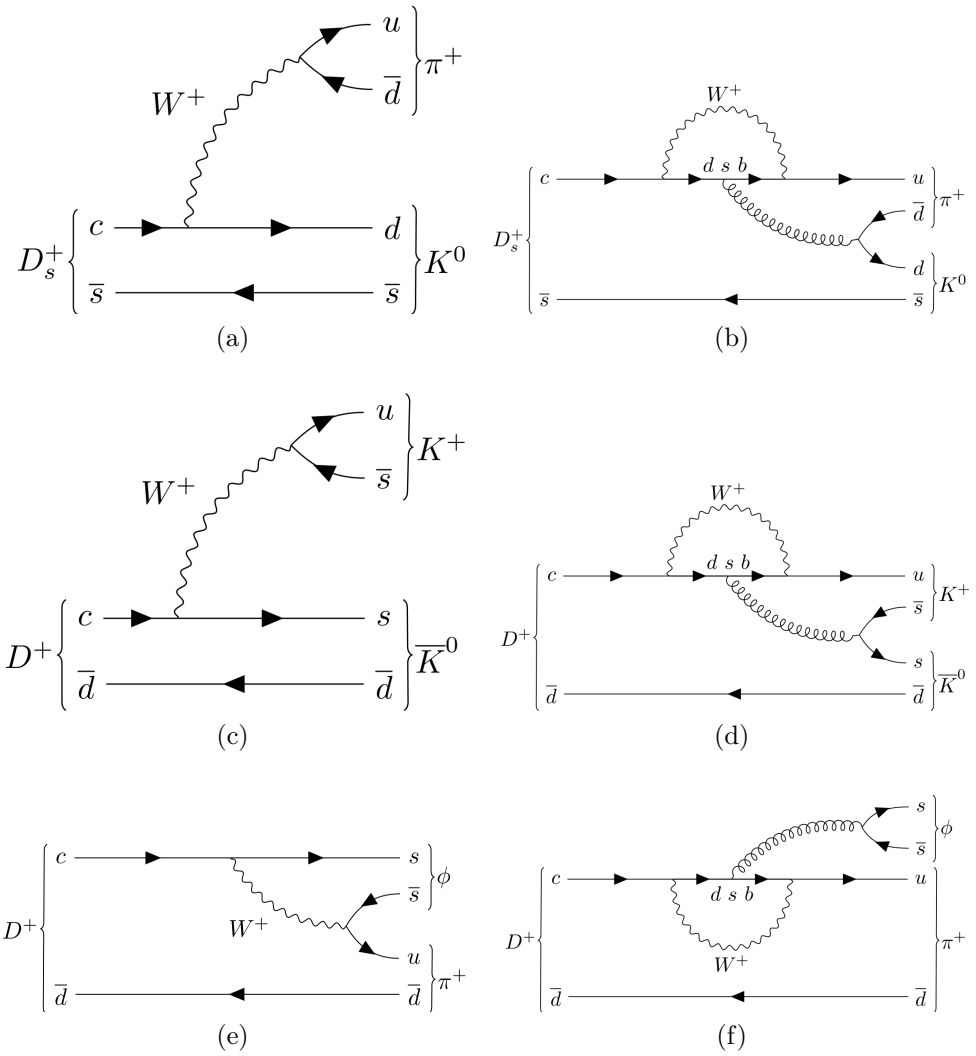


Figure 1.2: Feynman diagrams of the possible topologies of the $D_s^{\pm} \to K_S^0\pi^{\pm}$ (a, b), $D^{\pm} \to K_S^0K^{\pm}$ (c, d) and $D^{\pm} \to \phi\pi^{\pm}$ (e, f) decays. On the left: tree diagram. On the right: loop (penguin) diagram.

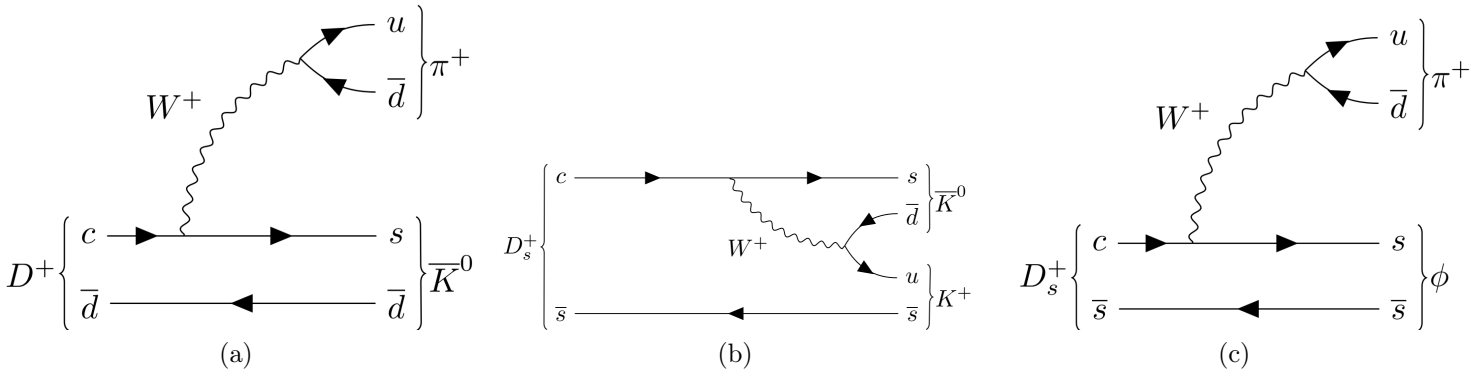


Figure 1.3: Feynman diagrams of the Cabibbo-favoured decays $D^{\pm} \to K^0_S\pi^{\pm}$ (a), $D^{\pm}_s \to K^0_S K^{\pm}$ (b) and $D^{\pm}_s \to \phi\pi^{\pm}$ (c) decays.

Chapter 2

The LHCb experiment at the LHC

The Large Hadron Collider beauty (LHCb) experiment [48] is dedicated to the study of heavy flavour physics at the Large Hadron Collider (LHC) [49] at CERN (Geneva). Its primary goal, as synthesized in the logo,

is to look for indirect evidence of New Physics (NP) in *CP*-violating processes and rare decays of beauty and charm hadrons.

The LHCb experiment is designed to exploit the great production cross-section of $b\bar{b}$ pairs in pp collisions at the LHC energies, measured to be $\sigma(pp \to b\bar{b}X) = (154.3 \pm 1.5 \pm 14.3)$ μb at $\sqrt{s} = 13$ TeV and within the LHCb acceptance [50]. The same characteristics that make LHCb a perfect experiment for b physics are ideal for the study of c physics as well, also because the $c\bar{c}$ production cross-section is even larger than that of $b\bar{b}$, namely $\sigma(pp \to c\bar{c}X) = (2369 \pm 3 \pm 152 \pm 118)$ μb at $\sqrt{s} = 13$ TeV and within LHCb acceptance [51].

The LHCb experiment approach to the search of NP is complementary to that used by the ATLAS and CMS collaborations. Indeed, those experiments are designed to measure high p_{T} decay products that could indicate the direct production of new particles, whereas the LHCb experiment wants to measure the existence of such new particles indirectly, by means of their virtual quantum effects. Furthermore, the research is also active in other fields, like the studies of heavy-flavour spectroscopy and production of gauge bosons, and searches for new exotic particles.

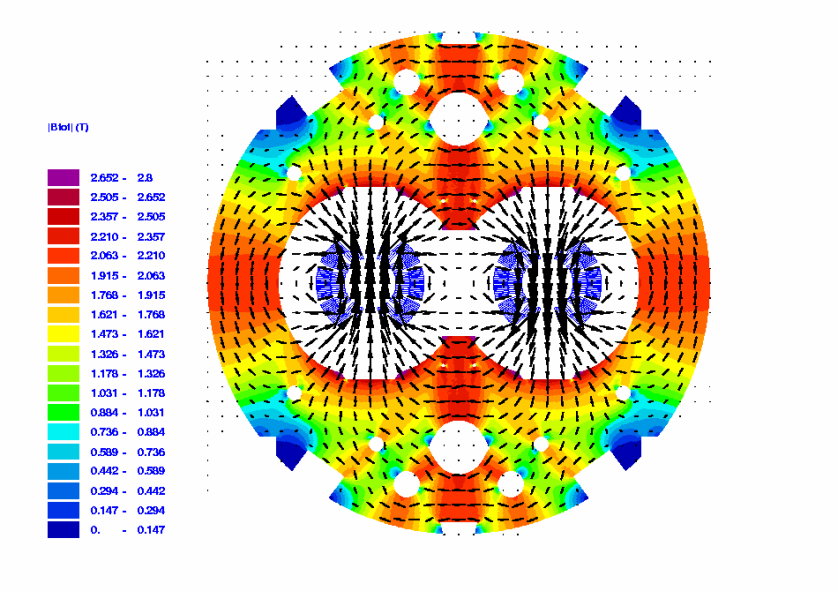


Figure 2.1: Illustration representing the typical magnetic field emitted by the dipole magnets of the LHC. The coloured portions of the diagram indicate the magnetic flux, or the amount of magnetic field passing through a given area. The arrows indicate the direction of the magnetic field. The two circles (in blue) in the centre of the diagram indicate the beam pipes for beams one and two. Notice how the arrows (direction of the magnetic field) point in opposite directions allowing to control two counter-rotating beams of protons in the same beam pipe [49].

2.1 The Large Hadron Collider

The LHC is a is a two-ring hadron collider housed in the 27 km tunnel previously constructed for the Large Electron Positron collider (LEP). It is placed 100 m underground across the border between France and Switzerland. Inside the accelerator, two high-energy beams travel at velocities close to the speed of light before they are made to collide. The beams travel in opposite directions in separate beam pipes, two tubes kept at ultra-high vacuum. The accelerator is designed to collide protons up to a centre-of-mass energy of 13 TeV with an instantaneous luminosity of 10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. In order to reach this energy a magnetic field with an intensity of 8.3 T is needed. This high field can only be achieved using superconducting material (NbTi) and by cooling the magnets in super-fluid helium at 1.9 K (-271.3 °C). The tunnel diameter is only 3.8 m, insufficient for the installation of two separate rings. The two rings are therefore incorporated into a single magnetic structure with two set of coils in a common yoke and cryostat. An example of the typical magnetic field emitted by the dipole magnets of the LHC is shown in Fig. 2.1.

The accelerator complex at CERN is a succession of machines that accelerate particles to increasingly higher energies. Each machine boosts the energy of a beam of particles, before injecting the beam into the next machine in the sequence. Most of the other accelerators in the chain have their own experimental halls where beams are used for experiments at lower energies (*e.g.* Isolde, Alpha).

The proton source is a simple bottle of hydrogen gas. An electric field is used to strip hydrogen atoms of their electrons to yield protons. As it is not possible to directly accelerate protons from their quasi-rest conditions up to 6.5 TeV, it is necessary to pre-accelerate them through a complex of machines, represented in Fig. 2.2. First, protons are injected in LINAC2, a linear accelerator that provides the Proton Synchrotron Booster (PSB) with proton bunches of 50 MeV energy. The PBS can accelerate protons up to 1 GeV; after this, the particles are injected in the Proton Synchrotron (PS), where they reach an energy of 26 GeV. Then, the PS passes them to the Super Proton Synchrotron (SPS), where they are accelerated for the last time up to an energy of 450 GeV, before being injected in the LHC via two tunnels, called T12 and T18. After the two rings are filled, the machine is ramped to its nominal energy over about 28 min. The two beams are brought into collision inside four detectors (ALICE, ATLAS, CMS and LHCb) where the total energy at the collision point is equal to 13 TeV.

In addition to accelerating protons, the accelerator complex can also accelerate lead ions. Lead ions are produced from a highly purified lead sample heated to a temperature of about 800°C. The lead vapour is ionized by an electron current. Many different charge states are produced with a maximum around ^{29}Pb. These ions are selected and accelerated to 4.2 MeV/u (energy per nucleon) before passing through a carbon foil, which strips most of them to ^{54}Pb. The beam is accumulated, then accelerated to 72 MeV/u in the Low Energy Ion Ring (LEIR), which transfers them to the PS. The PS accelerates the beam to 5.9 GeV/u and sends it to the SPS after passing it through a second foil where it is fully stripped to ^{82}Pb. The SPS accelerates the beam to 177 GeV/u and then sends it to the LHC, which accelerates it to 2.56 TeV/u. Heavy-ions collisions (Pb-Pb) happen with a peak luminosity of 10^{27} cm^{-2}s^{-1}. This value is possible by modifying the existing obsolete Low Energy Anti-proton Ring (LEAR) into a ion accumulator (LEIR) where electron cooling is applied.

At the nominal operation regime, the LHC rings store 2808 proton bunch per ring, each of them containing 1.1^{11} protons colliding with a frequency of 40 MHz. The LHC has performed very well in these years of data taking, allowing the LHCb experiment to cross the threshold of 6 fb^{-1} of integrated luminosity over LHC Run-1 and Run-2, collecting data with an efficiency of over 90%.

This implies that an unprecedented sample of D and B hadrons has been collected, allowing the LHCb collaboration to perform high precision measurements, improving previous results coming from BaBar, Belle and CDF collaborations and allowing the discovery of new effects in the charm and beauty sector.

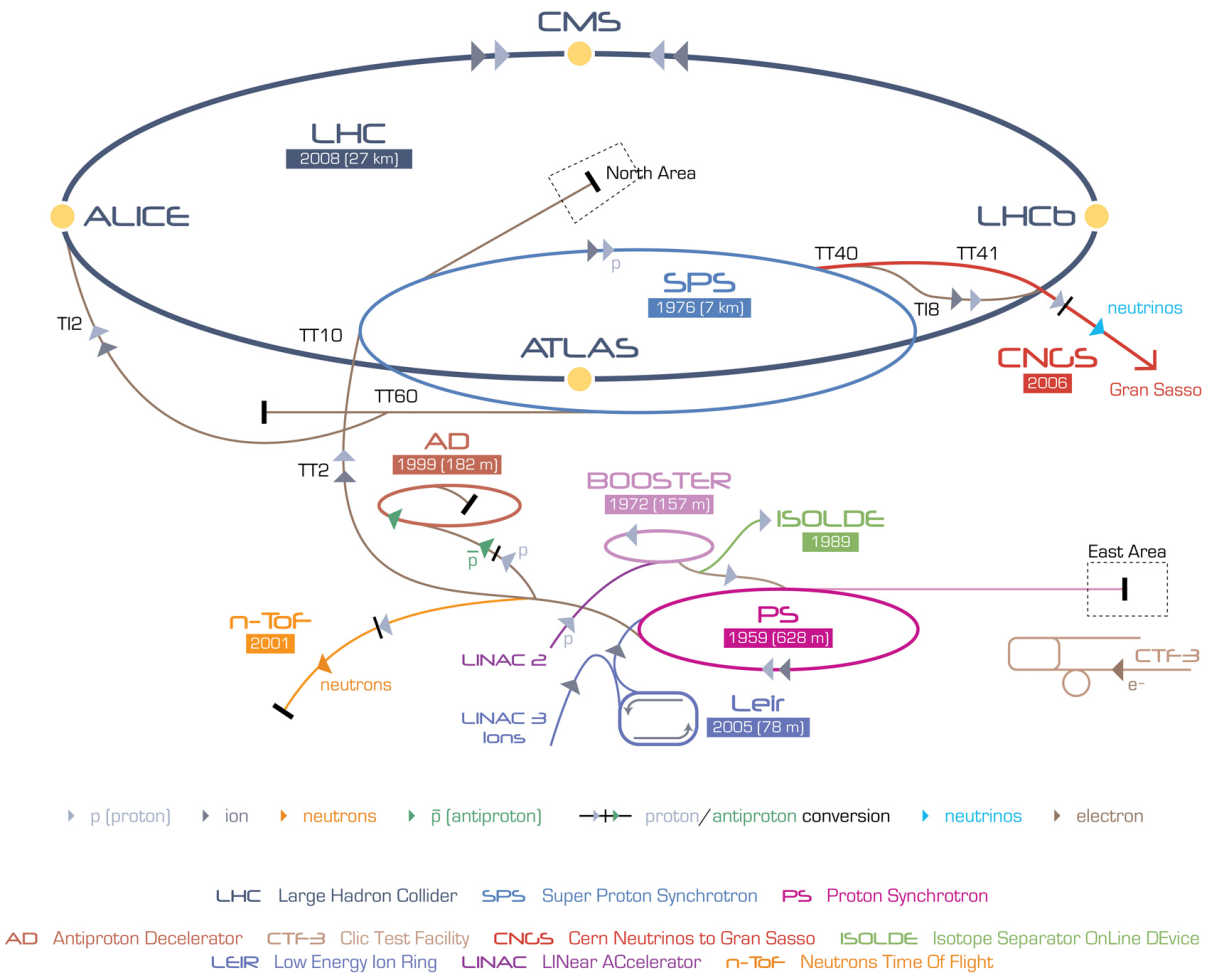


Figure 2.2: Scheme representing the CERN accelerator complex. The various machines employed to pre-accelerate the protons that will be injected in the LHC are LINAC2, PBS, PS, SPS.

2.2 The LHCb detector

The LHCb detector is a single-arm spectrometer with a forward angular coverage ranging from approximately 10 mrad to 300 (250) mrad in the bending (non-bending) plane. The choice of the detector geometry is justified by the fact that at high energies the B hadrons are predominantly produced in the same forward or backward cone. Indeed, the average imbalance in momentum of two partons that collide during a pp interaction means that the b quarks are produced with a strong boost along the beam line. The layout of the LHCb spectrometer is shown in Fig. 2.3.

To purse the LHCb physics program, the detector requires the following features.

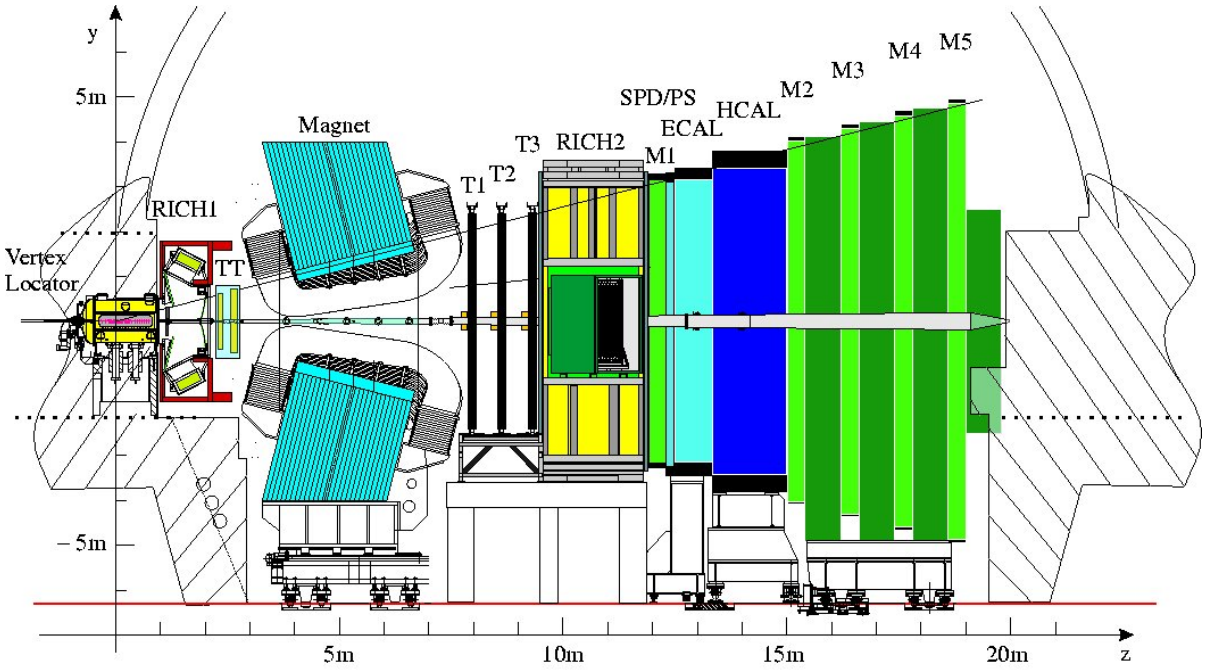


Figure 2.3: Design of the LHCb detector. The right-handed coordinate system adopted has the z axis along the beam, and the y axis along the vertical. From the left to the right all the sub-detectors are visible: VELO, RICH1, TT, Magnet, tracking stations, RICH2, ECAL, HCAL and Muon Stations.

- An efficient, robust and flexible trigger in order to cope with the harsh hadronic environment. The trigger must be sensitive to many different final states.
- An excellent precision in the reconstruction of the interaction vertices and of the B and D hadrons decay vertices. Indeed, to measure neutral mesons oscillation and CP violation it is fundamental to have a suitable proper-time resolution.
- A good identification of protons, kaons and pions in order to cleanly reconstruct many hadronic B and D meson decay final states.
- A data acquisition system with high bandwidth and powerful online data processing capability, needed to optimise the data taking.

A detailed description of all the sub-detectors is given in the next pages.

2.3 The LHCb tracking system

The tracking system is devoted to identify the interaction vertices, reconstruct the trajectories of charged particles and measure their momentum exploiting a magnetic field

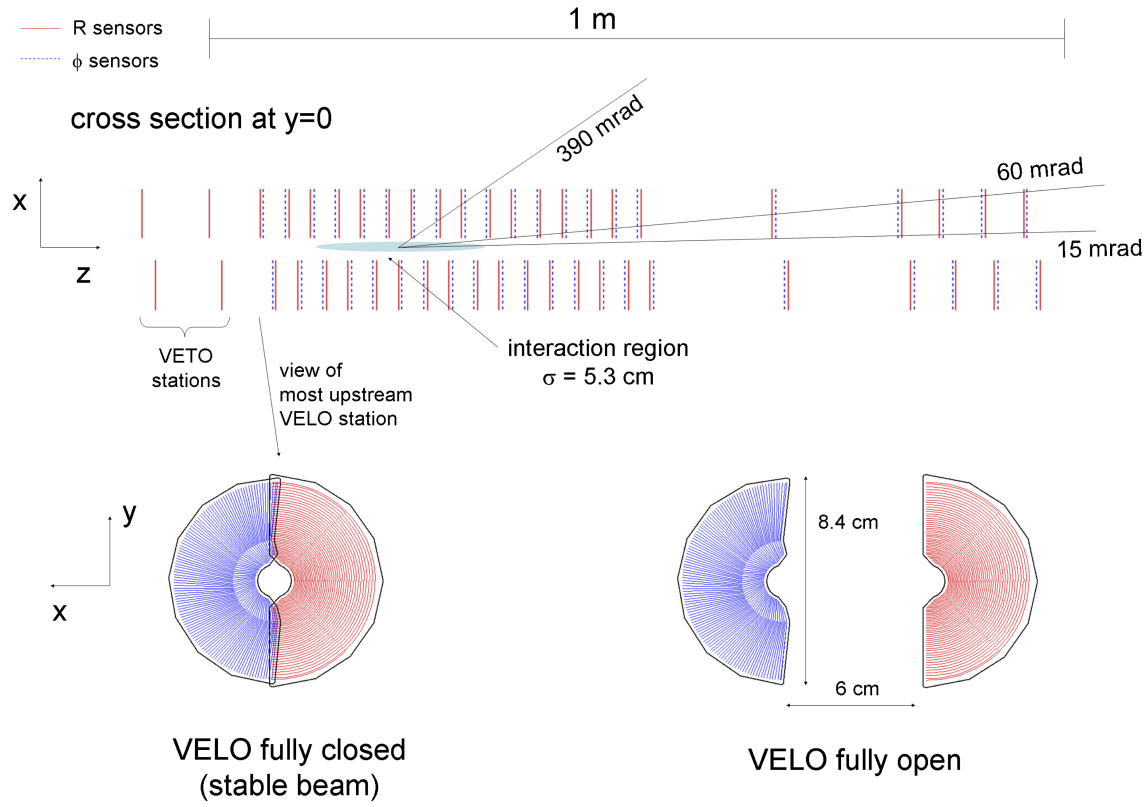


Figure 2.4: Top view of the VELO silicon sensors, with the VELO in the fully-closed mode (top). Frontal view of the modules in the closed and open mode (bottom left and bottom right, respectively) [48].

to bend them.

2.3.1 The Vertex Locator

Beauty (charm) hadrons at LHCb travel a distance of about 1 cm (0.3 cm) before decaying. The presence of a secondary vertex well displaced from the *pp* primary vertex is thus an important signature. For this reason and also due to the high track multiplicity in LHC collisions, it is imperative to have a vertex locator with micro-metric precision in order to select signal events and reject most of the background.

The VELO [52] is composed of 21 circular silicon modules, installed perpendicularly along the beam line, as shown in the top of Fig. 2.4. Each silicon modules is divided in two halves to allow the positioning of the VELO during the data taking phase (closed)

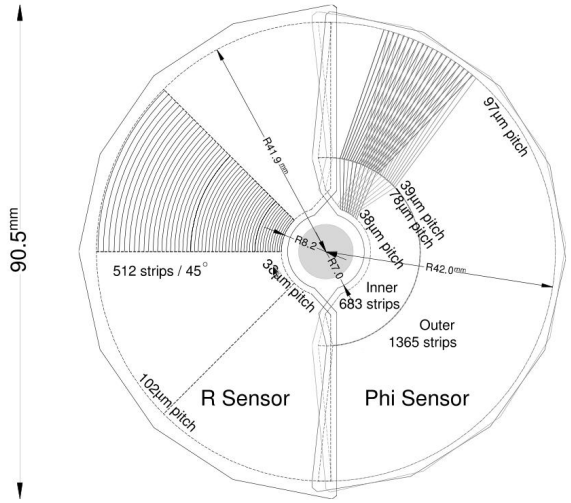


Figure 2.5: Sketch illustrating the R (left) and ϕ (right) geometry of the VELO sensors. In the ϕ-sensor, the strips on two adjacent modules are indicated, to highlight the stereo angle [48].

or during the beam stabilization phase (open), as can be seen in the bottom of Fig. 2.4. For this reason, the modules are installed on a movable device placed inside a cavuum vessel; it is important to note that the two halves of a module partly overlap in the closed VELO configuration, in order to achieve a better geometrical coverage. The modules are composed of two planes of 220 μm thick silicon micro-strip sensors able to measure the distance from the beam (radial distance, R) and the polar angle ϕ of hits generated by the ionizing particles that cross the VELO. The structure of such R and ϕ sensors is reported in Fig. 2.5. The third coordinate z is simply given by the module position.

The R sensors are divided into four parts per halve, each one covering an angle of about 45°; the micro-strips composing these parts are modelled in a semi-circular shape and their width increases as the distance from the centre becomes grater, because the majority of the particles is expected to be near the beam axis (*i.e.* in the high η[1] regions). The micro-strips width ranges from 40 μm near the centre to 92 μm far from the beam.

The ϕ sensors are divided in an inner and in an outer region. Those regions have different skew to the radial direction to improve pattern recognition: they are tilted by 20° and 10° respectively. Furthermore, to improve the track reconstruction, the longitudinally adjacent ϕ sensors have opposite skew to each other.

[1]The psudo-rapidity η is defined as a approximation for the Lorentz-invariant rapidity y:

$$\eta = -\ln\left(\tan\left(\theta/2\right)\right) \approx \frac{1}{2}\ln\left(\frac{\vec{p}+p_z}{\vec{p}-p_z}\right) = y$$

where θ is the angle between the particle and the beam line and p_z is the longitudinal momentum.

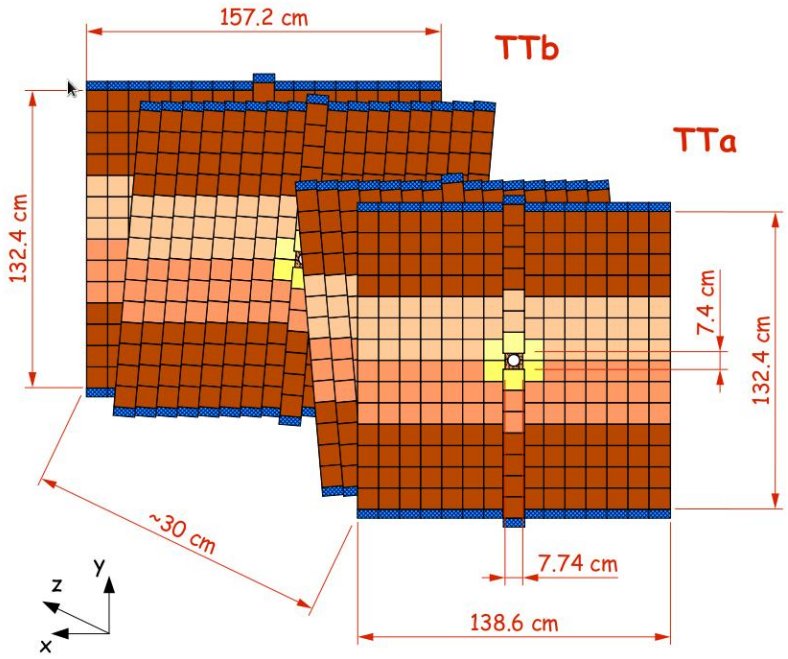


Figure 2.6: Design of the Trigger Tracker sub-detector. The first and the fourth stations have sensor parallel to the vertical plane, while the second and third stations have sensors tilted respectively respectively by +5° and -5° [54].

The resolution on the x and y coordinates ranges from 40 μm to 10 μm depending on the number of tracks fitted, while the resolution on the z coordinate ranges from 250 μm to 50 μm, for the same reason.

2.3.2 The Trigger Tracker

The Trigger Tracker (TT) [53] is placed after RICH1 and before the Magnet. The TT task is to provide reference segments used to combine the track reconstructed in the tracking stations with those reconstructed in the VELO, in order to improve the momentum and coordinate resolution. Since in the space between the VELO and the TT stations an integrated magnetic field of 0.15 Tm is present, the track transverse momentum can be estimated with a resolution of $\delta p_{\mathrm{T}}/p_{\mathrm{T}} = 25\%$ at $p_{\mathrm{T}} = 1$ GeV/c.

The system is composed by four stations, divided in two groups called respectively TTa and TTb, at a distance of about 30 cm one from the other and placed approximately 2.4 m after the beam interaction region. A detailed scheme of this sub-detector is shown in Fig. 2.6. Each of the four stations covers a rectangular region of about 120 cm in height and about 150 cm in width. A TT detector layer is composed of silicon micro-

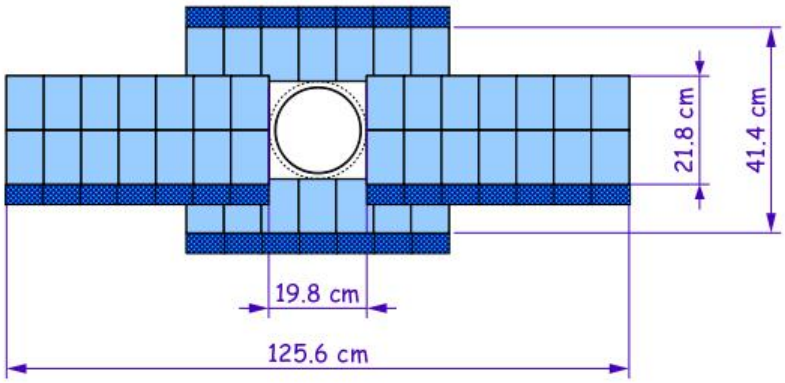


Figure 2.7: Inner Tracker layer with vertically oriented microstrip sensors. The four boxes are arranged around the beam pipe and the individual sensors inside the boxes are visible. The deep blue part of each box represents the readout plugs [54].

strip sensors with a 183 μm pitch, arranged in readout strips up to 38 cm long, to keep the number of readout channels low. In the first and fourth stations the strips are parallel to the vertical plane, while in the second and in the third stations they are tilted respectively by +5° and -5°. This is done to improve the precision of the track reconstruction.

2.3.3 The tracking stations

The three Tracking Stations T1, T2 and T3 are placed behind the magnet. They are divided in two main parts, depending on the distance from the beam pipe. The inner part of the Tracking Stations is called Inner Tracker (IT), while the outer part is called Outer Tracker (OT). They adopt different technologies to detect particles: the former is composed of silicon micro-strips sensors, while the latter consists of drift straw tubes.

The Inner Tracker [55] covers the region around the beam pipe and it is arranged in a cross-shaped geometry, that grants optimal coverage while conserving surface; each station consists of four independent boxes arranged as shown in Fig. 2.7. As for the TT, the first and fourth planes of the IT have the sensors parallel to the vertical plane, while the second and the third have the sensors tilted by +5° and -5°. The side boxes have two ladders of micro-strips, where the lower sensors are connected in series with the upper sensors to a single readout channel, while the top and the bottom boxes have only one micro-strip ladder. The total IT size is about 1.2 m in the bending plane and about 40 cm in the vertical plane.

The Outer Tracker [56] is a gas-filled straw tubes detector, covering about 99% of the summed surface of the T1-T3 tracker stations. For each tracking station there are

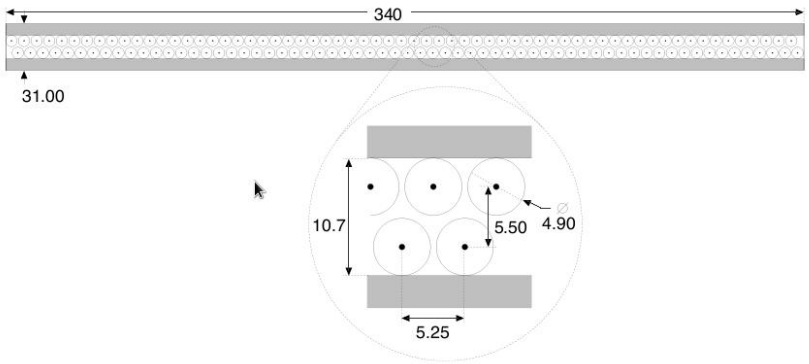


Figure 2.8: Cross section of a straw-tubes module [48].

four planes of straw tubes arranged in the same way as the TT and IT silicon microstrips. Moreover, each plane is composed of two rows of tubes, arranged in a honeycomb structure, as shown in Fig. 2.8. The straw tubes have a radius of 5 mm and are filled with a mixture of Ar-CF_4-CO_2. At the tube ends, locator pieces support and centre the anode wire with a precision better than 100 μm. Unlike other tracking detectors here described, the OT measures drift times rather than pulse heights. The readout time window exceeds a single LHC bunch crossing interval due to the limited drift speed of the gas mixture. The OT resolution is better than 200 μm.

2.3.4 The LHCb dipole magnet

All modern experiments measure particle momenta through their curvature in a given magnetic field. For this reason, the LHCb detector is provided with a warm (*i.e.* non superconducting) magnet dipole placed between the TT and the first tracking station T1, as can be seen in Fig. 2.3. The magnet is formed by two coils placed with a small angle with respect to the beam axis, to increase the opening window with z in order to follow the acceptance of the LHCb detector. The main component of the magnetic field is along the y-axis as shown in Fig. 2.9 and thus the xz-plane can be considered with good approximation as the bending plane. The maximum magnetic field strength is above 1 T, while its integral is about $\int \vec{B} \cdot d\vec{l} = 4$ Tm. All the tracking detectors are located outside the magnetic dipole, as shown in Fig. 2.9. The magnetic field is measured before the data-taking periods with Hall probes to obtain a precise map, which is crucial to have a good momentum resolution and consequently a good mass resolution.

Among the main LHC experiments, the LHCb detector has a unique feature consisting in the possibility to reverse the polarity of the magnetic field (*MagUp* or *MagDown*). This allows a precise control of the left-right asymmetries introduced by the detector. Indeed, particles hit preferentially one side of the detector, depending on their charge,

thus generating non-negligible asymmetries if the detector is not completely symmetric. If data samples collected with the two different polarities have approximately equal size and the operating conditions are stable enough, effects of charge asymmetries are expected to cancel. The magnet polarity is therefore reversed approximatively every two weeks to meet these constraints. In the measurement reported in Chapter 3, the polarity reversal is explicitly exploited since the two samples with different polarities are analysed together.

2.3.5 Tracking algorithm and performances

The tracks are divided in five categories, as reported in Fig. 2.9.

Long tracks: Particles generating hits in all tracking sub-detectors.

VELO tracks: Particles generating hits only inside the VELO since they have been produced with a wide angle with respect to the beam pipe; for this reason they exit from the detector geometrical acceptance just after the VELO.

Upstream tracks: These tracks are generated by particles with a low momentum, that produce hits in the VELO and in the TT, but are kicked off the geometrical acceptance of the detector by the magnetic field. However, the momentum of these particles can still be determined thanks to the residual magnetic field present in the VELO, even if the measurement is affected by a 20% relative uncertainty.

Downstream tracks: Long lived neutral particles can decay between the VELO and the TT, producing charged particles that generate hits only in the TT and in the tracking stations. These are the so-called downstream tracks.

T tracks: Particles generating hits only in the tracking stations.

The track reconstruction process is organized in a hierarchical way: the algorithm tries firstly to reconstruct long tracks and then it picks up unused segment to reconstruct downstream and upstream tracks. Long tracks are reconstructed with two algorithms: the first extrapolates VELO segment to the tracking stations, adding to the track the compatible hits in the TT. The second matches VELO and tracking station segments one to each other, extrapolating VELO segments in the forward direction and tracking station segments in the backward direction. Downstream tracks are reconstructed starting from tracking station segments and then adding the compatible hits in the TT to those segments. Upstream tracks are obtained extrapolating VELO segments to the TT, adding compatible hits and requiring a non-compatibility with any of the tracking station segments.

The reconstruction of tracks is divided in two steps: track finding and track fitting. The first starts with the definition of segments in the various sub-detectors: inside the

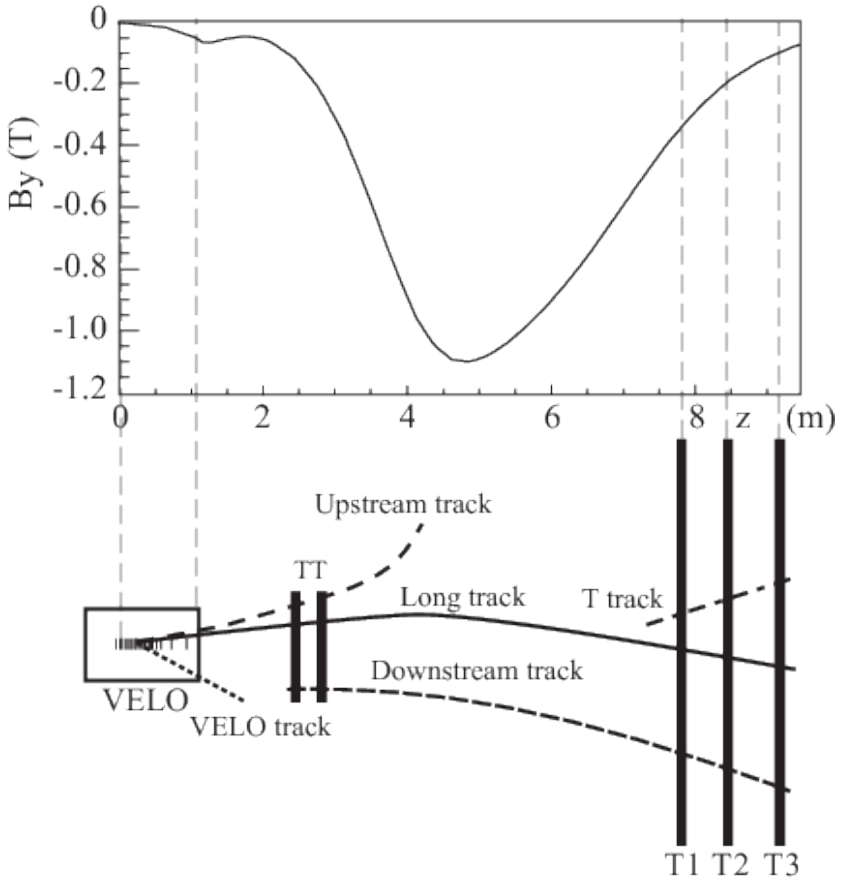


Figure 2.9: Sketch illustrating the various track types: long, VELO, upstream, downstream and T tracks. For reference the main B-field component (B_y) is plotted above as a function of the z coordinate [57].

VELO, segments are created matching all hits that lie on a straight line. In the tracking stations, a segment is created matching the hits contained in a region of the first and third station, using the information given only by one plane of vertically oriented micro-strips sensors. Then, under the hypothesis of a parabolic trajectory, the algorithm calculates the position of the hit in the middle stations and searches for compatible hits. If a signal is found, it is added to the segment and it is used to better determine parameters of the trajectory. Finally, the compatible hits coming from the other planes of sensors are also added, in order to have a 3-dimensional segment. Afterwards, a bi-directional Kalman filter [58] is applied to better determine the track parameters and then a clone-killer algorithm compares the reconstructed tracks, two by two: if a pair of tracks shares more than a fixed percentage of hits they are considered clones and only that with more hits (or the best χ^2 from the track fitting) is stored.

2.4 The LHCb particle identification system

In this section all the LHCb sub-detectors used for the particles identification (PID) are described. They consists in two Ring Imaging Cherenkov (RICH1 and RICH2) detectors, the electromagnetic calorimeter (ECAL), the hadronic calorimeter (HCAL) and finally the muon stations.

2.4.1 The RICH detectors

Particle identification is of fundamental importance in *CP* violation measurements. The LHCb experiment exploits two RICH detectors, the first on installed immediately after the VELO and the second one positioned after the tracking stations, to discriminate between charged pions, kaons and protons in a momentum range from few GeV/c up to about 150 GeV/c. Cherenkov light detectors exploit the light emitted by particles that travel in a medium faster than light in the same medium. The relation between the Cherenkov photon emission angle $\theta_{\check{c}}$ and the refraction index n of the radiator is

$$\cos\left(\theta_{\check{c}}\right) = \frac{1}{\beta n}, \tag{2.1}$$

where $\beta = v/c$ is the particle velocity relative to the speed of light in the vacuum. The Cherenkov light emission only occurs when the particles exceeds the threshold value of $\beta_{th} = 1/n$ (*i.e.* $\theta_{\check{c}} = 0$) while each radiator has a maximum emission angle $\theta_{\check{c}}^{max} = \arccos\left(1/n\right)$ which is obtained when $v = c$. It is evident that for particles approaching the speed of light the Chenkov angle will saturate to $\theta_{\check{c}}^{max}$ and it is therefore necessary to have different radiators in order to discriminate particles in a wide range of momenta.

The RICH1 [48] is optimized to identify tracks with a relatively low momentum, between 1 GeV/c and about 50 GeV/c. The structure of the apparatus is reported in the left part of Fig. 2.10. The geometrical acceptance (from 25 mrad to 330 mrad) of the device is enough to cover the whole LHCb detector acceptance. During Run-1, there were two different types of radiators inside RICH1: the first was a 5 cm thick Aerogel layer with $n = 1.03$, suitable for low momentum particles, while the second was gaseous (C_4F_{10}) with $n = 1.0015$ filling the remaining part of the detector and was employed to detect particles with higher momenta. The Aerogel radiator was removed in the operational shut down before Run-2 as its ability to provide particle ID was compromised by the total number of photons in RICH1 in such a high track multiplicity environment [59].

The structure of the RICH2 [48] sub-detector is reported in the right part of Fig. 2.10. Its geometrical acceptance, ± 120 mrad (horizontal) and ± 100 mrad (vertical), covers the region of the detector where most of high momentum particles are found. The radiator chosen is CF_4 with a refraction index $n = 1.00046$, optimal for the higher momentum, up to about 150 GeV/c.

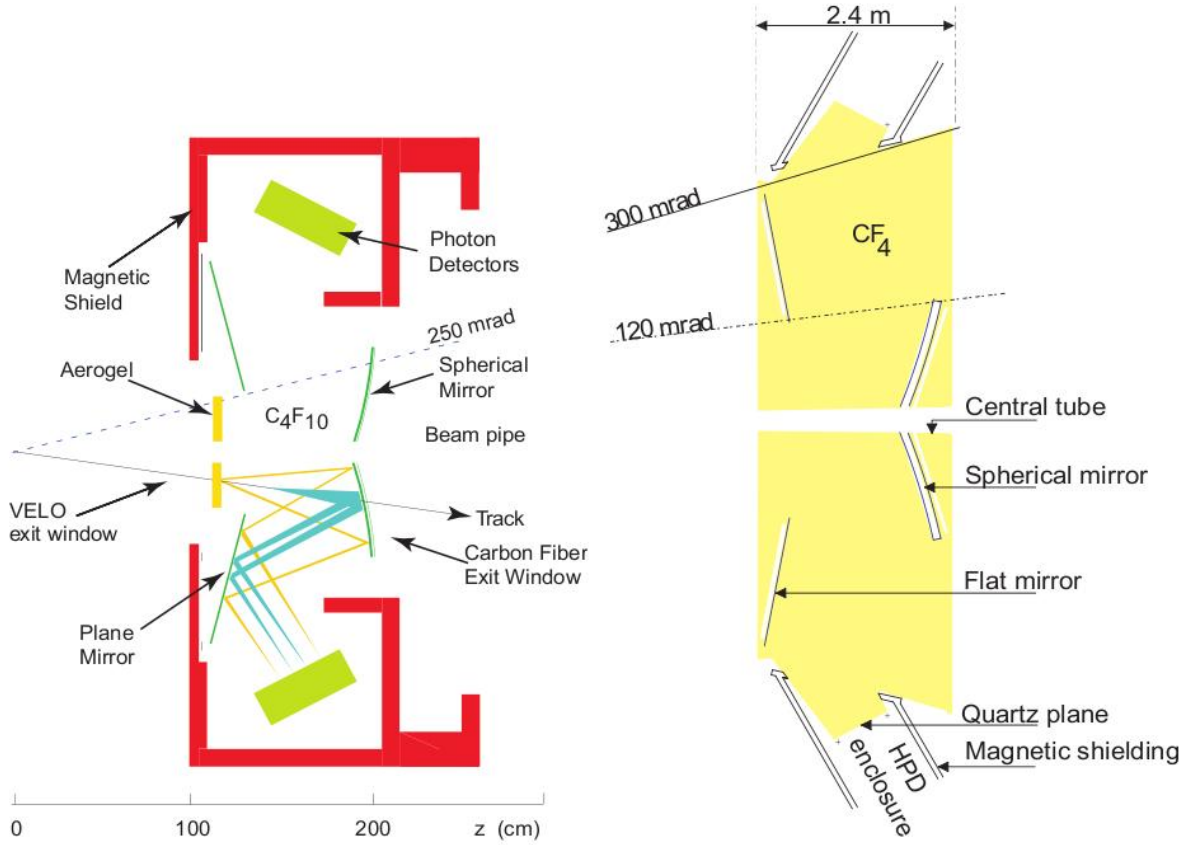


Figure 2.10: Left: schematic view of the RICH1 sub-detector [48]. The Cherenkov light is emitted with different angles from the Aerogel (yellow) and the C_4F_{10} (light blue) radiators. Right: schematic view of the RICH2 sub-detector, filled with CF_4 gas [48].

In both the detectors, the Cherenkov light is focused, through a system of spherical and plane mirrors, onto a lattice of photo detectors, the Hybrid Photon Detector (HPD). The HPDs are placed in both the RICH sub-detectors, outside the experiment acceptance and they are shielded against the residual magnetic field. Indeed, the photo-electrons created in the photomultipliers would be bent by the residual magnetic field reducing the HPD's performances.

Particle identification method

RICH detectors are able to discriminate between the various mass hypothesis for a given particle. Indeed, as shown in Fig. 2.11, the photon emission angle is related to the particle mass and to its momentum. Since the Cherenkov light emission covers the full solid angle, rings with radius proportional to $\theta_{\check{c}}$ are expected on the HPD plane.

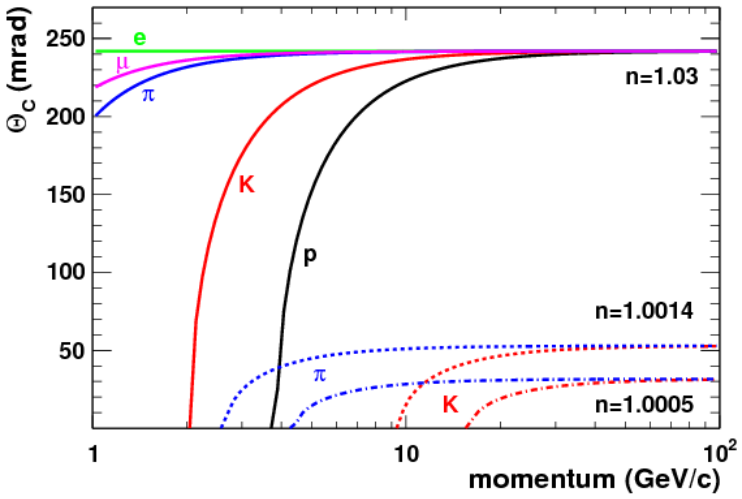


Figure 2.11: Cherenkov angles as a function of momentum for different particle species and for the three different values of the refractive index n corresponding to the three radiator materials used in the RICH setup [60].

Measuring the photons hit positions, it is then possible to discriminate the various mass hypotheses.

Due to an irreducible background, given by photons coming from other particles, and due to the complexity of the problem, the following approach has been chosen to achieve the best particle discrimination. For a given set of mass hypotheses, the probability for a single photon to be detected on a single HPD pixel is computed; then, the expected contribution from all sources is compared with the observed number of photons and a likelihood is calculated (the change in the likelihood value depends only on the mass hypothesis assigned to the tracks). Only five mass hypotheses are considered for the tracks detected: electron, muon, pion, kaon and proton. Since the computation of the likelihood for all tracks would be unfeasible, a different approach is adopted. In fact, the pion mass hypothesis is used for all the tracks detected and a first global likelihood is computed. Then the hypothesis is changed to e, μ, K and p for one particle at a time and the change in the global likelihood is computed. The chosen mass hypothesis is the one that returns the maximum improvement in the global likelihood. This process is repeated for all tracks, until no improvement is observed in the likelihood value.

The discriminating variable is the so-called $\Delta \log(\mathcal{L})_{X-\pi}$ which is the difference between the logarithm of the likelihood under the X (e, μ, K or p) and π hypothesis for the observed track:

$$\Delta \log(\mathcal{L})_{X-\pi} = \log(\mathcal{L}_X) - \log(\mathcal{L}_\pi). \tag{2.2}$$

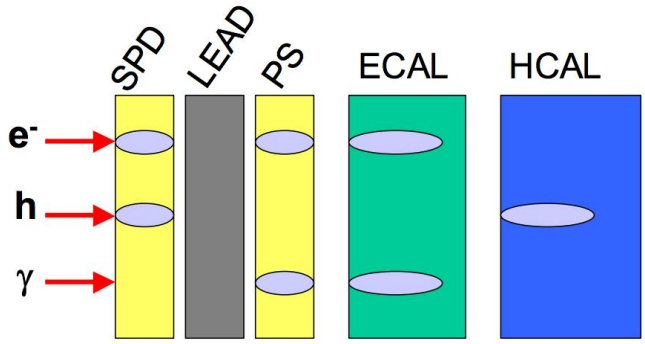


Figure 2.12: Signal deposited on the different parts of the calorimeter by an electron, a hadron and a photon [48].

For example, a large positive value of $\Delta \log (\mathcal{L})_{K-\pi}$ corresponds to a high probability that the particle is a kaon, while a large negative value corresponds to a high probability that the particle is a pion.

The efficiency of this discriminating method had been widely studied using real data sample with high purity final states selectable only using kinematical cuts, due to their particular kinematic characteristics (*e.g.* $K^0_S \to \pi^+\pi^-$, $\Lambda^0 \to p\pi^-$, and $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$).

2.4.2 The calorimeters system

The calorimeters system [61] is used to measure hadron, electron and photon energies, thus giving information for their identification. Moreover, it provides important information for the Level-0 trigger (L0), evaluating hadron, electron and photon transverse energy E_{T}. The calorimeters system is divided into four sub-detectors:

- Scintillator Pad Detector (SPD).
- Pre-Shower (PS).
- Electromagnetic Calorimeter (ECAL).
- Hadronic Calorimeter (HCAL).

A sketch of the calorimeters system and the response of each detector with the particle types is reported in Fig. 2.12. Each sub-detector is divided into regions where differently sized sensors are used. SPD, PS and ECAL are divided in three regions (inner, middle and outer), while HCAL is divided only in two regions (inner and outer). The sensor size increases as the distance from the beam pipe is greater to reach a compromise between occupancy and the number of readout channels.

The SPD and the PS are auxiliary sub-detectors of the Electromagnetic calorimeter and they are placed in front of it. The SPD is used to discriminate between charged and neutral particles, as the former emit light when crossing a scintillator material while the latter do not. The PS is instead used to obtain a better discrimination between electrons and pions. Both the sub-detectors consist of scintillating pads with a thickness of 15 mm, inter-spaced with a 2.5 radiation lengths[2] lead converter. The light produced by the scintillator material is collected using wavelength-shifting fibers (WLS). These fibers are used to transmit the light to multi-anode photomultipliers (MAPMTs) located outside the detector. The SPD and PS contain about 6000 pads each.

The ECAL is a sampling calorimeter separated in independent modules. In each module, the scintillation light is read out via WLS fibers running perpendicularly to the converter/absorber plates: this technique offer the combination of an easy assembly, good hermicity and fast time response. A sketch of the ECAL is given in Fig. 2.13. Each ECAL module is composed of 66 lead converter layers (2 mm thick), each one installed between two plastic scintillator layers 4 mm thick. In total, all the layers installed in the ECAL correspond to about 25 radiation lengths and 1.1 nuclear interaction lengths[3]. The WLS fibers bring the light produced by the scintillator material to the readout photomultipliers in the back part of the module. The module size and the number of readout channels differ depending on the region where the module is installed. In the inner region each module has a section of 4×4 cm^2 and 9 readout channels. Finally, the outer region is composed of 12×12 cm^2 modules with one channel each.

The HCAL main task is to measure the energies of hadronic showers. This information is fundamental for the Level-0 trigger. The HCAL structure is very similar to the ECAL structure, with the difference that each module is composed of scintillator layers 4 mm thick interleaved with steel layers 16 mm thick. This corresponds to roughly 5.6 nuclear interaction lengths in total. In the inner region modules have a section of 13×13

[2]The radiation length is defined as

$$X_0 = \frac{A \cdot 716.4\text{g/cm}^3}{Z(Z+1)\ln(287\sqrt{Z})}$$

where A is the mass number and Z is the atomic number of the considered material. The radiation length corresponds to the distance over which the energy of an electron is reduces by a factor $1/e$ only due to radiation loss.

[3]The nuclear interaction length is defined as

$$\lambda = \frac{A}{N_A \, \rho \, \sigma_{inel}}$$

where N_A is the Avogadro constant, A is the mass number and ρ is the density of the considered material while $\sigma_{inel} \sim \sigma_{pp} A^{2/3}$ is the inelastic cross section between the particle and the nucleus. The nuclear interaction length is the mean path length required to reduce the numbers of relativistic charged particles by the factor $1/e$ as they pass through matter.

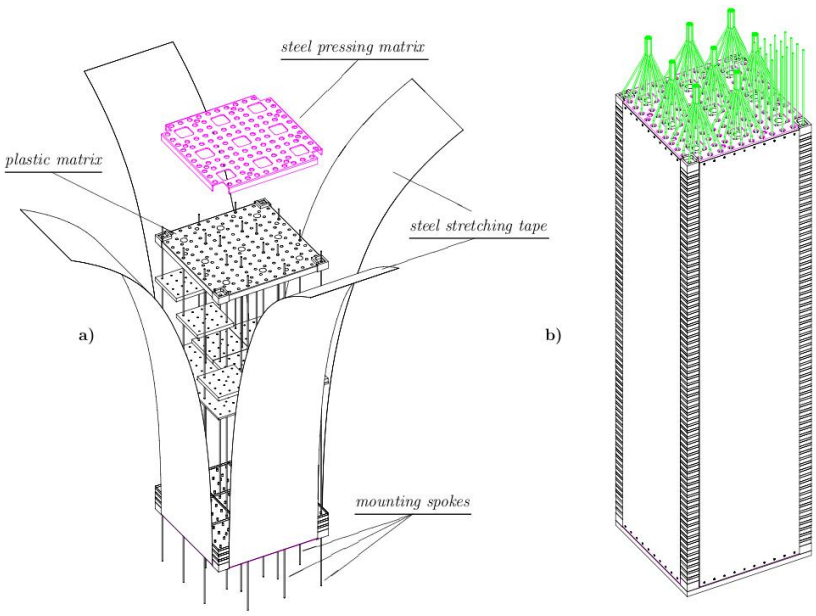


Figure 2.13: Left: representation of an ECAL module during the assembly phase. The lead/scintillator layers are clearly visible. Right: representation of an assembled ECAL module. The green lines connected to an end are the WLS fibers connecting the calorimeter to the photomultipliers [61].

cm^2, while in the outer region their dimensions are 26 $\times$ 26 cm^2.

Calorimeters system resolution

The calorimeters system performances have been evaluated from many test beams made before the start of the data taking. Energy resolutions are given by $\sigma(E)/E = \frac{(8.5-9.5)\%}{\sqrt{E}} \oplus$ 0.8% for ECAL and $\sigma(E)/E = \frac{(69\pm5)\%}{\sqrt{E}} \oplus (9 \pm 2)\%$ for HCAL. The ECAL calibration is achieved by reconstructing resonances decaying to two photons like $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$. Calibration of the HCAL can be realized by measuring the ratio E/p between the energy E measured in the calorimeter for a hadron with momentum p, measured by the tracking system.

2.4.3 The muon system

The final part of the LHCb detectors consists of five muon stations, that altogether form the muon sub-detector [62]. Muons with high p_{T} are very important particles since several final products of B-hadron decay chains contain muons. The five stations (M1-M5) cover an angular acceptance of $\pm$300 mrad in the horizontal plane and $\pm$200 mrad

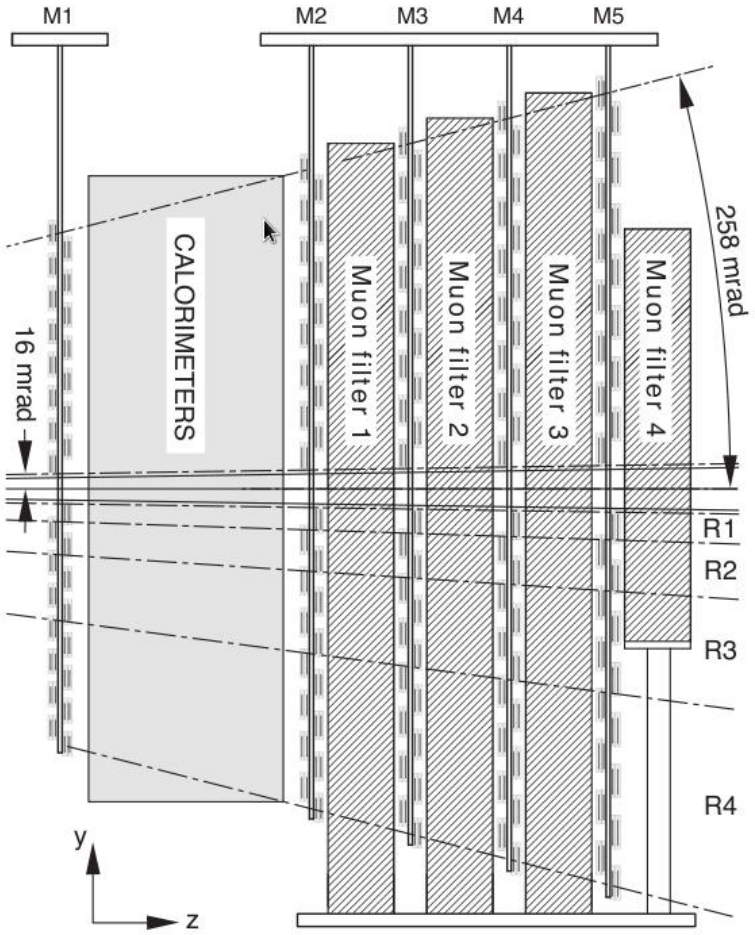


Figure 2.14: Lateral view of the muon detector. The first muon station M1 is placed before the calorimeters while the other stations are placed at the end of the LHCb detector [62].

in the vertical plane. The geometrical efficiency for the detection of muons coming from B-hadrons is nearly 46%. The first muon station M1 is placed before the calorimeters, to avoid possible muon multiple scattering effects, that could modify the particle trajectory. The remaining stations (M2-M5) are placed after the calorimeter system, at the end of the LHCb detector. A schematic view of the muon sub-detector is reported in Fig. 2.14.

Each muon station is divided into four regions (R1-R4) around the beam pipe. The dimensions of the chambers increase as they are more and more distant from the beam pipe. Moreover, also the segmentation of each region increases as the distance from the beam pipe becomes grater in a ratio 1:2:4:8. In this way, the charged particle occupancy is expected to be about the same in each region. All the chambers are Multi-Wire

Proportional Chambers (MWPCs), except for the inner region of the M1 station where Gas Electron Multipliers (GEMs) detectors are employed. The latter consist of three GEM foils sandwiched between anode and cathode planes. MWPCs have four overlapped gaps, each one 5 mm thick and with a distance between wires of about 2 mm. In total, the muon detector contains 1380 MWPCs.

Muon-ID algorithm performances

The algorithm of muon identification in the hardware trigger starts from the hits in the M3 station. For each hit, a straight line is extrapolated to the interaction region defining a "field of interest", that takes into account also the magnetic field kick, around such a trajectory. Hits coming from long and downstream tracks that are found around the extrapolated trajectory are fitted together to form a muon track. To consider the track as a muon it is requested to have hits in M1-M3 if the track momentum is between 3 and 3.5 GeV/c and in M1-M4 if the track momentum is between 3.5 and 4.5 GeV/c. Finally, it is required to have hits in all the five chambers if the track momentum is higher than 4.5 GeV/c. After this, complex algorithms compute the muon likelihood for each muon track, used as a particle identification discriminator.

2.5 The LHCb trigger

As already stated, the production cross-section of $b\bar{b}$ and $c\bar{c}$ pairs are quite large and together they account to about 10% of the total pp inelastic cross-section. This means that a good trigger system is needed in order to accept only the interesting events while rejecting at the same time most of the background events. The LHCb trigger has been developed to work at the bunch crossing frequency of the LHC. The only way to reach the desired performances is to divide the trigger into different levels, each processing the output of the previous. The LHCb trigger system is divided into three levels.

Level-0 (L0): this is the first trigger level and it is based on custom electronics. It is designed to perform a first filtering of the events, reducing the input rate of about 40 MHz to an output rate of only 1 MHz.

High Level Trigger 1 (HLT1): this is the second trigger level and it is software based. The task of HLT1 is to filter events in an inclusive way and to reduce the rate of accepted events to 50 KHz.

High Level Trigger 2 (HLT2): this is the last trigger level and it is completely software based. The HLT2 apply an exclusive selection of beauty and charm decays, performing a full reconstruction of the events which is finally sent to mass storage. At the beginning of Run-1 HLT2 operated with an output rate of about 3.5

kHz. Improvements have been made over the years and the output rate have been increased up to about 12.5 kHz.

2.5.1 The Level-0 trigger

The L0 trigger uses information coming mainly from the tracking system and from the calorimeter system. In fact, at this level, the trigger decides to keep or discard events based on measures of p_{T} and E_{T} of the particles composing the event. The system uses three independent systems running in parallel:

L0Photon/Electron This trigger uses the information given by the SPD, PS and ECAL detectors. Custom boards are programmed to measure the energy of electromagnetic showers. The event is accepted if there is at least one cluster with E_{T} greater than a certain threshold.

L0Hadron This trigger exploits the information given by the HCAL detector. The way in which it works is the same as the electron/photon trigger: the event is accepted if there is at least one cluster with enough transverse energy.

L0Muon It uses the information given by the five muon stations. Tracks are reconstructed defining field of interest around particle hits and then connecting hits in the same field of interest. Events are accepted if at least one muon candidate has a transverse momentum exceeding a given threshold. Moreover, the trigger contains a line to select muon pairs, asking that the sum of their transverse momentum is greater than a threshold.

2.5.2 High Level Trigger (HLT)

For events passing the L0 trigger, the full set of detector information is transferred to the Event Filter Farm (EFF), where the HLT runs asynchronously with respect to the bunch crossing. The HLT is a software based C++ application performing a fast full reconstruction of events. About 29 000 copies of the HLT application run concurrently in the EFF. The HLT consists of two stages. The first stage (HLT1) processes the full L0 rate (1 MHz) and uses partial event reconstruction to reduce the rate to 50 kHz. Then in the second stage (HLT2) a complete event reconstruction is performed with some simplifications that are needed to satisfy the time constraints.

HLT1 The VELO reconstruction is fast enough to perform a full 3D pattern recognition and primary vertex (PV) finding for all events entering the HLT. Only the VELO tracks having a significant impact parameter (IP) with respect to all PVs and a minimum p_{T} are passed to the forward tracking algorithm that searches for matching hits in the tracking stations. The tracks are required to be separated from

the primary vertex, to have a good track quality and a high transverse momentum to fire the trigger.

HLT2 The HLT2 stage performs a full event reconstruction in an almost offline-like reconstruction quality. Only tracks with $p > 5$ GeV/c and $p_{\mathrm{T}} > 300$ MeV/c are used due to time constraints. HLT2 is composed of inclusive and exclusive selections. A number of different event selection strategies can be used according to the topology of the decays of interest.

Trigger decisions are organized in *trigger lines*, which are sequences of reconstruction and selection algorithms to trigger a given category of events. Different trigger lines select decays with different signatures. Specific choices of the trigger selection used for the $\mathcal{A}_{CP}$ measurements described in this thesis are reported in § 3.4.1.

2.6 Data management and computing

The basic LHCb computing model is based on a series of distributed multi-tier regional centres of different dimensions. LHCb (as well as the other three major experiment at the LHC) requires a large amount of memory disks as well as CPU power in order to store and process the data coming from the detector and to perform analysis task (*e.g.* ntuple production). The computing system is divided in different tiers dedicated to specific duties. The Tier0 is the CERN data centre and provides to LHCb about 20% of the total resources required by the experiment and it is connected to the Tier1 centres via a private network of 10 Gbit/s optical-fiber links (LHCOPN). Moreover, Tier0 stores the RAW data, also providing a copy distributed among the Tier1 centres. There are 6 LHCb Tier1 centres worldwide that are responsible for storing a proportional share of raw and reconstructed data, as well as performing large-scale processing and storing the corresponding output. Furthermore, the Tier1 centres have to distribute the data to the Tier2 centres and to store a part of the simulated data coming from them (*i.e.* Tier2 centres mainly provide CPU resources). Each Tier1 is connected to a number of Tier2 centres, usually in the same geographical area. Finally, Tier3 resources consist of local clusters in a centre of research or a university department; they are dedicated to specific jobs needed by the research team who owns them. This system is collectively referred to as the World LHC Computing Grid (WLCG).

2.6.1 Data processing

The data processing involves several phases that normally follow each other in a sequential manner. The real raw data come from the detector and they are reconstructed via the online Event Filer Farm. Obviously, the first step is to collect the events of interest with an appropriate trigger system. The raw data are then processed using optimized

and highly specialized algorithms implemented by the HLTs. The software applies the necessary calibration corrections during the reconstruction of the properties of the particle and imposes requirements based on physics criteria. The raw events accepted by the trigger are then transferred to the CERN Tier0 centre in order to be archived and afterwards forwarded to the Tier1 centres for further processing.

For what concerns the simulated data, the events are generated from a simulation of the LHCb detector, that includes the best understanding of the LHCb detector response, trigger response and passive material budget. The format of this type of data is the same of raw data.

Whether the data are real or simulated, they must be reconstructed in order to provide physical meaningful quantities: for example, one must determine the energy of electromagnetic and hadronic showers measuring calorimeter clusters, or hits in the tracking system have to be associated to tracks. Furthermore, the information about PID coming from the RICH sub-detectors must also be reconstructed to provide particle identification. The reconstruction process produces a new type of data, the so-called Data Summary Tape (DST).

The information contained in the DST (tracks, energies, clusters, PID) is further analysed with specific algorithms, in order to identify candidates that could form composite particles. These algorithms are designed to select only certain categories of events (*e.g.* the B2HH algorithm selects only B candidates decaying to two hadrons) and are called *stripping lines*. Such lines are written for each channel of interest and they produce the output used for further analyses. The output of the stripping stage is referred to as full DST. In addition, an event tag is also created for faster reference to the selected events. The tag contains a little summary of the event characteristics together with the results of the pre-selection algorithms and a reference to the events contained in the DST dataset.

In Run-2, few changes to the data flow included the possibility to perform the full event reconstruction in the trigger, thus bypassing the offline reconstruction and discarding the raw event. This new strategy is particularly interesting for charm physics that mostly suffers the trigger output rate constraints. In the *Turbo stream* [63], the HLT directly writes out a DST containing all information necessary for analyses, and this allows an increased output rate and thus higher average efficiencies. Event pre-selection algorithms (lines) are used for data reduction and designed to identify specific decay channels. The sample of data used in this thesis comes from various Turbo stream lines, as described in § 3.4.

Chapter 3

Measurement of CP asymmetries in $D_s^\pm \to K_S^0\pi^\pm$, $D^\pm \to K_S^0 K^\pm$ and $D^\pm \to \phi\pi^\pm$ decays

3.1 Introduction

Cabibbo-suppressed (CS) charm decays are the main channels to search for direct *CP* violation (CPV) in the charm sector. In these decays, direct CPV will occur if tree and loop (penguin) processes interfere with different strong and weak phases. Furthermore, contributions from physics beyond the Standard Model may appear in the virtual loops [64], larger than *CP* violating effects predicted by the Standard Model $\sim \mathcal{O}(0.1\%)$ or below [35].

In this analysis, *CP* violation in charged Cabibbo-suppressed decays, namely $D_s^\pm \to K_S^0\pi^\pm$ and $D^\pm \to K_S^0 K^\pm$ (collectively referred to as $D_{(s)}^\pm \to K_S^0 h^\pm$), and $D^\pm \to \phi\pi^\pm$, is investigated. In the SM, *CP* violation in $D_{(s)}^\pm$ decays into a neutral kaon includes a contribution from *CP* violation in the K^0 mixing and has to be taken into account for the final measurement.

To research for CPV, the Cabibbo-favoured (CF) $D^\pm \to K_S^0\pi^\pm$, $D_s^\pm \to K_S^0 K^\pm$ and $D_s^\pm \to \phi\pi^\pm$ channels are used as control sample to constraint the detection and production asymmetries. These decays are sensitive to CPV via the interference with the doubly Cabibbo-suppressed (DCS) amplitude. However, the *CP* asymmetry in these channels is predicted to be at most $\mathcal{O}(0.01\%)$ in the Standard Model [65]. Therefore CPV in the CF decay is assumed to be negligible, if compared to the CS decays.

Several studies on *CP* violation in $D_{(s)}^\pm \to K_S^0 h^\pm$ and $D^\pm \to \phi\pi^\pm$ decays have been performed previously by CLEO [40], BaBar [41, 42], Belle [43–45] and LHCb using Run-1 data [46, 47]. As regards $D_{(s)}^\pm \to K_S^0 h^\pm$ decays, no previous result indicate a *CP* asymmetry greater than the one expected due to the neutral kaon final state. No

evidence for CP violation in $D^\pm \to \phi\pi^\pm$ decays was found.

3.2 Overview of the analysis

The experimental method is discussed in § 3.3, while the trigger requirements and cut-based selections for the $D^\pm_{(s)} \to K^0_S h^\pm$ and $D^\pm_{(s)} \to \phi\pi^\pm$ signals are presented in § 3.4.1 and § 3.4.2. In these sections, how to constrain the topology of the decays and to suppress the backgrounds arising from particle misidentification is also reported. In § 3.5 it is described how a multivariate analysis for $D^\pm_{(s)} \to K^0_S h^\pm$ decays is used to remove most of the combinatorial background. The signal yields and asymmetries are extracted by fitting the $K^0_S h^\pm$ and $\phi\pi^\pm$ invariant mass distributions. The fit method is introduced in § 3.6 and their outcomes are reported in § 3.7. The last section (§ 3.8) reports a qualitative introduction to the main systematic sources.

3.3 Experimental method

The CP asymmetry for charged meson decays is defined as

$$\mathcal{A}_{CP}^{X^\pm \to f^\pm} = \frac{\Gamma(X^+ \to f^+) - \Gamma(X^- \to f^-)}{\Gamma(X^+ \to f^+) + \Gamma(X^- \to f^-)}, \tag{3.1}$$

where Γ is the partial decay width, X denotes a charged meson and f is a final state.

The CP asymmetry in $D^\pm_{(s)} \to V^0 h^\pm$ decays, where V^0 is a short-lived kaon or a ϕ meson and h is a pion or a kaon, can be determined by measuring the raw signal-yield asymmetry

$$\mathcal{A}(D^\pm_{(s)} \to V^0 h^\pm) = \frac{N_{sig}(D^+_{(s)} \to V^0 h^+) - N_{sig}(D^-_{(s)} \to V^0 h^-)}{N_{sig}(D^+_{(s)} \to V^0 h^+) + N_{sig}(D^-_{(s)} \to V^0 h^-)}, \tag{3.2}$$

where N_{sig} is the number of observed signal decays. This asymmetry includes additional contributions other than $\mathcal{A}_{CP}$, such that, when considered asymmetries are small, it is possible to approximate

$$\mathcal{A}(D_s^\pm \to K_S^0\pi^\pm) \approx \mathcal{A}_{CP}^{D_s^\pm \to K_S^0\pi^\pm} + \mathcal{A}_{prod}(D_s) + \mathcal{A}_{det}(\pi) + \mathcal{A}(K^0), \tag{3.3}$$

$$\mathcal{A}(D^\pm \to K_S^0K^\pm) \approx \mathcal{A}_{CP}^{D^\pm \to K_S^0K^\pm} + \mathcal{A}_{prod}(D) + \mathcal{A}_{det}(K) + \mathcal{A}(\overline{K}^0), \tag{3.4}$$

$$\mathcal{A}(D^\pm \to \phi\pi^\pm) \approx \mathcal{A}_{CP}^{D^\pm \to \phi\pi^\pm} + \mathcal{A}_{prod}(D) + \mathcal{A}_{det}(\pi). \tag{3.5}$$

Here, $\mathcal{A}_{prod}(D_{(s)})$ is the asymmetry in the production of $D_{(s)^\pm}$ mesons in high-energy pp collisions in the forward region, and $\mathcal{A}_{det}(h)$ arises from the difference in detection

efficiencies between positively and negatively charged hadrons. The asymmetry

$$\mathcal{A}(K^0) \equiv \frac{N(K^0) - N(\overline{K}^0)}{N(K^0) + N(\overline{K}^0)} = -\mathcal{A}(\overline{K}^0), \tag{3.6}$$

where $N(K^0)$ is the number of K^0 mesons produced, takes into account the detection asymmetry between a K^0 and a $\overline{K}^0$ meson due to regeneration and the presence of mixing and *CP* violation in the $K^0 - \overline{K}^0$ system (see § 3.8.1). It is important to note that decays of $D^{\pm}_{(s)}$ mesons produce a precise flavour of $\overset{(-)}{K}{}^0$ as shown in Figs. 1.2 and 1.3. The contribution from the neutral kaon asymmetries is estimated in Ref. [66] and the value $\mathcal{A}(K^0) = (+0.0791 \pm 0.0047)\%$ is included as a correction to the measured asymmetries.

The Cabibbo-favoured decays can be used as control channels for the study of instrumental and production asymmetries, assuming negligible the *CP* violation. Their measured asymmetries can be approximated as

$$\begin{aligned}
\mathcal{A}(D^{\pm} \to K^0_S\pi^{\pm}) &\approx \mathcal{A}_{prod}(D) + \mathcal{A}_{det}(\pi) + \mathcal{A}(\overline{K}^0), &(3.7)\\
\mathcal{A}(D^{\pm}_s \to K^0_S K^{\pm}) &\approx \mathcal{A}_{prod}(D_s) + \mathcal{A}_{det}(K) + \mathcal{A}(\overline{K}^0), &(3.8)\\
\mathcal{A}(D^{\pm}_s \to \phi\pi^{\pm}) &\approx \mathcal{A}_{prod}(D_s) + \mathcal{A}_{det}(\pi). &(3.9)
\end{aligned}$$

Therefore, *CP*-violating asymmetries can be determined combining the raw asymmetries, as follows:

$$\begin{aligned}
\mathcal{A}_{CP}^{D^{\pm}_s \to K^0_S\pi^{\pm}} &\approx \mathcal{A}(D^{\pm}_s \to K^0_S\pi^{\pm}) - \mathcal{A}(D^{\pm}_s \to \phi\pi^{\pm}) - \mathcal{A}(K^0), &(3.10)\\
\mathcal{A}_{CP}^{D^{\pm} \to \phi\pi^{\pm}} &\approx \mathcal{A}(D^{\pm} \to \phi\pi^{\pm}) - \mathcal{A}(D^{\pm} \to K^0_S\pi^{\pm}) - \mathcal{A}(K^0), &(3.11)\\
\mathcal{A}_{CP}^{D^{\pm} \to K^0_S K^{\pm}} &\approx \mathcal{A}(D^{\pm} \to K^0_S K^{\pm}) - \mathcal{A}(D^{\pm} \to K^0_S\pi^{\pm}) \\
&\quad + \mathcal{A}(D^{\pm}_s \to \phi\pi^{\pm}) - \mathcal{A}(D^{\pm}_s \to K^0_S K^{\pm}) - \mathcal{A}(K^0). &(3.12)
\end{aligned}$$

Additional *CP*-violating asymmetries $\mathcal{A}'(D^{\pm}_{(s)} \to V^0 h^{\pm})$ are introduced to represents intermediate results in the analysis. They are defined only through the combination of the asymmetries measured in this analysis, *i.e.* following Eqs. 3.10-3.12 without the neutral kaon asymmetry correction.

In this note, the three *CP*-violating asymmetries are evaluated combining raw asymmetries from fits to the $D^{\pm}_{(s)} \to K^0_S\pi^{\pm}$, $D^{\pm}_{(s)} \to K^0_S K^{\pm}$ and $D^{\pm}_{(s)} \to \phi\pi^{\pm}$ invariant mass distributions.

3.4 Event selection

The analysis is performed using the complete dataset recorded by the LHCb detector during the 2015 and 2016 campaign. This corresponds to approximately 2 fb^{-1} of data: 328 pb^{-1} collected in 2015 and 1665 pb^{-1} in 2016.

Each decay channel sample comes from dedicated trigger lines, consisting in a full online selection (Turbo lines). The $D^\pm_{(s)} \to K^0_S\pi^\pm$ sample is collected using the `Hlt2CharmHadDp2KS0pip_KS0LLTurbo` Turbo line, while $D^\pm_{(s)} \to K^0_S K^\pm$ uses `Hlt2CharmHadDp2KS0Kp_KS0LLTurbo`. As regards the $D^\pm_{(s)} \to \phi\pi^\pm$ sample, it is splitted in two Turbo lines: `Hlt2CharmHadDpToKmKpPipTurbo` and `Hlt2CharmHadDspToKmKpPipTurbo` respectively for $D^\pm$ and $D_s^\pm$ channel. The cuts of these four lines are reported in Tab. 3.1 and Tab. 3.2. Variables used include: the flight distance χ^2 from the best primary vertex (BPVVDχ^2); the proper lifetime of the particle (BPVLTIME); the direction angle between the momentum and displacement vector from the PV (DIRA) and the impact parameter $\chi^2_{IP}(PV)$, which is the minimum change in χ^2 when the particle is included in the vertex fit to the PV. VZ is the decay vertex position of the K^0_S projected along the z-axis. $DLL_{K\pi}$ is the likelihood variable that describes the probability of a track is a kaon (the more positive value, more is the probability that the track is kaon). The information of this variable is based on the system of the RICH detector.

	$D^{\pm}_{(s)} \to K^0_S h^{\pm}$
$D^{\pm}_{(s)}$	
	1789 MeV/c^2 < M < 2049 MeV/c^2
	BPVVDχ^2 > 30
	BPVLTIME > 0.25 ps
	$DIRA$ < 17.3 mrad
	p_T > 2000 MeV/c
K^0_S	
	$m_{K^0_S}$ - 35 MeV/c^2 < M < $m_{K^0_S}$ + 35 MeV/c^2
	$\chi^2_{IP}(PV)$ > 9
	BPVLTIME > 0.5 ps
	-100 mm < V_Z < 500 mm
	p_T > 500 MeV/c
	V(χ^2/NDOF) < 30
bachelor hadron (π or K)	
	$DLL_{K\pi}$ < 5 (π) or > 5 (K)
	$\chi^2_{IP}(PV)$ > 36
	p_T > 200 MeV/c
K^0_S daughters (π)	
	track (χ^2/NDOF) < 3
	$\chi^2_{IP}(PV)$ > 36
	track type = *long*

Table 3.1: Requirements applied in the online selection for $D^{\pm}_{(s)} \to K^0_S\pi^{\pm}$ and $D^{\pm}_{(s)} \to K^0_S K^{\pm}$ decays.

	$D^\pm_{(s)} \to \phi\pi^\pm$
$D^\pm$	
	1789 MeV/c^2 < M < 1949 MeV/c^2
	BPVVD$\chi^2 > 150$
	BPVLTIME > 0.4 ps
	$DIRA < 10$ mrad
	Vertex $\chi^2/NDOF < 6$
	$p_T > 3000$ MeV/c
$D_s^\pm$	
	1889 MeV/c^2 < M < 2049 MeV/c^2
	BPVVD$\chi^2 > 100$
	BPVLTIME > 0.2 ps
	$DIRA < 14.1$ mrad
	Vertex $\chi^2/NDOF < 6$
	$p_T > 3000$ MeV/c
$D^\pm_{(s)}$ daughters	
	$DLL_{K\pi} < 5$ (π) or > 5 (K)
all	$\chi^2_{IP}(PV) > 4$
2 of 3	$\chi^2_{IP}(PV) > 10$
1 of 3	$\chi^2_{IP}(PV) > 50$
all	$p_T > 250$ MeV/c
2 of 3	$p_T > 400$ MeV/c
1 of 3	$p_T > 1000$ MeV/c

Table 3.2: Requirements applied in the online selection for $D^\pm_{(s)} \to \phi\pi^\pm$ decays.

3.4.1 Trigger requirements

Before applying the offline selection, the $D^{\pm}_{(s)} \to K^0_S h^{\pm}$ and $D^{\pm}_{(s)} \to \phi\pi^{\pm}$ events are required to satisfy the following L0 and HLT requirements:

L0 trigger: At least one $D^{\pm}_{(s)}$ daughter (Trigger On Signal TOS) must fire the the hadronic calorimeter or any tracks, not belonging to the selected signal candidate (Trigger Independent of Signal TIS) must fire the trigger system.

HLT1 trigger: One or two $D^{\pm}_{(s)}$ daughter (TOS) must satisfy the criteria of a selection based on a multivariate analysis.

These requirements are necessary to minimise any charge asymmetry biases in the trigger due to the alignment and homogeneity of the calorimeter (for L0) and the tracking system (for HLT1).

3.4.2 Cut-based selection

The events filtered by the trigger requirements are further processed by an offline selection, designed to reduce specific backgrounds, whilst retaining signal events as efficiently as possible. The offline selection is optimized comparing the signal and background high-purity distributions. Summaries of the selection criteria for the $D^{\pm}_{(s)} \to K^0_S h^{\pm}$ and $D^{\pm}_{(s)} \to \phi\pi^{\pm}$ decay channels are given in Tab. 3.3 and Tab. 3.4.

After this selection, the remaining background is considered to be mainly composed by candidates where a correctly reconstructed K^0_S or ϕ (not necessary from a $D^{\pm}_{(s)}$ decays) is accidentally paired to a random charged hadron from the primary vertex.

The $D^{\pm}_{(s)} \to K^0_S h^{\pm}$ signal channel

Signal selection

In the case of the $D^{\pm}_{(s)} \to K^0_S h^{\pm}$ selection, $D^{\pm}_{(s)}$ mesons are reconstructed from a $K^0_S \to \pi^+\pi^-$ decay candidate combined with a bachelor charged hadron. All the tracks are required to be categorized as *long* tracks and the K^0_S mass is constrained to its known value when the $D^{\pm}_{(s)}$ vertex is formed. For this reason the invariant mass of the $\pi^+\pi^-$ pair is required to be within three standard deviations of the invariant mass resolution model from the nominal value of the K^0_S mass.

Background rejection

Several criteria are applied to reduce specific decay backgrounds. The backgrounds considered, and the requirements used to reduce them, are outlined below.

$D^\pm_{(s)} \to h^\pm h^\mp h^\pm$: Background events of this type are rejected by taking advantage by the fact that the K^0_S meson flights a significant distance before decaying. A cut on the flight-distance with respect to the $D^\pm_{(s)}$ decay vertex is applied to the K^0_S candidate. The $\pi^+\pi^-$ mass distribution is show in Fig. 3.1 (a) where the distribution of the rejected candidates is in red while in Fig. 3.1 (b) the same distribution is reported after this selection.

$D^\pm \to K^0_S h^\pm$: It arises from the misidentification of the bachelor hadron, when the mass of the kaon (or pion) is wrongly assigned to the pion (or kaon). The dependence of the $K^0_S h^\pm$ invariant mass as a function of the momentum imbalance, defined as $(p_{K^0_S} - p_{h^\pm})/(p_{K^0_S} + p_{h^\pm})$, when the bachelor is misidentified is shown in Figs. 3.2 and 3.3 for the decay to a pion and a kaon, respectively. Superimposed on the figure the analytical expressions as calculated in Appendix A (Eqs. A.6 and A.7) are reported with a green-dashed line. As regards the $D^\pm_{(s)} \to K^0_S\pi^\pm$ signal decay, this background lies in the low mass region with respect to the $D^\pm_{(s)}$ peaks and can be neglected, therefore a simple PID cut for the pion identity is applied. On the contrary, in the case of the $D^\pm_{(s)} \to K^0_S K^\pm$ signal decay, it is situated under the $D^\pm_s$ signal region and a hard cut on $DLL_{K\pi}$ is needed.

$\Lambda_c^\pm \to K^0_S \overset{(-)}{p}$: In this case, the proton from the Λ_c decay is misidentified as a π or a K. This background is visible in Figs. 3.2 and 3.3 for the decay to a pion and a kaon, respectively. Superimposed on the figure the analytical expressions given by Eqs. A.8 and A.9 are reported with a red-dashed line. A cut on DLL_{pK} or $DLL_{p\pi}$ is applied to the bachelor hadron.

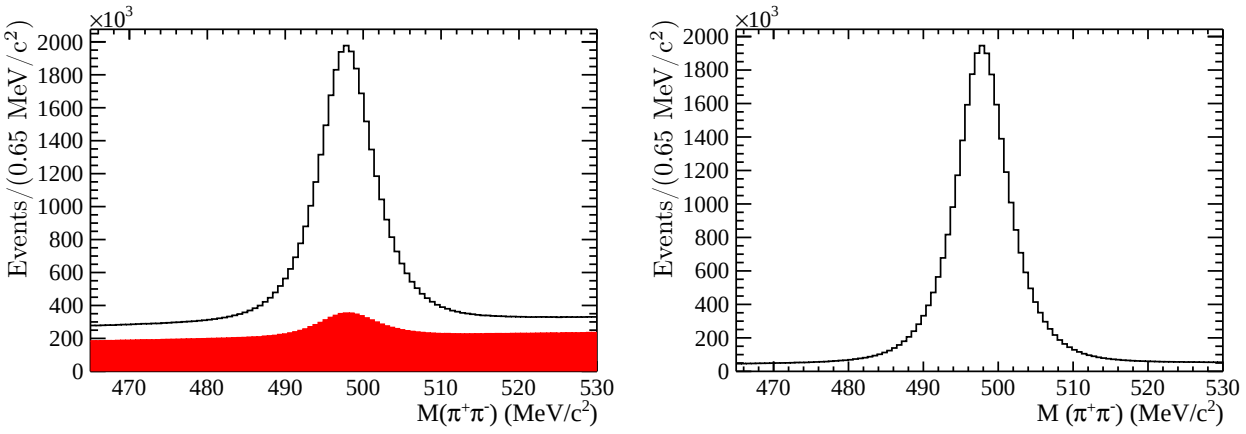
$\overset{(-)}{\Lambda^0} \to \overset{(-)}{p}\pi^\mp$: This background arises when the proton from the Λ^0 decay is misidentified as a K^0_S daughter. One can notice this feature of the background comparing the data in Fig 3.4 with the superimposed blue-dashed line for the analytical expression given by Eq. A.11. When the Λ^0 is produced in a $\Lambda_c^\pm \to \overset{(-)}{\Lambda^0}\pi^\pm$ decay, where the bachelor pion is correctly assigned, the background is analytically described by Eq. A.10 and can be found along the blue-dashed line superimposed on Fig. 3.2. The number of these undesired events can be reduced by the application of a $DLL_{p\pi}$ cut to both the K^0_S daughter pions, but this is not applied since the three standard deviation of the invariant mass resolution model cut on K^0_S mass is already efficient.

Non-prompt charm: Cuts are applied to ensure the impact parameter χ^2 of the $D^\pm_{(s)}$ candidate with respect to the primary vertex is small. However, a non-negligible fraction of non-prompt charm decays persists in the data. The $D^\pm_{(s)}$ mesons resulting from non-prompt decays can have significantly different production asymmetries that can result in a false *CP* violation signal. These backgrounds are discussed further in § 3.8.2

In Figs. 3.2 (b) and 3.3 (b), one can notice that, after the selection, residual misidentified decays are still present. This arises from the necessity that signal efficiency has not to be compromised by PID cuts. This backgrounds cross all the mass region and are not peaking, thus can be considered not dangerous for the measurement. In the invariant mass spectrum they behave as the combinatorial background.

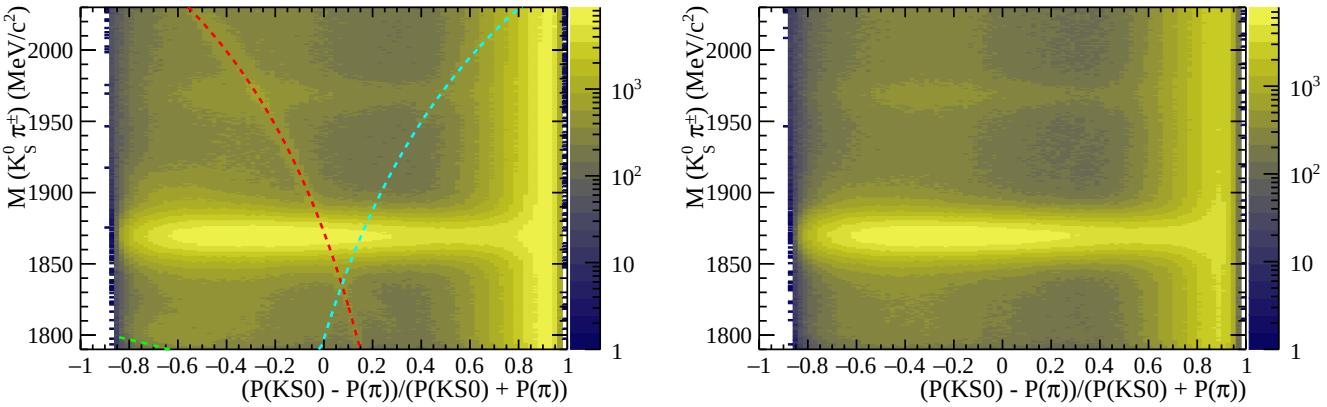
There is also an irreducible low-mass background from decays such as $D_s^{\pm} \to K_S^0 h^{\pm}\pi^0$ where the π^0 is not reconstructed. This background lies in a region well below the nominal value of the $D^{\pm}$ invariant mass and is take into account in the fit for $D_{(s)}^{\pm} \to K_S^0 K^{\pm}$ decays as described in § 3.6.

The figures 3.5 and 3.6 show the $K_S^0 h$ invariant mass distribution of the $D_{(s)}^{\pm}$ candidates that pass the selection described above. Looking at the lower-statistics channel, the $D_{(s)}^{\pm} \to K_S^0\pi^{\pm}$ decay, the combinatorial background under the two Cabibbo-suppressed signal peak is still large and need to be reduced in order to produce a competitive measurement on $\mathcal{A}_{CP}$. Therefore, we decided to apply a multivariate selection. As concerns the $D_{(s)}^{\pm} \to K_S^0 K^{\pm}$ decay, the combinatorial background is less dominant, but can be reduced with the same approach. Multivariate analysis is widely described in the next section.



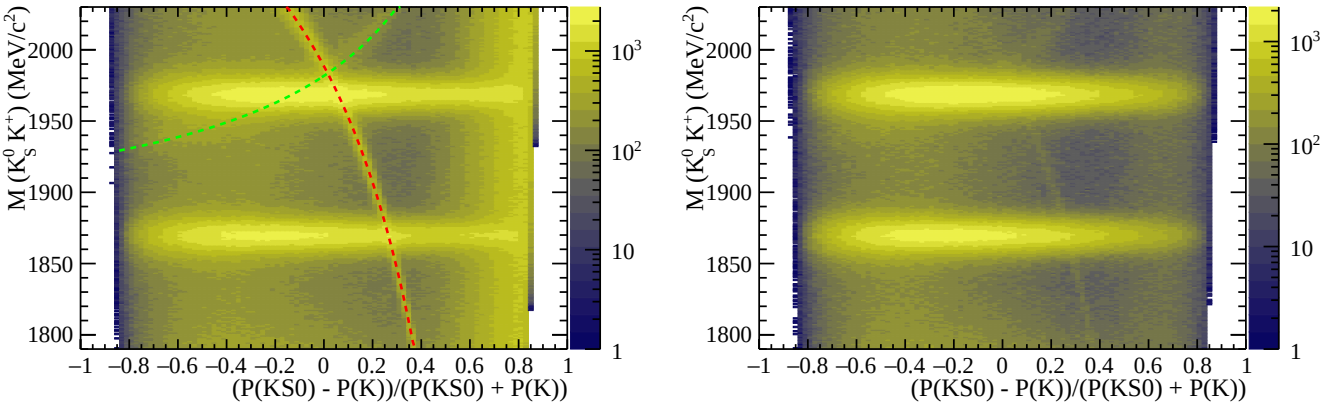
(a) (b)

Figure 3.1: Invariant mass distribution of reconstructed K_S^0 candidates before (a) and after (b) the K_S^0 flight-distance cut (b) for $D_{(s)}^{\pm} \to K_S^0 h^{\pm}$ decays. In (a) the distribution in red represents the candidates rejected from the K_S^0 flight distance selection. The MagDown 2016 sample $D_{(s)}^{\pm} \to K_S^0 K^{\pm}$ and $D_{(s)}^{\pm} \to K_S^0\pi^{\pm}$ are reported as an example in (a) and (b), respectively. However similar results can be observed for different decay mode, year and magnet polarity configuration.



(a) (b)

Figure 3.2: $K_S^0\pi^{\pm}$ invariant mass distribution as a function of the momentum imbalance of the daughters before (a) and after (b) the PID selection. The horizontal regions with high population represent the $D_{(s)^{\pm}}$ peaks. The green-dashed line represents the $D^{\pm}_{(s)} \to K_S^0 K^{\pm}$ decay where the kaon is wrongly identified as a pion. The red-dashed line represents the $\Lambda_c^{\pm} \to K_S^0 \overset{(-)}{p}$ decay where the proton is wrongly identified as a pion. The blue-dashed line represents the $\Lambda_c^{\pm} \to \overset{(-)}{\Lambda^0}\pi^{\pm}$ where $\overset{(-)}{\Lambda^0} \to \overset{(-)}{p}\pi^{\mp}$ decay is wrongly identified as a K_S^0 decay. These lines are calculated in the hypothesis of perfect invariant mass resolution model. The MagDown 2016 sample is here reported. Similar distributions can be observed for different year and magnet polarity configuration.



(a) (b)

Figure 3.3: $K^0_S K^\pm$ invariant mass distribution as a function of the momentum imbalance of the daughters before (a) and after (b) the PID selection. The horizontal regions with high population represent the $D_{(s)^\pm}$ peaks. The green-dashed line represents the $D^\pm_{(s)} \to K^0_S \pi^\pm$ decay where the pion is wrongly identified as a kaon. The red-dashed line represents the $\Lambda^\pm_c \to K^0_S \overset{(-)}{p}$ decay where the proton is wrongly identified as a kaon. These lines are calculated in the hypothesis of perfect invariant mass resolution model. The MagDown 2016 sample is here reported. Similar distributions can be observed for different year and magnet polarity configuration.

	$D^\pm_{(s)} \to K^0_S \pi^\pm$
$D^\pm_{(s)}$	$\chi^2_{IP}(PV) < 9$
K^0_S	497.614 - 15 MeV/c^2 < M < 497.614 + 15 MeV/c^2 flight distance > 20 mm
bachelor hadron (π)	$DLL_{K\pi} < 0$ $DLL_{p\pi} < 0$
bachelor hadron (K)	$DLL_{K\pi} > 8$ $DLL_{p\pi}$ - $DLL_{K\pi} < 0$
K^0_S daughters (π)	$DLL_{K\pi} < 10$

Table 3.3: Preliminary selection for $D^\pm_{(s)} \to K^0_S h^\pm$ candidates.

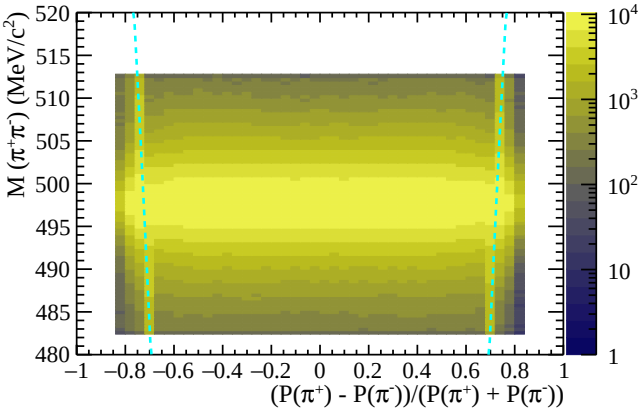
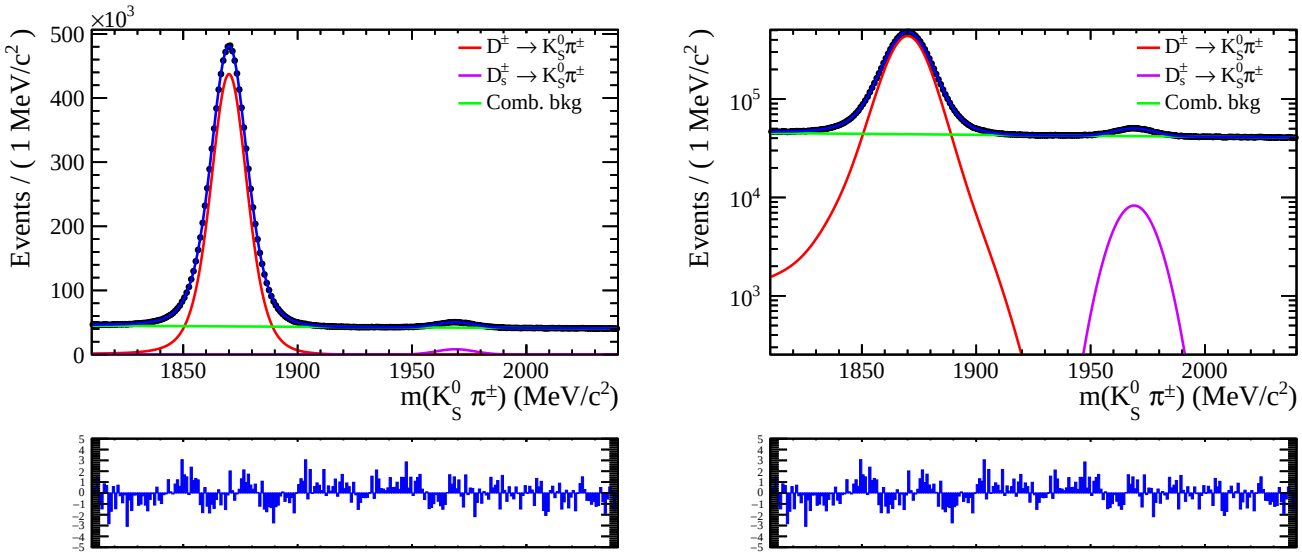


Figure 3.4: K_S^0 invariant mass distribution as a function of the momentum imbalace of the daughters after the offline selection. The blue-dashed lines represent the $\overline{\Lambda}^0 \to \overline{p}\pi^{\mp}$ decay where the proton is wrongly identified as a pion. These lines are calculated in the hypothesis of perfect invariant mass resolution model. The MagDown 2016 sample is here reported. Similar distributions can be observed in the $D^{\pm}_{(s)} \to K_S^0\pi^{\pm}$ channel as well as in a different year and magnet polarity configuration.



(a) (b)

Figure 3.5: $K_S^0\pi$ invariant mass distribution (a) and in log scale (b) with the result of the fit overlaid. At the bottom the residual distribution is reported. The MagDown 2016 sample is shown as an example. Similar results can be observed for different year and magnet polarity configuration.

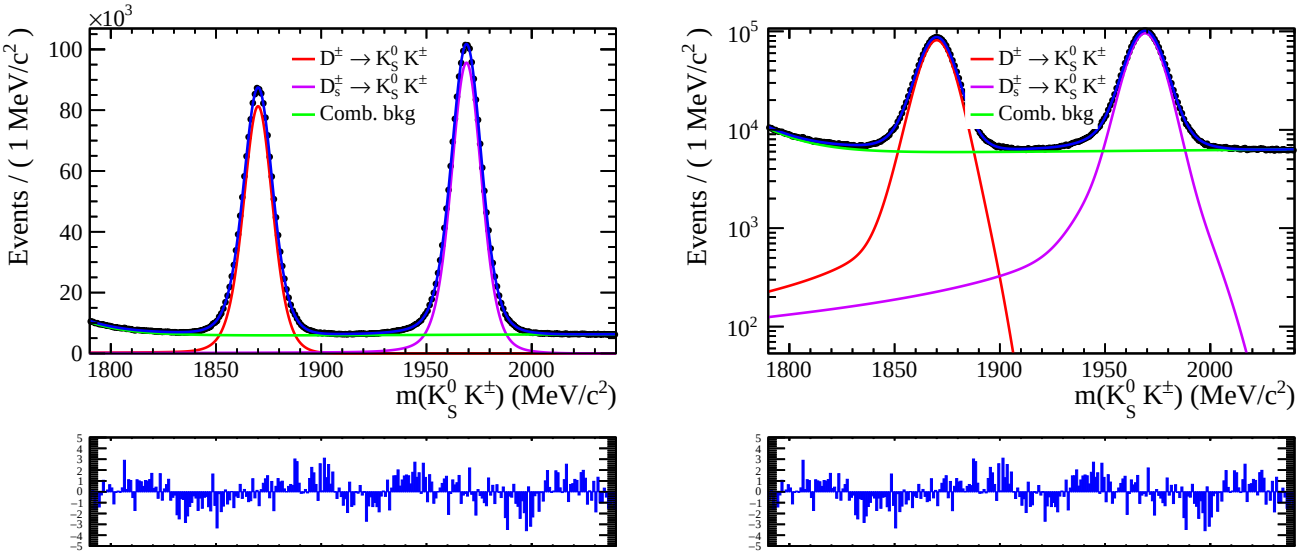


Figure 3.6: $K^0_S K$ invariant mass distribution (a) and in log scale (b) with the result of the fit overlaid. At the bottom the residual distribution is reported. The MagDown 2016 sample is shown as an example. Similar results can be observed for different year and magnet polarity configuration.

The $D^\pm_{(s)} \to \phi\pi^\pm$ signal channel

Signal selection

As already mentioned, the $D^\pm_{(s)} \to \phi\pi^\pm$ sample comes from different online selections for $D^\pm$ and $D_s^\pm$ described in Tab. 3.2. Therefore, it is necessary to fit the two samples jointly, introducing a step function at 1900 MeV/c^2 in the fit model as described in § 3.6.3 to take into account the different selection efficiencies in different invariant mass windows.

$D^\pm_{(s)}$ candidates are reconstructed from three charged tracks originating from a single vertex, since the ϕ does not travel an appreciable distance through the detector like the K_S^0. Therefore, the invariant mass of the kaons pair from the ϕ is required to be within three standard deviations of the invariant mass resolution model from the ϕ invariant mass nominal value. The K^+K^- invariant mass distribution before this selection is shown in Fig. 3.7 (a).

Background rejection

Several criteria are applied to reduce specific decay backgrounds. The backgrounds considered, and the requirements used to reduce them, are outlined below.

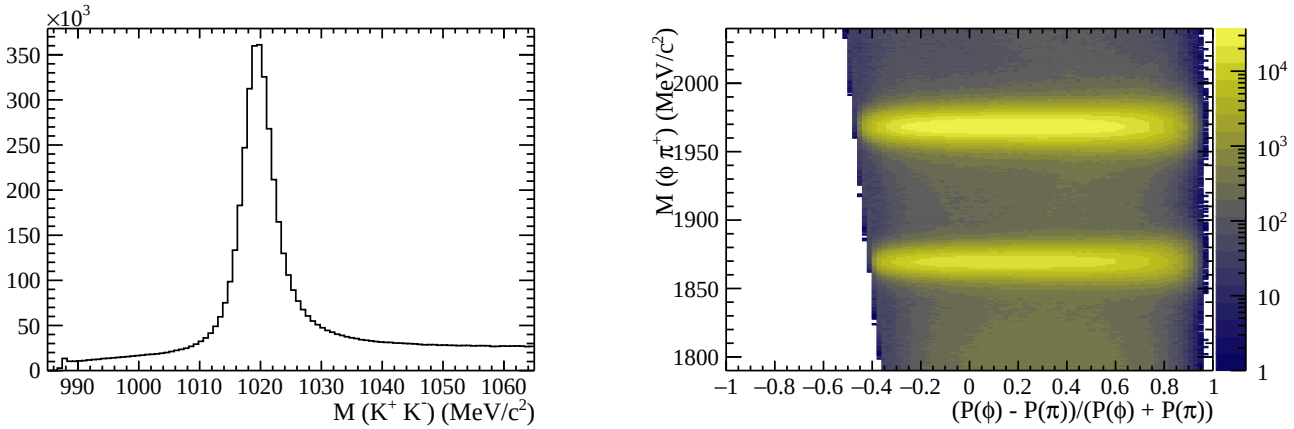
$\mathbf{D^\pm_{(s)} \to h^\pm h^\mp h^\pm}$: These background arises when an incorrect pion or kaon mass hypothesis is assigned. Decays such as $D^\pm_{(s)} \to K^\pm\pi^\mp\pi^\pm$ and $D^\pm_{(s)} \to \pi^\pm\pi^\mp\pi^\pm$ are already reduced at HLT2 with a hard cut in $DLL_{K\pi}$ to the ϕ daughters as described in Tab. 3.2. Besides, the three standard deviation of the invariant mass resolution model cut on ϕ mass is already efficient for the suppression of this backgrounds. As regards the $D^\pm_{(s)} \to \phi K^\pm$ decays, they are also rejected with a simple PID requirement for the pion.

Non-prompt charm: Cuts are applied to ensure that the impact parameter χ^2 of the $D^\pm_{(s)}$ candidate with respect to the primary vertex is small. These backgrounds are discussed further in § 3.8.2

A further irreducible background is taken into account in the fit model. It consists of the $D^\pm_{(s)} \to K^\pm K^\mp \pi^\pm \pi^0$ decays where the π^0 is not reconstructed. This background lies in a region well below the nominal value of the $D^\pm$ invariant mass as described in § 3.6.3.

Figure 3.7 (b) shows the $\phi\pi$-mass distribution as a function of the momentum imbalance between the ϕ and the π momentum after the selection. Here, misidentified decays are not visible thus the remaining background is considered as only combinatorial. Furthermore, one can notice that the phase-space is limited (*i.e.* $(p_\phi - p_\pi)/(p_\phi + p_\pi)$ covers only a fraction of the interval in [-1,1]). This is due to the small Q-value of the decay, of about 32 MeV/c^2.

The $\phi\pi$-mass distribution of the selected candidates is reported in Fig. 3.8. As one can see, this is a very clean sample and candidates do not have to meet any additional condition. Final results are presented in § 3.7.



(a) (b)

Figure 3.7: Invariant mass distribution of reconstructed ϕ candidates before the offline selection (a) and $\phi\pi$-mass invariant mass distribution as a function of the momentum imbalance after the offline selection (b). In (b), the horizontal regions with high population represent the $D_{(s)^\pm}$ peaks.

	$D^\pm_{(s)} \to \phi\pi^\pm$
$D^\pm_{(s)}$	
	$\chi^2_{IP}(PV) < 9$
ϕ	
	1019.461 - 10 MeV/c^2 < M < 1019.461 + 10 MeV/c^2
bachelor hadron (π)	
	$DLL_{K\pi} < 0$

Table 3.4: Offline selection for $D^\pm_{(s)} \to \phi\pi^\pm$ candidates.

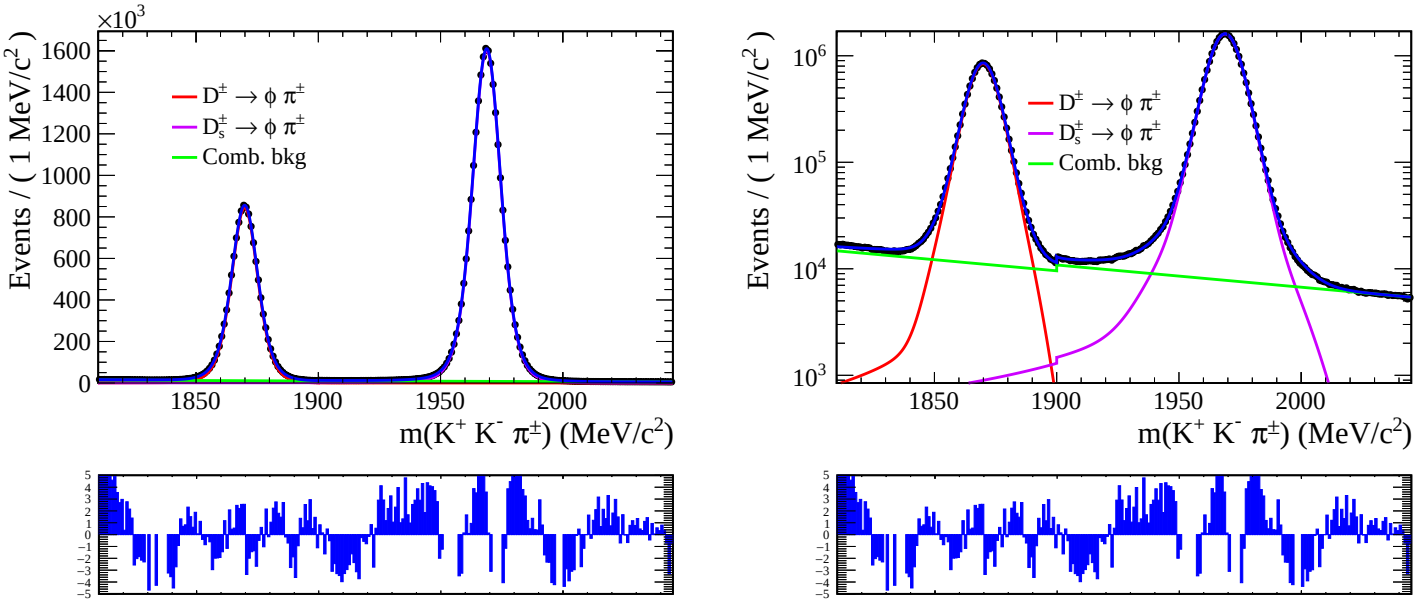


Figure 3.8: $\phi\pi$ invariant mass distribution (a) and in log scale (b) with the result of the fit overlaid. At the bottom the residual distribution is reported. The MagDown 2016 sample is shown as an example. Similar results can be observed for different year and magnet polarity configuration.

3.5 Multivariate analysis for $D^{\pm}_{(s)} \to K^0_S h^{\pm}$ decays

In $D^{\pm}_{(s)} \to K^0_S h^{\pm}$ decays, the combinatorial background under the Cabibbo-suppressed peaks can be further reduced with the application of a neutral network discriminator (later referred to as MVA). More in detail, the multivariate selection is implemented through the ROOT internal MLP class (*i.e.* TMlpANN) in the TMVA framework [67].

The MVA is trained with signal and background samples derived directly from data. In Appendix B a brief introduction to the techniques for data-driven multivariate analysis can be found. As regards the $D^{\pm}_s \to K^0_S\pi^{\pm}$ selection, the training is performed using the high-statistic $D^{\pm}$ decay as signal. This choice comes from the not optimal choice of training the MVA with the $D^{\pm}_s$ signal because of its very low statistic and by the fact that $D^{\pm}$ and $D^{\pm}_s$ decays have very similar kinematics. The main difference between the two signals consists in the decay time (the $D^{\pm}_s$ meson lifetime is about the half of that of the $D^{\pm}$). For this reason, training variables are required not to be correlated with the meson decay time. In the case of $D^{\pm} \to K^0_S K^{\pm}$ decays, instead, the training is realized on the signal itself.

For both the channels, the signal and background distributions are calculated applying *sWeights* [68] from fits to $K^0_S\pi$- and $K^0_S K$-mass distributions in the 1830 - 2040 MeV/c^2 range of the invariant mass. The results of these fits are given in Figs. 3.5 and 3.6. The MVA is trained separately per year of data taking on 40 000 signal and 40 000 background events.

A long list of input variables have been studied trying to identify those able to better discriminate signal and background, but also with a low correlation with the invariant mass[1] and $D^{\pm}_{(s)}$ mean lifetime. The final set of variables chosen is listed together with their separation power[2] in Tab. 3.5 and Tab. 3.6 for $D^{\pm}_{(s)} \to K^0_S\pi^{\pm}$ and $D^{\pm}_{(s)} \to K^0_S K^{\pm}$ decays, respectively. Input variables are related to the kinematic of the decay and to the decay topology. Their signal and background distributions together with their linear correlation matrices are shown in Figs. 3.9 and 3.10 (Figs. 3.11 and 3.12) as concerns the $D^{\pm}_{(s)} \to K^0_S\pi^{\pm}$ ($D^{\pm}_{(s)} \to K^0_S K^{\pm}$) decay channel.

The resulting MVA distributions for signal and background are shown in Figs. 3.13 and 3.14 as regards the decays with a pion and a kaon in the final state, respectively.

[1] It is a very standard requirement in multivariate analysis. In this context, this condition is even stronger since for the application of *sPlot* technique it is mandatory to have no correlation with the fitting variable as described in Appendix B.

[2] The discrimination between signal and background distributions of a variable y is estimated computing the statistical separation S^2, defined as

$$\langle S^2(y) \rangle = \frac{1}{2} \int \frac{[\mathrm{pdf}_S(y) - \mathrm{pdf}_B(y)]^2}{\mathrm{pdf}_S(y) + \mathrm{pdf}_B(y)} dy$$

where $\mathrm{pdf}_{S,B}$ are the signal and background probability density functions.

These figures also show that the algorithm satisfy the over-training checks: the distributions produced in the training are compatible with those obtained with the application of the MVA in an independent test sample.

To ensure no bias in the background mass shape after MVA selection, a check on the correlation of the MVA with the K_S^0h-mass has been performed. As shown in Fig. 3.15, the correlation is small.

The optimal MVA requirement is chosen in order to maximaze the figure of merit defined as $S/\sqrt{S+B}$ where the number of signal (S) and background (B) events are chosen to lie in the invariant mass window corresponding to [1930, 2010] ([1830, 1910]) MeV/c^2 for the $D_s^\pm \to K_S^0\pi^\pm$ ($D_s^\pm \to K_S^0K^\pm$) decay. Therefore, many MVA selections are tested. For each, the number of $D^\pm$, $D_s^\pm$ and background candidates is measured by fitting the obtained K_S^0h-mass distributions (details on the fit are presented in § 3.6). Figures 3.16 and 3.17 show the figure of merit for the Cabibbo-suppressed decays as a function of the MVA requirement. In the $D_s^\pm \to K_S^0\pi^\pm$ case, a maximum is visible for TMlpANN > 0.75 (TMlpANN > 0.81) for the 2016 (2015) sample, corresponding to an improvement of about two times with respect to the figure of merit where no MVA cut is applied. As regards the $D_s^\pm \to K_S^0K^\pm$ sample the gain is lower and a maximum in significance is found at TMlpANN > 0.19 (TMlpANN > 0.27) for the 2016 (2015) data sample.

The resulting $K_S^0\pi$- and K_S^0K-mass distributions are shown in § 3.7, together with results.

	Separation	
Variable	2015	2016
bachelor π p_T	33%	36%
$D_{(s)}$ cos(DIRA)	29%	28%
K_S^0 p	17%	20%
$D_{(s)}$ p_T	16%	16%
$D_{(s)}$ vertex χ^2	11%	11%
bachelor π Track $\chi^2/NDOF$	10%	10%

Table 3.5: Variables used as input for the MVA training, ordered by their separation power, for $D_{(s)}^\pm \to K_S^0\pi^\pm$ decay sample. The bachelor transverse momentum p_T, the cosine of the direction angle between the momentum and displacement vector from the PV (DIRA) of $D_{(s)}^\pm$ meson, the neutral kaon momentum, the $D_{(s)}^\pm$ transverse momentum, the χ^2 of the reconstructed $D_{(s)}^\pm$ decay vertex and finally the χ^2 per degrees of freedom of the reconstructed track of the bachelor hadron.

	Separation	
Variable	2015	2016
$D_{(s)}$ cos(DIRA)	39%	36%
$D_{(s)}$ p_T	18%	15%
bachelor K Track $\chi^2/NDOF$	9%	8%
bachelor K p_T	8%	7%
$D_{(s)}$ vertex χ^2	5%	6%
K^0_S p	4%	5%

Table 3.6: Variables used as input for the MVA training, ordered by their separation power, for $D^{\pm}_{(s)} \to K^0_S K^{\pm}$ decay sample. The cosine of the direction angle between the momentum and displacement vector from the PV (DIRA) of $D^{\pm}_{(s)}$ meson, the $D^{\pm}_{(s)}$ transverse momentum, the χ^2 per degrees of freedom of the reconstructed track of the bachelor hadron together with its transverse momentum p_T, the χ^2 of the reconstructed $D_{(s)}$ decay vertex and finally the neutral kaon momentum.

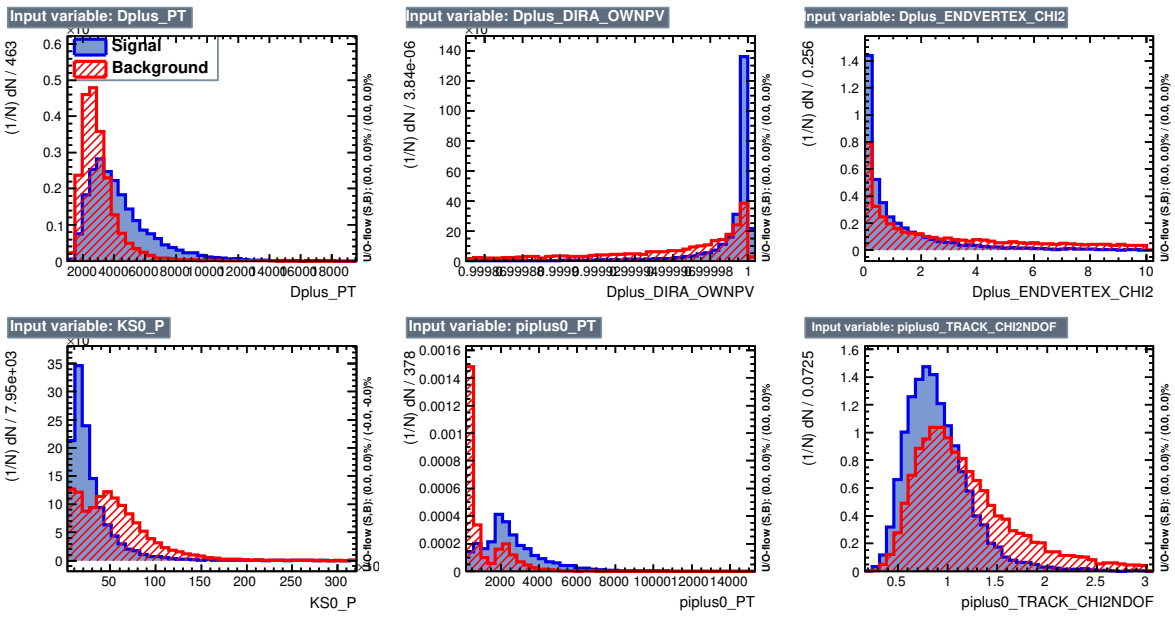
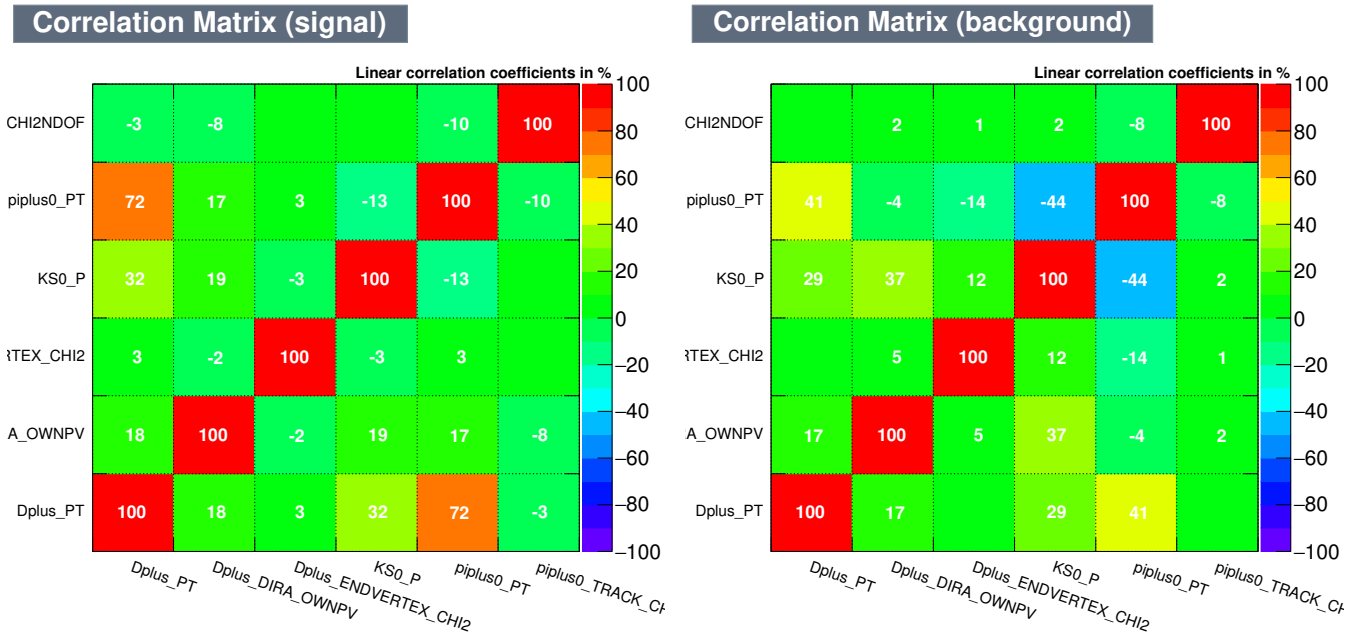


Figure 3.9: Signal and background distributions of the input variables chosen for the MVA selection in the $D^{\pm}_{(s)} \to K^0_S\pi^{\pm}$ sample. From the left to the right and from the top to the bottom, the variables are the $D^{\pm}_{(s)}$ transverse momentum, the cosine of the direction angle between the momentum and displacement vector from the PV (DIRA) of $D^{\pm}_{(s)}$ meson, the χ^2 of the reconstructed $D^{\pm}_{(s)}$ decay vertex, the neutral kaon momentum, the bachelor transverse momentum p_T, and finally the χ^2 per degrees of freedom of the reconstructed track of the bachelor hadron. Bins with negative value can be found, this feature is a consequence of the *sPlot* technique when negative weights are not locally offset by positive ones, but only globally. The 2016 sample is here reported, however similar distributions can be observed in 2015 dataset.



(a) (b)

Figure 3.10: Signal (a) and background (b) linear correlation matrix of the input variables chosen for the MVA selection in the $D^{\pm}_{(s)} \to K^0_S \pi^{\pm}$ sample. From the left/bottom to the right/top, the variables are the $D^{\pm}_{(s)}$ transverse momentum, the cosine of the direction angle between the momentum and displacement vector from the PV (DIRA) of $D^{\pm}_{(s)}$ meson, the χ^2 of the reconstructed $D^{\pm}_{(s)}$ decay vertex, the neutral kaon momentum, the bachelor transverse momentum p_T, and finally the χ^2 per degrees of freedom of the reconstructed track of the bachelor hadron. The 2016 sample is here reported, however similar distributions can be observed in 2015 dataset.

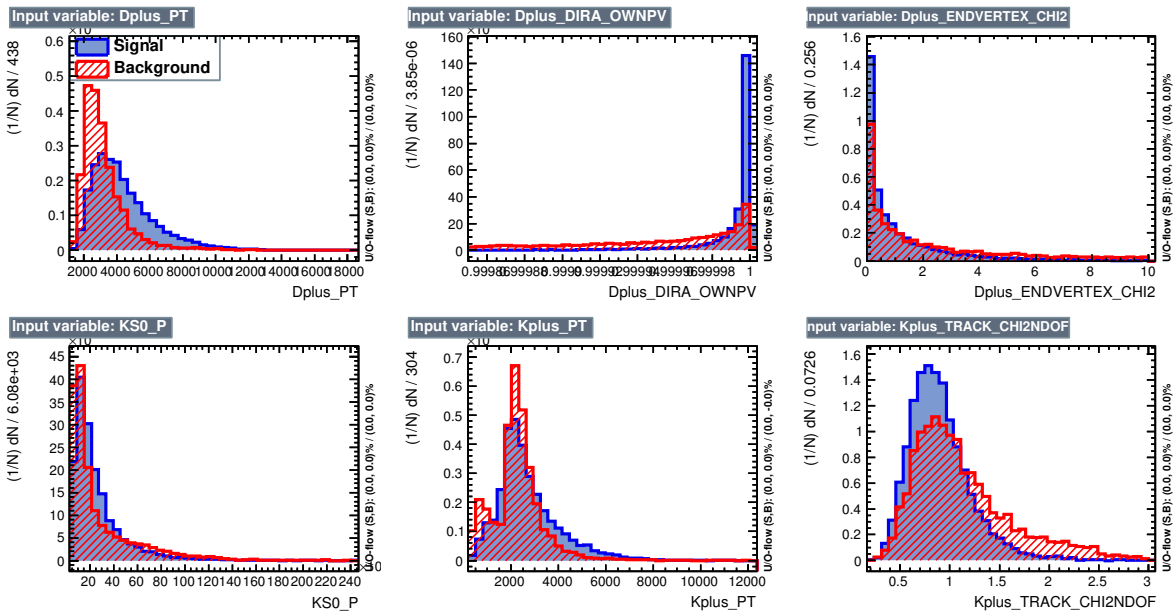
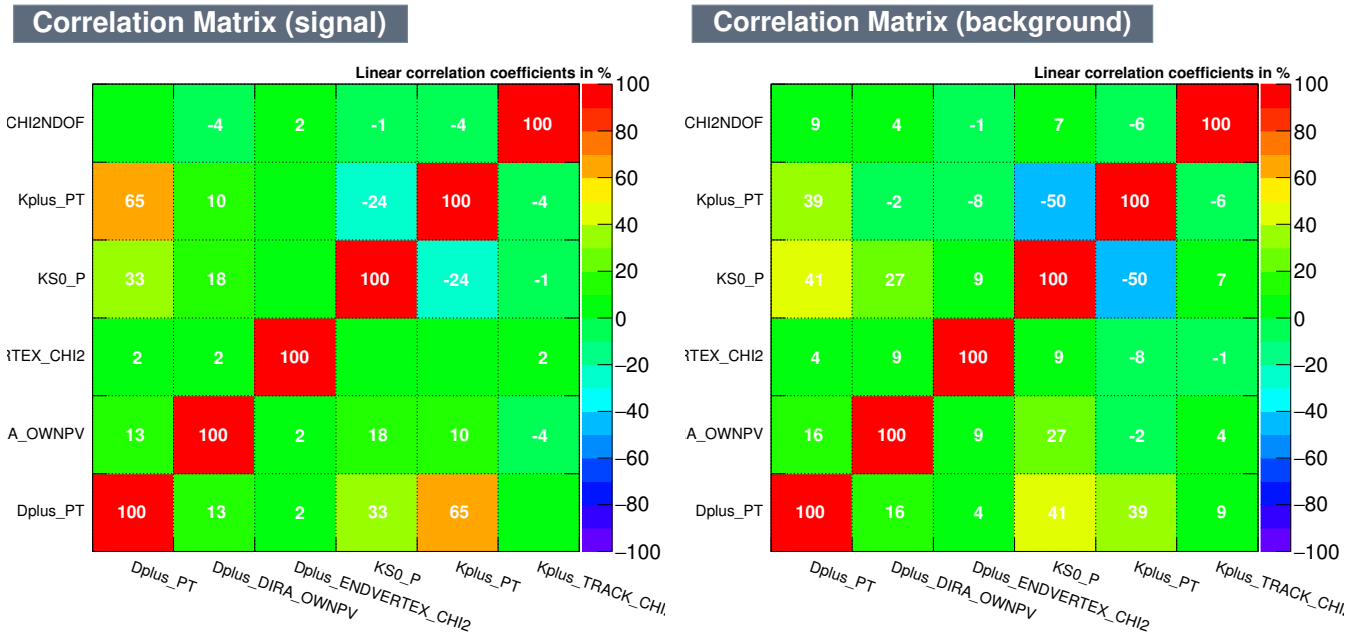
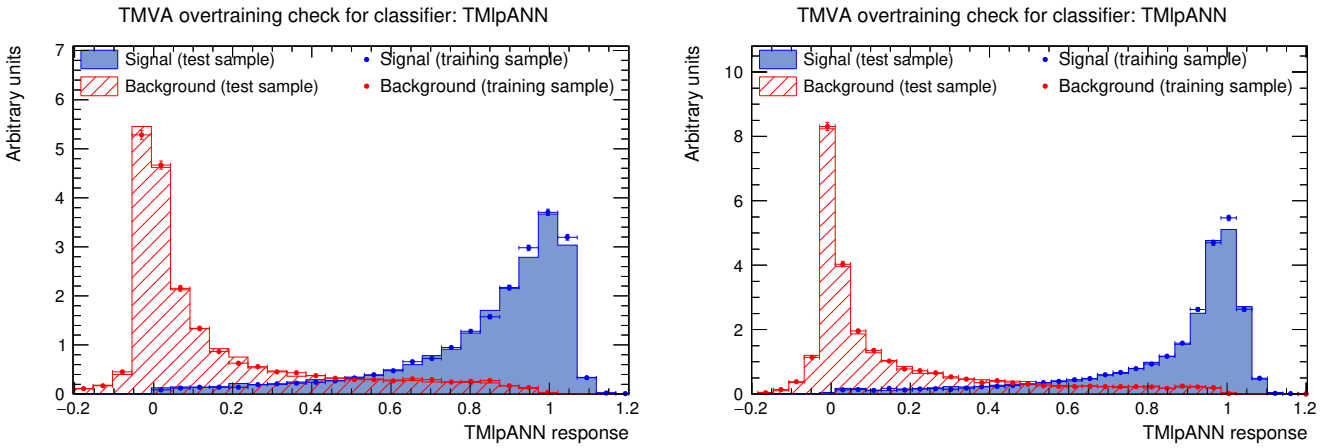


Figure 3.11: Signal and background distributions of the input variables chosen for the MVA selection in the $D^{\pm}_{(s)} \to K^0_S K^{\pm}$ sample. From the left to the right and from the top to the bottom, the variables are the $D^{\pm}_{(s)}$ transverse momentum, the cosine of the direction angle between the momentum and displacement vector from the PV (DIRA) of $D^{\pm}_{(s)}$ meson, the χ^2 of the reconstructed $D^{\pm}_{(s)}$ decay vertex, the neutral kaon momentum, the bachelor transverse momentum p_T, and finally the χ^2 per degrees of freedom of the reconstructed track of the bachelor hadron. Bins with negative value can be found, this feature is a consequence of the *sPlot* technique when negative weights are not locally offset by positive ones, but only globally. The 2016 sample is here reported, however similar distributions can be observed in 2015 dataset.



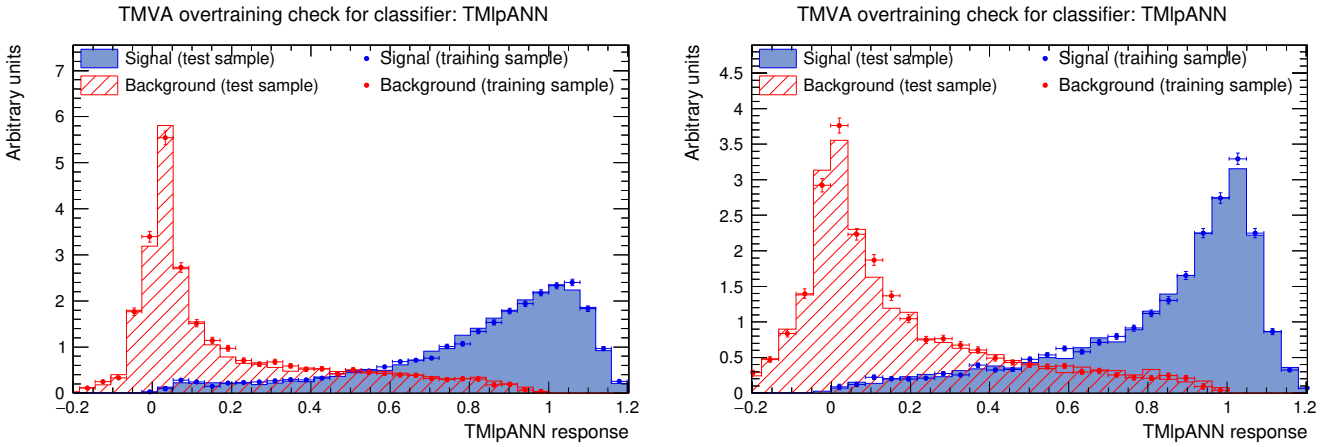
(a) (b)

Figure 3.12: Signal (a) and background (b) linear correlation matrix of the input variables chosen for the MVA selection in the $D^{\pm}_{(s)} \to K^0_S K^{\pm}$ sample. From the left/bottom to the right/top, the variables are the $D^{\pm}_{(s)}$ transverse momentum, the cosine of the direction angle between the momentum and displacement vector from the PV (DIRA) of $D^{\pm}_{(s)}$ meson, the χ^2 of the reconstructed $D^{\pm}_{(s)}$ decay vertex, the neutral kaon momentum, the bachelor transverse momentum p_T, and finally the χ^2 per degrees of freedom of the reconstructed track of the bachelor hadron. The 2016 sample is here reported, however similar distributions can be observed in 2015 dataset.



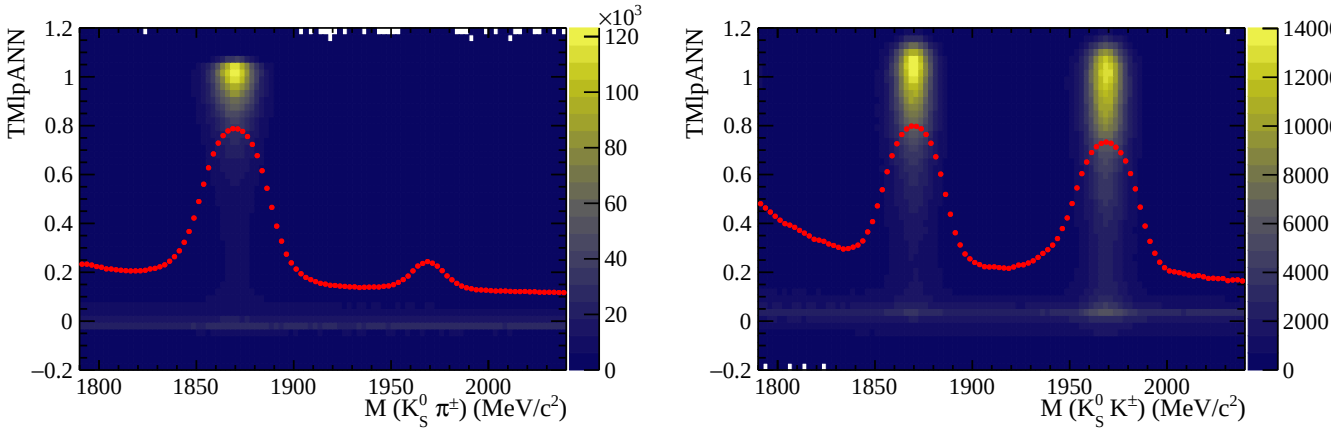
(a) (b)

Figure 3.13: MVA distributions for signal and background as results of the training (dots) and from an independent test sample (line) for $D_{(s)}^{\pm} \to K_S^0\pi^{\pm}$ decays in 2016 (a) and 2015 (b) sample.



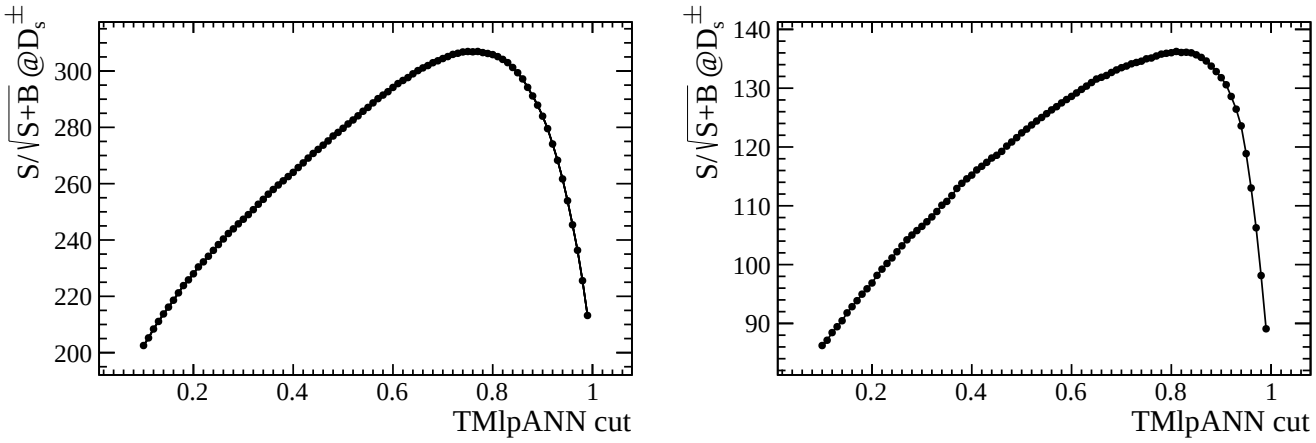
(a) (b)

Figure 3.14: MVA distributions for signal and backgrund as results of the training (dots) and from an independent test sample (line) for $D_{(s)}^{\pm} \to K_S^0 K^{\pm}$ decays in 2016 (a) and 2015 (b) sample.



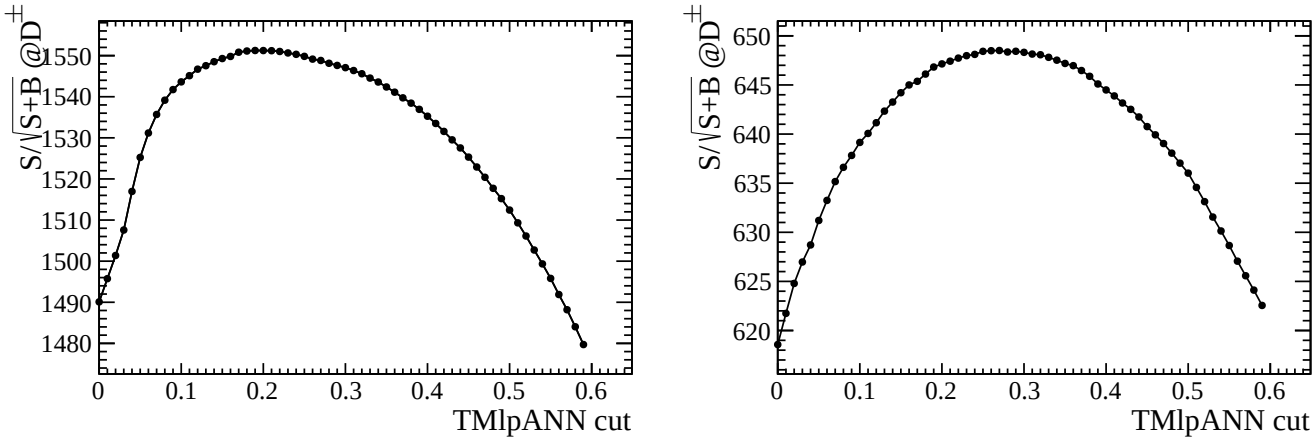
(a) (b)

Figure 3.15: MVA distribution and MVA average value ($< MVA >$) (red line) as a function of $M(K^0_S h)$ for $D^{\pm}_{(s)} \to K^0_S \pi^{\pm}$ (a) and $D^{\pm}_{(s)} \to K^0_S K^{\pm}$ (b) decays. The 2016 sample is here reported, however similar distributions can be observed in 2015 dataset.



(a) (b)

Figure 3.16: Significance of $D^{\pm}_s$ peak as a function of the MVA cut for 2016 (a) and 2015 (b) data sample.



(a) (b)

Figure 3.17: Significance of $D^{\pm}$ peak as a function of the MVA cut for 2016 (a) and 2015 (b) data sample.

3.6 Fit method

The raw *CP* asymmetries for the $D^{\pm}_{(s)} \to K^0_S\pi^{\pm}$, $D^{\pm}_{(s)} \to K^0_S K^{\pm}$ and $D^{\pm}_{(s)} \to \phi\pi^{\pm}$ decays are determined from three separate binned extended χ^2 fits to three respective invariant mass distributions. All fits are performed using RooFit in the 1790 - 2045 MeV/c^2 invariant mass region with a bin every 1 MeV/c^2.

A simultaneous fit is performed to the two $V^0h^{\pm}$ invariant mass distributions. The positive and negative charge invariant mass distributions are fitted using the same fit model but with independent shape parameters (*i.e.* shape parameters are doubled). This feature permits to readily absorb eventual systematic uncertainties coming from detection charge asymmetries. For the same reason, down and up magnet polarity samples are fitted together. The datasets are fitted separately for data taking year since small differences in LHC and LHCb running conditions and trigger configurations could have a small affect on fitting line-shapes. The main parameters of the fits are the two total $D^{\pm}_{(s)}$ signal yields and the asymmetries $\mathcal{A}(D^{\pm} \to V^0h^{\pm})$. In the fit, all the parameters are allowed to vary.

3.6.1 Fit to the $D^{\pm}_{(s)} \to K^0_S\pi^{\pm}$ signal channel

The probability density function for the $D^{\pm} \to K^0_S\pi^{\pm}$ signal peak is determined convolving the sum of three Gaussian $G(m)$ with the same mean value μ with a power law parametrizing the final-state QED radiation, as follows

$$\begin{aligned} f(m|\mu, s, \sigma_1, \sigma_2, \sigma_3, f_1, f_3) \propto \int_0^{+\text{inf}} (m')^s \Big[& f_1 \cdot G(m + m'|\mu, \sigma_1) \\ & + (1 - f_1 - f_3) \cdot G(m + m'|\mu, \sigma_2) \\ & + f_3 \cdot G(m + m'|\mu, \sigma_3) \Big] dm'. \end{aligned} \tag{3.13}$$

As regards the $D^{\pm}_s \to K^0_S\pi^{\pm}$ signal, the probability density function is given by the sum of two Gaussian with the same mean value convoluted with the same power law of the $D^{\pm}$ peak (*i.e.* the shape parameter s is shared by the two signal distributions):

$$\begin{aligned} f(m|\mu, s, \sigma_1, \sigma_2, f_1) \propto \int_0^{+\text{inf}} (m')^s \Big[& f_1 \cdot G(m + m'|\mu, \sigma_1) \\ & + (1 - f_1) \cdot G(m + m'|\mu, \sigma_2) \Big] dm'. \end{aligned} \tag{3.14}$$

The combinatorial background is modelled by an exponential with negative slope, *i.e.*

$$f(m|\lambda) \propto e^{\lambda x}. \tag{3.15}$$

3.6.2 Fit to the $D_{(s)}^\pm \to K_S^0 K^\pm$ signal channel

The probability density function for each $D_{(s)}^\pm \to K_S^0 K^\pm$ signal peaks is given by Eq. 3.13. As concerns the background, it is modelled by the sum of two exponentials with negative slope, as follows

$$f(m|\lambda_1, \lambda_2) \propto f_1 \cdot e^{\lambda_1 x} + (1 - f_1) \cdot e^{\lambda_2 x}. \tag{3.16}$$

This parametrization permits to embed the low-mass backgrounds, such as $D_s^\pm \to K_S^0 K^\pm \pi^0$, which are difficult to implement otherwise, into the combinatorial one.

3.6.3 Fit to the $D_{(s)}^\pm \to \phi\pi^\pm$ signal channel

The total probability density function ($F_{TOT}(m|\vec{\theta})$) has the same template of the $D_{(s)}^\pm \to K_S^0 K^\pm$ decay. Here, low-mass backgrounds, such as $D_{(s)}^\pm \to \phi\pi^\pm\pi^0$, are included in the combinatorial background parametrization. As described in § 3.4.2, the invariant mass distribution comes from different triggers. This is taken into account in the fit by introducing a step function at 1900 MeV/c^2 in the invariant mass distribution with a scale factor C:

$$F_{TOT}(m|\vec{\theta}, C) = \begin{cases} F_{TOT}(m|\vec{\theta}) & \text{m} \leq 1900 \text{ MeV/c}^2 \\ C \cdot F_{TOT}(m|\vec{\theta}) & \text{m} > 1900 \text{ MeV/c}^2 \end{cases} \tag{3.17}$$

3.7 Results

The fits to the $D_{(s)}^\pm \to K_S^0\pi^\pm$, $D_{(s)}^\pm \to K_S^0 K^\pm$ and $D_{(s)}^\pm \to \phi\pi^\pm$ invariant mass distributions are performed separately for the 2015 and 2016 data samples, corresponding to a total integrated luminosity of approximately 2 fb^{-1} (0.33 fb^{-1} in 2015 and 1.67 fb^{-1} in 2016). The fitting procedure produces raw asymmetries which give the baseline results, as described below.

The resulting $D_{(s)}^\pm \to K_S^0\pi^\pm$, $D_{(s)}^\pm \to K_S^0 K^\pm$ and $D_{(s)}^\pm \to \phi\pi^\pm$ fits to the 2016 and 2015 datasets are presented separately for positive and negative charges in Figs. 3.18, 3.19 and 3.20, respectively.

Figures 3.21, 3.22 and 3.23 represent the asymmetry calculated as $(N^+ - N^-)/(N^+ + N^-)$ for the number of candidates in each bin of invariant mass histograms as a function of the invariant mass with the fit result overlaid for the $D_{(s)}^\pm \to K_S^0\pi^\pm$, $D_{(s)}^\pm \to K_S^0 K^\pm$ and $D_{(s)}^\pm \to \phi\pi^\pm$ decays, respectively. Here, the irregular shapes of the distributions are given by the fact that positive and negative charges have different instrumental efficiencies. In addition, down and up magnet polarity samples are joined in order to reduce this effect as they exhibit similar but opposite patterns.

Looking at the residual distributions at the bottom of Figs. 3.18-3.20, one can notice that fits are not without defects. The difficulty in using the appropriate fitting model

increases with the statistics. However, Figs. 3.21-3.23 offer the possibility to check if we are actually measuring the integrated asymmetry accurately. As one can notice, asymmetries are well described by the fits.

A summary of the total signal yields obtained from the fits is given in Tab. 3.7 and the raw asymmetries, $\mathcal{A}(D^\pm \to K^0_S\pi^\pm)$, $\mathcal{A}(D^\pm_s \to K^0_S\pi^\pm)$, $\mathcal{A}(D^\pm \to K^0_S K^\pm)$, $\mathcal{A}(D^\pm_s \to K^0_S K^\pm)$, $\mathcal{A}(D^\pm \to \phi\pi^\pm)$ and $\mathcal{A}(D^\pm_s \to \phi\pi^\pm)$, are presented in Tab. 3.8.

The extracted values for $\mathcal{A}'(D^\pm_s \to K^0_S\pi^\pm)$, $\mathcal{A}'(D^\pm \to K^0_S K^\pm)$ and $\mathcal{A}'(D^\pm \to \phi\pi^\pm)$ from addition and subtraction of raw asymmetries without the K^0 detection asymmetry correction are presented in Tab. 3.9 .

year(s)	$D^\pm \to K^0_S\pi^\pm$	$D^\pm_s \to K^0_S\pi^\pm$
2016	13 249 998 ± 5 025	229 797 ± 1 007
2015	2 530 511 ± 2 205	43 930 ± 433
2016 and 2015	15 780 509 ± 5 488	273 727 ± 1 096

year(s)	$D^\pm \to K^0_S K^\pm$	$D^\pm_s \to K^0_S K^\pm$
2016	2 681 065 ± 3011	3 123 572 ± 3 486
2015	471 278 ± 1 088	560 042 ± 1 152
2016 and 2015	3 152 343 ± 3 202	3 683 614 ± 3 671

year(s)	$D^\pm \to \phi\pi^\pm$	$D^\pm_s \to \phi\pi^\pm$
2016	23 555 392 ± 7 061	47 535 022 ± 9 065
2015	3 864 094 ± 2 857	7 776 193 ± 3 677
2016 and 2015	27 419 486 ± 7 617	55 311 215 ± 9 782

Table 3.7: $D^\pm_{(s)} \to V^0 h^\pm$ signal yields.

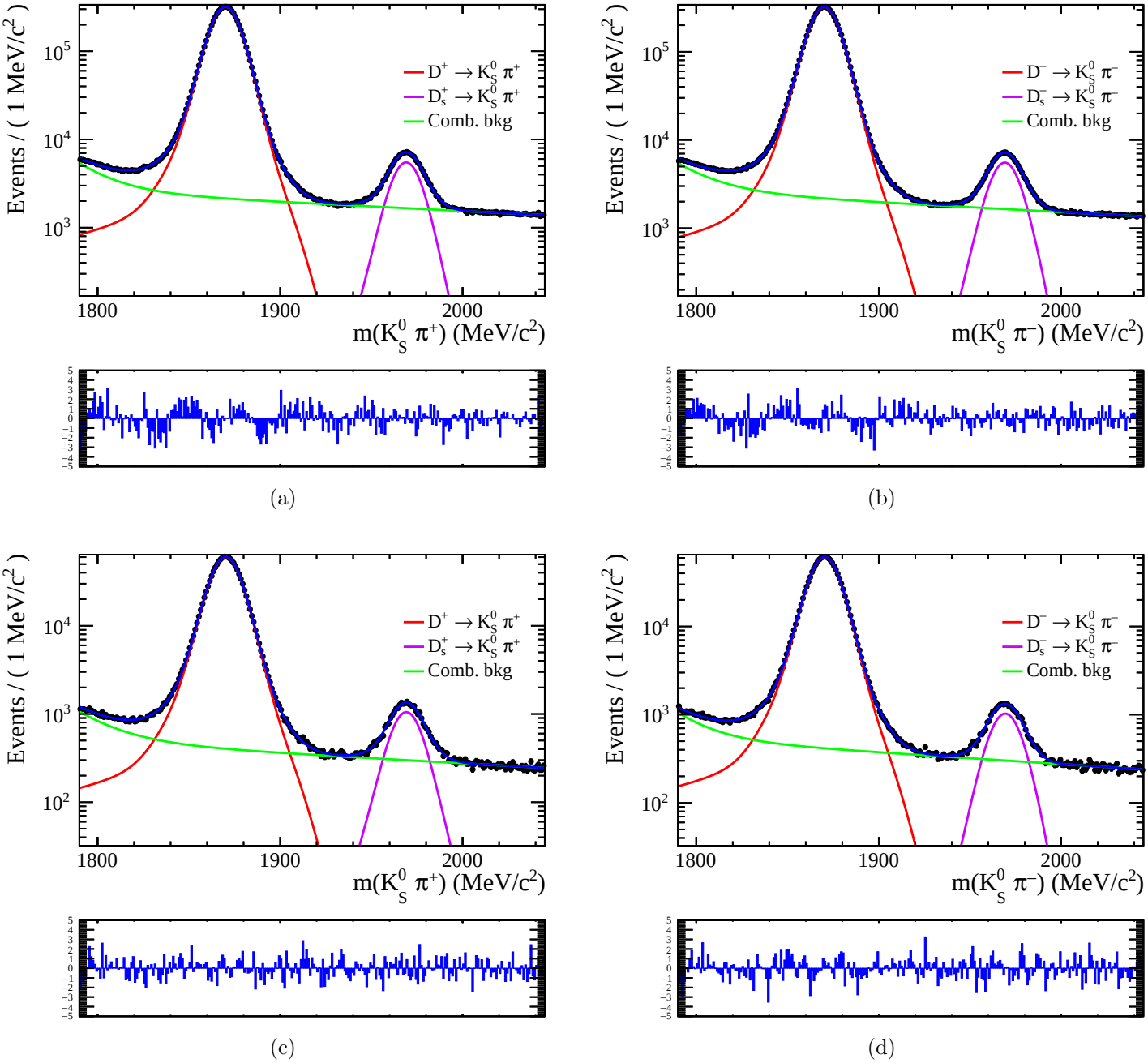


Figure 3.18: $K_S^0\pi^{\pm}$ invariant mass distributions and the results of the fit overlaid for the 2016 (top) and 2015 (bottom) year of data taking, and for the $D_{(s)}^+$ (left) and $D_{(s)}^-$ (right) candidates. At the bottom the residual distribution is reported.

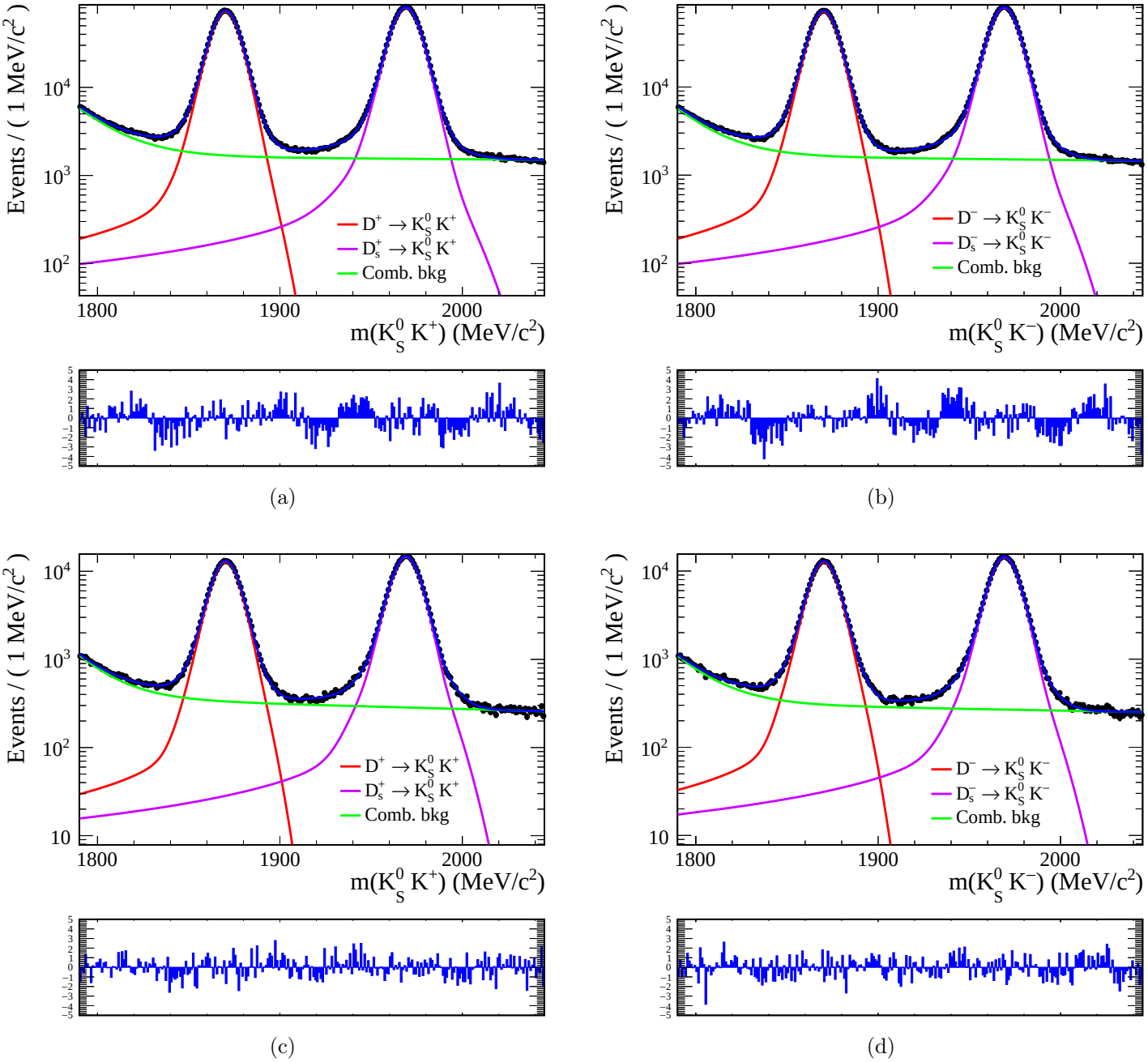


Figure 3.19: $K^0_S K^\pm$ invariant mass distributions and the results of the fit overlaid for the 2016 (top) and 2015 (bottom) year of data taking, and for the $D^+_{(s)}$ (left) and $D^-_{(s)}$ (right) candidates. At the bottom the residual distribution is reported.

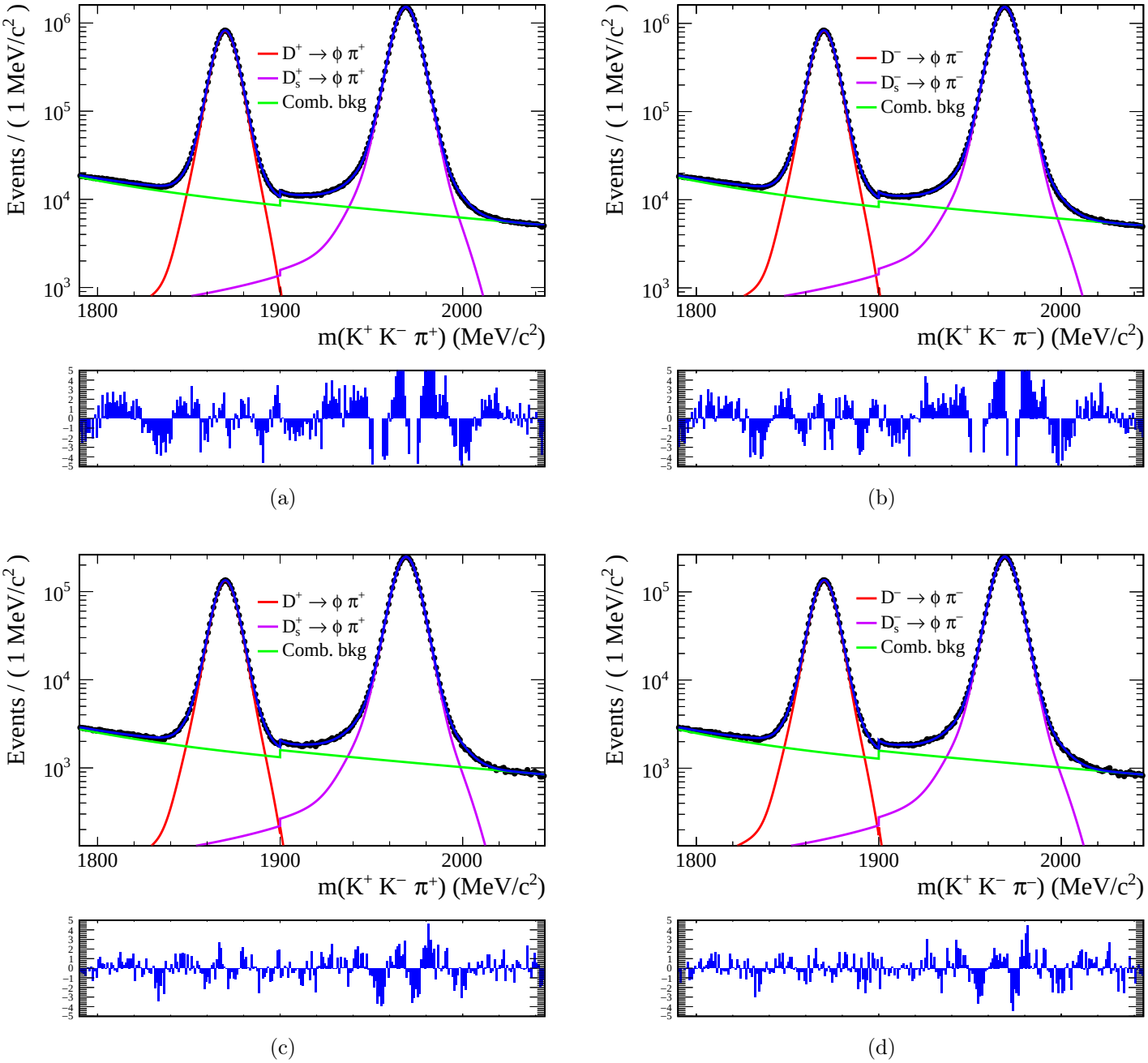
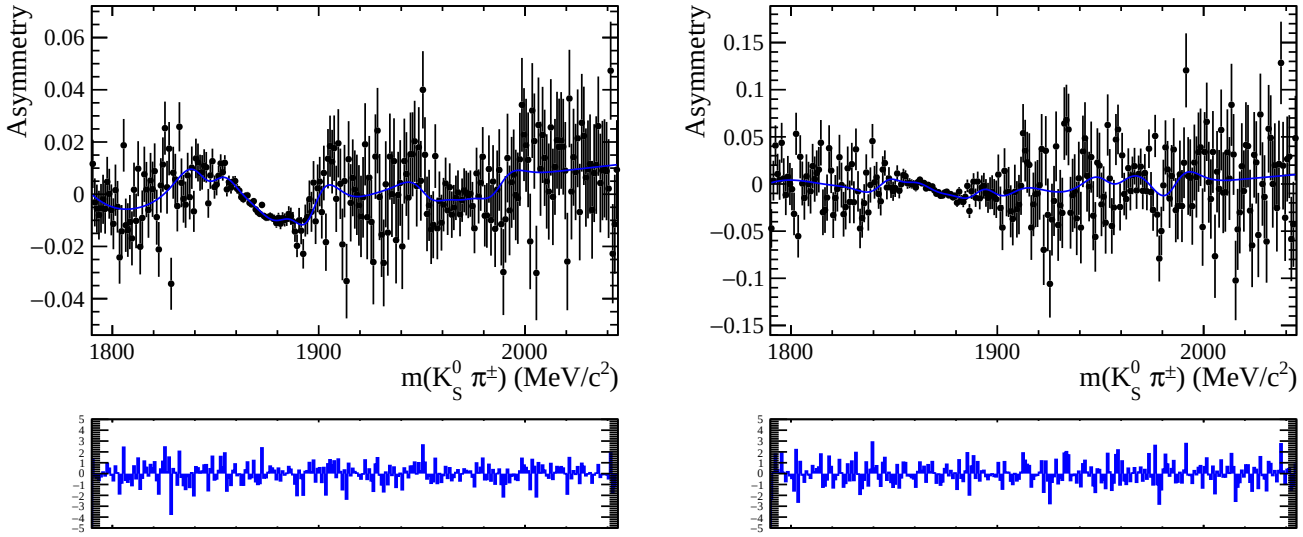
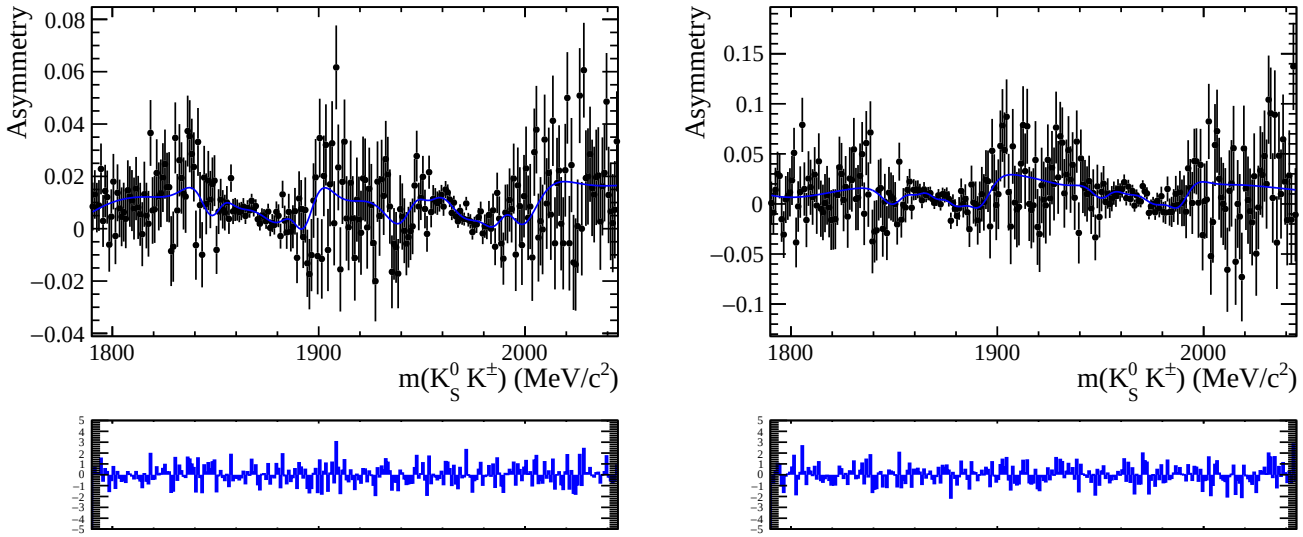


Figure 3.20: $\phi\pi^{\pm}$ invariant mass distributions and the results of the fit overlaid for the 2016 (top) and 2015 (bottom) year of data taking, and for the $D^+_{(s)}$ (left) and $D^-_{(s)}$ (right) candidates. At the bottom the residual distribution is reported.



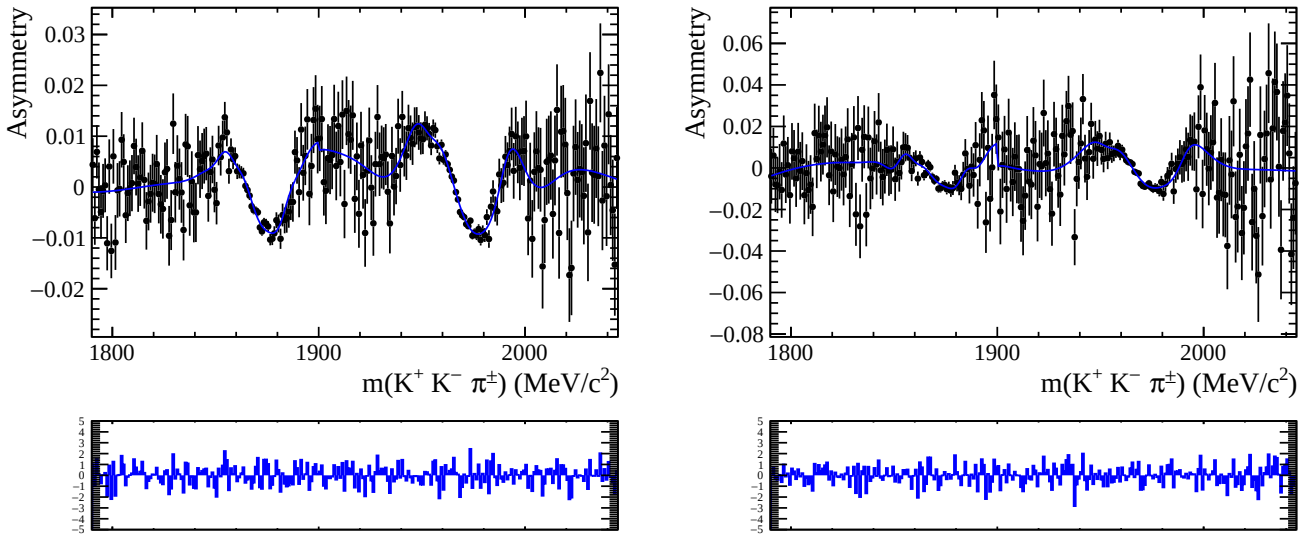
(a) (b)

Figure 3.21: $D^{\pm}_{(s)} \to K^0_S\pi^{\pm}$ asymmetries as a function of the $m(K^0_S\pi^{\pm})$ invariant mass for the 2016 (left) and 2015 (right) year of data taking.



(a) (b)

Figure 3.22: $D^{\pm}_{(s)} \to K^0_S K^{\pm}$ asymmetries as a function of the $m(K^0_S K^{\pm})$ invariant mass for the 2016 (left) and 2015 (right) year of data taking



(a) (b)

Figure 3.23: $D_{(s)}^{\pm} \to \phi\pi^{\pm}$ asymmetries as a function of the $m(\phi\pi^{\pm})$ invariant mass for the 2016 (left) and 2015 (right) year of data taking

year(s)	$\mathcal{A}(D^{\pm} \to K^0_S \pi^{\pm})$	$\mathcal{A}(D^{\pm}_s \to K^0_S \pi^{\pm})$
2016	(-0.38 ± 0.038)%	(-0.38 ± 0.44)%
2015	(-0.624 ± 0.087)%	(0.86 ± 0.98)%
2016 and 2015	(-0.418 ± 0.035)%	(-0.17 ± 0.40)%

year(s)	$\mathcal{A}(D^{\pm} \to K^0_S K^{\pm})$	$\mathcal{A}(D^{\pm}_s \to K^0_S K^{\pm})$
2016	(0.58 ± 0.11)%	(0.57 ± 0.11)%
2015	(0.20 ± 0.23)%	(0.36 ± 0.21)%
2016 and 2015	(0.50 ± 0.10)%	(0.520 ± 0.098)%

year(s)	$\mathcal{A}(D^{\pm} \to \phi \pi^{\pm})$	$\mathcal{A}(D^{\pm}_s \to \phi \pi^{\pm})$
2016	(-0.448 ± 0.030)%	(-0.281 ± 0.019)%
2015	(-0.536 ± 0.074)%	(-0.296 ± 0.047)%
2016 and 2015	(-0.461 ± 0.028)%	(-0.283 ± 0.018)%

Table 3.8: $D^{\pm}_{(s)} \to V^0 h^{\pm}$ raw asymmetries .

year(s)	$\mathcal{A}'(D_s^\pm \to K_S^0\pi^\pm)$
2016	(-0.10 ± 0.44)%
2015	(1.16 ± 0.99)%
2016 and 2015	(0.11 ± 0.40)%

year(s)	$\mathcal{A}'(D^\pm \to K_S^0 K^\pm)$
2016	(0.11 ± 0.16)%
2015	(0.17 ± 0.32)%
2016 and 2015	(0.12 ± 0.15)%

year(s)	$\mathcal{A}'(D^\pm \to \phi\pi^\pm)$
2016	(-0.069 ± 0.048)%
2015	(0.09 ± 0.11)%
2016 and 2015	(-0.045 ± 0.045)%

Table 3.9: $D_{(s)}^\pm \to V^0 h^\pm$ CP asymmetries without the neutral kaon asymmetry correction.

3.7.1 Final results for *CP*-violating asymmetries

Table 3.10 reports the three *CP* asymmetries with the correction $\mathcal{A}(K^0) = (+0.0791 \pm 0.0047)\%$ [66] included, calculated using Eqs. 3.10-3.12. Measurements of $\mathcal{A}_{CP}^{D^\pm \to K^0_S K^\pm}$ and $\mathcal{A}_{CP}^{D^\pm \to \phi\pi^\pm}$ have reached very high-precision. At this level the $\mathcal{A}(K^0)$ correction becomes important as its value is comparable with the statistical precision on the measured *CP* asymmetries. For this reason, it is fundamental to improve the knowledge of this quantity in order to reduce the systematic uncertainty due to this correction.

year(s)	$\mathcal{A}_{CP}(D_s^\pm \to K^0_S\pi^\pm)$
2016	(-0.18 ± 0.44)%
2015	(1.08 ± 0.99)%
2016 and 2015	(0.03 ± 0.40)%

year(s)	$\mathcal{A}_{CP}(D^\pm \to K^0_S K^\pm)$
2016	(0.03 ± 0.16)%
2015	(0.09 ± 0.32)%
2016 and 2015	(0.04 ± 0.15)%

year(s)	$\mathcal{A}_{CP}(D^\pm \to \phi\pi^\pm)$
2016	(-0.148 ± 0.049)%
2015	(0.01 ± 0.11)%
2016 and 2015	(-0.124 ± 0.045)%

Table 3.10: $D^\pm_{(s)} \to V^0 h^\pm$ *CP* asymmetries after neutral kaon asymmetry correction.

3.7.2 Comparison with Run-1 results

In Table 3.11, the reached statistical sensitivies to $\mathcal{A}_{CP}$ together with the collected yields in Cabbibbo-suppressed decays are compared with the results from Run-1. One can notice that the improvement on $\sigma(\mathcal{A}_{CP})$ is smaller than expected from the increased

yield[3]. This is due by the fact that we are currently using different shapes parameter for the positive and negative distributions while in previous analyses shape parameters were common between $D^+_{(s)}$ and $D^-_{(s)}$ datasets. This feature in the fitting shape has been mandatory in order to well describe the differential asymmetry, shown in Figs. 3.21-3.23. Furthermore, if one considers that $\sigma(\mathcal{A}_{CP})$ is only due to the statistical uncertainty of the Cabibbo-suppressed yields (this is true for $D_s^\pm \to K_S^0\pi^\pm$ decays), one can obtain the relation

$$\sigma(\mathcal{A})\sqrt{N} \approx \sigma(N)/\sqrt{N} \tag{3.18}$$

which is verified in our results.

	N	$\sigma(\mathcal{A}_{CP})$
$D_s^\pm \to K_S^0\pi^\pm$		
Run-2	273 727 ± 1 096	(0.40)%
Run-1	120 976 ± 692	(0.46)%
$D^\pm \to K_S^0 K^\pm$		
Run-2	3 152 343 ± 3 202	(0.15)%
Run-1	1 013 516 ± 1 379	(0.17)%
$D^\pm \to K_S^0 K^\pm$		
Run-2	27 419 486 ± 7617	(0.045)%
2011	1 576 900 ± 1 500	(0.14)%

Table 3.11: Collected yields and $\mathcal{A}_{CP}$ statistical uncertainties for Cabibbo-suppressed decays resulting from Run-1 and Run-2 analyses.

3.8 Systematic uncertainties

The measured asymmetries and $\mathcal{A}_{CP}$ results are subject to several sources of systematic uncertainty. An overview of the main sources is reported in this section.

3.8.1 K^0 interactions in the VELO material

K_S^0 and K_L^0 mesons are an admixture of K^0 and $\overline{K}^0$. K^0 and $\overline{K}^0$ mesons react asymmetrically in the detector material. $\overline{K}^0$ mesons can be absorbed in hyperon production

[3]One should expect $\sigma(\mathcal{A}_2) \approx \sqrt{N_1/N_2}\,\sigma(\mathcal{A}_1)$.

processes while K^0 mesons can only scatter elastically or undergo charge exchange. It is important to note that decays of $D^{\pm}_{(s)}$ mesons produce a precise flavour of $\overset{(-)}{K}{}^0$ as shown in Figs. 1.2 and 1.3. The initially pure K^0 and $\overline{K}{}^0$ states oscillate back forth, as they propagate through the VELO, *CP* eigenstates K^0_S and K^0_L are produced. Within the VELO, the propagating states have a non-zero chance of interacting with VELO material. The interaction of neutral kaons with matter is often called regeneration, because it results in the production of a small number of K^0_S mesons from an initially pure beam of K^0_L mesons. The regenerated K^0_S can then decay into two pions miming a *CP* violation.

In this analysis, we can reconstruct a candidate K^0_S decay from two charged pions if:

- There is a true K^0_S to two charged pions decay (*CP* conserving).
- The K^0_L states interacts with the VELO material and is generated as a K^0_S which then decays into two charged pions (*CP* conserving).
- The K^0_L states violate *CP* and decays to two charged pions.

As a consequence $D^+_{(s)}$ and $D^-_{(s)}$ mesons are not reconstructed with the same efficiencies. The size of this effect is $\mathcal{A}(K^0) = (+0.0791 \pm 0.0047)\%$ [66] for long-long pions reconstructed in 2016 using $D^{\pm} \to K^0_S\pi^{\pm}$ decays and following the method described in Ref. [69](§ 7).

It should be noted that the K^0 detection asymmetry could theoretically vary between different $D^{\pm}_{(s)} \to K^0_S h^{\pm}$ modes and selections as the momentum spectra of the K^0_S candidates are different. This effect is ignored in Eqs. 3.10-3.12 but need to be taken into account for systematics as suggested in § 3.7.1.

3.8.2 Secondary $D^{\pm}_{(s)}$ decays

In the offline selection of $D^{\pm}_{(s)}$ candidates (see § 3.4.2), the $\chi^2(IP)$ requirement on the $D^{\pm}_{(s)}$ removes the majority of background from secondary $D^{\pm}_{(s)}$ mesons originating from the decay of a b hadron. The remaining secondary $D^{\pm}_{(s)}$ mesons may introduce a bias in the measured *CP* asymmetries due to a difference in the production asymmetries for b hadrons and $D^{\pm}_{(s)}$ mesons. This bias can be taken into account by modifying the $D^{\pm}_{(s)}$ production asymmetries in Eqs. 3.3-3.5 and Eqs. 3.7-3.9 for $D^{\pm}_{(s)} \to V^0 h^{\pm}$ decays as

$$\mathcal{A}^{D^{\pm}_{(s)}}_{prod}(corr.) = \frac{\mathcal{A}^{D^{\pm}_{(s)}}_{prod} + \mathcal{A}^{B}_{prod}}{1+f}, \tag{3.19}$$

where f is the fraction of secondary $D^{\pm}_{(s)}$ candidates in a particular decay channel and $\mathcal{A}^{B}_{prod}$ is the corresponding b-hadron production asymmetry. The fraction f can be estimated from the measured $D^{\pm}$, $D^{\pm}_s$ and b hadron inclusive cross-sections, the inclusive

branching fractions $\mathfrak{B}(b \to D^\pm X)$ and $\mathfrak{B}(b \to D_s^\pm X)$, where X corresponds to any other particles in the final state, the exclusive branching fractions $\mathfrak{B}(D_{(s)} \to V^0 h^\pm)$ and the estimated efficiencies.

3.8.3 Differences in signal and control channels kinematics

The $D_{(s)}^\pm \to K_S^0 h^\pm$ and $D_{(s)}^\pm \to \phi\pi^\pm$ selections have been optimized in isolation from one another. However, the experimental methods outlined in § 3.3 require that the production asymmetries of the $D_{(s)}^\pm$ candidates and the detection asymmetries of the bachelor hadrons cancel between final states. Therefore, kinematic distributions of the $D_{(s)}^\pm$ and the bachelor hadron in the signal and control channel selections should correspond. Figures 3.24, 3.25 and 3.26 show the comparison of $D_{(s)}^\pm$ and $\pi^\pm$ or $K^\pm$ transverse momentum distributions for the signal and control channels. As one can notice, the distributions are not well aligned. Therefore, further criteria need to be applied in order to harmonize the $D_{(s)}^\pm \to K_S^0 h^\pm$ and $D_{(s)}^\pm \to \phi\pi^\pm$ selections and finally a kinematic reweighing should be used to obtain exactly the same distributions. This approach introduces systematics which need to be appropriately assigned to the resulting measurements.

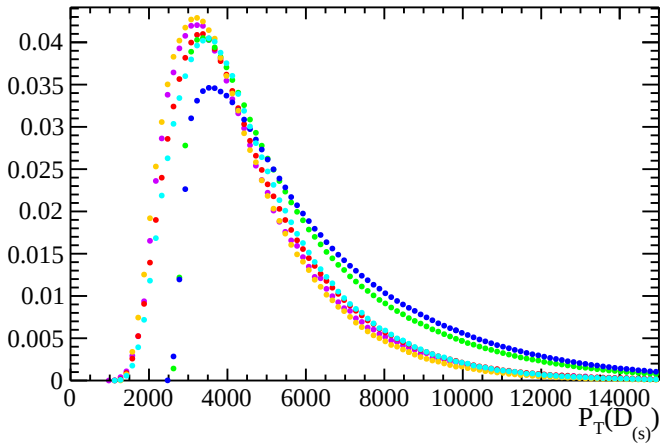


Figure 3.24: Normalized $D_{(s)}^\pm$ p_T distributions for all the decay modes: $D^\pm \to K_S^0\pi^\pm$ (red), $D_s^\pm \to K_S^0\pi^\pm$ (violet), $D^\pm \to \phi\pi^\pm$ (green), $D_s^\pm \to \phi\pi^\pm$ (blue), $D^\pm \to K_S^0 K^\pm$ (orange), $D_s^\pm \to K_S^0 K^\pm$ (cyan).

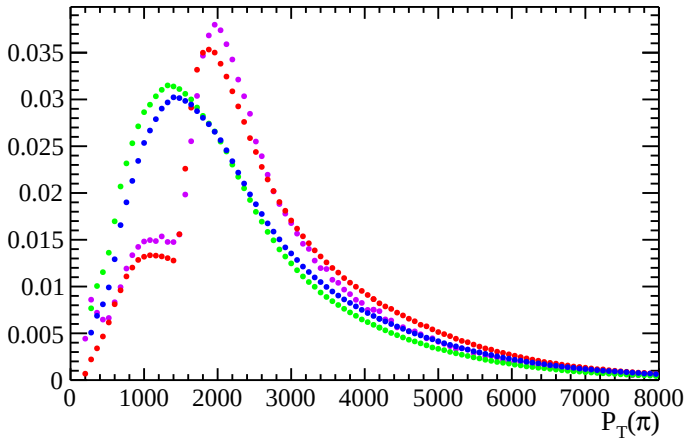


Figure 3.25: Normalized bachelor p_T distributions for modes including a bachelor pion: $D^{\pm} \to K^0_S\pi^{\pm}$ (red), $D^{\pm}_s \to K^0_S\pi^{\pm}$ (violet), $D^{\pm} \to \phi\pi^{\pm}$ (green), $D^{\pm}_s \to \phi\pi^{\pm}$ (blue).

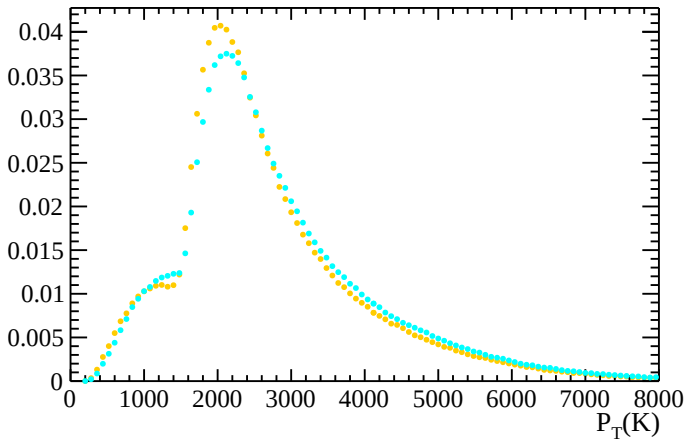


Figure 3.26: Normalized bachelor p_T distributions for modes including a bachelor kaon: $D^{\pm} \to K^0_S K^{\pm}$ (orange), $D_s \to K^0_S K^{\pm}$ (cyan).

3.9 Summary and conclusions

In summary, this work represents a search for direct CP violation in charm decays using $D_s^\pm \to K_S^0\pi^\pm$, $D^\pm \to K_S^0 K^\pm$ and $D^\pm \to \phi\pi^\pm$ decays recorded by LHCb during Run-2 (2015-2016) data taking period. The obtained results seem to be competitive with other available measurements based on these decays [40–47]. The separate fit for positive and negative distributions allows to take into account systematics biases due to the left-right detection asymmetry with a strong impact on statistical sensitivity which is then comparable with Run-1 results. However, all results presented have to be considered preliminary and unofficial since the systematic errors have not been estimated, but only discussed. The final result obtained are

$$\begin{aligned} \mathcal{A}_{CP}(D_s^\pm \to K_S^0\pi^\pm) &= (0.03 \pm 0.40\,(\text{stat.}))\%, \\ \mathcal{A}_{CP}(D^\pm \to K_S^0 K^\pm) &= (0.04 \pm 0.15\,(\text{stat.}))\% \text{ and} \\ \mathcal{A}_{CP}(D^\pm \to \phi\pi^\pm) &= (-0.124 \pm 0.045\,(\text{stat.}))\%. \end{aligned}$$

The sensitivity to the CP-violating asymmetries could be increased by updating the measurements already performed on Run-1 data, including the data collected during 2017 and by adding events where K_S^0 mesons are reconstructed downstream of the LHCb vertex locator. The expectation is to have results with a more than doubled precision, which can be published in the next months.

Conclusions

This thesis reports the results of a search for direct *CP* violation in charm decays, namely the singly-Cabibbo-suppressed $D_s^\pm \to K_S^0\pi^\pm$, $D^\pm \to K_S^0 K^\pm$ and $D^\pm \to \phi\pi^\pm$ decays. The analysis is realized by using the full 2015 and 2016 data sets of *pp* collisions collected with the LHCb detector, corresponding to an integrated luminosity of about 2 fb^{-1} .

The thesis work consists in the development of the best offline selection in order to obtain the maximum significance of the signal peaks and then extract signal yields from fit to the invariant mass distributions to measure the direct *CP*-violating observables. The selection includes rectangular cuts to constrain the topology of the decays and to suppress the backgrounds arising from particle misidentification, and finally the application of a neural network permits to suppress the combinatorial backgrounds. In the thesis work, effects induced by the detector and by proton-proton production mechanism, *i.e.* $D^\pm_{(s)}$ production asymmetries, are also evaluated by means of the control samples.

The individual *CP*-violating asymmetries are measured to be

$$\begin{aligned} \mathcal{A}_{CP}(D_s^\pm \to K_S^0\pi^\pm) &= (0.03 \pm 0.40\,(\text{stat.}))\%, \\ \mathcal{A}_{CP}(D^\pm \to K_S^0 K^\pm) &= (0.04 \pm 0.15\,(\text{stat.}))\%, \\ \mathcal{A}_{CP}(D^\pm \to \phi\pi^\pm) &= (-0.124 \pm 0.045\,(\text{stat.}))\%, \end{aligned}$$

assuming that *CP* violation in the Cabibbo-favoured decays is negligible. The measurements are consistent with and supersede those of the previous LHCb analyses. As the results are compatible with zero, no evidence of *CP* violation is found. However, the results are preliminary and additional studies are still ongoing. Indeed, there are still few tasks, such as the study of potential systematic effects, to be accomplished before this work can be turned into a publication. These further studies will be completed within the analysis activities of the Bologna-LHCb group in collaboration with the CERN-LHCb group.

Appendix A

Kinematic separation of $D^{\pm} \to K^0_S h^{\pm}$ and $K^0_S \to \pi^+\pi^-$ decays

The invariant mass of two particles X and Y of mass m_X and m_Y and momenta $\vec{p}_X$ and $\vec{p}_Y$ is defined as

$$M^2_{XY} = \left(\sqrt{p^2_X + m^2_X} + \sqrt{p^2_Y + m^2_Y}\right)^2 - (\vec{p}_X + \vec{p}_Y)^2. \tag{A.1}$$

If the two particles originate from the two-body decay of a common mother P of mass m_0, then $M_{12} = m_0$. If, instead, one wrongly assings the masses to the outgoing particles, the resulting invariant mass

$$M^2_{12} = \left(\sqrt{p^2_1 + m^2_1} + \sqrt{p^2_2 + m^2_2}\right)^2 - (\vec{p}_X + \vec{p}_Y)^2, \tag{A.2}$$

where the mass m_1 (m_2) is assigned to the particle with momentum $\vec{p}_X$ ($\vec{p}_Y$), is shifted with respect to m_0 as

$$M^2_{12} = m^2_0 - \Delta M^2, \tag{A.3}$$

with

$$\begin{aligned}
\Delta M^2 &= M^2_{XY} - M^2_{12} \\
&= (m^2_X - m^2_1) + (m^2_Y - m^2_2) \\
&\quad + 2p_X p_Y \left(\sqrt{1 + \left(\frac{m_X}{p_X}\right)^2} \sqrt{1 + \left(\frac{m_Y}{p_Y}\right)^2} - \sqrt{1 + \left(\frac{m_1}{p_X}\right)^2} \sqrt{1 + \left(\frac{m_2}{p_Y}\right)^2} \right) \\
&\approx (m^2_X - m^2_1) + (m^2_Y - m^2_2) \\
&\quad + p_X p_Y \left[\left(\frac{m_X}{p_X}\right)^2 + \left(\frac{m_Y}{p_Y}\right)^2 - \left(\frac{m_1}{p_X}\right)^2 - \left(\frac{m_2}{p_Y}\right)^2 \right] \\
&= (m^2_X - m^2_1)(1 + p_Y/p_X) + (m^2_Y - m^2_2)(1 + p_X/p_Y) \\
&= \frac{2}{1+\beta}(m^2_X - m^2_1) + \frac{2}{1-\beta}(m^2_Y - m^2_2),
\end{aligned} \tag{A.4}$$

where the approximation holds the first order in (m/p) and the charged momentum imbalance β is further defined as

$$\beta = \frac{p_X - p_Y}{p_X + p_Y}. \tag{A.5}$$

The above expression, for a $P^{\pm} \to X^0 Y^{\pm}$ decay, when using a charged $D_{(s)}$ meson, a neutral kaon and a charged pion or kaon as arbitrary mass assignments for the initial and the final state particles (*i.e.* a $D^{\pm}_{(s)} \to K^0_S h^{\pm}$ decay), becomes:

$$M^2(D^{\pm}_{(s)} \to K^0_S K^{\pm})[D^{\pm}_{(s)} \to K^0_S \pi^{\pm}] \approx m^2_{D^{\pm}_{(s)}} - \frac{2}{1-\beta}(m^2_K - m^2_\pi), \tag{A.6}$$

$$M^2(D^{\pm}_{(s)} \to K^0_S \pi^{\pm})[D^{\pm}_{(s)} \to K^0_S K^{\pm}] \approx m^2_{D^{\pm}_{(s)}} - \frac{2}{1-\beta}(m^2_\pi - m^2_K), \tag{A.7}$$

$$M^2(\Lambda^{\pm}_c \to K^0_S \overset{(-)}{p})[D^{\pm}_{(s)} \to K^0_S \pi^{\pm}] \approx m^2_{\Lambda^{\pm}_c} - \frac{2}{1-\beta}(m^2_p - m^2_\pi), \tag{A.8}$$

$$M^2(\Lambda^{\pm}_c \to K^0_S \overset{(-)}{p})[D^{\pm}_{(s)} \to K^0_S K^{\pm}] \approx m^2_{\Lambda^{\pm}_c} - \frac{2}{1-\beta}(m^2_p - m^2_K), \tag{A.9}$$

$$M^2(\Lambda^{\pm}_c \to \Lambda^0 \pi^{\pm})[D^{\pm}_{(s)} \to K^0_S \pi^{\pm}] \approx m^2_{\Lambda^{\pm}_c} - \frac{2}{1+\beta}(m^2_{\Lambda^0} - m^2_{K^0_S}). \tag{A.10}$$

For a $P^0 \to X^+ Y^-$ decay, when using a neutral kaon and a pair of charged pions as arbitrary mass assignments for the initial and the final state particles (*i.e.* a $K^0_S \to \pi^+\pi^-$ decay), it becomes:

$$M^2(\overset{(-)}{\Lambda^0} \to \overset{(-)}{p}\pi^{\mp})[K^0_S \to \pi^+\pi^-] \approx m^2_\Lambda - \frac{2}{1 \pm \beta}(m^2_p - m^2_\pi). \tag{A.11}$$

The above relations are graphically shown in Figs. 3.2 (a), 3.3 (a) and 3.4 (a).

Appendix B

Techniques for data-driven multivariate analysis

Sideband subraction

The training of multivariate classifiers (MVA) using real data, instead of data from simulation, is not trivial since usually no pure samples of signal and background events are available. If however in the distribution of a variable a signal peak is visible over some background, a training is possible after a *sideband substraction*. Indeed, different regions in the distribution of the discriminating variable need to be defined to determine which events are included as samples for signal and background.

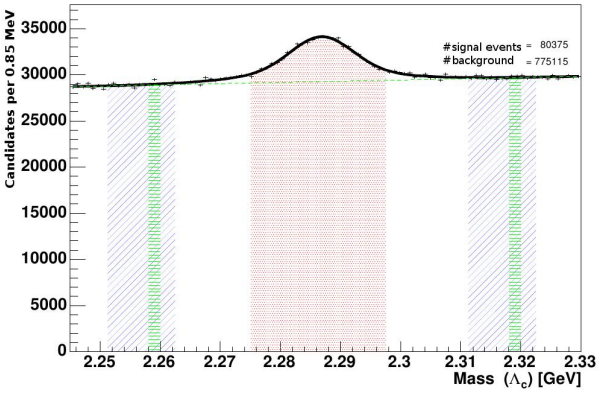


Figure B.1: Invariant mass distribution of Λ_c candidates. The signal region is highlightened in red, whereas the background regions (sidebands) are marked in blue and green [70].

Figure B.1 shows an example, where the invariant mass $m(\Lambda_c)$ is used as discriminating variable. The red region (signal region) can be used in the training as signal with weight 1. An estimate of the number of signal events (N_{sig}) and background events

(N_{bkg}) lying in this region can be done by means of a fit to the invariant mass distribution. The size of the blue region has to be chosen such that it contains the same number of background events as in the signal region. The sample thus selected can be used in the training as signal with weight -1. This has the effect that the background events from the red region with weight 1 are statistically subtracted from the signal events. This leads to a signal sample that statistically only consists of signal events. On the other hand, the green region can be chosen in a size such that it contains N_{sig} events that can be used in the training as background with weight 1.

The statistical subtraction of background events only works if the background behaves the same across the range of the discriminating variable. This does not mean that the density of background events has to be the same across the range, but that the variables used in the training have to be independent of the discriminating variable. As a consequence, the discriminating variable itself cannot be used; it is only used to prepare the training sample.

The *sPlot* technique

The *sPlot* technique can be used as an advanced sideband subtraction. It does not require a clean background region, but only measurable differences between the distributions of the fit components. With the *sPlot* technique, each event is assigned a weight (referred to as *sWeight*) for each fit component. The advantage over sideband subtraction is that with this approach no regions need to be defined. Everything is derived from the fit model and thus the data on the whole range can be used, given that there are no satellite resonances present within the range.

The *sPlot*[1] technique was introduced in Ref. [68] with the subtitle "A statistical tool to unfold data distributions". This means that based on the model of some discriminating variables, the distributions of other variables can be reconstructed for each fit model component. The *sPlot* technique allows to produce these distributions even if it is not clear for all events to which class they belong while accounting for the correlation between the classes.

In the following, a simple example is given to verify the validity of the technique. Let's consider a Monte Carlo (MC) dataset of 1 000 000 events divided in three classes, labelled as D, D_s and bkg. Two independent variables m and x are simulated for each class. The shapes of the variables distribution together with the number of events N for each class are reported in Tab. B.1. Note that this dataset is realized with the aim of reproducing the effect of the *sPlot* technique on the $D^{\pm}_{(s)} \to K^0_S\pi^{\pm}$ data sample where the m distribution aims to simulate the $K^0_S\pi^{\pm}$ invariant mass distribution, whereas x represents an ordinary variable that can be used as input variable in a MVA training.

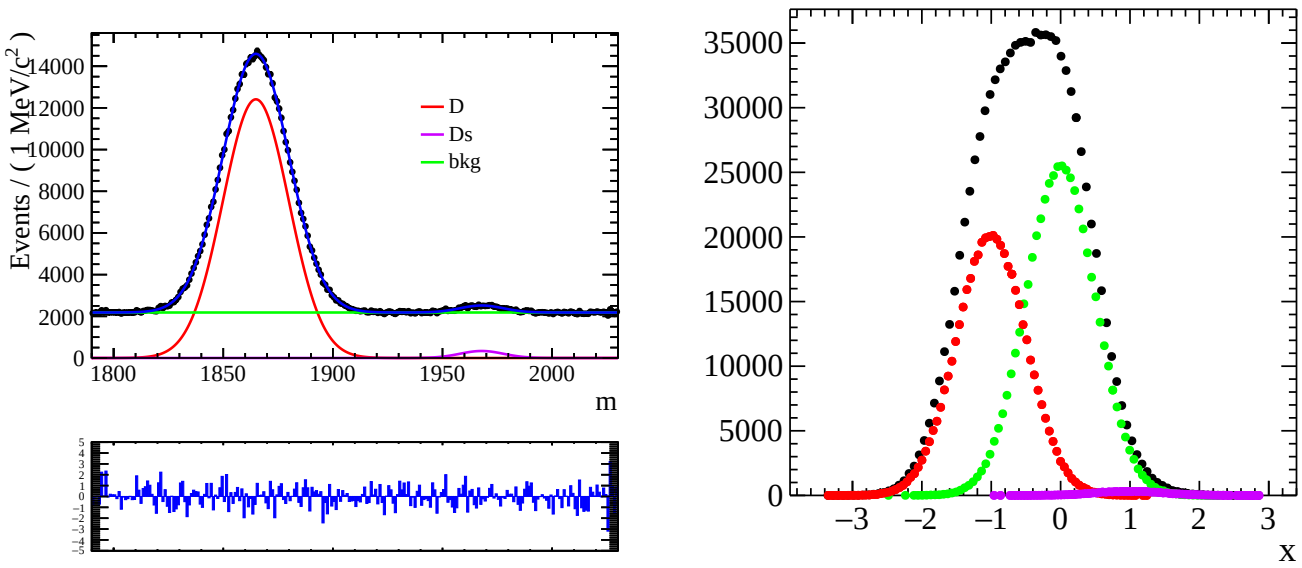
[1] According to the authors, the s in *sPlot* originates from the usage of the covariance in the calculation of the weights. The covariance is related to the variance which is in turn sometimes denoted with s^2.

However, qualitatively speaking, similar results can be obtained for different samples.

Class	m distribution	x distribution	N
D	Gaussian(1864.84,15)	Gaussian(-1,0.5)	468213
D_s	Gaussian(1968.30,10)	Gaussian(1,0.5)	8470
bkg	Uniform(1790,2030)	Gaussian(0,0.5)	523317

Table B.1: m and x variables distribution and the number of events N for the classes D, D_s and bkg. Gaussian(μ,σ) refers to a Gaussian distribution with mean μ and standard deviation σ while Uniform(a,b) represents a Uniform distribution between a and b.

Figure B.2 shows the m and x distributions of the whole dataset. Looking at the m distribution, it is possible to recognize the different contributions from the events belonging to different classes. Instead, in the x distribution, the contributes of the singular class are not visible. Therefore, one can chose m as the discriminating variable and then extract the *sWeights* from the fit for each class in order to unfold the corresponding x distribution.



(a) (b)

Figure B.2: Total m (a) and x (b) distributions from the MC toy (in black). Red, violet and green distributions refer to the D, D_s and bkg classes, respectively. In (a) the result of the fit is overlaid. In (b) the individual distribution for each class is represented overlaid.

By the application of the *sWeights* to the events, one can finally obtain the distributions of the x variable as shown in Fig. B.3. These results are compatible with the true distributions described in Tab. B.1.

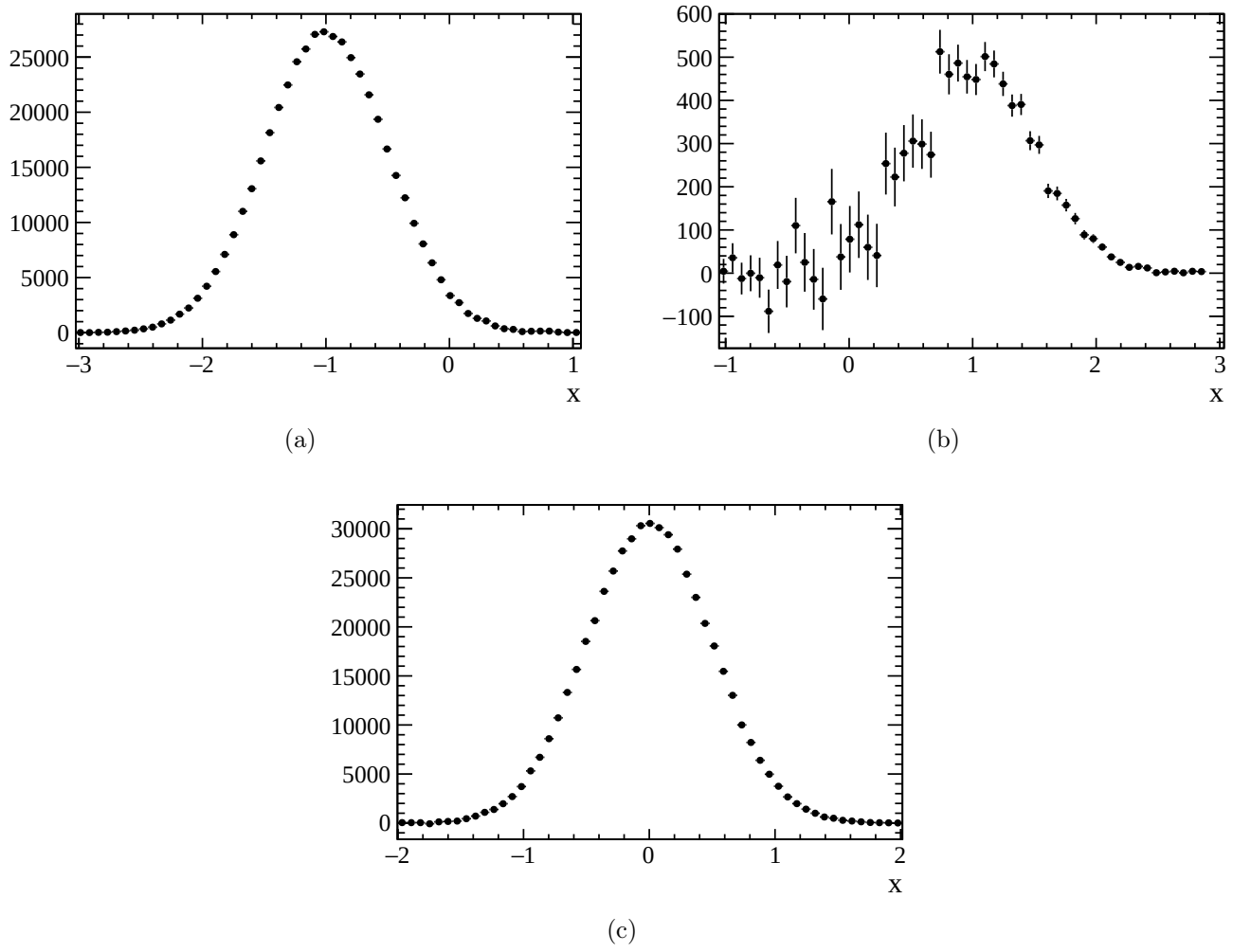


Figure B.3: x distributions after the application of the *sWeights* for D (a), D_s (b) and *bkg* (c) classes.

If one plots the distribution of the m variable by applying the *sWeights* of a specific class can notice that they have positive and negative values depending on the fact that we are looking at a signal or a background region with respect to the chosen class (see Fig. B.4). This feature highlights the fact that the contribution of background events under a peak can be removed through the application of negative weights to the events far from the peak (exactly as in the sideband subtraction). On the other hand, this figure shows that the m distribution after the application of the *sWeights* does not

corresponds to the true distribution. This means that the distribution of variables which are correlated with m can be distorted.

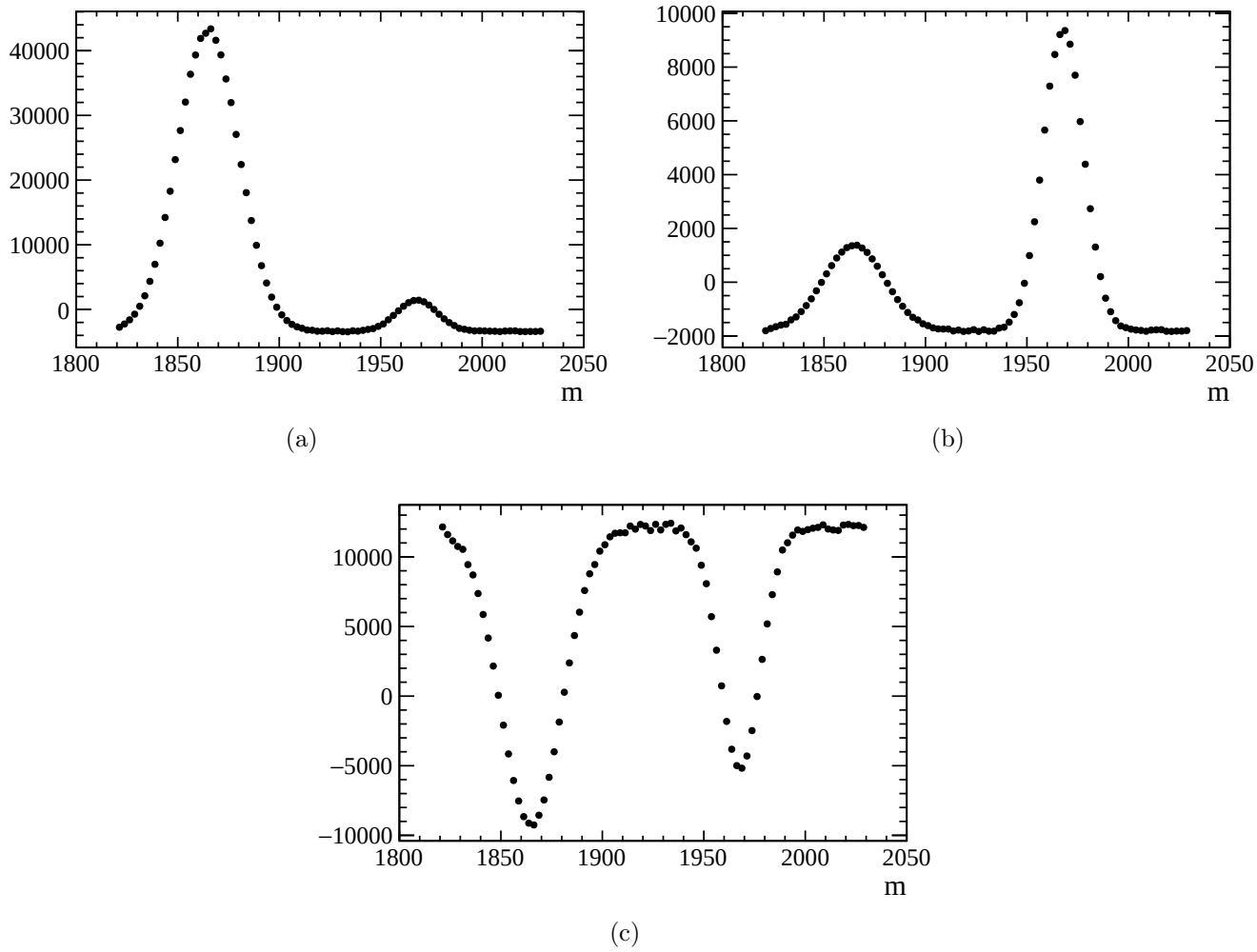


Figure B.4: m distributions after the application of the *sWeights* for D (a), D_s (b) and *bkg* (c) classes.

Bibliography

[1] T. D. Lee and C. N. Yang. Question of parity conservation in weak interactions. *Phys. Rev.*, 104:254–258, Oct 1956. doi: 10.1103/PhysRev.104.254. URL `https://link.aps.org/doi/10.1103/PhysRev.104.254`.

[2] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, Feb 1957. doi: 10.1103/PhysRev.105.1413. URL `https://link.aps.org/doi/10.1103/PhysRev.105.1413`.

[3] R. L. Garwin, L. M. Lederman, and M. Weinrich. Observations of the failure of conservation of parity and charge conjugation in meson decays: the magnetic moment of the free muon. *Phys. Rev.*, 105:1415–1417, Feb 1957. doi: 10.1103/PhysRev.105.1415. URL `https://link.aps.org/doi/10.1103/PhysRev.105.1415`.

[4] M. Goldhaber, L. Grodzins, and A. W. Sunyar. Helicity of neutrinos. *Phys. Rev.*, 109:1015–1017, Feb 1958. doi: 10.1103/PhysRev.109.1015. URL `https://link.aps.org/doi/10.1103/PhysRev.109.1015`.

[5] G. Luders. On the Equivalence of Invariance under Time Reversal and under Particle-Antiparticle Conjugation for Relativistic Field Theories. *Kong. Dan. Vid. Sel. Mat. Fys. Med.*, 28N5(5):1–17, 1954.

[6] G. Luders. Proof of the TCP theorem. *Annals Phys.*, 2:1–15, 1957. doi: 10.1016/0003-4916(57)90032-5. [Annals Phys.281,1004(2000)].

[7] 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`.

[8] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961. doi: 10.1016/0029-5582(61)90469-2.

[9] S. Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967. doi: 10.1103/PhysRevLett.19.1264. URL `https://link.aps.org/doi/10.1103/PhysRevLett.19.1264`.

[10] A. Salam. Weak and electromagnetic interactions. In W. Svartholm, editor, *Elementary Particle Theory*, page 367, Stockholm, 1968. Almquist and Wiksell.

[11] Nicola 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`.

[12] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak Interactions with Lepton-Hadron Symmetry. *Phys. Rev.*, D2:1285–1292, 1970. doi: 10.1103/PhysRevD.2.1285.

[13] M. Kobayashi and T. Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973. doi: 10.1143/PTP.49.652.

[14] G. S. Abrams et al. The Discovery of a Second Narrow Resonance in e+ e- Annihilation. *Phys. Rev. Lett.*, 33:1453–1455, 1974. doi: 10.1103/PhysRevLett.33.1453. [Adv. Exp. Phys.5,150(1976)].

[15] J. J. Aubert, U. Becker, P. J. Biggs, J. Burger, M. Chen, G. Everhart, P. Goldhagen, J. Leong, T. McCorriston, T. G. Rhoades, M. Rohde, S. C. C. Ting, S. L. Wu, and Y. Y. Lee. Experimental observation of a heavy particle j. *Phys. Rev. Lett.*, 33:1404–1406, Dec 1974. doi: 10.1103/PhysRevLett.33.1404. URL `https://link.aps.org/doi/10.1103/PhysRevLett.33.1404`.

[16] S. W. Herb et al. Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions. *Phys. Rev. Lett.*, 39:252–255, 1977. doi: 10.1103/PhysRevLett.39.252.

[17] F. Abe et al. Observation of top quark production in $\bar{p}p$ collisions. *Phys. Rev. Lett.*, 74:2626–2631, 1995. doi: 10.1103/PhysRevLett.74.2626.

[18] S. Abachi et al. Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev. Lett.*, 74:2422–2426, 1995. doi: 10.1103/PhysRevLett.74.2422.

[19] B. Aubert et al. Observation of CP violation in the B^0 meson system. *Phys. Rev. Lett.*, 87:091801, 2001. doi: 10.1103/PhysRevLett.87.091801.

[20] K. Abe et al. Observation of large CP violation in the neutral B meson system. *Phys. Rev. Lett.*, 87:091802, 2001. doi: 10.1103/PhysRevLett.87.091802.

[21] R. Aaij et al. First observation of CP violation in the decays of B_s^0 mesons. *Phys. Rev. Lett.*, 110(22):221601, 2013. doi: 10.1103/PhysRevLett.110.221601.

[22] R. Aaij et al. Observation of $D^0 - \overline{D}^0$ oscillations. *Phys. Rev. Lett.*, 110(10):101802, 2013. doi: 10.1103/PhysRevLett.110.101802.

[23] 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. [Usp. Fiz. Nauk161,61(1991)].

[24] G. R. Farrar and M. E. Shaposhnikov. Baryon asymmetry of the universe in the standard electroweak theory. *Phys. Rev.*, D50:774, 1994. doi: 10.1103/PhysRevD.50.774.

[25] P. Huet and E. Sather. Electroweak baryogenesis and standard model CP violation. *Phys. Rev.*, D51:379–394, 1995. doi: 10.1103/PhysRevD.51.379.

[26] M. B. Gavela, P. Hernandez, J. Orloff, and O. Pene. Standard model CP violation and baryon asymmetry. *Mod. Phys. Lett.*, A9:795–810, 1994. doi: 10.1142/S0217732394000629.

[27] Ikaros I. Y. Bigi. CP violation: An Essential mystery in nature's grand design. *Surveys High Energ. Phys.*, 12:269–336, 1998. doi: 10.1080/01422419808228861. [Proc. Int. Sch. Phys. Fermi137,451(1998)].

[28] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, 1998. doi: 10.1103/PhysRevLett.81.1562.

[29] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016. doi: 10.1088/1674-1137/40/10/100001.

[30] M. Bona et al. The 2004 UTfit collaboration report on the status of the unitarity triangle in the standard model. *JHEP*, 07:028, 2005. doi: 10.1088/1126-6708/2005/07/028.

[31] M. Bona et al. Model-independent constraints on $\Delta F = 2$ operators and the scale of new physics. *JHEP*, 03:049, 2008. doi: 10.1088/1126-6708/2008/03/049.

[32] L. Wolfenstein. Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 51:1945, 1983. doi: 10.1103/PhysRevLett.51.1945.

[33] C. Jarlskog. Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal CP nonconservation. *Phys. Rev. Lett.*, 55:1039–1042, Sep 1985. doi: 10.1103/PhysRevLett.55.1039. URL `https://link.aps.org/doi/10.1103/PhysRevLett.55.1039`.

[34] C. Dib, I. Dunietz, F. J. Gilman, and Y. Nir. Standard Model Predictions for CP Violation in B0 Meson Decay. *Phys. Rev.*, D41:1522, 1990. doi: 10.1103/PhysRevD.41.1522.

[35] S. Bianco, F. L. Fabbri, D. Benson, and I. Bigi. A Cicerone for the physics of charm. *Riv. Nuovo Cim.*, 26N7:1–200, 2003. doi: 10.1393/ncr/i2003-10003-1.

[36] B. Aubert et al. Evidence for $D^0 - \bar{D}^0$ Mixing. *Phys. Rev. Lett.*, 98:211802, 2007. doi: 10.1103/PhysRevLett.98.211802.

[37] M. Starič et al. Evidence for D^0-$\bar{D}^0$ mixing. *Phys. Rev. Lett.*, 98:211803, May 2007. doi: 10.1103/PhysRevLett.98.211803. URL `http://link.aps.org/doi/10.1103/PhysRevLett.98.211803`.

[38] R. Aaij et al. Measurements of charm mixing and *CP* violation using $D^0 \to K^{\pm}\pi^{\mp}$ decays. *Phys. Rev.*, D95(5):052004, 2017. doi: 10.1103/PhysRevD.96.099907,10.1103/PhysRevD.95.052004. [Erratum: Phys. Rev.D96,no.9,099907(2017)].

[39] F. Buccella, Maurizio Lusignoli, G. Mangano, G. Miele, A. Pugliese, and P. Santorelli. CP Violating asymmetries in charged D meson decays. *Phys. Lett.*, B302: 319–325, 1993. doi: 10.1016/0370-2693(93)90402-4.

[40] H. Mendez et al. Measurements of D Meson Decays to Two Pseudoscalar Mesons. *Phys. Rev.*, D81:052013, 2010. doi: 10.1103/PhysRevD.81.052013.

[41] J. P. Lees et al. Search for CP violation in the Decays $D^{\pm} \to K^0_S K^{\pm}$, $D^{\pm}_s \to K^0_S K^{\pm}$, and $D^{\pm}_s \to K^0_S \pi^{\pm}$. *Phys. Rev.*, D87(5):052012, 2013. doi: 10.1103/PhysRevD.87.052012.

[42] J. P. Lees et al. Search for direct CP violation in singly Cabibbo-suppressed $D^{\pm} \to K^+K^-\pi^{\pm}$ decays. *Phys. Rev.*, D87(5):052010, 2013. doi: 10.1103/PhysRevD.87.052010.

[43] B. R. Ko et al. Search for CP violation in the decays $D^+_{(s)} \to K^0_S\pi^+$ and $D^+_{(s)} \to K^0_S K^+$. *Phys. Rev. Lett.*, 104:181602, 2010. doi: 10.1103/PhysRevLett.104.181602.

[44] B. R. Ko et al. Search for CP Violation in the Decay $D^+ \to K^0_S K^+$. *JHEP*, 02:098, 2013. doi: 10.1007/JHEP02(2013)098.

[45] M. Staric et al. Search for CP Violation in $D^{\pm}$ Meson Decays to $\phi\pi^{\pm}$. *Phys. Rev. Lett.*, 108:071801, 2012. doi: 10.1103/PhysRevLett.108.071801.

[46] R. Aaij et al. Search for CP violation in $D^+ \to \phi\pi^+$ and $D^+_s \to K_S\pi^+$ decays. *JHEP*, 06:112, 2013. doi: 10.1007/JHEP06(2013)112.

[47] R. Aaij et al. Search for CP violation in $D^{\pm} \to K^0_S K^{\pm}$ and $D^{\pm}_s \to K^0_S \pi^{\pm}$ decays. *JHEP*, 10:025, 2014. doi: 10.1007/JHEP10(2014)025.

[48] A. Augusto Alves, Jr. et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008. doi: 10.1088/1748-0221/3/08/S08005.

[49] L. Evans. The large hadron collider. *New Journal of Physics*, 9(9):335, 2007. URL `http://stacks.iop.org/1367-2630/9/i=9/a=335`.

[50] R. Aaij et al. Measurement of the b-quark production cross section in 7 and 13 tev pp collisions. *Phys. Rev. Lett.*, 118:052002, Feb 2017. doi: 10.1103/PhysRevLett.118.052002. URL `https://link.aps.org/doi/10.1103/PhysRevLett.118.052002`.

[51] R. Aaij et al. Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 03:159, 2016. doi: 10.1007/JHEP03(2016)159,10.1007/JHEP09(2016)013,10.1007/JHEP05(2017)074. [Erratum: JHEP05,074(2017)].

[52] P. R. Barbosa-Marinho et al. *LHCb VELO (VErtex LOcator): Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2001. URL `http://cds.cern.ch/record/504321`.

[53] R. Antunes-Nobrega et al. *LHCb reoptimized detector design and performance: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2003. URL `https://cds.cern.ch/record/630827`.

[54] Silicon Tracker UZH group. Lhcb silicon tracker - material for publications. URL `http://lhcb.physik.uzh.ch/ST/public/material/index.php`.

[55] P. R. Barbosa-Marinho et al. *LHCb inner tracker: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2002. URL `https://cds.cern.ch/record/582793`. revised version number 1 submitted on 2002-11-13 14:14:34.

[56] R. Arink et al. Performance of the LHCb Outer Tracker. *JINST*, 9(01):P01002, 2014. doi: 10.1088/1748-0221/9/01/P01002.

[57] R. Aaij et al. LHCb Detector Performance. *Int. J. Mod. Phys.*, A30(07):1530022, 2015. doi: 10.1142/S0217751X15300227.

[58] R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth.*, A262:444–450, 1987. doi: 10.1016/0168-9002(87)90887-4.

[59] A. Papanestis and C. D'Ambrosio. Performance of the LHCb RICH detectors during the LHC Run II. Performance of the LHCb RICH detectors during the LHC Run II. Technical Report LHCB-PUB-2017-012. CERN-LHCb-PUB-2017-012, CERN,

Geneva, Mar 2017. URL https://cds.cern.ch/record/2255885. Updated authors' details and DOI.

[60] C. Lippmann. Particle identification. *Nucl. Instrum. Meth.*, A666:148–172, 2012. doi: 10.1016/j.nima.2011.03.009.

[61] S. Amato et al. *LHCb calorimeters: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2000. URL https://cds.cern.ch/record/494264.

[62] A. A. Alves, Jr. et al. Performance of the LHCb muon system. *JINST*, 8:P02022, 2013. doi: 10.1088/1748-0221/8/02/P02022.

[63] S. Benson, V. V. Gligorov, M. A. Vesterinen, and J. M. Williams. The LHCb Turbo Stream. *J. Phys. Conf. Ser.*, 664(8):082004, 2015. doi: 10.1088/1742-6596/664/8/082004.

[64] Y. Grossman, A. L. Kagan, and Y. Nir. New physics and *cp* violation in singly cabibbo suppressed *d* decays. *Phys. Rev. D*, 75:036008, Feb 2007. doi: 10.1103/PhysRevD.75.036008. URL https://link.aps.org/doi/10.1103/PhysRevD.75.036008.

[65] Ikaros I. Y. Bigi and H. Yamamoto. Interference between Cabibbo allowed and doubly forbidden transitions in $D \to K_S, K_L + \pi$'s decays. *Phys. Lett.*, B349:363–366, 1995. doi: 10.1016/0370-2693(95)00285-S.

[66] A. Davis, L. Dufour, F. Ferrari, S. Stahl, M. A. Vesterinen, and J. Van Tilburg. Measurement of the $K^-\pi^+$ two-track detection asymmetry in Run 2 using the Turbo stream. Technical Report LHCb-INT-2017-023. CERN-LHCb-INT-2017-023, CERN, Geneva, Sep 2017. URL https://cds.cern.ch/record/2284097.

[67] A. Hoecker et al. TMVA - Toolkit for Multivariate Data Analysis. *ArXiv Physics e-prints*, March 2007.

[68] M. Pivk and F. R. Le Diberder. SPlot: A Statistical tool to unfold data distributions. *Nucl. Instrum. Meth.*, A555:356–369, 2005. doi: 10.1016/j.nima.2005.08.106.

[69] S. Stahl, J. Van Tilburg, E. Gersabeck, H. Dijkstra, T. Ruf, and M. Vesterinen. Search for CP violation in $D^0 \to K^-K^+, \pi^-\pi^+$ using semileptonic B decays on 3 fb^{-1}. May 2014. URL https://cds.cern.ch/record/1556318.

[70] B. Lipp. sPlot-based Training of Multivariate Classifiers in the Belle II Analysis Software Framework, 2015. URL https://ekp-invenio.physik.uni-karlsruhe.de/record/48717. KIT, Bachelorarbeit, 2015.

Ringraziamenti

Vorrei usare questo piccolo spazio per esprimere la mia gratitudine nei confronti di tutte le persone che hanno contribuito in un modo o nell'altro questa tesi.

Inanzitutto vorrei ringraziare i miei due angeli, di nome e di fatto, per avermi seguito in questi mesi. Mi ritengo molto fortunata ad aver incontrato Angelo D.C. poiché mi ha fatto appassionare a una fisica che ritenevo noiosa (della serie "chi male capisce, peggio risponde"). Ritengo che il tempo passato assieme sia stato, oltre che piacevole, professionalizzante. Angelo C. mi ha accolto calorosamente nel gruppo di LHCb-Bologna e si è preoccupato che noi laureandi avessimo il massimo dall'esperienza di tirocinio. Lo ringrazio perché, nonostante si fosse messo in disparte durante i lavori, è sempre stato disponibile per ogni tipo di questione. Ringrazio Fabio per il grande aiuto nella fase iniziale, quando dovevo ancora inquadrare il lavoro e nella gestione di Ganga, DaVinci e compagnia bella. Oltre a questo si è aggiudicato il titolo di campione di *debugging*. Sullo stesso podio ci sono anche i ragazzi dell'ufficio, Stefano e Federico. A quest'ultimo va anche il gradito premio per "migliore manutenzione padelle". Ringrazio anche Pietro per gli spunti interessanti, sia tecnici che dal punto di vista dell'organizzazione dell'analisi. Grazie anche ai miei compagni di corso, in particolare ad Alessandra x2 per avermi ospitato più volte e avere permesso a questa povera pendolare un po' di vita bolognese.

A metà (?) fra il lavoro e gli affetti, ringrazio il mio *yesman* Michele per rispondere ai miei no con un sorriso e ancora di più per regalarlo a me. Grazie alle mie migliori amiche Claudia, Catherine e Ilaria e alla "Balera". Infine, grazie ai miei genitori e ai miei fratelli.