Università degli Studi di Ferrara

DIPARTIMENTO DI FISICA E SCIENZE DELLA TERRA
Corso di Laurea Magistrale in Fisica

Tesi di Laurea Magistrale

Test of Lepton Flavour Universality using Bs semileptonic decays

Candidato:
Alessandro Scarabotto

Relatore:
Prof. Massimiliano Fiorini

Correlatore:
Dr.ssa Stefania Vecchi

Anno Accademico 2019-2020

A Caro e ai miei genitori

Abstract

The Large Hadron Collider beauty experiment (LHCb) focuses its studies on measurements of the decay properties of heavy-flavour hadrons, i.e. mesons or baryons containing charm (c) or beauty (b) quarks. LHCb is a single-arm spectrometer situated on the LHC ring with a forward angular acceptance, built this way to detect the majority of the $b\bar{b}$ quarks pair produced in the pp collisions.

One interesting topic LHCb has been studying is related to Lepton Flavour Universality of the Standard Model (LFU) testing. In case this symmetry is violated experimentally, this would be a clear sign of New Physics with contributions that favor couplings with one lepton family rather than the others. LFU tests can be performed in b-hadron decays by comparing branching fractions to final states with different lepton species and also by checking the angular distributions of the decays of interest.

Experimental results on LFU tests performed by Belle, BaBar and LHCb show some tensions with respect to the SM. In particular, measurements of branching fraction ratios of decays involving $b \to c\ell^-\bar{\nu}_\ell$ transitions, as $R(D)$ and $R(D^*)$ show a limited compatibility with the SM predictions of about 3.08σ. Similarly, measurements of branching fraction ratios of decays involving $b \to s\ell^-\ell^+$ rare decays, show a compatibility with the SM predictions of about 2.5σ.

To contribute in clarifying this interesting picture, LHCb is planning and currently performing measurements to test LFU using different b-hadrons decays produced at LHC, including different decay modes. This thesis focuses on the $R(D_s)$ measurement, defined as:

$$R(D_s) = \frac{\mathcal{B}(B_s^0 \to D_s^- \tau^+ \nu_\tau)}{\mathcal{B}(B_s^0 \to D_s^- \mu^+ \nu_\mu)} .$$

The $B_s^0 \to D_s^- \tau^+ \nu_\tau$ decay is reconstructed through the τ three-prong hadronic decay, $\tau^+ \to \pi^+\pi^-\pi^+(\pi^0)\nu_\tau$, while the D_s^- meson is reconstructed through the $D_s^- \to K^+K^-\pi^-$ decay chain, so the visible final state consists of six charged tracks.

In order to achieve the most precise experimental result, the measurement is performed by introducing a normalization channel similar to the signal and redefining the ratio as:

$$R(D_s) = \frac{\mathcal{B}(B_s^0 \to D_s^- \tau^+ \nu_\tau)}{\mathcal{B}(\text{norm})} \times \frac{\mathcal{B}(\text{norm})}{\mathcal{B}(B_s^0 \to D_s^- \mu^+ \nu_\mu)} = \mathcal{K} \times \alpha .$$

Here the ratio α is completely determined by external inputs, i.e. measurements of branching fractions by independent analyses, while the ratio $\mathcal{K}$ is to be determined from the analysis of the signal and normalization channels as:

$$\mathcal{K} = \frac{N_{\text{sig}}}{\epsilon_{\text{sig}}} \frac{\epsilon_{\text{norm}}}{N_{\text{norm}}} \frac{1}{\mathcal{B}(\tau^+ \to \pi^+\pi^-\pi^+(\pi^0)\nu_\tau) \times \mathcal{B}(D_s^- \to K^+K^-\pi^-)} .$$

Here, N_{sig} (N_{norm}) and ϵ_{sig} (ϵ_{norm}) represent, respectively, the signal (normalization) yield and selection efficiency.

The main purpose of this thesis is to choose among possible normalization channels that are suitable for the $R(D_s)$ measurement. As a general requirement, the normalization and the signal channels should be reconstructed and selected in a similar way so that, when measuring the ratio $\mathcal{K}$, any systematic uncertainties related to the reconstruction and selection efficiency will be strongly reduced. The normalization channels studied in this thesis are $B_s^0 \to D_s^- \pi^+ \pi^+ \pi^-$ with $D_s^- \to K^+ K^- \pi^-$ and $B^0 \to D^- \pi^+ \pi^+ \pi^-$ with $D^- \to K^+ \pi^- \pi^-$.

Chapter 1 is an introduction to the LHCb experiment and its apparatus. In Chapter 2, an overview of the LFU and of the experimental tests performed so far are presented, leading to the motivations for the $R(D_s)$ measurement, that is discussed in Chapter 3. Chapter 4 details several aspects of the analysis concerning the signal and normalization channels selection, the efficiency calculation and a quantitative evaluation of the main systematic uncertainties contributing to the $R(D_s)$ measurement. Final conclusions are drawn at the end.

Contents

List of Figures

Chapter 1

The LHCb experiment

1.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the largest particle collider in the world. It was built by the Conseil Européen pour la Recherche Nucléaire (CERN) between 1998 and 2008. The CERN organisation includes more than 10,000 scientists from hundreds of universities and laboratories from more than 100 countries [1].

The accelerator is placed in a tunnel 27 km long, 3.8 m wide and with a depth ranging from 50 to 175 m under the France-Switzerland border near Geneva. The tunnel was previously hosted by the Large Electron-Positron Collider (LEP). In the tunnel two adjacent parallel beam lines, except for the interaction points, are present, each containing a proton beam . LHC has four crossing points at which the proton collisions are performed, around which are positioned four detectors, each of them designed to perform dedicated researches in the particle physics field.

A total of 1232 dipole magnets are used to keep the beams on their circular trajectory, while 392 quadrupole magnets are necessary to keep the beams focused, with stronger quadrupole magnets placed near to the intersection points used to increase the beam density, hence increasing the interaction cross-sections between the particles.

The LHC started its operation in 2010 delivering beams at 3.5 TeV of energy and reaching peaking instantaneous luminosity of $10^{30}\text{cm}^{-2}\text{s}^{-1}$. During the period named Run2 (2015-2018), LHC increased the beam energy to 6.5 TeV (13 TeV total center of mass energy). At the end of 2018, it entered a two-year shutdown period during which many experiments will perform upgrades.

When running at the energy of 6.5 TeV, the protons have a Lorentz factor of about 6,930 and move at about 0.999999990c, which means 11,245 revolutions per second. The protons are grouped together into bunches and up to 2,808 bunches are injected in LHC for each run (which lasts between 5 and 24 hours), with 115×10^9 protons in each bunch. The bunch collision rate is 40 MHz and lead to a maximum instantaneous luminosity of $2 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$.

The particles are accelerated in different steps, as shown in Figure 1.1. The first system is the linear particle accelerator LINAC 2 which accelerates the protons to 50 MeV. In the Proton Synchrotron Booster (PSB) the protons are accelerated to 1.4 GeV and then they are injected into the Proton Synchrotron (PS), where their energy reaches 26 GeV. Finally the Super Proton Synchrotron (SPS) increases their energy to 450 GeV before they are injected into the LHC ring.

Four main experiments are located at LHC: the ATLAS and the Compact Muon Solenoid (CMS) experiments are large general-purpose particle detectors. A Large Ion Collider Experiment (ALICE) and Large Hadron Collider beauty experiment (LHCb) have more specific detector design. The ALICE experiment is focussed on Quantum Chromo-

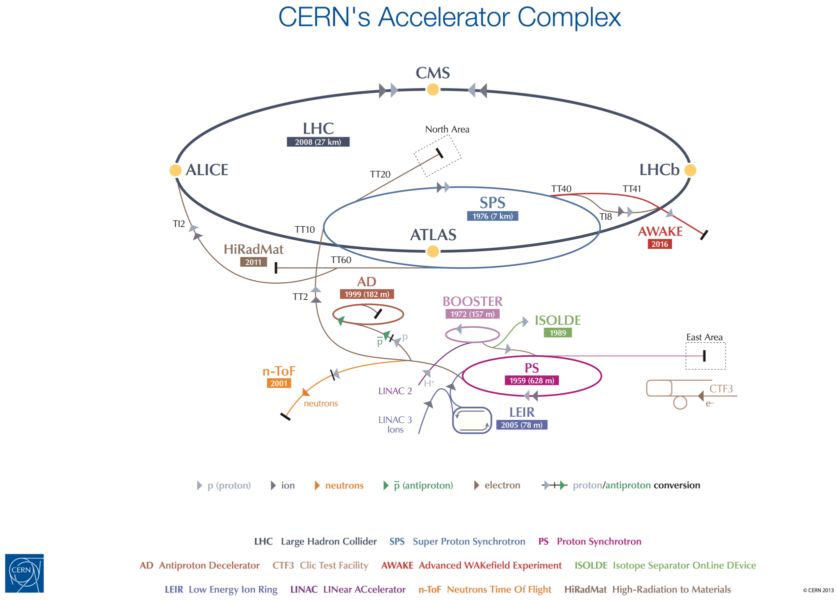


Figure 1.1: Scheme of the CERN accelerator complex [2].

dynamics (QCD) studies, and it has been designed to address the physics of strongly interacting matter and the quark-gluon plasma at extreme values of energy density and temperature in nucleus-nucleus collisions. Besides running with Pb ions, the physics program includes collisions with lighter ions, lower energy running and dedicated proton-nucleus runs. The LHCb experiment is mainly dedicated to the study of the asymmetry between matter and anti-matter present in the universe through precise measurements of decay properties of heavy-flavour hadrons, i.e. mesons or baryons containing charm (c) or beauty (b) quarks.

1.1.1 b-quark production at LHC/LHCb

At the LHC inelastic pp interactions occur in beam pipe. From these collisions pairs of $b\bar{b}$ quarks are produced in several ways, mainly from the fusion of sea gluons and quarks through the "gluon-gluon" and "quark-quark" fusion mechanisms (Figure 1.2).

The inelastic pp cross-section at a centre-of-mass energy of $\sqrt{s} = 7$ TeV and 13 TeV is respectively $\sigma_{\text{inel}} = 66.9 \pm 2.9 \pm 4.4$ mb [3] and $\sigma_{\text{inel}} = 75.4 \pm 3.0 \pm 4.5$ mb [4], expressed with their statistical and systematic uncertainties.

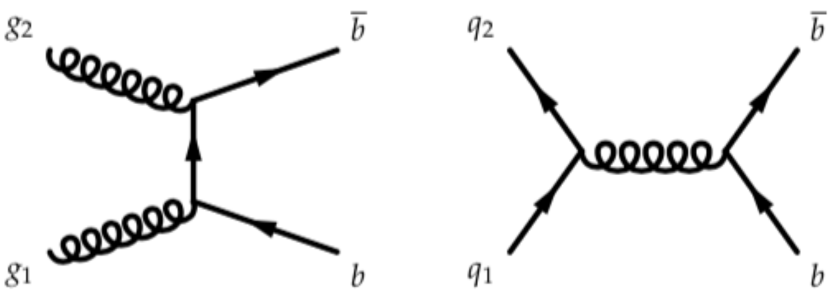


Figure 1.2: Feynman diagrams for the production of a pair $b\bar{b}$ quarks at in pp collisions at LHC.

The peculiarity of the gluon fusion mechanism is that the momenta of the incoming partons are strongly asymmetric in the laboratory frame. As a result, the $b\bar{b}$ quarks are produced inside a narrow cone around the beam in either the forward or backward directions. It is crucial for b-physics studies to design a detector located in either the forward or backward cone, as the $b\bar{b}$ pair will have a great probability to be in the detector acceptance. Figure 1.3 shows the $b\bar{b}$ quarks pair azymuthal angle distribution and the red part of the distribution corresponds to the LHCb detector acceptance: about a forth of the total number of pairs is produced here.

The number of events with a $b\bar{b}$ pair can be calculated as:

$$N_{b\bar{b}} = \sigma(pp \to b\bar{b}X) \cdot \mathcal{L}_{\text{int}}$$

where σ is the cross section and $\mathcal{L}_{\text{int}}$ is the integrated luminosity. The cross-section $\sigma(pp \to b\bar{b}X)$ for 7 and 13 TeV collisions as a function of the pseudorapidity η in the range $2 < \eta < 5$ (covered by the LHCb acceptance) are respectively $72.0 \pm 0.3 \pm 6.8$ μb and $144 \pm 1 \pm 21$ μb [5].

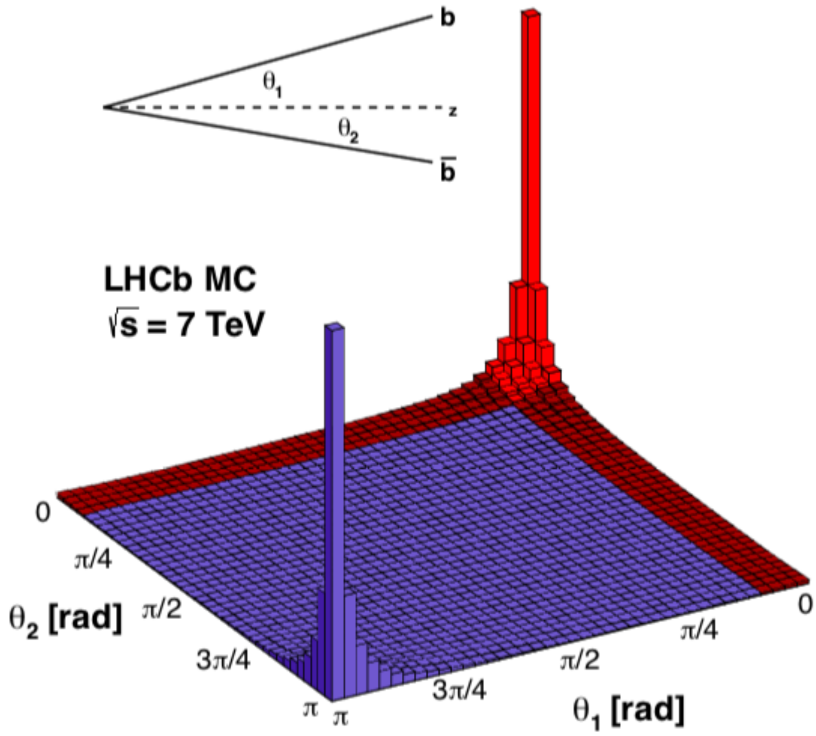


Figure 1.3: Azimuthal angle distribution of the $b\bar{b}$ quark pairs. The red part of the distribution corresponds to the geometrical acceptance of the LHCb detector.

1.2 The LHCb detector

LHCb is a single-arm spectrometer with a forward angular acceptance which covers the range $10-300$ ($10-250$) mrad in the bending (non-bending) plane. The specific geometry of the detector is optimized to study heavy-flavour hadron decays which, at high energy, are produced mainly in a small cone around the beam line [6]. The detector covers a pseudorapidity (η) range of $2 < \eta < 5$.

In Figure 1.4 a schematic drawing of the LHCb experiment is shown. The right-handed coordinate system is a (O, x, y, z) reference frame centered at the collision point, the z axis is along the beam line and the y axis is along the vertical.

The LHCb spectrometer requires a dipole magnet for charged particles momentum measurement. The magnet provides an integrated field of 4 T·m that deflects charged particles in the horizontal plane. Three dipole magnets are used to compensate the LHCb dipole effect on the LHC proton beams and to ensure a closed orbit for them. The tracking system is composed of many sub-detectors: the VErtex LOcator (VELO), a silicon-strip vertex detector surrounding the beam interaction region; the Tracker Turicensis (TT), a large-area silicon-strip detector located upstream of the LHCb dipole magnet; lastly, three stations of silicon-strip detectors placed downstream of the magnet, T1, T2, T3, each of them composed of two separate sub-detector, the Inner Tracker (IT), a silicon micro-strips detector, and the Outer Tracker (OT), a drift-tube gas detector. The charged particle momentum is measured by a high-precision tracking system, with a relative uncertainty which goes from 0.5% at the lowest measurable momentum to 1.0% at 200 GeV/c. The impact parameter, that is the minimum distance of a track to a primary vertex, is measured with a resolution of $(15 + 29/p_{\mathrm{T}})$ μm, where p_{T} is the transverse (with respect to the beam) momentum component (in GeV/c) [7].

Charged hadron identification is performed by two Ring Imaging Cherenkov Detectors (RICH), RICH1 and RICH2. RICH1 is installed upstream the magnet while RICH2 is placed downstream of it. Both RICH detectors uses gas as a radiator and they are mainly designed to separate π from K contribution in the range of momenta p, $2 - 100$ GeV/c, typical for the decays of interest. The ability for LHCb to be able to separate pions from kaons in b-hadron decays is essential to the experiment. A calorimeter system is placed downstream of RICH2 and provides information on the energy of electrons, photons, charged and neutral hadrons, as well as their spatial coordinates. These information are mainly used for particle identification purposes. Lastly, 5 muon stations are installed in the most downstream region of LHCb to detect muons produced in the collisions. Part of the information coming from the tracking, particle identification and calorimetry detectors is quickly processed by a low-level hardware trigger (L0) with dedicated electronics and then by an high-level software trigger (HLT) running on a computer farm. The trigger allows a first selection of the interesting events.

In order to optimise the physics performances, during Run1 and Run2 data taking the LHCb experiment run at an instantaneous luminosity of $2 - 4 \times 10^{32}\,\mathrm{cm}^2\mathrm{s}^{-1}$, one order of magnitude lower than the one achieved for the general purpose experiments (ATLAS and CMS). Figure 1.5a shows the instantaneous luminosity delivered to LHCb. This was achieved by "levelling" the luminosity to an almost constant value during the LHC fill, by adjusting the beam overlap. This allows to maximize the integrated luminosity collected during the fill, to maintain the same trigger configuration and to reduce systematic uncertainties. The integrated luminosity collected by the LHCb to date is about 9 fb^{-1}, and the split per each year of the data taking is presented in Fig. 1.5b.

During the second long shutdown of the LHC (early 2019) the LHCb experiment has started a major upgrade of its detectors with the aim to increase the data taking capabilities. This will be possible by running at an increased instantaneous luminosity of $2 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and improve the the trigger efficiency. To accomplish this, the LHCb hardware-based L0 trigger will be removed. The detector will be read-out at the full rate of 40 MHz and the event selection will be performed by a flexible software trigger running on a hybrid farm of processors (GPUs and CPUs). All sub-detectors of LHCb need to be re-designed to comply with a higher luminosity (5 times the Run2 luminosity) and a higher data acquisition rate.

A Phase-II Upgrade is proposed for the LHCb experiment in order to take full advan-

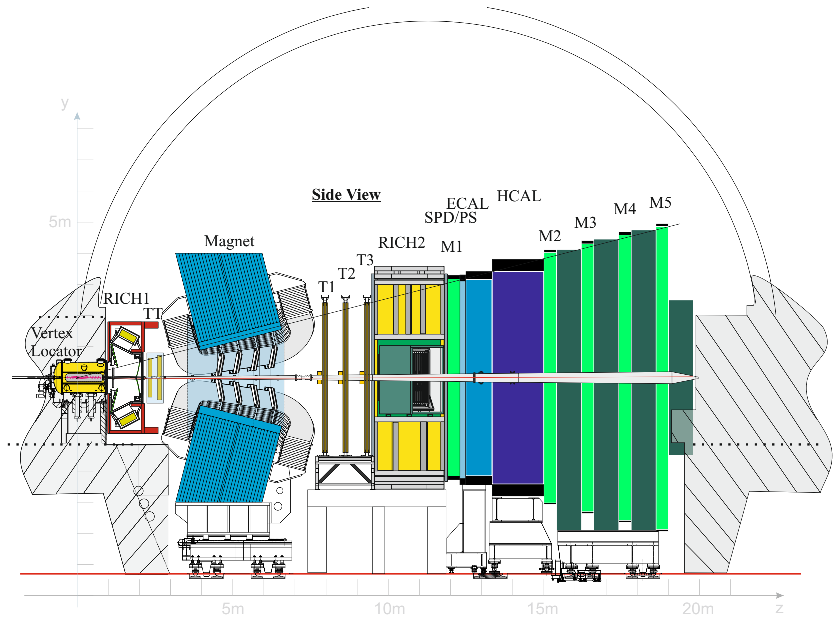


Figure 1.4: Side view of the LHCb experiment.

tage of the flavour- physics opportunities at the High Luminosity LHC, and other topics that can be studied with a forward spectrometer. This Upgrade, which will be installed in Long Shutdown 4 of the LHC (2030), will build on the strengths of the current experiment and the Phase-I Upgrade, but will consist of re-designed sub-systems that can operate at a luminosity of $2 \times 10^{34} \mathrm{cm}^{-2}\mathrm{s}^{-1}$, ten times that of the Phase-I Upgrade detector.

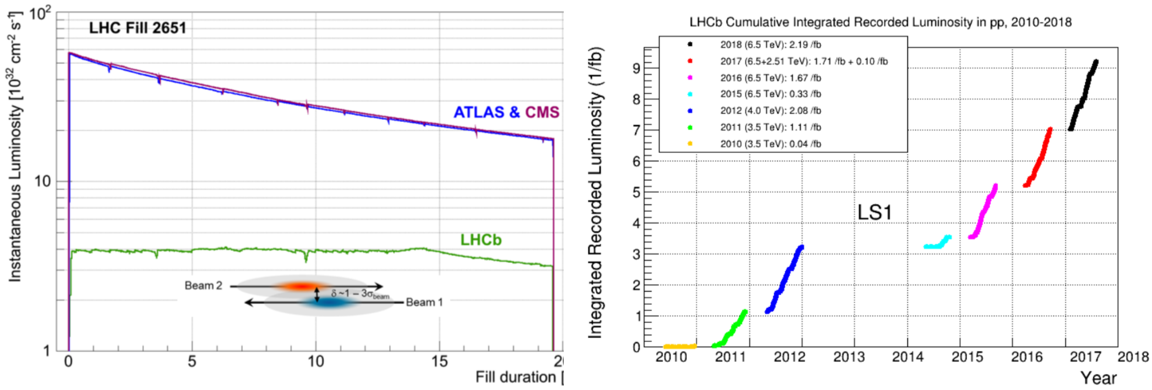
1.2.1 The Tracker

An overview of the LHCb tracking system is shown in Fig. 1.6. It consists of tracking sub-detectors, positioned both downstream and upstream of the magnet in order to measure the momenta of the charged particles.

LHCb uses a dipole magnet composed by two symmetric non-superconductive aluminum coils arranged above and below the beam axis. It has a bending power of about 4 T·m and a peak strength of 1.1 T (see Fig. 1.7); its polarity can be switched in order to minimize systematic effects in the tracking. A charged particle is bent when passing through a magnetic field depending on its velocity. The curvature of the trajectory allows to determine the charge and the momentum of the particle. A preliminary measurement of the transverse momentum of the tracks can be already done using the residual magnetic field present in the TT stations.

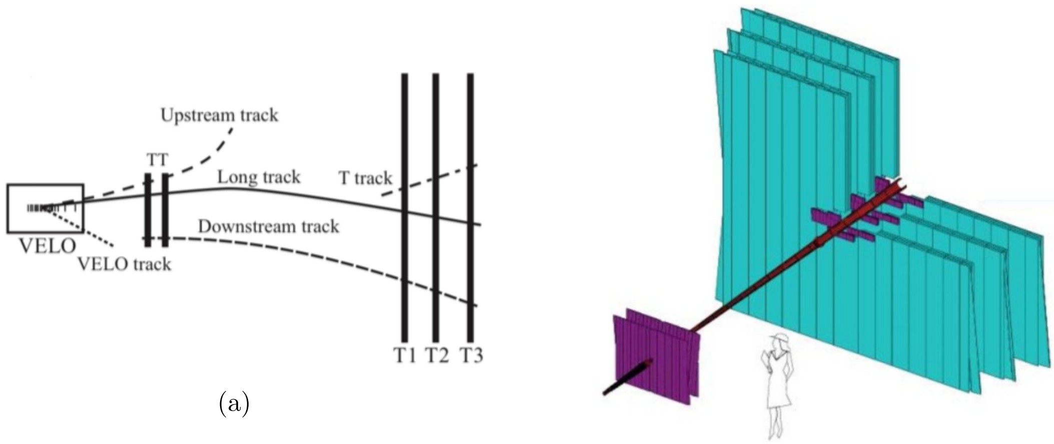
1.2.1.1 The VELO detector

The VELO was installed inside the LHC vacuum vessel to achieve the best vertex resolution possible, and the sensors are separated from the LHC beams only by a 300 μm thin aluminum foil. This is done in order to minimize the material traversed by charged particles before they cross the sensor while protecting the sensors from the radio frequency (RF) waves produced by the beam. The track coordinates are used to reconstruct vertices from long-lived beauty and charmed hadrons (e.g. B and D mesons) as well as the primary pp collision vertex to allow for measurements of decay times and impact parameters. This



(a) (b)

Figure 1.5: (a) Instantaneous luminosity for ATLAS, CMS and LHCb during one long LHC fill in Run1 [6]. (b) Integrated luminosity recorded by the LHCb detector during each year of Run1 and Run2 [6].



(a) (b)

Figure 1.6: (a) Scheme of the LHCb tracking system and track types. (b) Tracking stations of the LHCb detector. Both TT and IT are coloured in violet and the OT is shown in cyan.

information is crucial to discriminate signatures of heavy-flavor decays from background. An excellent vertex resolution is required for CP violation measurements to study the fast $B_s^0 - \bar{B}_s^0$ mesons oscillations.

The VELO covers a length of 1 m and consists of 42 double sided modules supported by a carbon fibre paddle stan [8]. Each module is composed of two layers of silicon micro-strip sensors, one measuring the radial distance from the beam axis (R-sensor) and the other measuring the azimuthal angle (Φ-sensor) see Fig. 1.8.

During the injection of the LHC beams, the VELO modules are moved by 29 mm in the horizontal direction opposite to the beam pipe and are subsequently moved back once the beam is declared stable. Once the VELO is closed, the closest active strip is at 8.2 mm from the beam. This fully automated procedure has been developed to prevent fatal radiation damages to the VELO.

The global performance of the VELO are characterized by a signal noise ratio greater than 14, efficiency of 99 % from tracks that traverse at least 4 sensors, and a spatial resolution of 5 μm for 100 mrad tracks. The resolution on the position of a vertex is

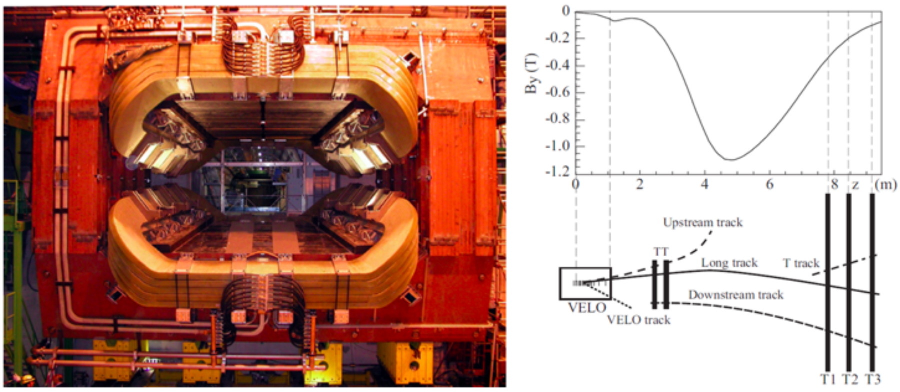
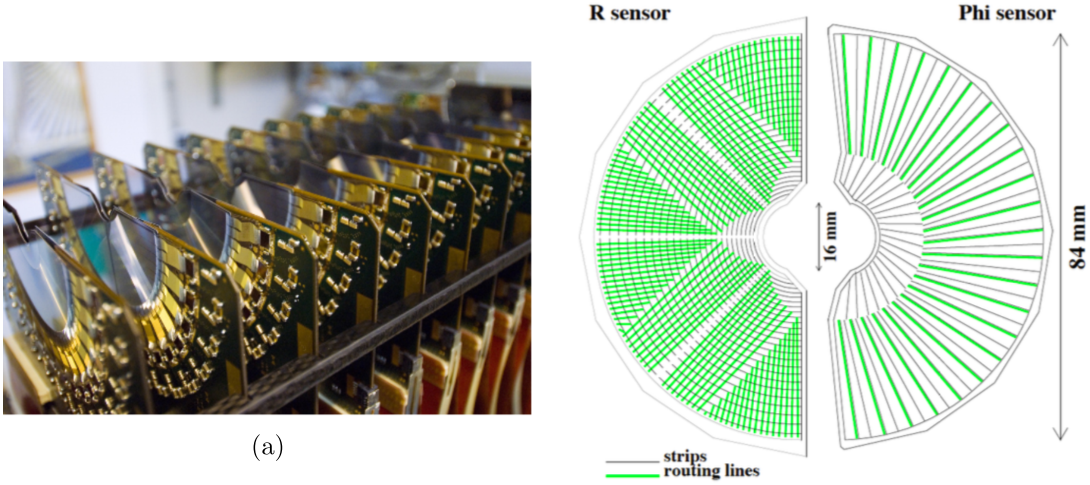


Figure 1.7: (Left) Picture of the LHCb dipole magnet and (right) magnetic field strength plot over the full detector length.



(a)

(b)

Figure 1.8: (a) Picture of the VELO silicon modules. (b) Schematic for R and Φ sensor for a VELO module.

13 μm in the (x, y) plane and 71 μm along the z axis. The resolution $\sigma_{\rm IP}$ on the impact parameter (the minimum distance of a track to a primary vertex) is a function of the transverse momentum of the track, $p_{\rm T}$, as shown in Fig. 1.9.

1.2.1.2 The Trigger Tracker

The upstream detector is the Trigger Tracker (TT) and consists of four planar layers of silicon micro-strips detectors (Figure 1.10). A total of 143 thousand strips are present, with a length of 40, 30, 20 or 10 cm and with a pitch of 183 μm [9]. The strips will run vertically, since the tracks are bent in the horizontal plane by the dipole magnet. Shorter strips were employed in the inner region of each detection layer, where particle densities are highest, while longer strips were used in the outer region. The TT total active area is approximately 8m^2, 130 cm high and 150 cm wide. The four detection layers are arranged in a $x-u-v-x$ layout, with vertical strips in each of the two x layers, and strips rotated by a stereo angle of $-5°$ and $+5°$ in the u and v layers, respectively.

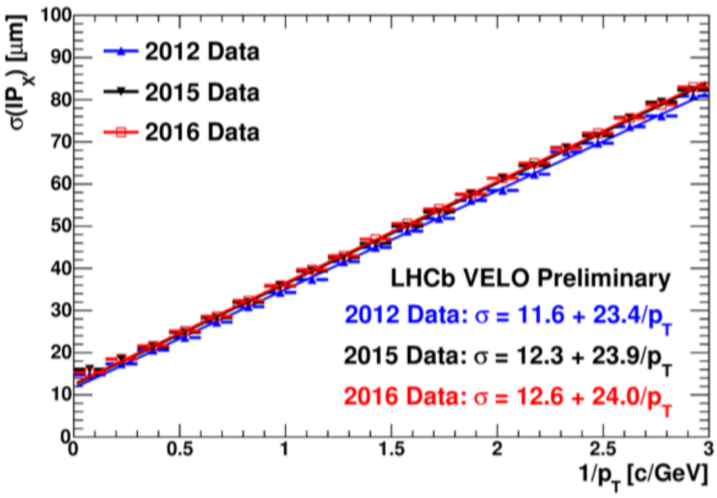


Figure 1.9: Resolution of the x-coordinate of the impact parameter vs. $1/p_{\mathrm{T}}$ tracks measured using 2012, 2015 and 2016 data.

1.2.1.3 The T-Stations

Three tracking stations (T1-T3) are located downstream of the magnet. As the particle flux is much higher in the central region than in the outer one, two detector technologies are employed. The IT is a silicon micro-strip detector (Figure 1.10) with strips of 198 μm pitch. The IT detector covers a high cross-shaped region in the center of the three tracking stations T1-T3, with width of 125.6 cm and height of 41.4 cm, for a total of approximately $4\,\mathrm{m}^2$. The OT (Figure 2.8) is a straw drift-tube detector consisting of approximately 200 gas-tight modules with drift-time read-out. The OT has four layers placed in an $x{-}u{-}v{-}x$ layout, with a total active area of $597 \times 485\,\mathrm{cm}^2$. It is present in the outer region of the tracking stations. OT and IT have four detection layers per station, both following the $x - u - v - x$ layout.

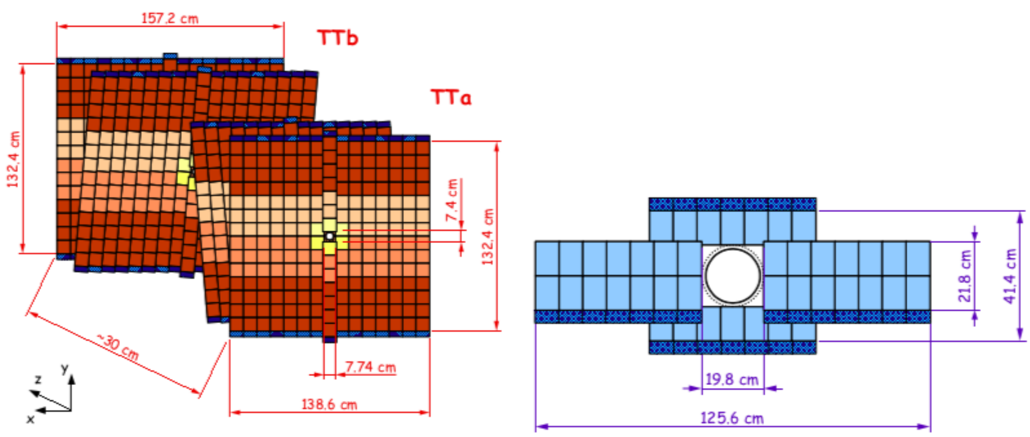


Figure 1.10: The two silicon tracker detector composing the tracker system: (left) TT and (right) IT.

Charged particles require a minimum momentum of 1.5 GeV/c to reach the tracking stations T1-T3. Tracking performances were measured in data during Run1 [6]. The momentum resolution is measured to be $\sim 0.4\%$ at 5 GeV/c up to 0.6% at 100 GeV/c

and the track reconstruction efficiency is above 96% in both 2011 and 2012 data taking periods (see Fig. 1.11).

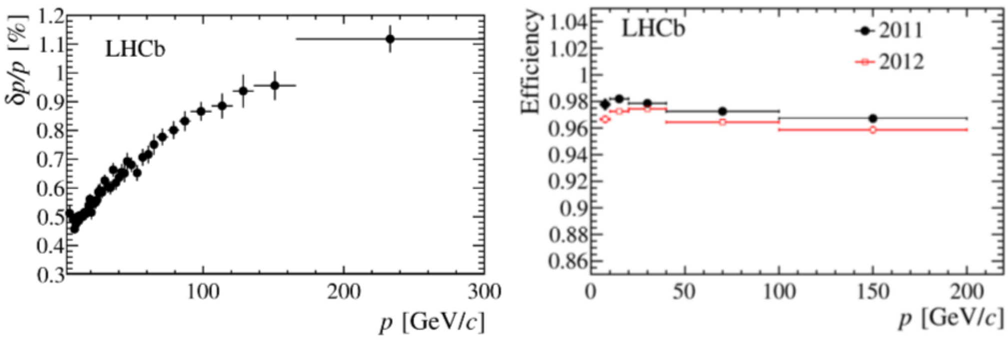


Figure 1.11: Relative momentum resolution for long tracks (left) and tracking efficiency (right) as a function of the momentum.

Upgrade

For the LHCb Upgrade the VELO detector will be replaced by a a pixel detector (55×55 μm^2 silicon pixel sensors) [10]. The upgraded VELO will use thinner sensors and a thinner aluminum cover per each detector half. Also, in the closed position it will be 5.1 mm distant to the beam, 3.1 mm closer with respect to the Run1-2 position. With these features, the detector will have a more robust track reconstruction performance and an improved resolution with respect to the Run1-2 VELO.

The TT has to be replaced during the LHCb Upgrade. The currently employed silicon sensors would not sustain the high radiation dose in the post-upgrade LHCb environment, especially the inner detector region. With the actual read-out strip geometry the detector occupancy would be too high in the foreseen luminosity conditions. The new Upstream Tracker (UT) will be installed during the upgrade. With this detector it will be possible to reconstruct long lived particles decaying after the VELO. The UT will be composed by 4 tracking layers employing silicon strip technology. The same $x - u - v - x$ layout as in TT will be used. In the outer region, the strips will be 99.5 mm long with 180 μm pitch. The strips in the central region will be as long as the outer ones but 95 with μm pitch. Lastly, the closest strips to the beam will be 51.5 mm long, with 95 μm pitch.

In order to cope with the higher occupancy in the post-upgrade running conditions the OT stations will be replaced by three stations of Scintillating Fibre Tracker (SciFi Tracker/SFT). The main requirements on the SciFi tracker are: a hit detection efficiency of larger than 98 % with a noise cluster rate of less than 10 %of the signal cluster rate and a spatial resolution for single hits better than 100 μm.

1.2.2 RICH detectors

The Ring Imaging Cherenkov (RICH) detectors provide particle identification information that is particularly important to distinguish between different types of charged hadrons, i.e. pions, kaons and protons.

RICH detectors rely on the Cherenkov effect: a particle propagating through a material (radiator) with a refractive index n, with a velocity v larger than the speed of light in this medium, c/n, emits photons in a cone of aperture θ around the particle's direction of propagation. The value of θ is directly related to the velocity of the particle ($\beta = v/c$)

through the formula:

$$cos(\theta) = \frac{1}{\beta n}.$$

As the tracking system provides the momentum estimation, the measurement of θ gives access to the velocity of the particle and thus to its mass (its identification).

Two Ring Imaging Cherenkov detectors (Figure 1.12) are responsible for charged hadron identification in the momentum range 2 to 100 GeV/c [11]. The upstream detector, RICH1, is designed to detect low momentum particles, from about 2 to 60 GeV/c, and uses C_4F_10 gas as radiator. The downstream detector, RICH2, is designed to cover the high momentum range, from 15 to 100 GeV/c, and uses CF_4 gas as radiator medium. The angular acceptance of RICH1 is from 25 mrad to 300 mrad (horizontal) and 250 mrad vertical, covering the LHCb acceptance. The angular acceptance of RICH2 is limited, covering the range from 15 mrad to 120 mrad (horizontal) and 100 mrad (vertical). This difference is due to fact that only high momentum particles reach RICH2 as soft particles are mainly deflected by the magnet out of the acceptance. In both RICH1 and RICH2 the Cherenkov photons produced in the gas are reflected by a system of plane and spherical mirrors and focused in ring images onto the photon detection planes, where the Hybrid Photon Detectors (HPDs) are installed. The HPDs are position-sensitive in order to reconstruct the typical ring images: measuring the radius of the rings allows the Cherenkov θ evaluation.

Figure 1.13 shows the particle identification performance of the RICH detectors to different particle species.

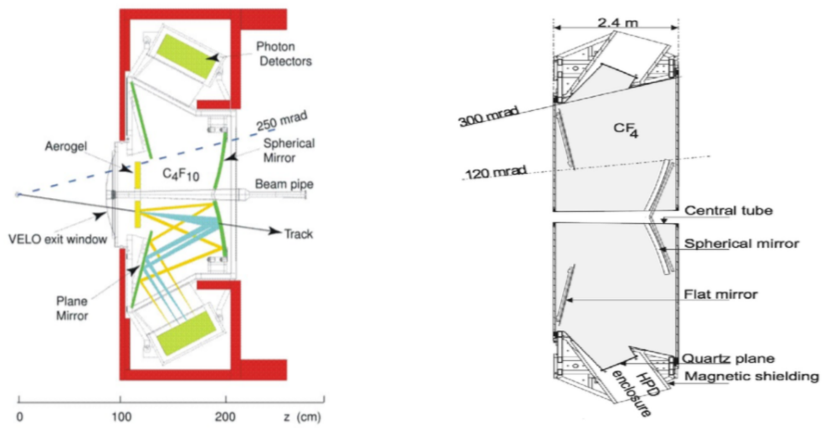


Figure 1.12: Schematic for (left) RICH1 and (right) RICH2 detectors in LHCb.

Upgrade

For the upgrade, the LHCb RICH photodetectors will be replaced, since the actual Hybrid Photodetectors (HPDs) have the 1 MHz read-out electronics embedded in the tube. The HPDs will be replaced with commercial Multi-Anode Photomultiplier tubes (MaPMTs) coupled to external read-out electronics. In addition, for the RICH1 a new optical system will be installed. This new system will increase the Cherenkov rings projection area onto the photodetectors planes and reduce the peak occupancy. For RICH2 other small mechanical mounting and electronics will be adjusted to comply with the increased luminosity.

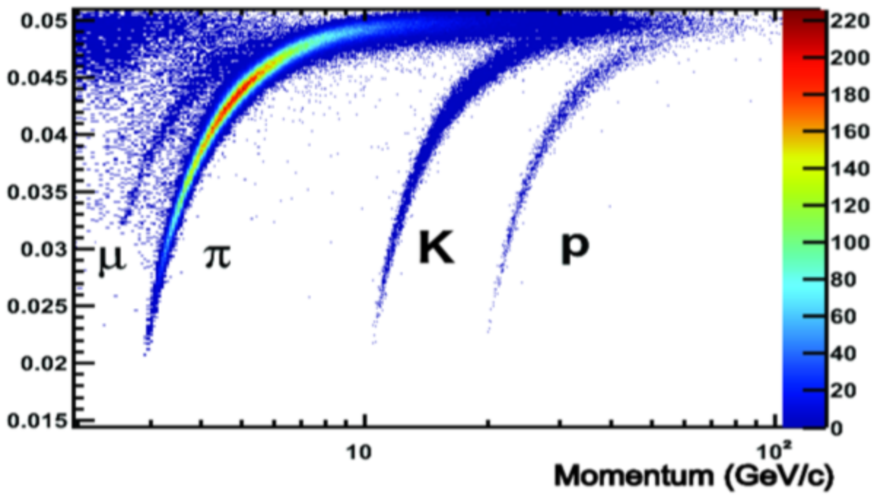


Figure 1.13: Reconstructed Cherenkov angle as a function of the charged track momentum in the C_4F_{10} gas radiator.

1.2.3 The Calorimeter System

The calorimeter system provides the measurement of energies and positions of electrons, photons and hadrons which is used by the L0 trigger, to select candidates with high transverse energy [12], and for particle indentification. It is composed of several subdetectors: Scintillating Pad Detector (SPD), PreShower (PS), electromagnetic calorimeter (ECAL), and hadronic calorimeter (HCAL).

The ECAL is a heterogeneous lead/scintillator sampling calorimeter, employed to measure the energy of photons and electrons. It is located after the RICH 2 but before the HCAL. The separation between photons and electrons is made by the SPD and PS. In fact, the electrons produce a shower when traversing the 12 mm lead layer between the SPD and PS, while photons do not. Once the particles have been identified by the SPD and the PS, their energies are determined by the ECAL with an energy resolution of:

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 1\%\,,$$

where the symbol $\oplus$ indicates sum in quadrature of the two terms.

In the HCAL neutral and charged hadrons deposit their remaining amount of energy. It is made by alternating iron absorbers and scintillating tiles. The energy resolution is given by:

$$\frac{\sigma_E}{E} = \frac{70\%}{\sqrt{E}} \oplus 10\%\,.$$

As the expected particle flux varies by two orders of magnitude between the region surrounding the beam pipe and the outermost region, different segmentation schemes were engineered. They are both presented in Fig. 1.14. Both SPD, PS and ECAL detectors are divided in three inner, middle and outer regions whereas the HCAL is only divided in two inner and outer regions.

Upgrade

The actual ECAL and HCAL granularity is sufficiently high to operate at the post-upgrade luminosity conditions, therefore both systems (modules, HV system and PMTs) will be kept as they are now. On the contrary, the PS/SPD systems, together with the lead

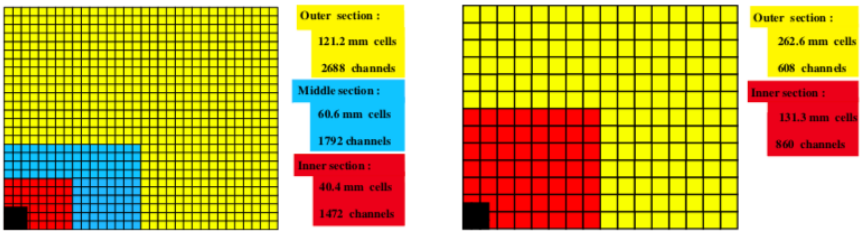


Figure 1.14: Modules dimension and granularity of the (left) ECAL and (right) HCAL. Only the top right quarter is shown. The black area corresponds to the empty space occupied by the beam pipe.

converter, will be removed. The main role of the PS/SPD systems is to provide fast particle identification information to L0 trigger. With the upgrade, L0 will be removed hence the preshower detector can be removed.

1.2.4 The Muon Detector

Muon identification is crucial to LHCb experiment, as muons are present in the final states of many B decays. Moreover, muons from semileptonic b-hadron decays provide a tag of initial state flavor of accompanying neutral B mesons, which is a key ingredient for measurements of time-dependent CP violation. In addition, the study of rare B decays such as flavor-changing neutral current decays, may reveal new physics beyond the standard model.

The muon system of LHCb consists of five stations M1 to M5 installed along the beam axis and provides information for the high transverse momentum muon trigger at level 0 and muon identification for the high-level trigger and offline analysis [13]. The five stations are made of Multi Wire Proportional Chambers (MWPC) with the exception of the inner part of M1, where the particle flow is higher, which is made of triple Gas Electron Multiplier (GEM). The M1 stage, placed in front of the calorimeters, improves the transverse momentum measurement in the trigger. The muon trigger relies on muon track reconstruction and requires a hit in all five stations. The other stations are located at the end of the line just behind the HCAL. The M2 to M5 stations are interleaved with 80 cm thick iron absorbers in order to stop hadrons. The total thickness of the full muon system corresponds to 20 nuclear interaction lengths which means that only muon with a minimum momentum of 6 GeV/c can cross the five stations. Stations M2 to M4 have a higher spatial resolution in the x coordinate which is used to define track direction and calculate the transverse momentum of muon candidates with a resolution of 20 %. Figure 1.15 shows the organization of the five stations. Each detector is split into rectangular logical pads whose dimensions define the x and y resolution.

Upgrade

The LHCb Muon detector Upgrade will allow 40 MHz detector readout and events will be selected by means of a very flexible software-based trigger. The muon system will be upgraded in two phases. In the first phase, the off-detector readout electronics will be redesigned to allow complete event readout at 40 MHz. In a second phase, higher-granularity detectors will replace the ones installed in highly irradiated regions, to guarantee efficient muon system performances in the upgrade data taking conditions.

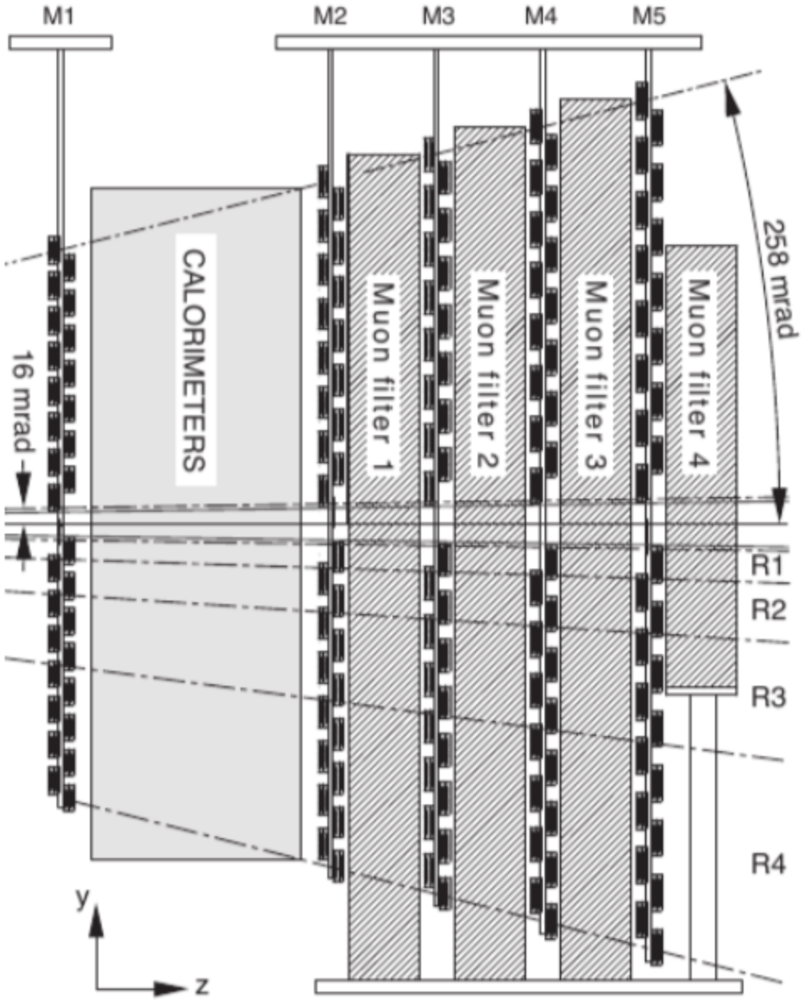


Figure 1.15: Side view of the five stations of the muon detector.

1.2.5 Particle Identification variables

Information gathered by the tracking system, the two RICH detectors, the calorimeter system and the muons chambers are combined by dedicated algorithm to provide a particle hypothesis for a given track.

The hadron identification relies on the calorimeter system and is further refined to separate between π and K using the two RICH detectors. Discrimination between electromagnetic particles γ and $e^{\pm}$ exploits the information provided by SPD, PS and ECAL detectors and muons are identified by using the muons detector. The RICH information is also used to identify electrons and muons.

Two different methods are used to provide particle identification variables in offline analyses: the first uses a likelihood-based technique and the second relies on a Machine Learning algorithm.

For each particle, the likelihood for a given particle hypothesis is evaluated with the following formula:

$$\mathcal{L}(h) = \mathcal{L}^{RICH}(h) \times \mathcal{L}^{CAL}(\text{non e}) \times \mathcal{L}^{MUON}(\text{non}\,\mu)\,,$$

where $\mathcal{L}^{CAL}(\text{non e})$ and $\mathcal{L}^{MUON}(\text{non}\,\mu)$ are respectively the likelihood given by the calorimeter system for a given track to not be identified as an electron, and the likelihood provided by the muon system for the track to be discarded as a muon.

The particle hypotheses are compared to the π one, the most abundant particle at LHCb. The difference of the log-likelihood are computed as follows:

$$\mathrm{PID}h = ln\mathcal{L}(h) - ln\mathcal{L}(\pi) = ln\frac{\mathcal{L}(h)}{\mathcal{L}(\pi)},$$

where h stands for either p, K, π, e or μ hypotheses.

The LHCb PID performances for the Run1 data taking period are excellent, with PIDK>0 giving an average K efficiency (K $\rightarrow$ K) of 95 % and a π misidentification ($\pi \rightarrow$ K) of 5 % (see Fig.1.16) [6].

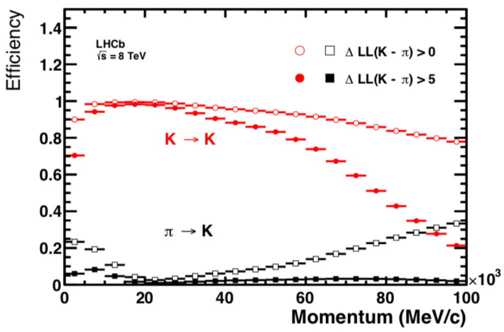


Figure 1.16: PID performances for Kaons during 2012 data taking. Kaon efficiency (K $\rightarrow$ K) and π misidentification ($\pi \rightarrow \pi$) for two different PIDK selection, PIDK >0 with open markers and PIDK >5 with filled ones.

A Neural Network (NN) algorithm is also used to further optimize particle identification performances. The PID information is combined in such algorithm and trained separately for each type of particle to be identified with simulation samples. The output of each Neural Network, called ProbNN, is a probability of a given track to be identified as the particle associated with this particular Neural Network. There are then ProbNN(K, e, π, p, μ) variables to be used in offline analyses. Such variables are able to take into account correlations between different PIDh variables and the information from the different sub-detectors resulting in better performances than PIDh as illustrated in Fig. 1.17 for both μ and p.

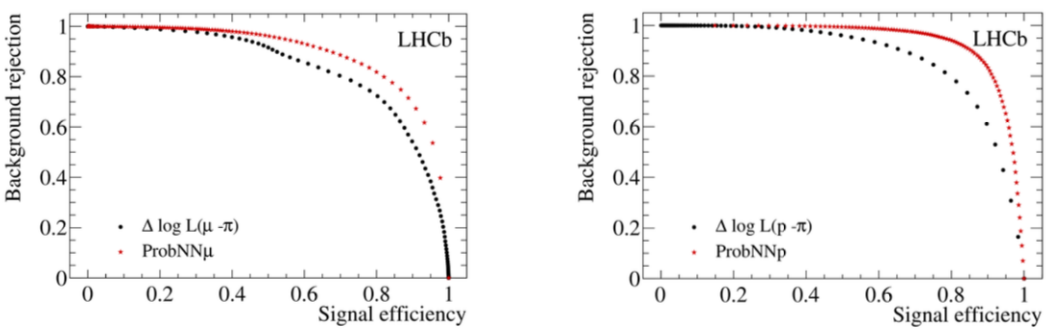


Figure 1.17: Comparison between PIDh ($\Delta ln\mathcal{L}(h-\pi)$) and ProbNN$h$ variables for both (left) μ and (right) protons.

1.2.6 The LHCb Trigger System

The LHCb trigger system is used to lower the data output rate from the beam crossing rate of 40 MHz down to a few kHz. Its architecture consists of two levels, the first level trigger (L0) and the High Level Trigger (HLT). The L0 trigger is hardware implemented, while the HLT is a software application implemented in a dedicated computing farm, where the event rate is reduced to around 12.5 kHz. The full trigger system is sketched in Fig. 1.18 (for Run1).

The L0 trigger receives information from the calorimeter and muon system to reduce the event rate from the bunch-crossing rate of 40 MHz to 1.1 MHz, at which frequency the entire detector is readout.

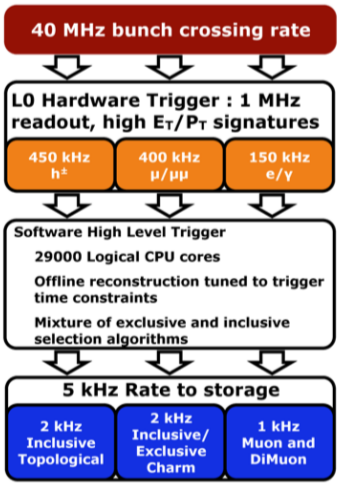


Figure 1.18: Logical sketch of the LHCb trigger for Run1.

To achieve the goal of rate reduction a decision has to be reached by the hardware implementation in less than 4 μs. The L0 trigger decisions are based exclusively on information from the muon system and calorimeters, as these are the only pieces of information available at such high rate. Events with either high p_{T} muons or large transverse energy deposits in the calorimeter are selected by the L0 trigger.

The muon trigger selects events containing either a single muon with $p_{\mathrm{T}} > 1.76$ GeV/c or a pair of muons with $p_{\mathrm{T}1} \times p_{\mathrm{T}2} > (1.6\ \mathrm{GeV}/c)^2$. The output rate of the L0 muon trigger is about 400 kHz.

The hadron trigger selects events in which a significant amount of transverse energy ($E_{\mathrm{T}} > 3.68$ GeV) is deposited in the hadronic calorimeter. Electrons and photons trigger the event selection by a deposit of $E_{\mathrm{T}} > 3$ GeV in the electromagnetic calorimeter. The output rates of the hadron and electromagnetic triggers are 450 kHz and 150 kHz, respectively.

In the L0 trigger there are several types of objects, e, γ, h, μ and $\mu\mu$ that can fire for the event to be triggered on. The corresponding L0 trigger lines are L0Electron, L0Photon, L0Hadron, L0Muon and L0Dimuon.

The events are selected in two ways by the hardware trigger:

- TOS: the signal candidate triggers the L0 trigger lines (trigger on signal, TOS);

- TIS: Any particle in the event, irrespective if it is part of the signal candidate or not, passes any of the L0 trigger lines (trigger independent of signal, TIS).

The software trigger of LHCb, referred to as the HLT, is made of two stages: HLT1 and HLT2. Whereas a partial reconstruction is performed at the HLT1 level using information from the vertexing and tracking system, a full reconstruction of the event is made at HLT2 level using information from all the sub-detectors.

The HLT1 performs the reconstruction of the primary vertices (PVs) and the search for secondary vertices. A preliminary track reconstruction is also performed. PVs are reconstructed by searching for at least five tracks coming from the same point with an uncertainty of $\pm$ 300 μm along the z axis.

The track impact parameter with respect to the primary vertex can be determined and used to trigger events. A typical signature of a particle track originating from the decay of a b-meson is a large impact parameter with respect to the primary vertex of the particle. The impact parameter (IP) is the distance of closest approach between a reconstructed track and the true origin of the particle. The vertex corresponding to the minimum IP value is set as the origin vertex of the track. Therefore, it is also possible to determine if the track is originating from a primary or a secondary vertex.

At HLT1 level, the track reconstruction is performed only partially, mainly on high p_{T} tracks: these are usually the ones produced from heavy particles as the b-hadrons. A preliminary measurement of the transverse momentum of the tracks is done using the residual magnetic field present in the TT stations.

A full reconstruction of the event is done at the HLT 2 level which is completely configurable. A large number of selection lines in the HLT2 trigger selection are tailored for specific physics analyses. There are exclusive lines dedicated to a particular decay as well as inclusive lines based on topological requirements. The selections applied by these lines are either simple cuts or more complex methods using multivariate analyses such as Boosted Decision Trees. Overall, the HLT-trigger reduces the incoming 1 MHz rate at the output of the L0-trigger to a few kHz rate which is recorded on disk.

Upgrade

In order to be able to read-out the LHCb experiment at the 40 MHz bunch-crossing rate, the L0 trigger has to be removed during the upgrade [14]. It will be replaced with a software based Low Level Trigger (LLT) which will use limited information from muon and calorimeter systems to select high p_{T} candidates in the decays of interest. The output rate of LLT will be in the range 15-30 MHz. A fully software based HLT will be responsible for event reconstruction and inclusive and exclusive selections based on the decays kinematics. The HLT will also provide information to perform a run-by-run spatial alignment and calibration of the LHCb detectors.

1.3 The LHCb Software

The detection of pp collisions in LHCb is only the first step to perform physics analysis. In the following some information are given about the simulation, the reconstruction and other technicalities useful to perform analyses within the LHCb experiment.

The production of simulated Monte Carlo (MC) samples involves several phases such as generation, digitization, reconstruction and several software packages are involved. The first three steps are unique of the MC production, the reconstruction is in common with the real data processing.

The software framework Gaudi [15] is used for the generation and data processing. It ensures availability of all LHCb tools and algorithms across the experiment and enforces global consistency of LHCb analyses.

The main components of the LHCb data processing can be separated in (see Fig. 1.19):

- Generation: The LHCb software package Gauss [16] is responsible for the generation of simulated events. In this step, the primary *pp* interactions are simulated using PYTHIA [17], while the decays of *b*-hadrons are simulated with the EvtGen package [18].

- Simulation: The interaction of the generated particles with the detector is simulated with Geant4 toolkit [19].

- Digitization: The response of the LHCb detector to the generated particles is emulated by the Boole application [20]. The digitized output is the MC-equivalent of real data. From this point on, both data and MC pass through the same reconstruction chain.

- Reconstruction: The Brunel application [21] handles the reconstruction of the digitized output of the detector responses. Signal hits in the detector are clustered and associated tracks are constructed. Particle properties such as momentum and PID information are evaluated. The output of Brunel is a full reconstructed dataset stored in "Data Storage Tape" (DST) files.

- Analysis: Offline analyses are performed using the DaVinci software package [22]. The DaVinci application allows for reconstructed particles to be combined and determines primary and secondary vertices as well as kinematic quantities, such as invariant masses of decayed particles and their distance of flight or decay time. It also allows the reconstruction and selection of the decays of interest as well as offline analysis tools.

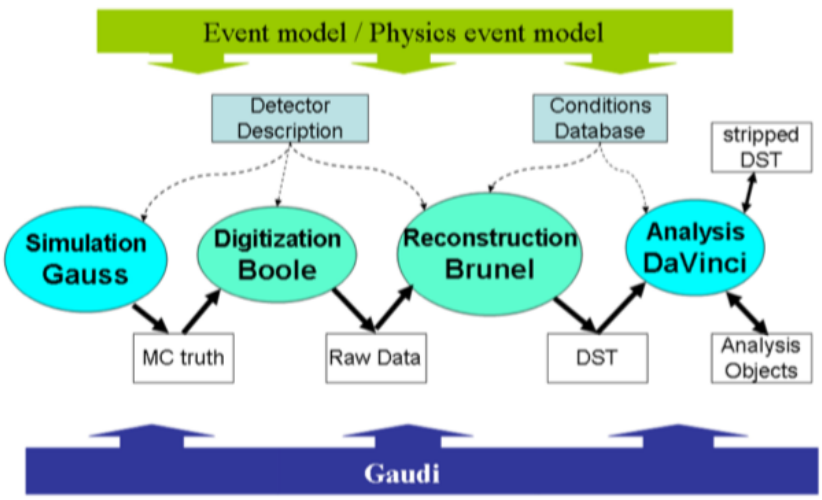


Figure 1.19: The LHCb data processing and data flow. Underlying all the applications is the Gaudi framework and the event model describes the data. The arrows represent the input/output data.

1.3.1 Track reconstruction

The tracks (the trajectories of charged particles) are reconstructed in LHCb combining the hits (measured positions of the intersections between tracks and detectors) coming from different tracking detectors (VELO, TT, IT, and OT).

The tracks are classified as (see Fig. 1.6a):

- VELO tracks: signatures only in the VELO detector. They are useful to determine the primary vertex and are reconstructed as straight line since they do not pass through the magnetic field.
- Upstream tracks: they contain hits in the VELO and in the TT detectors. They are usually low momentum tracks that are bent out of the LHCb detector acceptance by the magnetic field.
- Downstream tracks: tracks reconstructed with the TT and the T1-T3 tracking stations. They are useful to reconstruct long lived particles such as K_S and Λ which often decay after the VELO.
- Long tracks: they have hits in all sub-detectors, from the VELO to the T1-T3 stations. They have the most precise momentum measurement and thus they are the most important for B physics measurements.

The reconstruction of the tracks is divided in different steps:

1. The first step of the track reconstruction is the **pattern recognition** in which independently in the VELO and in the T-stations sequences of hits are grouped together and identified as coming from the same track. These VELO tracks and T tracks are then used as input to find long, upstream and downstream tracks. The long tracks are found either by extrapolating the VELO tracks into the T station and looking for matching hits, or by matching directly the VELO tracks with T tracks; TT hits are then added. Upstream tracks are found by extrapolating VELO tracks into the TT while the downstream tracks combine T tracks with TT information.
2. The second step of the track reconstruction is the **track fitting** that in LHCb is done with a Kalman filter [23] that takes into account effects from multiple scattering and energy loss due to ionization.
3. Mis-identified tracks (composed of pseudo-random combinations of hits), called ghost tracks, are then removed based on the **ghost probability**, the output of a neural network that takes as input various variables such as the track χ^2 and the number of hits in each subdetector.
4. The last step of the track reconstruction is the **clone killing** which consists of removing tracks that are also sub-tracks of other tracks, for example a VELO track used to build a long track.

The reconstruction efficiency for charged tracks that pass through the full tracking system varies as function of the kinematics of the tracks and the occupancy of the detector. It is above 95 % as shown in Fig. 1.20.

1.3.2 Selection

In an LHCb analysis the selection usually involves three steps:

- Trigger selection: decides which events are interesting enough to be saved for later analysis.
- Stripping selection: is the centralized selection of interesting events after the reconstruction that is applied to store only interesting events and save disk space. A stripping line contains the instructions for reconstructing the decay of interest from the reconstructed stable particles and the corresponding selections to be applied. A group of stripping lines selecting similar events is collected into a stream and their content are stored in a common file.

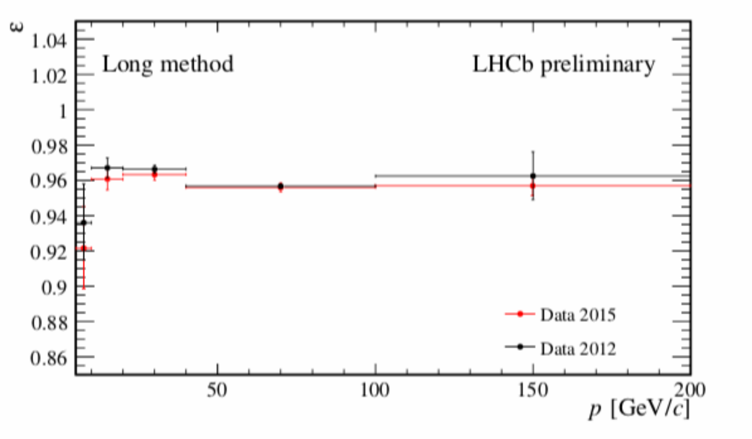


Figure 1.20: Track reconstruction efficiency for long tracks for 2012 and 2015.

- Offline selection: is the term used to refer to the end-user selection on the events which passed trigger and stripping selections.

1.4 Physics at LHCb

The successful operation of the Large Hadron Collider has led to a transformation of the high energy physics landscape [24]. The Higgs boson discovery completes the Standard Model (SM), a self-consistent theory that in principle could be valid up to the Planck scale [25, 26]. The calculation of the Higgs mass using the Quantum Field Theory of the Standard Model shows that it receives contributions from all energy scales, all the way up to the highest energy scale λ at which the Standard Model is valid. The most obvious choice is thus the Plank mass. The difference of many orders of magnitude between $M_{W,Z,H} \sim 100$ GeV and $M_{\mathrm{Planck}} \sim 10^{19}$ GeV is the hierarchy referred to in the name of "The Hierarchy Problem". This vast difference in scale, on the order of $\sim 10^{17}$ GeV, has been predicted in theories, like the super-symmetry, that have not been discovered yet.

Measurements of the Higgs boson's properties, like its quantum numbers, production and decay modes, are consistent with the SM predictions, validating the model even more. The precision constraints measurements of W boson and top quark parameters like masses and effective weak mixing angle are in agreement with the SM. The current experimental studies are consistent with the SM and the CKM quark mixing matrix includes the possibility of CP violation, with a sole matter-antimatter asymmetry source found [27, 28]; lastly, all determinations of the Unitarity Triangle properties fit together within the SM with the current level of precision. The Unitarity Triangle is the graphical representation on the complex plane of the unitarity constraints on the off-diagonal terms of the CKM matrix.

However, the search for physics beyond the SM has to continue. The hierarchy problem may be resolved in a theoretically way, at an energy scale so high that could not have been probed by any experiment. There is the need to understand the 95 % of the Universe energy density that is not ordinary baryonic matter. The matter-antimatter asymmetry cannot be fully explained in the SM.

The LHCb experiment achieved amazing results during the Run1 and Run2 data taking. LHCb is a specialized b-physics experiment, designed primarily to measure the parameters of CP violation in the interactions of b hadrons. Its flexible data-collection strategies have expanded its physics program, for example allowing studies of dark sectors and collecting data from heavy ion and fixed target collisions. LHCb gave also an exceptional contribution in the hadronic spectroscopy field, discovering new states as the tetra- and penta-quark.

Possible unknown particles could contribute to the decay processes that govern the heavy-flavour decays so that measured parameters like decay rates and CP violation asymmetries would be different from the SM predictions. Processes for which the SM contribution occurs through rare loop diagrams, that is flavour changing neutral currents (FCNC), are considered golden channels for eventual discoveries of physics beyond the SM. Processes like $B^0 - \bar{B}^0$, $B_s^0 - \bar{B}_s^0$ and $D^0 - \bar{D}^0$ mixing, as well as rare decays like $B \to \ell\ell$, $B \to X\ell\ell$ and $B \to X\gamma$ (where ℓ is a lepton and X is an hadronic system) are good candidates. Several anomalies reported by the LHCb experiment in these type of processes might anticipate a Beyond Standard Model (BSM) discovery. In fact, New Physics offshell particles could contribute in loop diagrams and interfere with SM processes modifying its observables as CP symmetry, branching fractions, angular distributions, ...

The angular distribution of $B^0 \to K^{*0}\mu^+\mu^-$ decay products exibits a deviation from the SM expectations and branching fractions of $B^0 \to X\mu^+\mu^-$ decays are consistently below the SM predicted values [29,30].

In the charm sector, one of the most interesting observables is referred to as A_Γ. It quantifies the difference in the inverse effective lifetimes between D^0 and $\bar{D}^0$ decays to CP eigenstates like K^+K^- and $\pi^+\pi^-$. A nonzero value of A_Γ would be a sign of CP violation in charm mixing, which is expected to be at the $\mathrm{O}(10^{-4})$ level, or less, in the SM. Differences between $A_\Gamma(D^0 \to K^+K^-)$ and $A_\Gamma(D^0 \to \pi^+\pi^-)$ would indicate time-integrated direct CP violation [31,32].

An example of indirect CP violation study is the mixing and CP-violation parameters measurement in $D^0 - \bar{D}^0$ oscillations which can be accessed through the comparison of the decay-time-dependent ratio of $D^0 \to K^+\pi^-$ to $D^0 \to K^-\pi^+$ rates with its corresponding ratio of the charge-conjugate processes [33,34].

The Lepton Flavour Universality (LFU) is a SM symmetry and measurements of its violation could be a clear sign of New Physics with contributions that favor couplings with one lepton family rather than the others. LFU tests could be performed comparing branching fraction ratios but also by checking the angular distributions of the decays of interest. The next chapter will deepen this topic.

The current LHCb physics program mainly focuses on:

- B decays to charmonium: studies of b-hadron decays to final states containing (usually leptonically triggered) charmonia ($b \to J/\psi X$), focusing on time-dependent and CP violation measurements;

- B decays to open charm ($B \to DX$) for CP violation and spectroscopy studies;

- Charmless b-hadron decays, for CP violation and spectroscopy studies;

- b-hadrons and quarkonia: production and spectroscopy of b-hadrons and quarkonia (and quarkonia-like) states;

- Charm physics: CP violation and lepton universality;

- QCD, Electroweak and Exotica: production processes including both soft and hard QCD and exotic new particles;

- Rare decays: searches for and studies of non-hadronic decays of flavoured particles (except charm);

- Semileptonic b-hadron decays: CP violation and lepton flavour universality;

- Ions and Fixed Target: for production and spectroscopy measurements with ion beams or gaseous targets.

The Phase-II Upgrade will be capable of a broad spectrum of important flavour-physics measurements. The key goals are as follows:

- A comprehensive measurement program of observables in a wide range of $b \to s\ell^+\ell^-$ and $b \to d\ell^+\ell^-$ transitions;
- Measurements of the CP-violating phases γ and ϕ_s with a precision of 0.4° and 3 mrad, respectively;
- Measurement of $R = \mathcal{B}(B_d^0 \to \mu^+\mu^-)/(B_s^0 \to \mu^+\mu^-)$ with an uncertainty of 20% and the first precise measurements of the observables associated to $B_s^0 \to \mu^+\mu^-$ decay;
- A wide-ranging set of lepton-universality tests in $b \to c\ell^-\bar{\nu}_\ell$ decays, exploiting the full range of b-hadrons;
- CP-violation studies in charm with 10^{-5} precision.

Chapter 2

Lepton Flavour Universality tests

The Standard Model (SM) of particle physics organizes the twelve elementary fermions (and their antiparticles) known to date into three families (or generations) and successfully describes their strong and electroweak interactions through the exchange of gauge bosons (see Fig. 2.1).

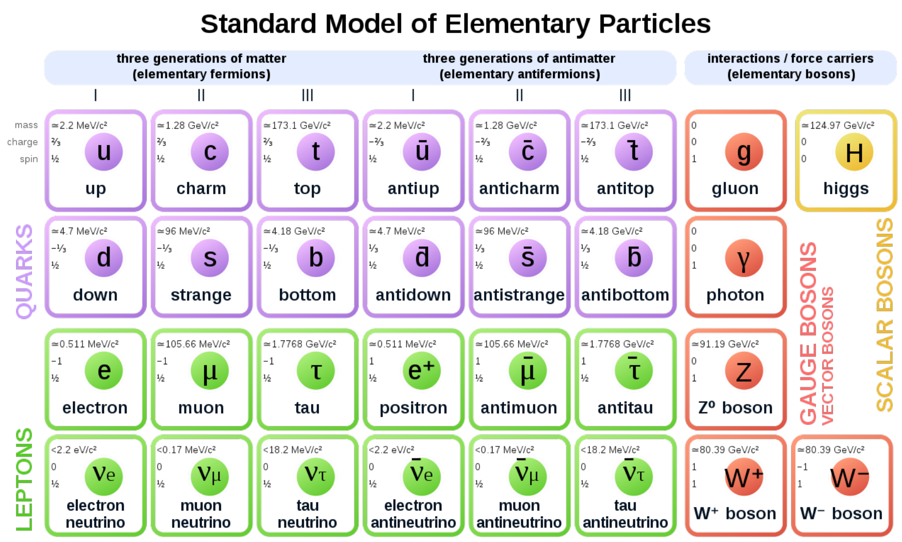


Figure 2.1: Building blocks of the elementary particles Standard Model.

Despite its tremendous success in describing all present measurements, the SM can only be regarded as the low-energy, effective, incarnation of a more global theory. For example, the SM cannot fully describe the matter-antimatter asymmetry currently present in the Universe, it does not provide a Dark Matter candidate and it does not explain its own gauge group structure, the charge assignment of the fermions or the mass hierarchy between the different families. A more global theory that extends the SM at higher energies and shorter distances could provide an answer to some of these questions, which are at the core of modern particle physics.

Searches for signs of New Physics (NP) existing beyond the SM are performed in two ways. The first one looks for the direct production of new particles. The key ingredient for this so-called "relativistic path" is the amount of energy available in the collision, which drives the maximum mass range that can be probed. For this reason, the vast majority of the most stringent limits on the mass of NP particles have been obtained at the Large Hadron Collider (LHC) by the ATLAS and CMS experiments, as shown for example in Ref. [35].

The second method, the so-called "quantum path", exploits the presence of virtual states in the decays of SM particles. Due to quantum mechanics, these intermediate states can be much heavier than the initial and final particles, and can affect the rate of specific decay modes as well as their angular distributions. The most evident example is the β decay of the neutron that probes physics at the W-boson scale.

Flavour physics, which studies the properties of s-, c- and b-hadrons, has been extremely successful in providing information on the virtual particles involved in the processes examined. A famous example is the observation of CP violation in K^0 decays [36] that could be interpreted as a sign of the existence of three quark families in the SM [28], several years before the discovery of the third generation [37, 38]. Similarly, the first observation of the $B^0 - \bar{B}^0$ mixing phenomenon [39] indicated that the t-quark mass was much larger than anticipated, eight years before its direct measurement [40, 41].

The investigation of the effects induced by virtual particles on SM decays requires very large data samples, since, for a given coupling to NP, the larger the NP scale, the smaller the influence on the process is. While the relativistic path probes the scale of potential NP contributions in a direct way through specific signatures, the quantum path provides constraints correlating the scale and the coupling to NP.

In the SM, the leptonic parts of the three families of fermions are identical, except for the different constituent mass of the particles. In particular, the photon, the W and the Z bosons couple in exactly the same manner to the three lepton generations. This very peculiar aspect of the SM, known as Lepton Universality (LU), can be examined to challenge its validity, since any departure from this identity (once the kinematic differences due to the different leptons masses have been corrected for) would be a clear sign that virtual NP particles are contributing to SM decays.

2.1 Lepton Flavour Universality: theory

The importance of LFU tests is strongly related to the very structure of the SM, which is based on the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$ (corresponding to strong and electroweak interactions). This group is broken down to $SU(3)_C \times U(1)_{em}$ (i.e. QCD and QED) through the non-vanishing vacuum expectation value of the Higgs field.

Three main features are relevant for the tests of LFU. Firstly, the fermion fields are organized in three generations with the same charge assignments leading to the same structure of couplings in all three generations (universality). Secondly, the Higgs mechanism for the breakdown of the electroweak gauge symmetry does not affect the universality of the couplings (including electromagnetism). Finally, the only difference between the three families comes from the Yukawa interaction between the Higgs field and the fermion fields. The diagonalisation of the mass matrices yields mixing matrices between weak and mass eigenstates occurring in the coupling of fermions to the weak gauge boson $W^{\pm}$, which are the only source of differences between the three generations: the CKM matrix, V, for the quarks and the PMNS matrix, U, for the leptons.

This last point is the key element to consider in order to design observables testing LFU within the SM. For a given transition in the SM, quarks and leptons stand on a different footing. The flavour of the quarks involved in a transition can be determined experimentally (mass and charge of the fermions involved), so that the CKM matrix elements involved can be determined unambiguously for a given process.

Purely leptonic decays can test LFU by measuring ratios of $Z \to \ell^+\ell^-$ and $W^- \to \ell^-\bar{\nu}_\ell$ partial widths. In the SM the couplings of the W and Z bosons to all lepton species are identical, so the measured ratios should be consistent with one. Leptonic decays of pseudoscalar mesons ($K^- \to \ell^-\bar{\nu}_\ell$ and $\pi \to \ell^-\bar{\nu}_\ell$), charmed mesons ($D_s^- \to \ell^-\bar{\nu}_\ell$)and quarkonia ($J/\psi \to \ell^+\ell^-$) also enable powerful tests of LFU.

In the case of Flavour-Changing Charged-Current (FCCC) transitions like "tree-level" $b \to c\ell^-\bar{\nu}_\ell$ decays (Fig. 2.2), where ℓ represents any of the three charged leptons, a single CKM matrix element is involved (V_{cb}).

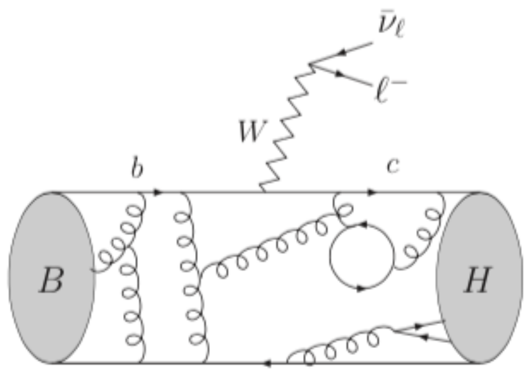


Figure 2.2: Illustration of a $b \to c\ell^-\bar{\nu}_\ell$ transition in the SM, as seen at the hadronic level, in the case of a B meson decaying into a generic H meson [42].

In the case of Flavour-Changing Neutral-Current (FCNC) processes, like "loop-level" $b \to s\ell^-\ell^+$ decays (Fig. 2.3), the CKM matrix elements at play depend on the flavour of the quark running through the loop and have the form $(V_{ib}V^*_{is})$, where $i = t, c, u$. For the lepton part, the PMNS matrix plays no role in these decays and it is ignored in most computations.

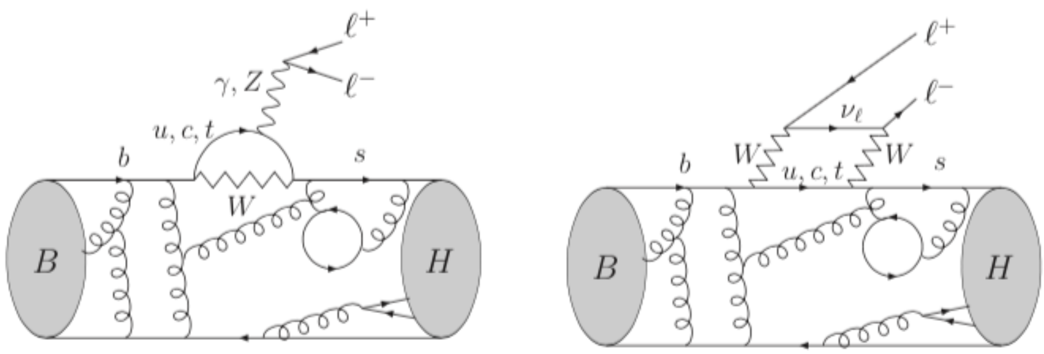


Figure 2.3: Illustration of $b \to s\ell^-\ell^+$ transitions in the SM, as seen at the hadronic level, in the case of a B meson decaying into an unspecified H meson [42].

Therefore, these characteristics explain why either purely leptonic or semileptonic processes that involve leptons of different generations, but with the same quark transition, are preferred to test LFU, so that there are no PMNS matrix elements and the CKM ones can cancel out in ratios.

It is thus natural to compare the same observable for processes differing only in the lepton flavours involved: this can be, for instance, the ratio of branching fractions for two decays ($b \to c\tau^-\bar{\nu}_\tau$ vs $b \to ce^-\bar{\nu}_e$ and/or $b \to c\mu^-\bar{\nu}_\mu$), or the difference of the same angular observable describing the kinematics of the two decays ($b \to s\mu^-\mu^+$ vs $b \to se^-e^+$). From a theoretical point of view, it is often beneficial to choose observables based on ratios, so that common hadronic form-factors cancel out in the numerator and the denominator:

the hadronic uncertainties due to the theoretical evaluation of form factors are much more reduced in the ratio rather than in the individual contributions. From an experimental point of view, the use of ratios is also advantageous since it alleviates the dependence on the absolute knowledge of efficiencies and thus reduces the size of the systematic uncertainties.

2.2 Lepton Flavour Universality: experimental results

2.2.1 Tests beyond the B sector

As already discussed, in the SM the couplings of the W and Z bosons to all lepton species are identical and a large number of experiments where the electroweak bosons are directly produced have tested this property. The most precise results were obtained by experiments running at e^+e^- colliders (LEP1 and SLC, running at the Z pole, or LEP2, where the centre-of-mass energy enabled the direct production of W boson pairs), and at the $p\bar{p}$ (Tevatron) and pp (LHC) colliders.

The measurements of the $Z \to e^+e^-$, $Z \to \mu^+\mu^-$ and $Z \to \tau^+\tau^-$ partial widths are in excellent agreement among each other [35], resulting in ratios of the leptonic partial widths [43]:

$$\frac{\Gamma_{Z\to\mu^+\mu^-}}{\Gamma_{Z\to e^+e^-}} = 1.0009 \pm 0.0028\,,$$
$$\frac{\Gamma_{Z\to\tau^+\tau^-}}{\Gamma_{Z\to e^+e^-}} = 1.0019 \pm 0.0032\,,$$

as expected from the SM in case of massless leptons (equal partial widths).

The experiments at LEP, Tevatron and LHC have also performed precise measurements using W boson decays to test the validity of LFU [44]. A naive average of all experimental results obtained by assuming that all uncertainties are fully uncorrelated results in a value consistent with one:

$$\frac{\mathcal{B}(W^- \to e^-\bar{\nu}_e)}{\mathcal{B}(W^- \to \mu^-\bar{\nu}_\mu)} = 1.0004 \pm 0.0008\,.$$

Tests involving the third lepton family are less precise due to the more challenging reconstruction of final states with τ leptons. Assuming that LFU holds between the first and the second families, a slightly more precise test can be performed by the following ratio:

$$\frac{2\Gamma_{W^-\to\tau^-\bar{\nu}_\tau}}{\Gamma_{W^-\to e^-\bar{\nu}_e} + \Gamma_{W^-\to\mu^-\bar{\nu}_\mu}} = 1.066 \pm 0.025\,,$$

which manifests a tension with the SM expectation at the level of 2.6σ [45].

Leptonic decays of pseudoscalar mesons also enable powerful tests of LFU because they are helicity suppressed in the SM. The most stringent constraints derive from the study of leptonic decays of charged pions or kaons. For kaons decays, the measurement of the ratio [46]:

$$\frac{\Gamma_{K^-\to e^-\bar{\nu}_e}}{\Gamma_{K^-\to\mu^-\bar{\nu}_\mu}} = (2.488 \pm 0.009) \times 10^{-5}\,,$$

is in agreement with the SM prediction of $(2.477 \pm 0.001) \times 10^{-5}$.

The equivalent ratio for charged pion decays has also been measured [47]:

$$\frac{\Gamma_{\pi^-\to e^-\bar{\nu}_e}}{\Gamma_{\pi^-\to\mu^-\bar{\nu}_\mu}} = (1.230 \pm 0.004) \times 10^{-4}\,,$$

with the SM computation being $(1.2352 \pm 0.0001) \times 10^{-4}$. In this sector LFU holds at the 0.2% level.

Analogous tests can be performed in the charmed-meson sector [48]:

$$\frac{\Gamma_{D_s^- \to \tau^- \bar{\nu}_\tau}}{\Gamma_{D_s^- \to \mu^- \bar{\nu}_\mu}} = 9.95 \pm 0.61\,,$$

with the SM prediction being 9.76 ± 0.10 and the LFU verified at a level of 6%.

Leptonic decays of quarkonia resonances can also be used to probe LFU. The most precise test is obtained from the ratio of the $J/\psi \to e^+e^-$ and $J/\psi \to \mu^+\mu^-$ partial widths [35]:

$$\frac{\Gamma_{J/\psi \to e^+e^-}}{\Gamma_{J/\psi \to \mu^+\mu^-}} = 1.0016 \pm 0.0031\,,$$

which is in good agreement with LFU with a precision of 0.31%.

2.2.2 Tests in $b \to c\ell^-\bar{\nu}_\ell$ transitions

Precision tests of LFU can be performed by studying semileptonic $b \to c\ell^-\bar{\nu}_\ell$ decays. New physics could enhance the coupling to the third generation of leptons given the large tau mass. As a result semileptonic B decays ($b \to c\tau^-\bar{\nu}_\tau$) could reveal deviations from the SM. Observables to probe NP contributions to such decays are ratios of branching fractions between the third and the first and second generation leptons:

$$R(H_c) = \frac{\mathcal{B}(H_b \to H_c\tau^-\bar{\nu}_\tau)}{\mathcal{B}(H_b \to H_c\ell^-\bar{\nu}_\ell)}\,.$$

As already discussed, in the ratio large part of the theoretical ($|V_{cb}|$ and form factors) and experimental (branching fractions and reconstruction efficiencies) uncertainties cancel.

Measurements of $R(Hc)$ ratios can be performed using different τ decay modes. Figure 2.1 reports τ decays with the largest branching fractions that have been used to perform measurements.

Decay	Mode	$\mathcal{B}$[%]
$\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$	leptonic	17.39 ± 0.04
$\tau^- \to e^-\bar{\nu}_e\nu_\tau$	leptonic	17.82 ± 0.04
$\tau^- \to \pi^-\pi^0\nu_\tau$	hadronic 1-prong	25.49 ± 0.09
$\tau^- \to \pi^-\nu_\tau$	hadronic 1-prong	10.82 ± 0.05
$\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$	hadronic 3-prongs	9.02 ± 0.05
$\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$	hadronic 3-prongs	4.49 ± 0.05

Table 2.1: Branching fractions of τ decays that have been used to perform measurements in semitauonic H_b decays.

The BaBar and Belle collaborations have used different τ decay modes to measure the $R(D)$ and $R(D^*)$ ratios. A summary of the measurements performed by the two experiments is reported in Table 2.2. Both collaborations have exploited the fact that the $\Upsilon(4S)$ resonance decays exclusively into a $B\bar{B}$ pair, either B^+B^- or $B^0\bar{B}^0$. One of the B mesons, referred as B-tag, is reconstructed using hadronic (Had) or semileptonic (SL) decays. The hadronic tagging algorithms search for an exclusive hadronic B decay from a list of more than one thousand decays, while semileptonic tagging algorithms search for $B \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays, exploiting their large branching fractions. The hadronic tagging is most widely used because in this case the B-tag decay is fully reconstructed and all remaining particles in the event are coming from the other (signal) B decay. This allows the precise measurement of the invariant mass of the undetected particles in the signal decay.

LHCb has performed measurements of the ratios $R(D^*)$ and $R(J/\psi)$ using leptonic τ decays. The $R(J/\psi)$ measurement is unique at LHCb because it exploits the charmed B_c meson production, which is not possible at the B-factories, and the fairly easily selection of the candidates due to the $J/\psi \to \mu^+\mu^-$ decay. $R(D^{*-})$ was also measured by using the three-prong τ decay [49].

The results are summarized in Table 2.2. The three analyses are based on the LHCb Run1 data.

Experiment (year)	B-tag	τ decay	$R(D)$	$R(D^*)$	Ref.
BaBar (2012)	Had.	$\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau$	$0.440 \pm 0.058 \pm 0.042$	$0.332 \pm 0.024 \pm 0.018$	[50, 51]
Belle (2015)	Had.	$\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau$	$0.375 \pm 0.064 \pm 0.026$	$0.293 \pm 0.038 \pm 0.015$	[52]
Belle (2016)	SL	$\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau$	/	$0.302 \pm 0.030 \pm 0.011$	[53]
Belle (2017)	Had.	$\tau^- \to \pi^-(\pi^0)\mu_\tau$	/	$0.270 \pm 0.035\ ^{+0.21}_{-0.16}$	[54]
Belle (2019)	SL	$\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau$	$0.307 \pm 0.037 \pm 0.016$	$0.283 \pm 0.018 \pm 0.014$	[55]
LHCb (2015)	/	$\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$	/	$0.336 \pm 0.027 \pm 0.030$	[56]
LHCb (2017)	/	$\tau^- \to \pi^-\pi^+\pi^-(\pi^0)\nu_\tau$	/	$0.291 \pm 0.019 \pm 0.029$	[49, 57]
Experiment (year)	B-tag	τ decay	$R(J/\psi)$		Ref.
LHCb (2017)	/	$\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$	$0.71 \pm 0.17 \pm 0.18$		[58]

Table 2.2: Experimental results from BaBar, Belle and LHCb on FCCC decays branching fraction ratios.

The measurements of $R(D)$ and $R(D^*)$ are represented in Fig. 2.4 together with their combination and the very precise expected values of the SM. They exceed the SM predictions by 1.4σ and 2.5σ respectively. Considering the $R(D) - R(D^*)$ correlation, the overall difference with the SM predictions corresponds to about 3.08σ.

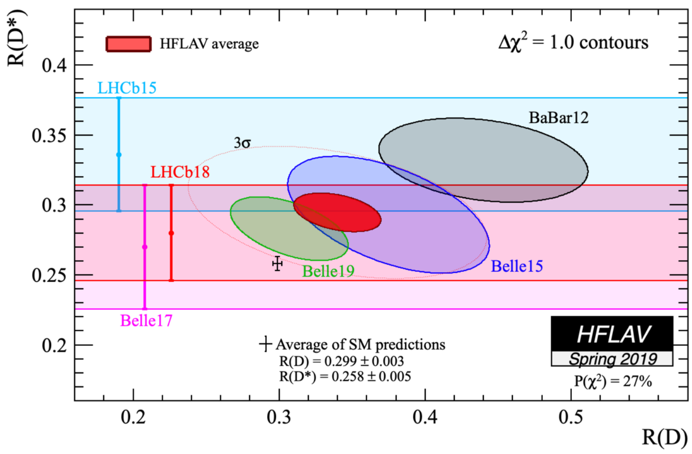


Figure 2.4: Experimental results on $R(D)$ and $R(D^*)$ and comparison with the SM predictions [42].

LHCb has the unique opportunity to test LFU using different species of b-hadrons decays, compared to the B-factories. Some tests have already been performed, as the $R(J/\psi)$ with B_c decays and $R(pK)$ with Λ_b decays (next section). At the moment, these measurements are less precise, because of higher experimental errors (low statistics) and theoretical knowledge (high form factors uncertainty), but in the future, in case the LFU anomalies are confirmed, these ratios will provide strong indications on the specific NP model involved in these processes.

2.2.3 $b \to s\ell^-\ell^+$ tests

Decays of b-hadrons involving a loop-level transition such as $b \to s\ell^-\ell^+$ provide an ideal laboratory to test LFU. This class of decays presents different challenges compared to the more frequent, leading-order, $b \to c\ell\bar{\nu}_\ell$ decays, both experimentally and theoretically. Since there are no tree-level contributions in the SM, FCNC decays are highly suppressed and hence provide an increased sensitivity to the possible existence of NP. The leading Feynman diagrams of a $b \to s\ell^-\ell^+$ transition proceed through the exchange of either a Z/γ penguin, or a W^+W^- box, as illustrated in Fig. 2.3.

Sensitive probes of LFU that exploit $b \to s\ell^-\ell^+$ decays are ratios of the type:

$$R(H_s) = \frac{\int_{q^2_{\min}}^{q^2_{\max}} \frac{d\Gamma(H_b \to H_s\mu^+\mu^-)}{dq^2} dq^2}{\int_{q^2_{\min}}^{q^2_{\max}} \frac{d\Gamma(H_b \to H_s e^+e^-)}{dq^2} dq^2},$$

where H_b represents a hadron containing a b-quark, such as a B^+ or a B^0 meson, H_s represents a hadron containing a s-quark, such as a K or a $K*$ meson, and q^2 is the invariant mass squared of the dilepton system integrated between $q^2_{\min}$ and $q^2_{\max}$ [59]. Also in this case, ratios of branching fractions allow very precise tests of LFU, as hadronic uncertainties in the SM predictions are largely reduced.

The BaBar and Belle experiments performed the first measurements of the $R(K^{(*)})$ ratios, which are found to be in agreement with the SM expectation within their limited precision [60, 61].

The LHCb experiment performed the first measurements of the $R(K^+)$ and $R(K^{*0})$ ratios at a hadron collider using the Run1 dataset [62, 63]. Both the results are found to be respectively 2.6σ and $2.1-2.5\sigma$ lower than the SM expectation. The $R(K^+)$ has been updated recently (2019) using also Run2 data and represents the most precise measurement of $R(K^+)$ [64]. It is compatible with the SM at the level of 2.5 standard deviations. A first test of lepton universality with b-baryons has been performed with Λ_b particles [65]. The value of the $R(pK)$ measurement is compatible within one standard deviation from the SM. A summary of the measurements is presented in Table 2.3.

Experiment (year)	H_s type	q^2 range [GeV$^2/c^4$]	Value	Ref.
Belle (2009)	K	0.0 - kin. endpoint	$1.03 \pm 0.19 \pm 0.06$	[61]
Belle (2009)	K^*	0.0 - kin. endpoint	$0.83 \pm 0.17 \pm 0.08$	[61]
BaBar (2012)	K	0.10 - 8.12	$0.74\ ^{+0.40}_{-0.31} \pm 0.06$	[60]
BaBar (2012)	K	> 10.11	$1.43\ ^{+0.65}_{-0.44} \pm 0.12$	[60]
BaBar (2012)	K^*	0.10 - 8.12	$1.06\ ^{+0.48}_{-0.33} \pm 0.08$	[60]
BaBar (2012)	K^*	> 10.11	$1.18\ ^{+0.55}_{-0.37} \pm 0.11$	[60]
LHCb (2014)	K	1.0 - 6.0	$0.745\ ^{+0.090}_{-0.074} \pm 0.036$	[63]
LHCb (2017)	K^{*0}	0.045 - 1.1	$0.66\ ^{+0.11}_{-0.03} \pm 0.05$	[62]
LHCb (2017)	K^{*0}	1.1 - 6.0	$0.69\ ^{+0.11}_{-0.07} \pm 0.05$	[62]
LHCb (2019)	K	1.1 - 6.0	$0.846\ ^{+0.060}_{-0.054}\ ^{+0.016}_{-0.014}$	[64]
LHCb (2019)	pK	0.1 - 6.0	$1.17\ ^{+0.18}_{-0.16} \pm 0.07$	[65]

Table 2.3: Summary of the $R(K^{(*)})$ measurements performed at the B-factories and by the LHCb experiment. The first $R(pK)$ measurement, unique from LHCb, is also included. The first uncertainty is statistical and the second is systematic.

The experimental picture that emerges from these measurements become even more intriguing considering that additional anomalies arose in other observables concerning $b \to s\ell^+\ell^-$ transitions. Among them, the most relevant deviation was observed in two bins of q^2 for the coefficient P_5' describing the angular distributions of $B^0 \to K^{*0}\mu^+\mu^-$ decays [66].

If such anomalies are confirmed with more precise measurements, it would be a clear sign of new physics contributions. Several global fits have been performed trying to obtain a consistent interpretation of the different anomalies observed introducing new physics contributions in a model-independent approach. The outcome of these studies are promising, but further efforts are still required, both from the experimental and theoretical side, to clarify the present scenario.

Chapter 3

$\mathbf{R(D_s)}$ measurement with τ three-prong decay

The current status of b-quark semileptonic decays requires new measurements to clarify the various deviations from the SM expectations that have been observed in recent analysis. If these were confirmed, more precise measurements and new observables would allow to determine more accurately their pattern and source. These sets of new measurements are already the core of the research program of the LHCb experiment, which, as explained in the previous chapter, have published papers probing Lepton Flavour Universality.

Tensions with respect the SM expectations at the level of $\sim 3\sigma$ are observed in $b \to c\ell^-\bar{\nu}_\ell$, which are mediated at tree level through a $W^\pm$ boson in the SM, when the branching ratios with $\ell = \tau$ and $\ell = e, \mu$ are compared. The ratio definition cancels a large part of the theoretical ($|V_{cb}|$ and form factors) and experimental (branching fractions and reconstruction efficiencies) uncertainties. The main contributions to the uncertainty is of systematic nature: they are due to the modelling of the fit components and to the limited size of the simulated samples.

LHCb is planning and currently performing measurements to test LFU using different b-hadrons decays produced at LHC, including different decay modes. This thesis focuses on the $R(D_s)$ measurement, defined as:

$$R(D_s) = \frac{\mathcal{B}(B_s^0 \to D_s^- \tau^+ \nu_\tau)}{\mathcal{B}(B_s^0 \to D_s^- \mu^+ \nu_\mu)}\,.$$

The $B_s^0 \to D_s^- \tau^+ \nu_\tau$ decay is reconstructed through the τ three-prong hadronic decay, $\tau^+ \to \pi^+\pi^-\pi^+(\pi^0)\nu_\tau$, while the D_s^- meson is reconstructed through the $D_s^- \to K^+K^-\pi^-$ decay chain, so the visible final state consists of six charged tracks (see Fig. 3.1). The branching fraction of the $B_s^0 \to D_s^- \mu^+ \nu_\mu$ decay is known by other measurements.

The three-prong τ decay modes have different features with respect to leptonic τ decays, leading to measurements with a better signal-to-background ratio and statistical significance. The absence of charged leptons in the final state avoids backgrounds originating from semileptonic decays of b or c hadrons. The three-prong topology enables the precise reconstruction of the τ decay vertex detached from the B_s^0 vertex due to the non zero τ lifetime, thereby allowing the discrimination between signal decays and background due to $B_s^0 \to D_s^- 3\pi$ decays. Moreover, because only one neutrino is produced from the τ decay vertex, the kinematics of the B_s and τ decays can be reconstructed (though with some approximation).

The $R(D_s)$ measurement with 3-prong tau decay will follow the analysis strategy of the hadronic $R(D^{*-})$ measurement [49], though with some differences related to the different decay considered (different background sources and systematic uncertainties involved). Moreover the $B_s^0 \to D_s^- \tau^+ \nu_\tau$ does not suffer as badly from cross-feed backgrounds from

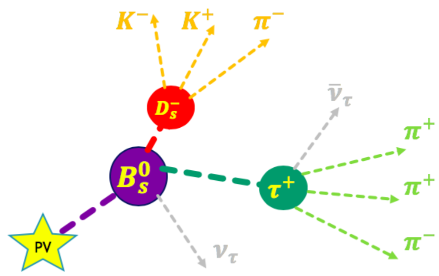


Figure 3.1: Scheme of the signal decay chain. The B_s^0 meson decays to $D_s^-\tau^+\nu_\tau$. The visible final states consists of six charged tracks coming from the D_s^- meson ($K^+K^-\pi^-$) and from the τ^+ ($\pi^+\pi^-\pi^+$).

other mesons, unlike, the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay, where both B^- and B_s^0 contribute to the $D^{*-}\pi^+\pi^-\pi^+X$ final states. At quark level, both $B^0 \to D^{*-}\tau^+\nu_\tau$ and $B_s^0 \to D_s^-\tau^+\nu_\tau$ transitions are dominated by the same diagram. NP contributions responsible of the small measured anomalies, if existent, should be noticeable in both decays.

The theoretical prediction of $R(D_s)$ has been recently published [67]:

$$R(D_s) = 0.2971 \pm 0.0034\,,$$

and provide already a very precise estimate of the SM value ($\sim 1\%$ uncertainty).

3.1 $\mathbf{R(D_s)}$ measurement technique

The presence of different final states in the numerator and denominator of the branching fraction ratio $R(D_s)$ represents a potential source of systematic uncertainties which complicate the measurement. This problem can be mitigated by redefining the ratio using external inputs, in particular introducing a normalization channel, "norm":

$$R(D_s) = \frac{\mathcal{B}(B_s^0 \to D_s^-\tau^+\nu_\tau)}{\mathcal{B}(norm)} \times \frac{\mathcal{B}(norm)}{\mathcal{B}(B_s^0 \to D_s^-\mu^+\nu_\mu)} = \mathcal{K} \times \alpha\,,$$

with

$$\mathcal{K} = \frac{\mathcal{B}(B_s^0 \to D_s^-\tau^+\nu_\tau)}{\mathcal{B}(norm)} \quad \text{and} \quad \alpha = \frac{\mathcal{B}(norm)}{\mathcal{B}(B_s^0 \to D_s^-\mu^+\nu_\mu)}\,.$$

In this way, the α ratio is completely determined by external inputs, i.e. measurements of branching fractions by independent analyses, while the ratio $\mathcal{K}$ needs to be determined from the analysis:

$$\mathcal{K} = \frac{N_{\text{sig}}}{\epsilon_{\text{sig}}}\frac{\epsilon_{\text{norm}}}{N_{\text{norm}}}\frac{1}{\mathcal{B}(\tau^+ \to \pi^+\pi^-\pi^+(\pi^0)\nu_\tau) \times \mathcal{B}(D_s^- \to K^+K^-\pi^-)}\,.$$

Here, N_{sig} (N_{norm}) and ϵ_{sig} (ϵ_{norm}) represent, respectively, the signal (normalization) yield and selection efficiency. The $B_s^0 \to D_s^-\tau^+\nu_\tau$ decay is intended as signal. The ratio $\mathcal{K}$ also includes external inputs as the measured branching fractions of $\mathcal{B}(\tau^+ \to \pi^+\pi^-\pi^+(\pi^0)\nu_\tau)$ and $\mathcal{B}(D_s^- \to K^+K^-\pi^-)$, whose world averages are taken from the PDG [35].

A common selection needs to be implemented for signal and normalization to obtain an efficiency ratio as close to one as possible and in this way reduce the systematic uncertainties on the $R(D_s)$ measurement.

The choice of the normalization decay channel will be the main focus of the next sections.

3.1.1 Choice of the normalization channel

In the previous section, the measurement technique has been presented introducing a normalization channel to simplify the analysis. The topology of the normalization mode needs to be as similar as possible to the signal $B_s^0 \to D_s^- \tau^+ \nu_\tau$ and, in particular, it requires similar/same final states. In this case, the visible final states are, as discussed before, six charged charged tracks: $K^+K^-\pi^-$ from the D_s^- meson decay and $\pi^+\pi^-\pi^+$ from the τ^+ decay.

By choosing a normalization channel that is as similar as possible to the signal and that is reconstructed in a similar way, a common selection can be implemented. Hence, when measuring the ratio $\mathcal{K}$, the efficiency ratio should be similar (except for effects due to the different kinematics of the final states) and consequently the systematic uncertainties should be strongly reduced.

Three normalization channels are proposed here:

1. $B_s^0 \to D_s^- \pi^+\pi^+\pi^-$ with $D_s^- \to K^+K^-\pi^-$
2. $B^0 \to D^- \pi^+\pi^+\pi^-$ with $D^- \to K^+K^-\pi^-$
3. $B^0 \to D^- \pi^+\pi^+\pi^-$ with $D^- \to K^+\pi^-\pi^-$

The first decay is the most similar to the signal channel and differs only in the τ flight distance from the B_s^0 and the related two neutrinos. As shown in Fig. 3.2, the six visible final state particles are exactly the same as those of the signal.

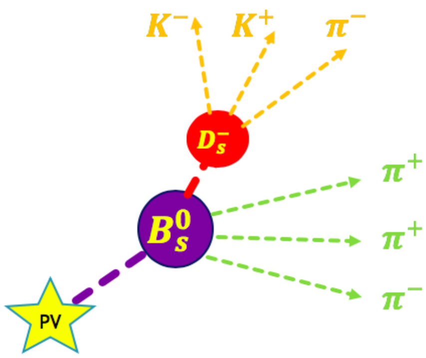


Figure 3.2: Scheme of the $B_s^0 \to D_s^-(\to K^+K^-\pi^-)\pi^+\pi^+\pi^-$ normalization channel. The visible final state particles are exactly the same as those of the signal: $K^+K^-\pi^-$ and $\pi^+\pi^-\pi^+$.

The second and third decay chains consider the same $B^0 \to D^-\pi^+\pi^+\pi^-$ decay but different D^- decays. They both have similar topologies to the signal decay one. The main difference is that the decaying particle is the B^0 instead of the B_s^0 meson and the involved c-hadron is the D^- instead of the D_s^- meson. Due to the different masses of the particles involved, a difference in the kinematics of the decay is expected. Among these two decays, the $B^0 \to D^-(\to K^+K^-\pi^-)\pi^+\pi^+\pi^-$ channel has exactly the same six visible final state particles as those of the signal, but is suppressed compared to the $B^0 \to D^-(\to K^+\pi^-\pi^-)\pi^+\pi^+\pi^-$ decay by almost a factor 10. From this consideration and also from some technical problems related to the stripping selection, this normalization channel will not be considered in the analysis reported in this thesis.

The $B^0 \to D^-(\to K^+\pi^-\pi^-)\pi^+\pi^+\pi^-$ normalization channel has one different final state particle compared to the signal (see Fig. 3.3), i.e. a pion instead of a kaon, but similar topology. The advantage of this normalization channel is that it is better known and studied than the first normalization channel, so the relative uncertainties coming from the external inputs will be smaller.

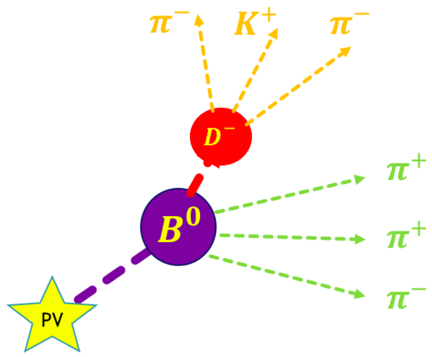


Figure 3.3: Scheme of the $B^0 \to D^-(\to K^+\pi^-\pi^-)\pi^+\pi^+\pi^-$ normalization channel. Both the visible final state particles and particle on top of the decay chain are different from the signal. However, the channel topology is very similar to that of the signal.

3.1.2 Normalization channels comparison

The choice of the normalization channel is based on the contribution to the uncertainties on $R(D_s)$, which we want to minimize in order to achieve the best precision. To this aim, a preliminary comparison of the the uncertainties due to the external inputs for each normalization channel considered above is useful.

By selecting the $B_s^0 \to D_s^-(\to K^+K^-\pi^-)\pi^+\pi^+\pi^-$ normalization channel, the $R(D_s)$ ratio is evaluated as:

$$R(D_s) = \mathcal{K} \times \alpha = \frac{\mathcal{B}(B_s^0 \to D_s^-\tau^+\nu_\tau)}{\mathcal{B}(B_s^0 \to D_s^-3\pi)} \times \frac{\mathcal{B}(B_s^0 \to D_s^-3\pi)}{\mathcal{B}(B_s^0 \to D_s^-\mu^+\nu_\mu)} =$$

$$= \frac{N_{\text{sig}}}{\epsilon_{\text{sig}}}\frac{\epsilon_{\text{norm}}}{N_{\text{norm}}} \frac{\mathcal{B}(D_s^- \to K^+K^-\pi^-)}{\mathcal{B}(\tau^+ \to \pi^+\pi^-\pi^+(\pi^0)\nu_\tau) \times \mathcal{B}(D_s^- \to K^+K^-\pi^-)} \times \frac{\mathcal{B}(B_s^0 \to D_s^-3\pi)}{\mathcal{B}(B_s^0 \to D_s^-\mu^+\nu_\mu)}\,.$$

In this case, since the same D_s meson decay into $K^+K^-\pi^-$ is used in both signal and normalization channels, the branching fraction $\mathcal{B}(D_s^- \to K^+K^-\pi^-)$ and related uncertainty cancel in ratio. By combining the other external inputs, reported in Table 3.1, one can compute the contribution to the uncertainty on the $R(D_s)$ measurement as:

$$\left.\frac{\delta R(D_s)}{R(D_s)}\right|_{\text{ext}} = \pm 19.10\%\,.$$

The same exact procedure can be applied to the $B^0 \to D^-(\to K^+\pi^-\pi^-)\pi^+\pi^+\pi^-$ normalization channel. In this case, since the normalization channel is a B_d^0 the ratio of fragmentation fractions f_s/f_d needs to be included:

$$R(D_s) = \mathcal{K} \times \alpha = \frac{\mathcal{B}(B_s^0 \to D_s^-\tau^+\nu_\tau)}{\frac{f_s}{f_d}\cdot\mathcal{B}(B^0 \to D^-3\pi)} \times \frac{\mathcal{B}(B^0 \to D^-3\pi)}{\mathcal{B}(B_s^0 \to D_s^-\mu^+\nu_\mu)} =$$

$$= \frac{N_{\text{sig}}}{\epsilon_{\text{sig}}}\frac{\epsilon_{\text{norm}}}{N_{\text{norm}}} \frac{\mathcal{B}(D^- \to K^+\pi^-\pi^-)}{\mathcal{B}(\tau^+ \to \pi^+\pi^-\pi^+(\pi^0)\nu_\tau) \times \frac{f_s}{f_d}\mathcal{B}(D_s^- \to K^+K^-\pi^-)} \times \frac{\mathcal{B}(B^0 \to D^-3\pi)}{\mathcal{B}(B_s^0 \to D_s^-\mu^+\nu_\mu)}\,.$$

In the calculation of the contribution to the relative uncertainty of $R(D_s)$ due to the external inputs it should be considered that branching fraction measurement $\mathcal{B}(B_s^0 \to D_s^-\mu^+\nu_\mu)$

from Ref. [68] is correlated to the measurement of $\frac{f_s}{f_d} \times \mathcal{B}(D_s^- \to K^+K^-\pi^-)$ with a correlation factor $\rho = -0.581$. Accounting for this, the relative uncertainty on $R(D_s)$ due to external inputs is:

$$\left.\frac{\delta R(D_s)}{R(D_s)}\right|_{\text{ext}} = \pm 14.3\% \,.$$

External input	Value [%]	Relative uncertainty [%]	Ref.
$\mathcal{B}(\tau^+ \to \pi^+\pi^-\pi^+\nu_\tau)$	9.31 ± 0.05	0.54	[35]
$\mathcal{B}(\tau^+ \to \pi^+\pi^-\pi^+\pi^0\nu_\tau)$	4.62 ± 0.05	1.08	[35]
$\mathcal{B}(B_s^0 \to D_s^- 3\pi)$	0.61 ± 0.10	16.4	[35]
$\mathcal{B}(B_s^0 \to D_s^- \mu^+\nu_\mu)$	2.49 ± 0.24	9.64	[35]
$\mathcal{B}(B^0 \to D^- 3\pi)$	0.60 ± 0.07	11.7	[35]
$\mathcal{B}(D^- \to K^+\pi^-\pi^-)$	9.38 ± 0.16	1.71	[35]
$\frac{f_s}{f_d} \times \mathcal{B}(D_s^- \to K^+K^-\pi^-)$	1.26 ± 0.05	3.97	[68]

Table 3.1: External inputs for the $R(D_s)$ measurement.

From this preliminary study, a first conclusion can be drawn: the contribution to the uncertainty of the $R(D_s)$ measurement by the external inputs is 19.10% in the case of the $B_s^0 \to D_s^-(\to K^+K^-\pi^-)\pi^+\pi^+\pi^-$ normalization channel and 14.3% in the case of the $B^0 \to D^-(\to K^+\pi^-\pi^-)\pi^+\pi^+\pi^-$ normalization channel. From these considerations only, one would choose as reference normalization channel the second one, leading to a more precise measurement.

However, one should consider that the final $R(D_s)$ measurement is affected additional sources of uncertainty, e.g. those coming from the $\frac{N_{\text{sig}}}{\epsilon_{\text{sig}}}\frac{\epsilon_{\text{norm}}}{N_{\text{norm}}}$ part of the ratio. In particular, one should consider:

- Statistical uncertainty from Signal and Normalization yields (depending on the values of the yields);
- Systematic uncertainties from Signal and Normalization yields (fluctuations dependent on the model used to extract them);
- Statistical uncertainty on Signal and Normalization efficiencies (depending on the selection implemented);
- Systematic uncertainty on the efficiencies (relative to the MC used and how well it represents the data).

These items depend on the signal selection, which will be discussed in the next chapter. They all concur to the measurement precision in a way that may depend on the normalization channel.

Chapter 4

Selection criteria and Normalization channels study

4.1 Data and Monte Carlo samples

In this thesis, the analysis is performed on the dataset collected by LHCb during 2012 (part of Run1 of LHC) data corresponding to an integrated luminosity of $\mathcal{L} = 2\text{fb}^{-1}$ at $\sqrt{s} = 8$ TeV. Once the full analysis strategy is defined, the full statistics of Run1 (adding 2011 data) and Run2 (2016-2018) will be added. Both the signal and the normalization channels candidates are selected by dedicated stripping lines. The selection is further refined offline as explained in the following.

Monte Carlo (MC) simulated samples are used to study the characteristics of signal and normalization channels, develop the offline selection, based also on multivariate classifiers, and evaluate the selection efficiencies of the various steps of the analysis. The production of MC samples involves several phases as explained in section 1.3. For this analysis, the MC samples are:

- a sample of $B_s^0 \to D_s^- \tau^+ \nu_\tau$ (with $D_s^- \to K^+ K^- \pi^-$ and $\tau^+ \to \pi^+ \pi^- \pi^+ (\pi^0) \nu_\tau$) decays that will be used to train the multivariate classifiers and evaluate the signal efficiencies,

- a sample of $B_s^0 \to D_s^- \pi^+ \pi^+ \pi^-$ with $D_s^- \to K^+ K^- \pi^-$ decays to study the first normalization channel and from which the reference B_s^0 mass fit model and the normalization efficiencies are determined,

- a $B^0 \to D^- \pi^+ \pi^+ \pi^-$ with $D^- \to K^+ \pi^- \pi^-$ sample to study the second normalization channel and from the reference B^0 mass fit model and the normalization efficiencies are determined.

These samples are generated, reconstructed and filtered following the 2012 detector conditions. To optimize the simulation processing, the final state particles are required to be in the LHCb acceptance.

The data and MC samples are required to pass the stripping selection criteria "`Bs2DsTauNuForB2XTauNu`" for the B_s^0 decays and "`B0d2DTauNuForB2XTauNu`" for the B^0 decays. Table 4.1 shows the complete set of cuts required for the B_s^0 and B^0 stripping selections. The stripping cuts are the same for the two lines, expect for a different reconstruction D/D_s vertex χ^2 cut. The different final states particles also introduce some diversities, generated by different cuts for kaons and pions.

A preliminary selection, focused on the candidates selection, will then refine the stripping requirements with more selective cuts to reduce the background sources.

Cut	Value
$B^0(_s)$	
$\Delta(M)$	(-2579)-300 or 720-1721 MeV/c^2
Max. DOCA	$<$ 0.15 mm
DIRA	$>$ 0.995
$D(_s)$	
p_{T}	$>$ 1600 MeV/c
$\mid M - M_{D(s)} \mid$	<40.0 MeV/c^2
DIRA	$>$ 0.995
Vertex distance χ^2	$>$ (36.0 if D_s)(50 if D)
Vertex χ^2 / NDOF	$<$ 10
IP χ^2	$>$ 10
K from $D(_s)$	
p_{T}	$>$ 1500 MeV/c
Track χ^2 / NDOF	$<$ 30
IP χ^2	$>$ 10
Ghost Probability	$<$ 0.4
$PIDK$	$>$ 3
π from $D(_s)$	
p_{T}	$>$ 150 MeV/c
Track χ^2 / NDOF	$<$ 3
Ghost Probability	$<$ 0.4
IP χ^2	$>$ 10
$PIDK$	$<$ 50
τ	
m($\pi\pi\pi$)	400-3500 MeV/c^2
$m(\pi_1\pi_2)$ or $m(\pi_2\pi_3)$	$<$ 1670 MeV/c^2
Max. DOCA	$<$ 0.15 mm
DIRA	$>$ 0.99
Vertex χ^2	$<$ 25
max 1 pion with p_{T}	$<$ 300 MeV/c
min 2 pions with IP χ^2	$>$ 5
π from τ	
p_{T}	$>$ 150 MeV/c
IP χ^2	$>$ 4
Track χ^2 / NDOF	$<$ 4
Ghost Probability	$<$ 0.4
$PIDK$	$<$ 8

Table 4.1: Cuts used in the stripping selection "`Bs2DsTauNuForB2XTauNu`" and "`B0d2DTauNuForB2XTauNu`". The variables DIRA, Max. DOCA, IP χ^2 correspond respectively to the cosine of the angle between the momentum of the particle and the direction of flight from the best PV to the decay vertex, the maximum distance of closest approach between all possible pairs of particles and the χ^2_{IP} on the related PV.

4.1.1 Trigger Selection

The LHCb trigger consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, in which all charged particles with $p_{\mathrm{T}} > 300$ MeV/c are reconstructed and further considered for the selection. At the hardware stage (L0), trigger decisions are based exclusively on information from the muon system and calorimeters and events with either high p_{T} muons or large transverse energy deposits in the calorimeter are selected. The events can be selected by the trigger as TOS (trigger on signal), if the signal candidate triggers the L0 trigger lines or as TIS (trigger independent of signal), if any particle in the event, irrespective if it is part of the signal candidate or not, passes any of the L0 trigger lines (see section 1.2.6 for more details). In this thesis, given that both the signal and the normalization decays don't involve any muons in the final state, the L0 trigger requirement is TOS Hadron trigger (called L0_Hadron_TOS) or TIS trigger (called L0_Global_TIS) on the B^0_s candidate.

The LHCb software trigger stage (HLT) is divided into HLT1 and HLT2. In this thesis, only HLT1 is used to select the event candidates. The software stage allows a primary vertices reconstruction and a search for secondary vertices. Also a preliminary track reconstruction is performed. The HLT1 requirement executed for all events accepted by L0 is that at least one track satisfies a number of track quality criteria called Hlt1TrackAllL0Decision, triggered on signal (TOS), required on the stripping line.

The trigger selection implemented has a 96.9% efficiency on the $B_s^0 \to D_s^- \tau^+ \nu_\tau$ and $B_s^0 \to D_s^- 3\pi$ MC samples and 97.3% efficiency on the $B^0 \to D^- 3\pi$ MC sample.

4.2 Selection criteria

A common selection needs to be implemented for signal and normalization decay modes in order to obtain a selection efficiency ratio as close as possible to one. In this way, when measuring the ratio $\mathcal{K}$, main focus of the $R(D_s)$ measurement, most of the uncertainties will cancel out. In particular the theoretical and experimental systematics will be strongly reduced.

There are different kinds of background to take into account:

- b-hadron decays with similar signal topology but with additional particles that are not reconstructed, partially reconstructed or with different final state particles (misidentified),

- combinations of particles that do not belong to the same b-hadron but from other sources, for example directly from the primary vertex or from a different b/c-hadron in the event (combinatorial background).

The main background in this analysis is combinatorial. To suppress it, the selection is implemented applying a multivariate analysis (MVA) classification to discriminate it from the signal. For this analysis, the ROOT Toolkit for Multivariate Analysis (TMVA) [69] is used, in particular, the Boosted Decision Tree (BDT) machine learning technique. The BDT is trained on samples of events, for which the desired separation (signal vs background) is already known. The goal is to determine a mapping classifier function that can be applied to any set of data to discriminate signal from background at a given efficiency.

A Decision Tree is a binary tree structured classifier similar to the one sketched in Fig. 4.1. Repeated left/right (yes/no) decisions are taken on one single variable so that the phase space is split into many regions that are classified as signal or background, depending on the majority of training events that end up in the final leaf node. The Boosting of a Decision Tree extends this concept from one tree to several trees which form a forest. The outputs of the trees is at the end combined into a single classifier which gives a weighted average of the individual decision trees. Boosting stabilizes the response of the decision trees with respect to fluctuations in the training sample and is able to considerably enhance the performance instead of using a single tree.

Figure 4.2 shows a scheme of the implemented selection. First, a selection on the final state particles is performed using two BDTs, one for the three pions and one for the $D_{(s)}$ meson selection. The classifiers outputs of the two BDTs are used to implement a third BDT for the $B^0_{(s)}$ candidate selection.

The training events variables of the three multivariate analysis must be, of course, discriminating between signal and background, but also congruent between the signal and normalization decay channels. Variables that have similar distributions between signal and normalization will be selected to achieve a overall common selection.

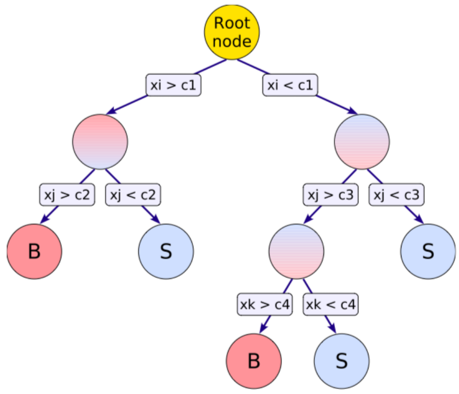


Figure 4.1: Schematic view of a decision tree. Starting from the root node, a sequence of binary splits using the discriminating variables x_i is applied to the data. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The same variable may thus be used at several nodes, while others might not be used at all. The leaf nodes at the bottom end of the tree are labeled "S" for signal and "B" for background depending on the majority of events that end up in the respective nodes.

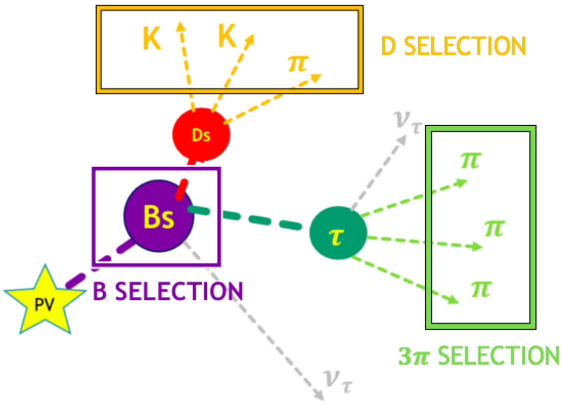


Figure 4.2: Scheme of the three BDT implemented.

4.2.1 Preliminary selection

Candidate $B_{(s)}$ decays are formed by combining a candidate $D_{(s)}$ with three particles originating from a common vertex, displaced from any PV, and consistent with the pion hypothesis. Similarly, the $D_{(s)}$ candidate is formed selecting combinations of three particles originating from a common vertex, displaced from any PV, and consistent with the hypothesis of two kaons and one pion (for the D_s case) or one kaon and two pions of the same charge (for the D case) and with an invariant mass within ± 40 MeV$/c^2$ the known $D_{(s)}$ mass value. The $D_{(s)}$ and the three-pion system should also form a common vertex.

Prior to the multivariate selection, candidate $B_{(s)}$ decays from the stripping selection are further selected. In particular, the purity of selected signal $D_{(s)}$ mesons is improved by applying tighter invariant mass requirements (± 20 MeV$/c^2$ around the known value) and PID requirements on the final state particles depending on the dynamics of the decay.

For the $D_s^+ \to K^+K^-\pi^+$ decay, the dynamics is dominated by the production of the $\phi(1020)$ and the $K^*(892)^0$ intermediate resonances, as shown in the Dalitz plot in Fig. 4.3a.

In the region around the narrow $\phi(1020)$ peak the kaons are required to pass loose PID requirements, while in the region around the wider $K^*(892)^0$ peak, and in the rest of the phase space (where there are no resonant structure), tighter PID requirements are applied (see Table 4.2).

For the $D^+ \to K^-\pi^+\pi^+$ decay the main resonance contribution is coming from $K^*(892^0)$ and $K^*(1430^0)$ (see Fig. 4.3b). At the moment, the PID cuts already present in the stripping line are sufficient to reject wrong D candidates because the $B_d^0 \to D^- 3\pi$ decay channel suffers less from background contamination than the $B_s^0 \to D_s^- 3\pi$ channel.

Decay hypothesis	Mass condition [MeV/c^2]	PID condition
$D_s^+ \to \phi(1020)\pi^+$	$\mid m(K^+K^-) - 1020 \mid\leq 12$	K^+ PIDK > -2 K^- PIDK > -2
$D_s^+ \to K^*(892)^0 K^+$	$\mid m(K^+K^-) - 1020 \mid> 12$ $\mid m(K^-\pi^+) - 892 \mid< 50$	K^+ PIDK > -2 K^- PIDK > 5
$D_s^+ \to K^+K^-\pi^+$ (non res.)	$\mid m(K^+K^-) - 1020 \mid> 12$ $\mid m(K^-\pi^+) - 892 \mid> 50$	K^+ PIDK > 5 K^- PIDK > 5 π^+ PIDK < 10

Table 4.2: $D_s^+ \to K^+K^-\pi^+$ PID preliminary requirements selection in different Dalitz plot regions.

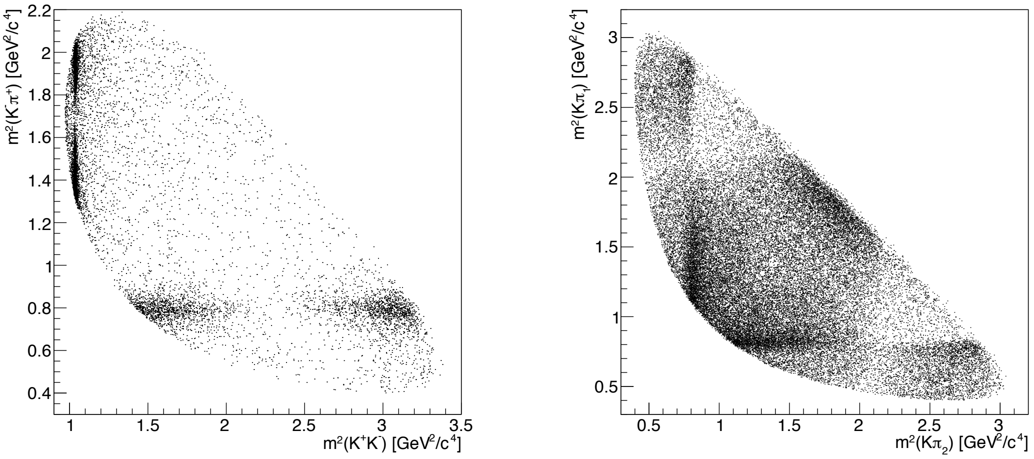
Furthermore, additional cuts ("vetoes") are applied to suppress other c-hadron decays that are selected as $D_{(s)}$ candidates due to mis-identification of the final state particles. For example a $D^+ \to K^+\pi^-\pi^+$ decay with a pion misidentified as a kaon, or a $\Lambda_c^+ \to pK^-\pi^+$ decay with the kaon misidentified as a proton could be selected as a D_s^+ candidate. To suppress these background the invariant mass of the three final state particles is recomputed with alternative assignment of the mis-identified kaon (m') and the candidate is rejected if the mass is consistent with the D^+ or Λ_c^+ mass and the mis-identified particle satisfy additional PID requirements (see Table 4.3). Similar cuts are applied to veto D_s^+ and Λ_c^+ in $D^+ \to K^-\pi^+\pi^+$ reconstruction (see Table 4.4). Lastly, $D^0 \to K^+K^-$ decays where a kaon is misidentified as a pion and an additional pion or kaon are considered to form a $D_{(s)}^+$ are also excluded by requiring the recomputed K^+K^- invariant mass is below 1840 MeV/c^2.

Decay hypothesis	Mass condition [MeV/c^2]	PID condition
$D^+ \to K^-\pi^+\pi^+$	$\mid m' - 1869 \mid< 30$	K^+ PIDK < 10
$\Lambda_c^+ \to pK^-\pi^+$	$\mid m' - 2280 \mid< 30$	K^+ (PIDK - PIDp) < 5

Table 4.3: Veto cuts to suppress misidentified D^+ and Λ_c^+ decays in $D_s^+ \to K^+K^-\pi^+$ candidate selection. The "and" of the two requirements are considered to reject the candidates.

Decay hypothesis	Mass condition [MeV/c^2]	PID condition
$D_s^+ \to K^+K^-\pi^+$	$\mid m' - 1968 \mid< 30$	π^+ PIDK < 0
$\Lambda_c^+ \to pK^-\pi^+$	$\mid m' - 2280 \mid< 30$	π^+ PIDp < 0

Table 4.4: Veto cuts to suppress misidentified D_s^+ and Λ_c^+ decays in $D^+ \to K^-\pi^+\pi^+$ candidate selection. The "and" of the two requirements are considered to reject the candidates.



(a) (b)

Figure 4.3: Dalitz plot of the $D_s^+ \to K^+K^-\pi^+$ decay (a) and of the $D^+ \to K^-\pi^+\pi^+$ decay (b). Figure (a) shows the $m^2(K^+K^-)$ vs $m^2(K^-\pi^+)$ plot. The events selected are within $\pm$ 50 MeV/c^2 around the B_s^0 mass. Figure (b) shows the $m^2(K^-\pi_1^+)$ vs $m^2(K^-\pi_2^+)$ plot. The events selected are within $\pm$ 50 MeV/c^2 around the B^0 mass.

4.2.2 BDT selections

The data and MC samples are first filtered with the trigger and $D_{(s)}$ preliminary selection, then a multivariate classifier analysis is performed. The first step consists in selecting the candidates using three BDTs, called in Fig. 4.2 "D selection", "3π selection" and "B selection". The variables selected to train the BDTs must be discriminating between signal and background, but must have similar distributions between signal and normalization to achieve a common selection. This is why variables containing intrinsic characteristics related to signal-normalization differences must be avoided even if they show similar distributions. The BDTs are trained on samples of events for which the desired separation (signal vs background) is known a-priori. To test the performance of the BDT, each sample is split in two. One is used for training, the other to test the performance and check whether the outcome is affected by overtraining problems. To avoid problems related to the correlation of the input variables, the training of the different BDTs is performed on statistically independent samples.

4.2.2.1 D selection

The "D selection" BDT is implemented using the following training samples:

- D_s candidates from the MC signal decay $B_s^0 \to D_s^-\tau^+\nu_\tau$ (with $D_s^- \to K^+K^-\pi^-$ and $\tau^+ \to \pi^+\pi^-\pi^+(\pi^0)\nu_\tau$) as **training signal**. The D_s candidates are required to be associated to the "true" D_s from the signal decay chain.

- D_s candidates from the B_s data sample selected events with $m(B_s^0) > 5200$ MeV/c^2 and in the D_s mass sidebands $\mid m(D_s) - 1968 \mid > 30$ MeV/c^2 as **training background**. With these requirements, the D_s candidate is just a random combination of tracks.

The variables selected showing a good discriminating power between signal and background are (see Fig. 4.4):

- variables related to the D_s candidate:

  - the transverse momentum (p_{T}),
  - the pseudorapidity (η),
  - the flight distance from the primary vertex (FD_OWNPV),
  - the flight distance from the B_s^0 vertex (FD_ORIVX),
  - the vertex reconstruction χ^2 (ENDVERTEX_CHI2),

- variables related to the D_s decay products:
  - the transverse momentum (p_{T}),
  - the impact parameter reconstruction χ^2 from the primary vertex (IP_CHI2_OWNPV),
  - the track ghost probability.

Figure 4.4 shows the distributions of the input variables for the signal (black) and background (red). In order to develop a common selection for the signal and the normalization channels, the input variables were chosen among those that show a similar distribution between signal and MC normalizations $B_s^0 \to D_s^- 3\pi$ (green) and $B^0 \to D^- 3\pi$ (blue).

The output of the BDT classifier is shown in Figure 4.5a. The separation of the distributions for the background (red histogram) with respect to the signal (black) is clearly visible, indicating the success of the BDT definition. Moreover, the similarities of the distributions for the signal and normalization channels is also indicating that the BDT satisfies the requirement of performing a common selection to both signal and normalization decays.

The signal selection efficiency as a function of the BDT cut is plotted in Fig. 4.5b together with the background selection efficiency. By choosing a cut at the BDT output value of -0.0876 it is possible to select 95% of the signal and reject a sizeable fraction (77.5%) of the background.

4.2.2.2 3π selection

The "3π selection" BDT is implemented using the following training events:

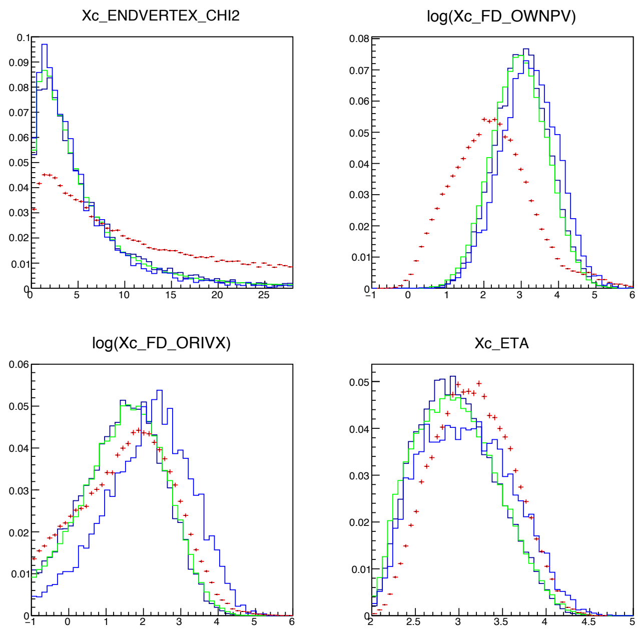
- three-pions candidates from the MC signal decay $B_s^0 \to D_s^- \tau^+ \nu_\tau$ (with $D_s^- \to K^+ K^- \pi^-$ and $\tau^+ \to \pi^+ \pi^- \pi^+ (\pi^0) \nu_\tau$) as **training signal**. The three pions are required to be associated to the "true" pions from the τ decay in the signal decay chain,

- three-pions candidates from the B_s data sample selected events with $m(B_s) > 5450$ MeV/c^2 as **training background**. In this mass region the contamination of b-hadron decays is negligible, so the background is mainly due to combinatorial.

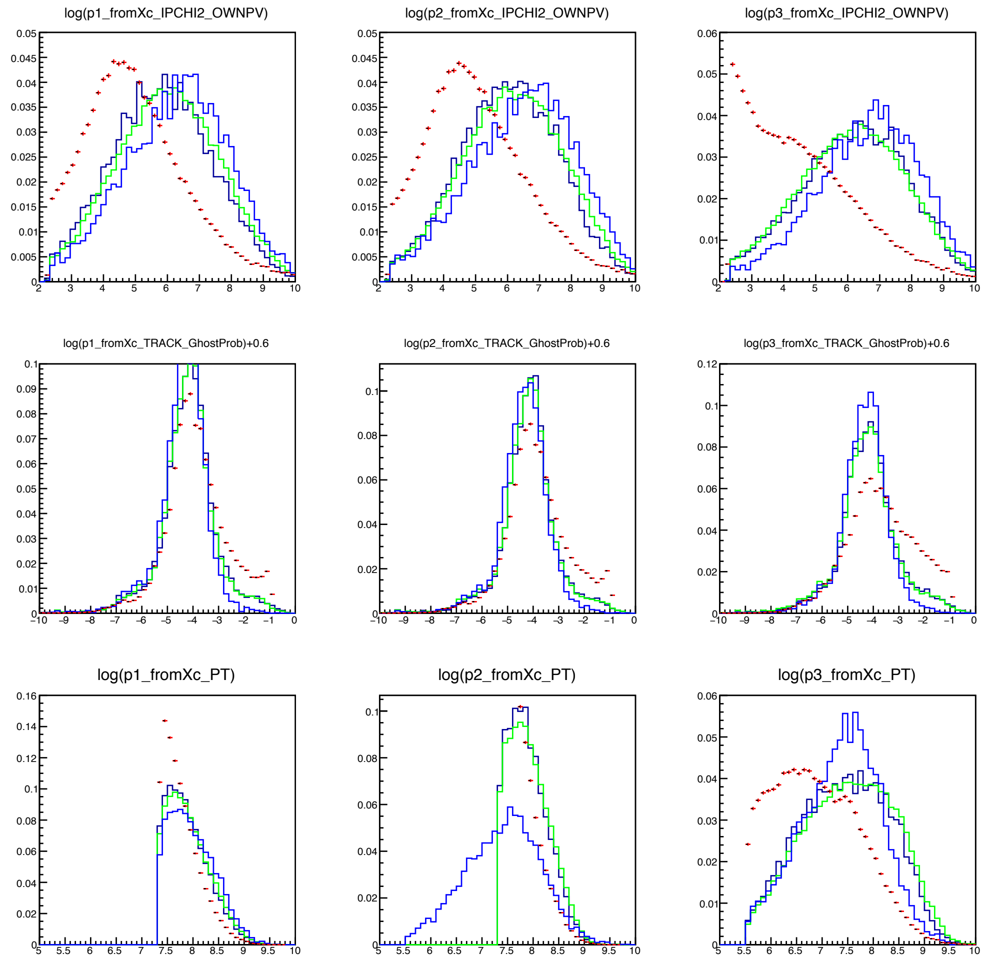
The variables selected showing a good discriminating power between signal and background are (see Fig. 4.6):

- variables related to the 3π vertex:
  - the distance of closest approach (DOCA) of one of the tracks compared to the intersection point of the other two
  - the vertex reconstruction χ^2 (ENDVERTEX_CHI2),

- variables related to the three pions:
  - the pseudorapidity (η),
  - the impact parameter reconstruction χ^2 from the primary vertex (IP_CHI2_OWNPV),

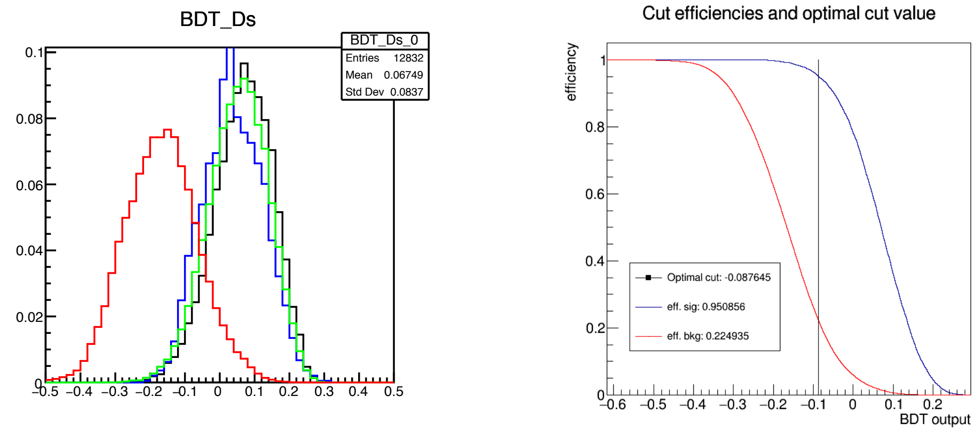


(a)



(b)

Figure 4.4: Variables used to train the BDT for the "D selection". The black, green and blue distributions represent respectively the MC samples of $B_s^0 \to D_s^- \tau^+ \nu_\tau$ (signal channel), $B_s^0 \to D_s^- 3\pi$ and $B^0 \to D^- 3\pi$ (normalization channels). The red distribution shows the B_s data sample with m$(B_s) > 5200$ MeV/c^2 and the sidebands of the D_s mass |m(D_s) - 1968| > 30 MeV/c^2.



(a) (b)

Figure 4.5: Output of the BDT implemented for the "D selection". A specific cut on the output of the BDT classifier is chosen.

– the track ghost probability.

Figure 4.6 shows the distributions of the input variables for the signal (black) and background (red). Also in this case the input variables are chosen in order to perform a common selection for the signal and normalization channels.

As a consequence, variables such as flight distance of the $\tau/3\pi$ candidate are not chosen as input for the BDT training because this would originate a difference between the signal and normalization channels.

The output of the BDT classifier output is shown in Fig. 4.7a. Like for the BDT for the "D selection", the separation of the distributions for the background (red histogram) with respect to the signal (black) is clearly visible, indicating the success of the BDT definition. Moreover, the similarities of the distributions for the signal and normalization channels are also indicating that the BDT satisfies the requirement of performing a common selection to both signal and normalization decays.

The signal selection efficiency as a function of the BDT cut is plotted in Fig. 4.5b together with the background selection efficiency. By choosing a cut at the BDT output value of -0.0727 it is possible to select 95% of the signal and reject a sizeable fraction (87.85%) of the background.

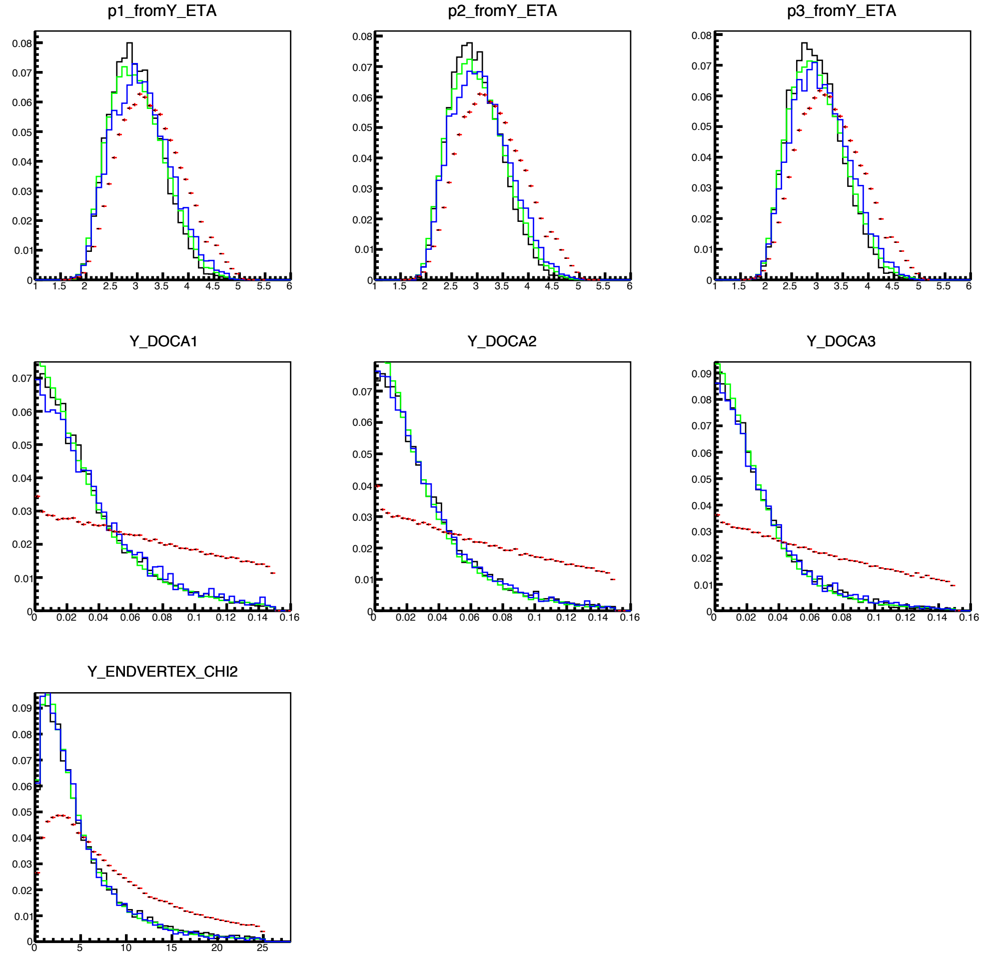
4.2.2.3 B selection

As a final step, a third multivariate analysis is performed to select signal $B^0_{(s)}$ candidates. To gain signal-to-background discrimination power, a cut on the output of the BDT classifiers for the "D selection" and "3π selection" is applied prior the training of the BDT for the "B selection". The training samples are:

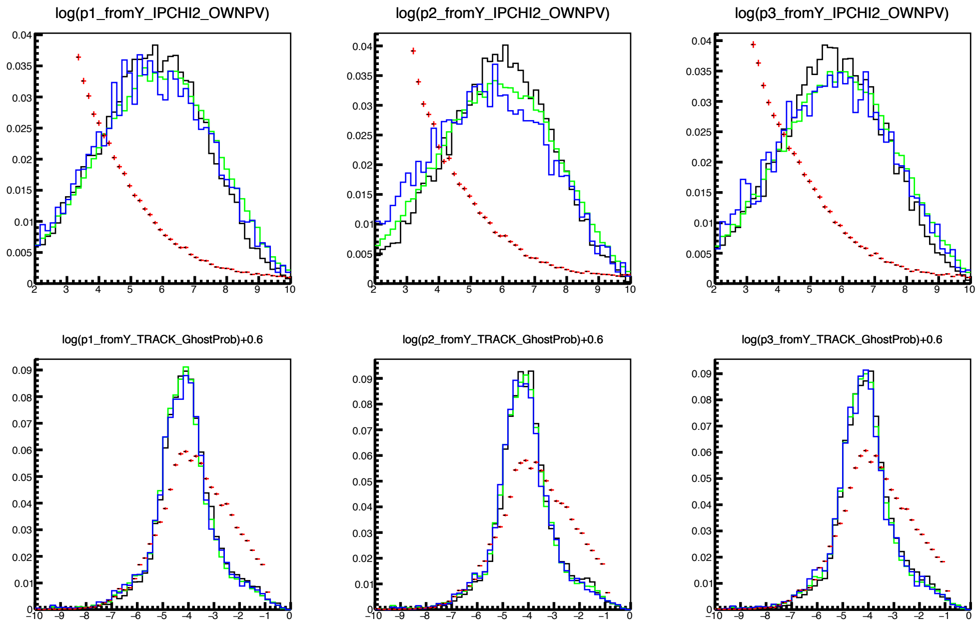
- B^0_s candidates from the MC signal decay $B^0_s \to D^-_s \tau^+ \nu_\tau$ (with $D^-_s \to K^+ K^- \pi^-$ and $\tau^+ \to \pi^+\pi^-\pi^+(\pi^0)\nu_\tau$) as **training signal**. The B^0_s candidate is required to be associated to the "true" signal.

- B^0_s candidates from the B_s data sample selected events with $m(B^0_s) > 5450$ MeV$/c^2$ as **training background**

- B^0_s candidates from the B_s "wrong-sign" data sample selected events with $m(B_s) >$ 5300 MeV$/c^2$ as additional **training background**. These candidates are selected

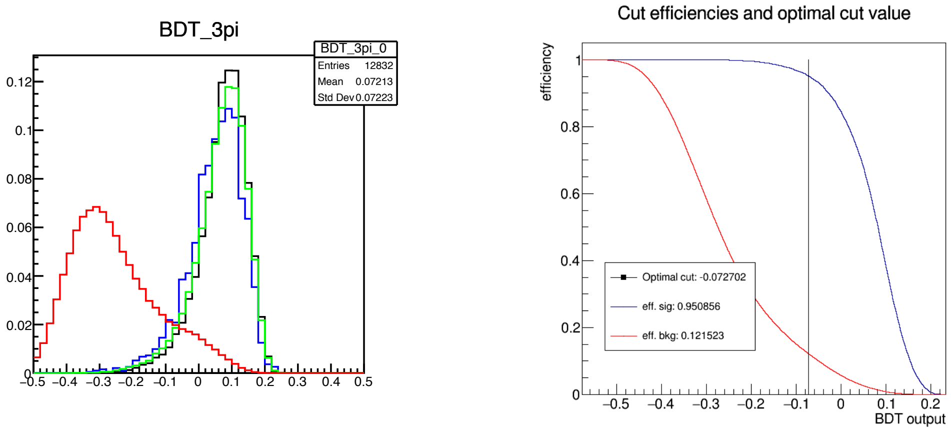


(a)



(b)

Figure 4.6: Variables used to train the BDT for the "3π selection". The black, green and blue distributions represent respectively the MC samples of $B_s^0 \to D_s^- \tau^+ \nu_\tau$ (signal channel), $B_s^0 \to D_s^- 3\pi$ and $B^0 \to D^- 3\pi$ (normalization channels). The red distribution shows the B_s data sample with $m(B_s) > 5450$ MeV/c^2.



(a) (b)

Figure 4.7: Output of the BDT implemented for the "3π selection". A specific cut on the output of the BDT classifier is chosen.

with the same criteria applied to the signal but are obtained from a dedicated stripping line ("`Bs2DsTauNuWSForB2XTauNu`") that combines D_s and 3π candidates of the same (wrong) charge. This sample show similar features to those of the combinatorial background of the "right-sign" combination and is added to training background sample to enrich statistics, as this will help the multivariate analysis to achieve a better classification.

The variables selected showing a good discriminating power between signal and background are (see Fig. 4.8):

- the B_s^0 pseudorapidity (η),
- the B_s^0 flight distance from the primary vertex (FD_OWNPV),
- output of the BDT for the "D selection" output,
- output of the BDT for the "3π selection".

The output of the BDT classifier output is shown in Fig. 4.9a. In this case the separation of the distributions for the background (red histogram) with respect to the signal (black) is less pronounced with respect to the previous cases, but still present. The distributions of the BDT output for the signal and the B_s normalization channels are very similar, while some differences are visible for the B normalization channel. This effect can be attributed to the differences in the kinematics and in the massed of the B_s and B decays.

The signal selection efficiency as a function of the BDT cut is plotted in Fig. 4.9b together with the background selection efficiency. By choosing a cut at the BDT output value of -0.0655 it is possible to select 90% of the signal and reject 65.63% of the background.

4.2.3 Signal and Normalization separation

After the BDT selections, the final selection step consists in separating the normalization from the signal to evaluate the normalization yield and efficiency. The main difference between the signal and the two normalization channels concerns the properties of the three pions: in the signal decay they originate from the τ decay, which is displaced from

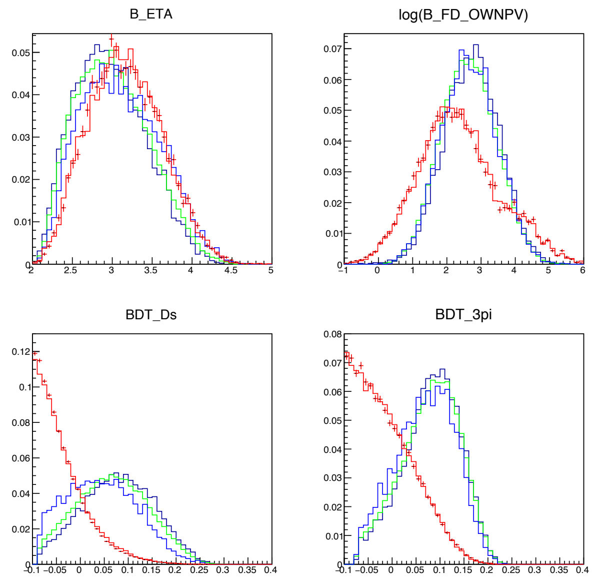
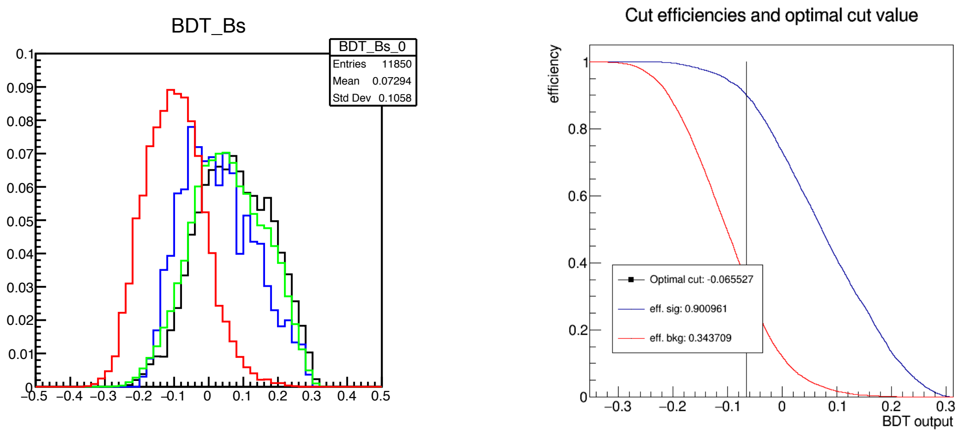


Figure 4.8: Variables used to train the BDT for the "B selection". The black, green and blue distributions represent respectively the MC samples of $B_s^0 \to D_s^- \tau^+ \nu_\tau$ (signal channel), $B_s^0 \to D_s^- 3\pi$ and $B^0 \to D^- 3\pi$ (normalization channels). The red distribution shows the B_s data sample with $m(B_s) > 5450$ MeV$/c^2$ and the red distribution with error bars corresponds to the B_s "wrong-sign" data sample with $m(B_s) > 5300$ MeV$/c^2$.



(a) (b)

Figure 4.9: Outputs of the BDT implemented for the "B selection". A specific cut on the BDT classifier output is chosen.

the B_s^0 decay vertex, while in the normalization channels they originate exactly from the $B_{(s)}^0$ decay vertex. As a consequence, the reconstructed displacement of the τ/three-pions vertex with respect the B vertex, namely the flight distance, can be used to separate the signal from the normalization channels.

The chosen cut selects 90% of the normalization decays (see Fig. 4.10).

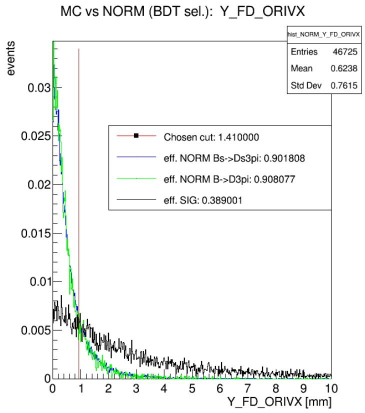


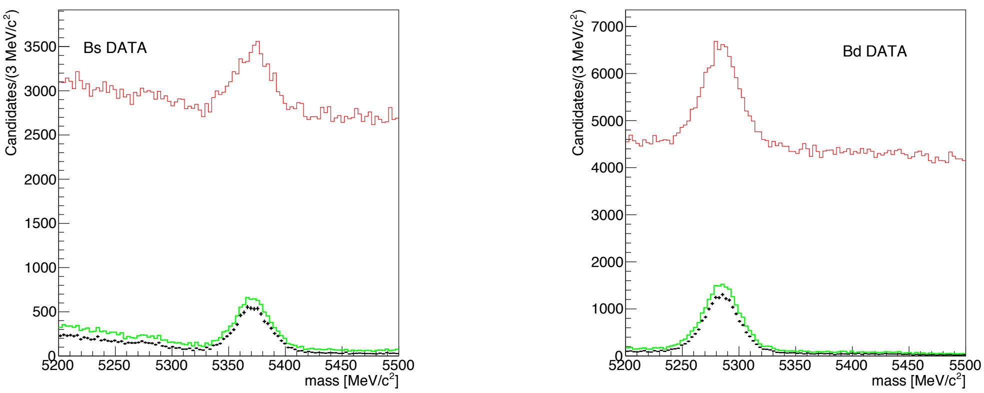
Figure 4.10: Distribution of the flight distance of the τ/three-pions system with respect to the $B_{(s)}$ decay vertex for the signal $B_s^0 \to D_s^- \tau^+ \nu_\tau$ (black) and the normalizations channels $B_s^0 \to D_s^- 3\pi$ (green) and $B^0 \to D^- 3\pi$ (blue). The chosen cut (red) selects 90% of the normalization decays.

4.3 Results

4.3.1 Selected data

The mass distributions of the B_s^0 and B_d^0 data after the $D_{(s)}$ preselection (red), the BDT selections (green) and with requirement on the three-pion flight distance from the $B_{(s)}$ decay vertex (black) is shown in Fig. 4.11, where the impact of the BDT selection on the background suppression is evident. Candidates from the normalization channel contribute to the peaking structure around the nominal $B_{(s)}^0$ mass, while the signal candidates have a continuous distribution at values lower than the B_s^0 mass, down to 3000 MeV/c^2, due to the inclusive final state reconstruction (missing neutrinos).

Contrary to the normalization channel, candidate signal B_s^0 decays needs to be further selected to suppress the different sources of background from several partially reconstructed b-hadron decays. A detailed study of such selection will be done in the future and is not the purpose of this thesis.



(a) (b)

Figure 4.11: Invariant mass distribution of the (a) B_s^0 (b) B_d^0 candidates after the (red) $D_{(s)}$ preselection, (green) the BDT selections and (red) the flight distance of the three-pions system with respect to the $B_{(s)}$ decay vertex.

4.3.2 Determination of the $\mathcal{K}$ ratio

In order to determine the ratio $\mathcal{K}$ both the yields of signal and normalization channels needs to be determined, as well as the corresponding efficiencies. As already mentioned, the signal selection must be further defined, therefore the determination of its yield is premature. Nevertheless, for the purpose of this thesis it is only needed to compare the quantities related to the two normalization channels, in particular the contribution to the uncertainty on $R(D_s)$.

4.3.2.1 Yield of normalization channels and fit model

The yield of normalization decays are determined from a fit to the invariant mass distributions of candidates selected by the full selection. The fit consists in an unbinned maximum-likelihood adaption to data of a probability density function (PDF) that describes several contributions:

- the peaking distribution due to the normalization decays. Its model is obtained from a fit to the normalization MC samples and consists in the sum of a Gaussian and Crystal Ball [70] functions;
- the combinatorial background, due to random combinations of one or more tracks not associated to b-hadron decays. Its model is described by an exponential function;
- partially reconstructed and mis-identified b-hadron decays, with dedicated functions.

The fit code is developed as `C++` code and exploits the PDF libraries of `RooFit` available within the ROOT package [71].

4.3.2.2 Fit of the B_s^0 normalization channel

Figure 4.12 shows the mass distribution of the $B_s^0 \to D_s^- 3\pi$ normalization decays in MC after the full selection. It is well represented by the sum of a Gaussian function, with parameters μ and σ_{gauss}, and a Crystal Ball function, with parameters μ, σ_{CB} and n_{CB}. The two distributions have a common mean μ and a relative fraction f_{gauss}. The parameter n_{CB} is fixed to 2. With the exception of the mean, all the parameters that best fit the MC distribution are then fixed to their central values while fitting to data. A common scale factor on the resolution parameters is included to account for possible differences between data and MC.

The fit to data requires the description of the contributing background components. The combinatorial background is described by an exponential function. In absence of MC samples for specific b-hadron decays that can contribute to the distribution of the selected candidates a fast and basic simulation based on the `TGenPhaseSpace` class from the ROOT package [72] is used. Several b-hadron decays with topologies similar to the normalization channel were considered. The simulation accounts for the observed kinematic distribution of the b-hadron decay and the experimental resolutions that affect the final state particle measurements. Only the part of the selection using kinematic cuts is applied.

Figure 4.13 shows the reconstructed $D_s 3\pi$ invariant mass for several b-hadron decays considered. The only background contributions in the mass range 5250-5500 MeV/c^2 of interest are:

- the $B_s^0 \to D_s^- K^+ \pi^- \pi^+$, where a the final state kaon is misidentified as a pion. This background is be modeled by a Crystal Ball function;
- the $B_s^0 \to D_s^{*-}(D_s^- \gamma) 3\pi$ decay, where the γ is not reconstructed. This background is modelled by an analytic function;

- the $B_d^0 \to D^-(\to K^+\pi^-\pi^-)3\pi$, where a π^- from the D^- decay is misidentified as a kaon. This background is modelled by the sum of a Novosibirsk and a Crystal Ball functions. Due to the "vetoes" applied (see section 4.2.1 this background is highly suppressed (a preliminary estimate using MC gives a relative contribution less than 1%) and indeed is found to be consistent with zero in the fit.

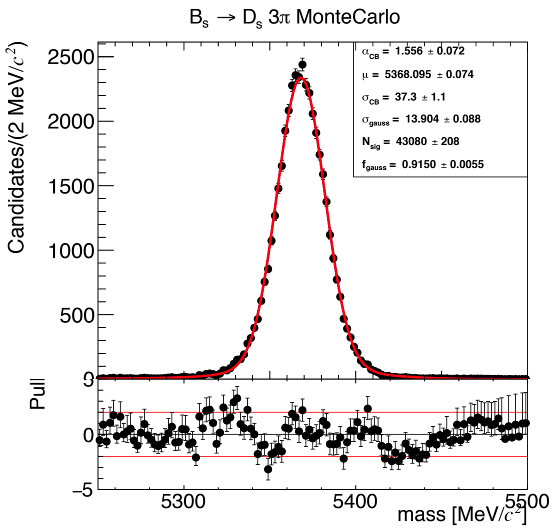


Figure 4.12: Invariant mass distribution of the $B_s^0 \to D_s^- 3\pi$ decays in MC after the full selection. The red function corresponds to the best fit to data. The fit parameters are shown in the legend. The pulls in the bottom plot are evaluated in each bin as the difference between the data and the fit value over the error and are all within ± 3, indicating a good fit.

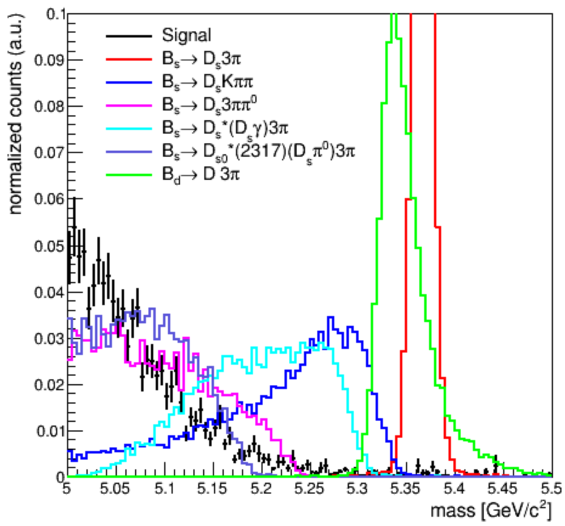


Figure 4.13: `TGenPhaseSpace` simulation of several B^0 decays in the 5000-5500 MeV/c^2 mass range. The normalization channel $B_s^0 \to D_s^- 3\pi$ is plotted in red. Possible background distributions are plotted in different colors as explained in the legend.

The data mass spectrum is fitted in the range 5250-5500 MeV/c^2 from the B_s^0 data sample to evaluate the yield under the peaking normalization decay $B_s^0 \to D_s^- 3\pi$. The mass fit model includes a signal model, with parameters fixed from the MC sample fit and a background model which involves combinatorial and known decays with free parameters.

Figure 4.14 shows the B_s^0 data mass spectrum fit with the full selection applied. The final yield for the normalization channel $B_s^0 \to D_s^- 3\pi$ is:

$$N_{\rm norm} = 6542 \pm 142$$

which contributes to the statistical uncertainty on $\mathcal{K}$ of 2.2%.

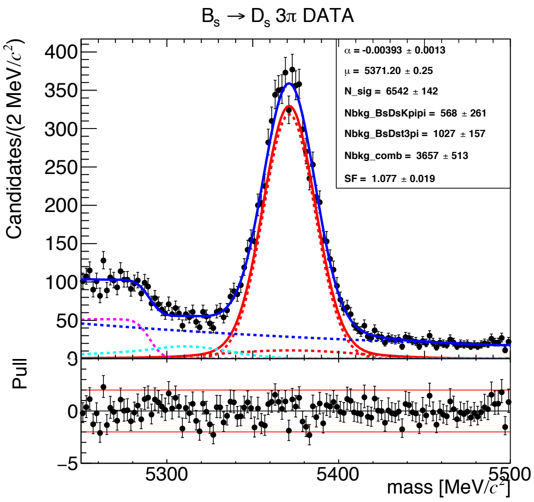


Figure 4.14: B_s^0 data sample mass spectrum fit in the range 5250-5500 MeV/c^2 with the full selection applied. The signal and background yields are reported in the legend.

4.3.2.3 Fit of the B_d^0 normalization channel

Figure 4.15a shows the mass distribution of the $B_d^0 \to D^- 3\pi$ normalization decays in MC after the full selection. The same fit model used for the B_s^0 normalization channel is used. With the exception of the mean, all the parameters that best fit the MC distribution are then fixed to their central values while fitting to data. A common scale factor on the resolution parameters is included to account for possible differences between data and MC.

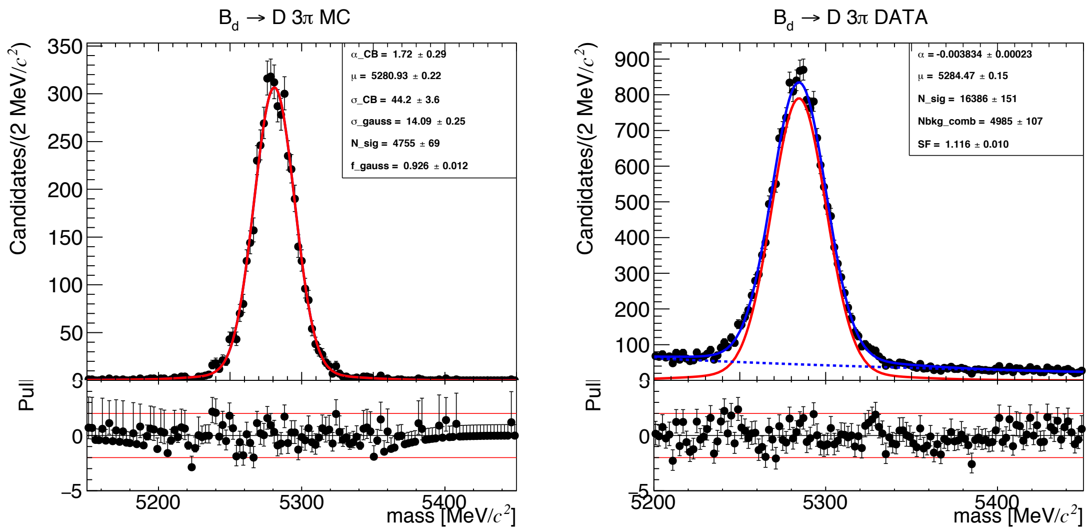
To describe the data distribution in the considered range of mass 5150-5400 MeV/c^2, the combinatorial background is sufficient. Figure 4.15b shows the B_d^0 data mass spectrum fit with the full selection applied. The final yield for the normalization channel $B_d^0 \to D^- 3\pi$ is:

$$N_{\rm norm} = 16386 \pm 151$$

which contributes to the statistical uncertainty on $\mathcal{K}$ of 0.9%.

4.3.2.4 Efficiencies

The signal and normalization efficiencies are evaluated respectively from the MC samples $B_s^0 \to D_s^- \tau^+ \nu_\tau$ ($\epsilon_{\rm sig}$), $B_s^0 \to D_s^- 3\pi$ and $B^0 \to D^- 3\pi$ ($\epsilon_{\rm norm}$). The computation first needs to take into account the generation, filtering and stripping efficiencies for the produced MC



(a) (b)

Figure 4.15: Invariant mass distributions of the $B_d^0 \to D^- 3\pi$ decays in (a) MC and (b) data after the full selection. The red function corresponds to the signal component, while the dashed blue line is the combinatorial background. The fit parameters are shown in the legend.

samples, then the various steps of the offline selection discussed in the previous sections. Table 4.5 reports the efficiencies values at each step of the selection.

The efficiency ratio can be evaluated for both normalization channels. For the $B_s \to D_s 3\pi$ channel, the value is:

$$\frac{\epsilon_{\text{norm}}}{\epsilon_{\text{sig}}} = 5.79 \pm 0.08\,.$$

This computation includes the statistical uncertainty due to the size of the MC samples which gives a contribution to the uncertainty of the $\mathcal{K}$ measurement of 1.4%. Though both the signal and the normalization channels are selected in the same way, the value of the ratio deviates from one. This result is due to the different kinematic of the two decays.

Decay	$B_s \to D_s \tau\nu$	$B_s \to D_s 3\pi$	$B \to D3\pi$
	ϵ_{sig} [%]	ϵ_{norm} [%]	ϵ_{norm} [%]
Generation	14.878 ± 0.017	3.414 ± 0.005	15.40 ± 0.02
Filtering	0.245 ± 0.001	2.459 ± 0.007	0.47 ± 0.003
Stripping	51.9 ± 0.2	75.67 ± 0.12	70.5 ± 0.3
MC truth	61.9 ± 0.3	75.17 ± 0.14	68.1 ± 0.3
Trigger	96.89 ± 0.14	96.90 ± 0.07	97.32 ± 0.14
D-D_s sel.	82.6 ± 0.3	81.36 ± 0.15	55.31 ± 0.43
BDTs	83.5 ± 0.3	78.14 ± 0.18	67.6 ± 0.5
FD cut	61.0 ± 0.5	90.07 ± 0.14	90.1 ± 0.4
Total	(4.57 ± 0.06)×10^{-5}	(26.43 ± 0.15)×10^{-5}	(11.39 ± 0.17)×10^{-5}

Table 4.5: Selection efficiencies evaluated on MC samples for each step of the selection, from generation level cuts to the final selection. Each efficiency cut is computed with respect to the previous cut. Uncertainties are statistical only.

In particular, due to the presence of additional neutrinos in the signal decay, the charged final state particles have lower momenta with respect the normalization channel.

For the $B \to D3\pi$ channel, the value is:

$$\frac{\epsilon_{\text{norm}}}{\epsilon_{\text{sig}}} = 2.49 \pm 0.05\,.$$

The contribution to the uncertainty of the $\mathcal{K}$ ratio is 2%. This uncertainty relies on the analysis of a partial MC sample for the B_d normalization channel that was available during the thesis. With the full sample the uncertainty on the efficiency ratio is expected to reduce to 1.3%. In this case the signal and the normalization channels are not exactly selected in the same way, due to the different stripping selections. As a result the ratio deviates from that obtained with the B_s normalization channel.

In the next section, the possible effect of the systematic uncertainties will be reviewed and their contribution to the determination of $\mathcal{K}$ will be computed.

4.4 Systematics

The systematic uncertainties on $\mathcal{K}$ are mainly divided in three categories, related to:

- the model used to fit the mass distribution in data, that affects the yield of normalization decays;
- uncertainty on the efficiency calculation related to the selection;
- the possible differences between the data and MC samples;

4.4.1 Uncertainties related to the fit model

The choice of the fit model can introduce a possible bias in the yield evaluation. In particular, both the modeling of the background sources and of the normalization contribution can affect the results.

Figure 4.16(a) shows the reconstructed $D3\pi$ invariant mass for several b-hadron decays considered, simulated with the `TGenPhaseSpace` tool. The only background contributions in the mass range 5200-5450 MeV/c^2 of interest are:

- the $B^0 \to D^- K^+ \pi^- \pi^+$, where a the final state kaon is misidentified as a pion. This background is be modeled by a Crystal Ball function;
- the $B^0 \to D^{*-}(D^- \gamma)3\pi$ decay, where the γ is not reconstructed. This background is modelled by an analytic function;
- the $B^0_s \to D^-_s(K^+K^-\pi^-)3\pi$, where the K^- is misidentified as a π^-. This background is modelled by the sum of a Novosibirsk and a Crystal Ball functions.

The data mass spectrum is fitted in the range 5200-5450 MeV/c^2 from the B^0 data sample with the new model, which includes the possible background sources discussed above. The background contamination fraction of $B^0_s \to D^-_s 3\pi$ is evaluated using the $\frac{f_s}{f_d}$ production ratio, the branching fractions of the two decay modes and the efficiencies measured in MC. The fraction is fixed in the fit to a value of 1.8% to avoid any yield bias.

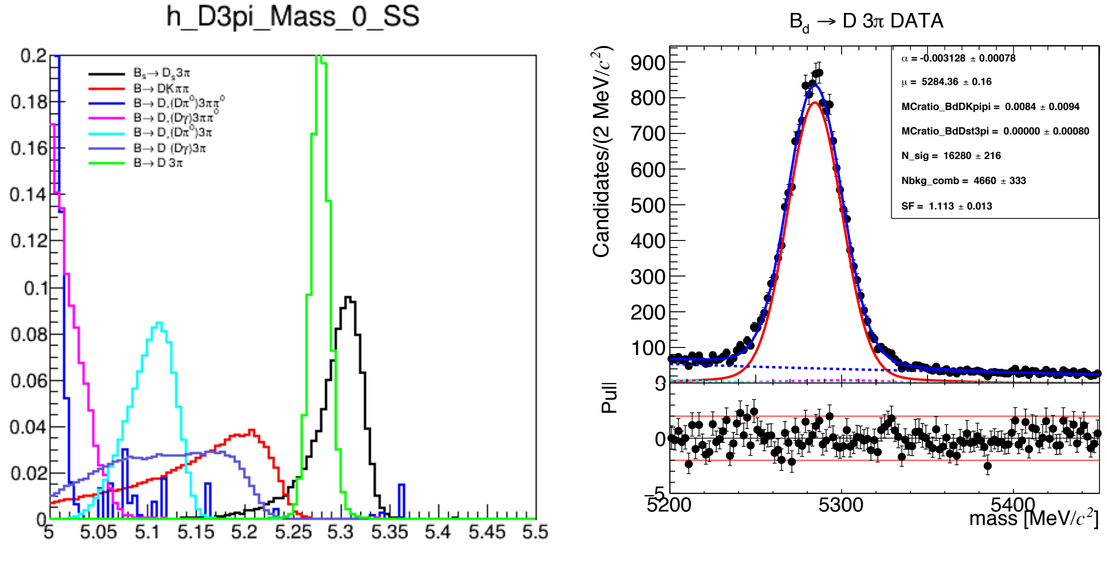
The new yield found for the normalization $B^0 \to D^- 3\pi$ is:

$$N_{\text{norm}} = 16280 \pm 216\,,$$

compared to the yield found with just the combinatorial background,

$$N_{\text{norm}} = 16386 \pm 151\,.$$

The two values are consistent with each other and, as shown in Fig. 4.16(b), the fractions of the newly introduced backgrounds are consistent with 0 within 1σ. Conservatively, since the uncertainties are correlated, the difference in the normalization yield of 0.7% is assigned as systematic uncertainty due background composition in the fit model.



(a) (b)

Figure 4.16: (a)`TGenPhaseSpace` simulation of several B^0 decays in the 5000-5500 MeV/c^2 mass range. The normalization channel $B^0 \to D^-3\pi$ is plotted in green. Possible background distributions are plotted in different colors as explained in the legend. (b) Invariant mass distribution of the $B_d^0 \to D^-3\pi$ decays in data after the full selection. The red function corresponds to the signal component, while the dashed blue line is the combinatorial background. The background decays studied in the `TGenPhaseSpace` simulation are added to the model. The background contamination fraction of $B_s^0 \to D_s^-3\pi$ is fixed.

In a similar way for the B_s data sample fit, the background contamination fraction of $B^0 \to D^-3\pi$ is evaluated using the $\frac{f_s}{f_d}$ production ratio, the branching fractions of the D and D_s decays and the efficiencies measured in MC. The fraction is fixed in the fit to a value of 0.8% to avoid any yield bias.

The new yield found for the normalization $B_s^0 \to D_s^-3\pi$ is (see Fig. 4.17):

$$N_{\rm norm} = 6506 \pm 141\,,$$

compared with the yield found with the nominal fit,

$$N_{\rm norm} = 6542 \pm 141\,.$$

The two results are consistent with each other and correspond to similar fit qualities as shown in Figure 4.17. Conservatively, since the uncertainties are correlated, the difference in the normalization yield of 0.6% is assigned as systematic uncertainty due to the fit model.

To evaluate possible systematic uncertainties related to the parametrization of normalization contribution, an alternative method is used. The fit is performed on the invariant mass obtained by a fit ("Decay Tree Fitter", DTF) which implements kinematic constraints on the $D_{(s)}$ mass and the requirement that the origin $B_{(s)}$ vertex coincides with the PV. An algorithm takes a complete decay chain, parameterizes it in terms of vertex positions,

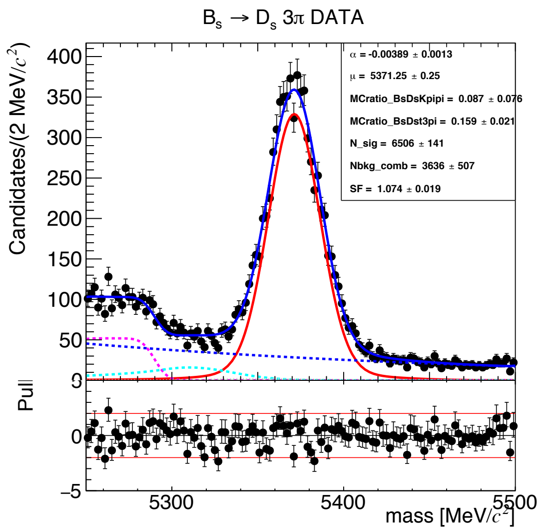


Figure 4.17: Invariant mass distribution of the $B_s^0 \to D_s^- 3\pi$ decays in data after the full selection. The background contamination fraction of $B^0 \to D^- 3\pi$ is fixed.

decay lengths and momentum parameters, and then fits these parameters simultaneously, taking into account the relevant constraints.

The $B_{(s)}$ mass observables obtained in this way can be used to evaluate the normalization yield. The fit model is retrieved from MC with DTF variables in the same way as in section 4.3.2.1.

The new yields found for the normalization $B_s^0 \to D_s^- 3\pi$ and $B^0 \to D^- 3\pi$ (see Fig. 4.18) are respectively:

$$N_{\text{norm}} = 6478 \pm 100\,,$$

and

$$N_{\text{norm}} = 16464 \pm 147\,,$$

The difference in the normalization $B_s^0 \to D_s^- 3\pi$ and $B^0 \to D^- 3\pi$ yields of respectively 1.0% and 0.5 % are assigned as systematic uncertainty due to the fit model for the normalization channels.

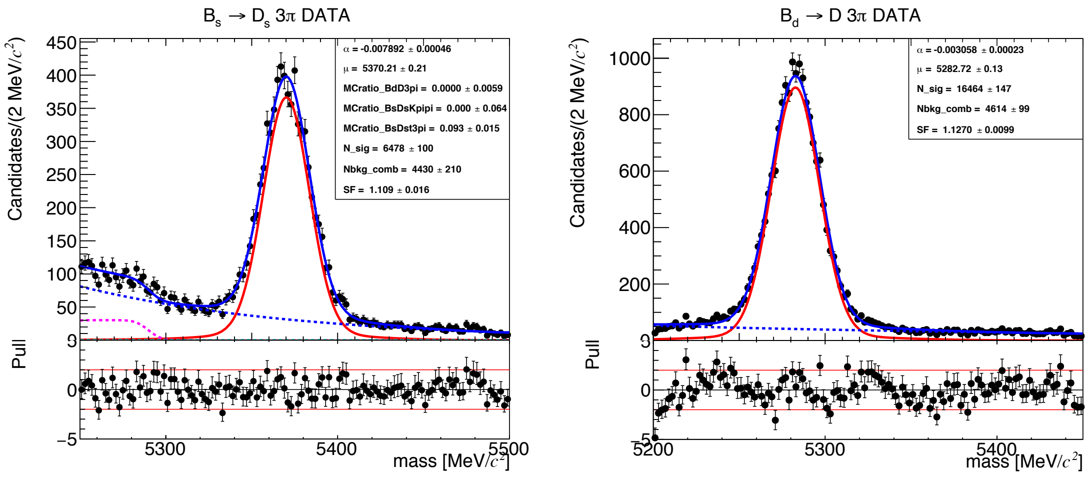
The total systematic uncertainty due to the fit model, including both the uncertainties due to background contributions and those due to the modelling of normalization channels, is 1.2% for the $B_s^0 \to D_s^- 3\pi$ and 0.9% for the $B^0 \to D^- 3\pi$ normalization channels.

4.4.2 Uncertainties related to the selection

Systematics related to the implemented selection affect both the signal and normalization efficiencies. Various techniques are employed to evaluate them, as explained in the next sections.

4.4.2.1 Particle Identification

MC simulation is corrected in order to match the performance of particle identification criteria measured in data. For this purpose, the particle identification variables are corrected using the "PIDCalib" package [73]. In particular, the set of scripts "PIDGen" allows a resampling of the PID response. The resampling approach uses PID distributions



(a) (b)

Figure 4.18: Invariant mass distributions of the (a) $B_s \to D_s^- 3\pi$ and (b) $B_d \to D^- 3\pi$ decays in data obtained with constraints on the $D_{(s)}^-$ mass and on the origin decay vertex requirements. The fit model for the normalization channel is determined on the same variables on MC.

from calibration data as probability density functions (PDFs) and generates random PID variables from those distributions based on the kinematic region to which the simulated event belongs. To account for the dependence of PID distribution on the kinematics of the track and event multiplicity, the PDFs for the PID response are calculated in bins of momentum p_{T} and pseudorapidity η of the track, and the number of tracks in the event N_{tracks}. This correction is used to evaluate the reference efficiencies of the D/D_s preliminary selection . To evaluate the systematic uncertainties related to the PID cut, an alternative calculation is performed. It is based on the tables of efficiencies for the PID cuts applied, that are determined from the PIDCalib package in bins of momentum p, pseudorapidity η of the track, and the number of tracks in the event N_{tracks}. Table 4.6 shows the D/D_s preliminary selection efficiencies on the MC samples computed with the two methods described, which affect both signal and normalization efficiencies.

The difference of the two efficiencies is assigned as systematic uncertainty and propagated to the efficiency ratio $\frac{\epsilon_{\mathrm{norm}}}{\epsilon_{\mathrm{sig}}}$. The relative systematic uncertainty is 0.3%, for the $B_s \to D_s 3\pi$ normalization channel, and 4.4% for the $B \to D3\pi$ normalization channel.

Decay	ϵ with PIDGen [%]	ϵ with PIDCalib [%]	relative difference [%]
$B_s \to D_s \tau\nu$	82.6 ± 0.3	79.2 ± 0.3	4.3
$B_s \to D_s 3\pi$	81.36 ± 0.15	78.25 ± 0.15	3.9
$B \to D3\pi$	55.31 ± 0.43	55.38 ± 0.40	0.1

Table 4.6: Comparison between the D/D_s selection efficiencies with two different methods to compute the PID efficiencies.

4.4.2.2 Trigger efficiency

The trigger selection efficiencies are computed on MC (see Table 4.5). Due to differences between data and MC these values could be affected by systematic uncertainties. However, since the computed efficiencies for the signal and normalizations MC samples are consistent among each other and very high ($\sim 97\%$), the expected deviations from their nominal values of the order of 1%, estimated on previous analyses [49], is associated as a systematic uncertainty to the trigger efficiency.

4.4.2.3 Momentum scale calibration

An accurate estimate of the charged momentum scale is a critical ingredient to achieve an optimal performance in forward spectrometers such the LHCb detector. If the momentum scale is not well calibrated the measured four vectors of charged particles will be biased. If the bias is large and depends on the particle kinematics the mass resolution may also be degraded. A deviation of the measured mass from the expected value, can be converted into an estimate of the momentum scale bias, referred as α. This is defined such that the measured mass would become equal to the expected value if all particle momenta were multiplied by 1-α. The accuracy of the momentum scale is determined by the knowledge of the magnetic field map, the alignment of the tracking detectors and the detector material. Corrections are performed that account for the variation of the momentum scale with time, particle kinematics and fix the global momentum scale. They are obtained by requiring the mass of the narrow peaks such as J/ψ, K_s, Λ, Υ and ψ(2S), match their nominal masses, which are known very precisely.

A scale factor is introduced during the fit to the mass spectrum in data to account for residual differences in the mass resolution between data and simulation due to the momentum scale calibration (Fig.s 4.14 and 4.15b). As a result the selection efficiency could be affected. To evaluate the corresponding systematic uncertainty, this correction is included in the selection by scaling the range of the $D_{(s)}$ mass cut (4.2.1) by a factor of 1.09. The overall efficiencies for the signal and normalization channels are re-computed with the correction and their deviations with respect to the nominal values are assigned as systematic uncertainties. As a result, the systematic uncertainties assigned to the efficiency ratio $\frac{\epsilon_{\text{norm}}}{\epsilon_{\text{sig}}}$ are $< 0.1\%$ and 0.4% respectively choosing the $B_s \to D_s 3\pi$ or $B_d \to D3\pi$ normalization channel.

4.4.3 Uncertainties related to data-MC samples differences

Simulation reproduces quite well the data distributions both in terms of resolutions, efficiencies, kinematics of the decay and the properties of the event. However some small deviations exists and can affect the evaluation of the total selection efficiency.

A first check is done by comparing the distributions of the BDTs output between data and MC for the normalization channels. To cancel the background contribution data, the "`sPlot`" method [74] is used. This technique assigns to each candidate a "signal weight" based on the result of the fit to the mass distribution. The weighted distributions of the BDTs output in data are compared to those obtained in MC in Fig 4.19. As a first visual control, the distribution shapes are in reasonable good agreement with each other.

As a cross check, the efficiencies for the BDT selections and the requirement on the three-pion vertex displacement from the B decay are computed on data by comparing the yield of normalization channel after each selection. This is not possible for the signal channel, as there is not a simple peaking mass distribution to fit and the background composition needs to be studied in detail. As previously showed in Fig.s 4.6, 4.4 and 4.8, the signal channel $B_s \to D_s \tau \nu$ present very similar distributions to those of the normalization decay $B_s \to D_s 3\pi$. As a consequence, it is reasonable to assume that any efficiency variation on the signal decay on data with respect to MC are comparable to those observed in the $B_s \to D_s 3\pi$ normalization channel.

Table 4.7 shows a comparison between the offline selection efficiencies evaluated on the MC and data samples for the two normalization channels. The relative difference between

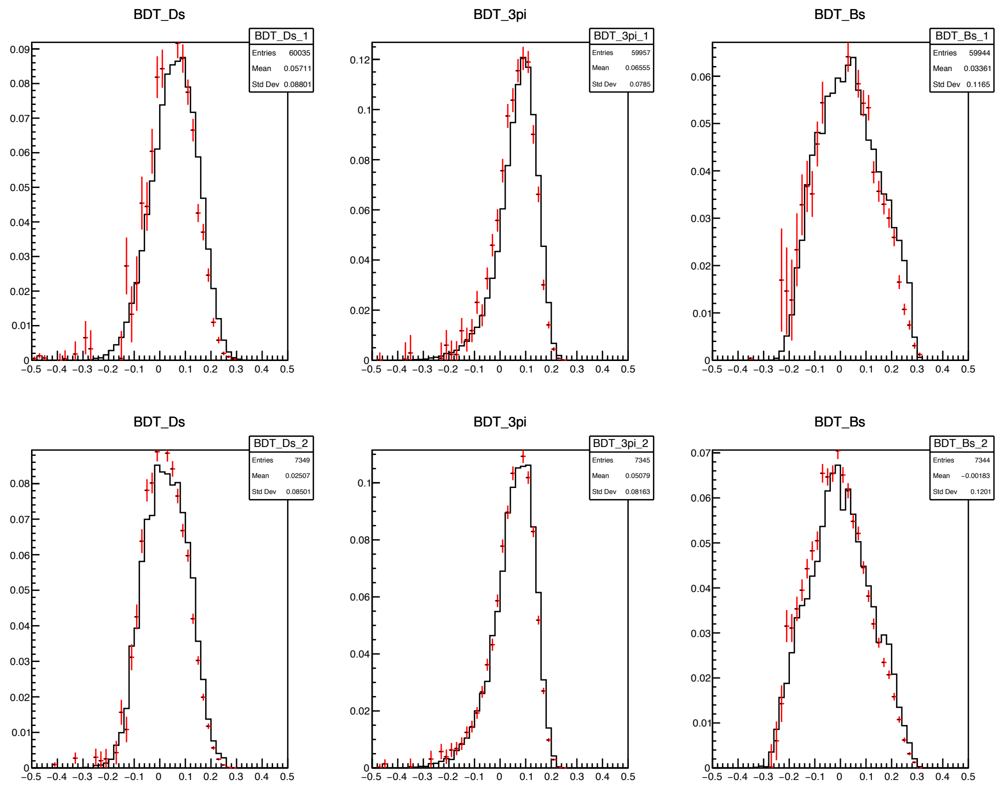


Figure 4.19: Comparison between MC (black) and signal-weighted data (red) distributions of BDTs classifier outputs for (top) $B_s \to D_s 3\pi$ and (bottom) $B \to D3\pi$ channels.

MC and data is assigned as systematic uncertainty. The values are respectively 6.4% and 8.3% for the $B_s \to D_s 3\pi$ and $B \to D3\pi$ normalization channels.

Assuming a systematic uncertainty of 6.4% for the signal, the resulting systematic uncertainty on the efficiency ratio $\frac{\epsilon_{\rm norm}}{\epsilon_{\rm sig}}$ is $< 0.1\%$ for the $B_s \to D_s 3\pi$ and 1.9% for the $B_d \to D3\pi$ normalization channels, respectively.

Decay	MC [%]	DATA [%]	relative difference [%]
$B_s \to D_s 3\pi$	70.43± 0.19	66.2 ± 2.1	6.4
$B \to D3\pi$	58.89 ± 0.6	54.4 ± 0.9	8.3

Table 4.7: Comparison between the efficiencies on MC and data for the normalization channels after the cuts on the BDT outputs and on the displacement of the three-pion vertex with respect the B decay vertex.

4.4.4 Summary of the systematic uncertainties

Table 4.8 summarizes the systematic uncertainties on the measurement of the $\mathcal{K}$ ratio. The total values are 1.2% and 5.0 % respectively choosing the $B_s \to D_s 3\pi$ or $B_d \to D3\pi$ normalization channel.

Contribution	B_s norm. channel [%]	B_d norm. channel [%]
Particle Identification	0.3	4.4
Trigger	1.0	1.0
Momentum scale calibration	< 0.1	0.4
Fit model	1.2	0.9
Data-MC differences	<0.1	1.9
Total	1.6	5.0

Table 4.8: Contribution of each systematic uncertainty to $\mathcal{K}$ ratio choosing the $B_s \to D_s 3\pi$ or $B_d \to D3\pi$ normalization channel.

4.5 Final results

The total uncertainty on the $R(D_s)$ measurement includes several factors summarized in Table 4.9:

- statistical uncertainties on the normalization yield measurement and on the selection efficiency ratio evaluated on MC;
- systematic uncertainties, discussed in the previous sections;
- external contributions due to the uncertainties on the branching fractions (section 3.1.2);

The total relative uncertainty on the $R(D_s)$ measurement is 19.3% choosing the $B_s \to D_s 3\pi$ normalization channel and 15.2% choosing the $B_d \to D3\pi$ normalization channel. As a consequence, the $B_d \to D3\pi$ is the best normalization channel for this analysis because it contribute to a smaller total relative uncertainty on the $R(D_s)$ measurement. The study of the signal selection will be deepened for the knowledge and suppression of the different background contributions. Nevertheless, the final result of this study gives important directions on how to define the overall analysis strategy and to the possible uncertainties related to the normalization channels.

It could be possible to include both normalization channels in the measurement, weighting the final result on the two channels. This could be done carefully measuring all the possible correlations between the analysis results.

Contribution	B_s norm. channel [%]	B_d norm. channel [%]
Normalization yield	2.2	0.9
MC statistics	1.4	1.3
Systematic uncertainties	1.6	5.0
External contributions	19.1	14.3
Total	19.3	15.2

Table 4.9: Summary of the overall uncertainties on the $R(D_s)$ measurement due to the selection implemented and the external branching fractions inputs.

Conclusions

In recent years the experimental results testing the validity of the Lepton Flavour Universality of the Standard Model electroweak interaction has become more precise and revealed hints of violation. If this picture is confirmed with more precise measurements, it will represent an indication that New Physics is also contributing to the underlying processes.

For this purposes, the LHCb experiment is performing and foreseeing several tests of LFU, exploiting both the different b-hadron species that are produced in pp collisions at LHC and different b-hadron decay modes.

This thesis provides an important contribution to the test of LFU in semileptonic decays $B_s^0 \to D_s^- \ell^+ \nu$ and the measurement of $R(D_s)$, defined as:

$$R(D_s) = \frac{\mathcal{B}(B_s^0 \to D_s^- \tau^+ \nu_\tau)}{\mathcal{B}(B_s^0 \to D_s^- \mu^+ \nu_\mu)} .$$

The $B_s^0 \to D_s^- \tau^+ \nu_\tau$ decay is reconstructed through the τ three-prong hadronic decay, $\tau^+ \to \pi^+ \pi^- \pi^+ (\pi^0) \nu_\tau$, while the D_s^- meson is reconstructed through the $D_s^- \to K^+ K^- \pi^-$ decay chain, so the visible final state consists of six charged tracks.

To achieve the best precision, $R(D_s)$ is determined by measuring the signal branching fraction with respect to a normalization channel, $\mathcal{K}$, which is then corrected by a combination of branching fraction measured by independent analyses, α as follows:

$$R(D_s) = \mathcal{K} \times \alpha = \frac{\mathcal{B}(B_s^0 \to D_s^- \tau^+ \nu_\tau)}{\mathcal{B}(\text{norm})} \times \frac{\mathcal{B}(\text{norm})}{\mathcal{B}(B_s^0 \to D_s^- \mu^+ \nu_\mu)} ,$$

where

$$\mathcal{K} = \frac{N_{\text{sig}}}{\epsilon_{\text{sig}}} \frac{\epsilon_{\text{norm}}}{N_{\text{norm}}} \frac{1}{\mathcal{B}(\tau^+ \to \pi^+ \pi^- \pi^+ (\pi^0) \nu_\tau) \times \mathcal{B}(D_s^- \to K^+ K^- \pi^-)} .$$

Here, N_{sig} (N_{norm}) and ϵ_{sig} (ϵ_{norm}) represent, respectively, the signal (normalization) yield and selection efficiency.

This thesis has studied different normalization channels suitable for the $R(D_s)$ measurement and identified the one that allows the best precision. Two normalization channels have been considered: $B_s^0 \to D_s^- \pi^+ \pi^+ \pi^-$ decays with $D_s^- \to K^+ K^- \pi^-$ and $B^0 \to D^- \pi^+ \pi^+ \pi^-$ decays with $D^- \to K^+ \pi^- \pi^-$. Both of them have similar topology to the signal and can be selected by a common selection. These features are particularly important to strongly reduce the systematic uncertainties on the ratio $\mathcal{K}$.

The analysis presented in this thesis has covered in detail all the aspects of the $R(D_s)$ measurement related to the normalization channel.

Firstly, an investigation of the experimental results on branching fraction measurements that contribute to the external measurements has been performed. Then a complete selection of the signal candidates has been developed. The selection provides a significant reduction of the background contribution in different steps. A simple cut-based selection, as well as the use of classifiers based on multivariate analyses, discriminate the signal from the main background source, due to combinatorial. The selection was developed by using samples of simulated signal and normalization channels decays, as well as background samples of data. The selection has been established in order to ensure the same performance on both the signal and the normalization channels.

The selection efficiencies ϵ_{sig} and ϵ_{norm} have been determined from the simulation samples, while the yield of the normalization channels by fitting the mass distribution of candidates passing the full selection in data.

The systematic uncertainties were finally evaluated. Their main contributions are due to the model used to fit the mass distribution in data, uncertainties on the efficiency calculation related to the selection and to possible differences between the data and MC samples.

All these studies allowed the determination of the contribution to the uncertainty on $R(D_s)$ due to the choice of the normalization channel.

The results indicate that the $B^0 \to D^-\pi^+\pi^+\pi^-$ normalization channel is favored with respect to the $B_s^0 \to D_s^-\pi^+\pi^+\pi^-$ decays decay as it contributes to the $R(D_s)$ measurement with an overall smaller relative uncertainty (15.2% instead of 19.3%).

The study of the signal selection will be deepened for the knowledge and suppression of the different background contributions. However, the final results of this study provide important directions on the possible uncertainties related to the normalization channels and on how to define the overall analysis strategy.

Bibliography

[1] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** (2008) S08001.

[2] J. Haffner, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, , General Photo.

[3] LHCb collaboration, R. Aaij *et al.*, *Measurement of the inelastic pp cross-section at a centre-of-mass energy of $\sqrt{s} = 7$ TeV*, JHEP **02** (2015) 029, arXiv:1412.2500.

[4] LHCb collaboration, R. Aaij *et al.*, *Measurement of the inelastic pp cross-section at a centre-of-mass energy of $\sqrt{s} = 13$ TeV*, JHEP **06** (2018) 100, arXiv:1803.10974.

[5] LHCb collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13 TeVpp collisions*, Phys. Rev. Lett. **118** (2017) 052002, Erratum ibid. **119** (2017) 169901, arXiv:1612.05140.

[6] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[7] O. S. and, *LHCb upgrades*, Journal of Physics: Conference Series **1271** (2019) 012010.

[8] LHCb collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021, 2013.

[9] LHCb collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001, 2014.

[10] K. Hennessy, *LHCb VELO Upgrade*, Nucl. Instrum. Meth. A **845** (2017) 97, arXiv:1604.05045.

[11] LHCb collaboration, *LHCb PID Upgrade Technical Design Report*, CERN-LHCC-2013-022, 2013.

[12] Y. Guz, *LHCb Calorimeter Upgrade*, .

[13] A. Cardini, *The LHCb muon upgrade*, Journal of Instrumentation **9** (2014) C02014.

[14] LHCb collaboration, *LHCb Trigger and Online Technical Design Report*, CERN-LHCC-2014-016, 2014.

[15] G. Barrand *et al.*, *Gaudi: A software architecture and framework for building hep data processing applications*, Computer Physics Communications **140** (2001) 45 , CHEP2000.

[16] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[17] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.

[18] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[19] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[20] *The boole project*, `http://lhcb-comp.\web.\cern.\ch/lhcb-comp/digitization/`.

[21] *The brunel project*, `http://lhcb-comp.\web.\cern.\ch/lhcb-comp/reconstruction/`.

[22] *The davinci project*, `http://lhcb-release-area.\web.\cern.\ch/lhcb-release-area/doc/davinci/`.

[23] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, `arXiv:physics/0503191`.

[24] LHCb collaboration, *Physics case for an LHCb Upgrade II — Opportunities in flavour physics, and beyond, in the HL-LHC era*, `arXiv:1808.08865`.

[25] A. collaboration, *Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc*, Physics Letters B **716** (2012) 1 .

[26] C. collaboration, *Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc*, Physics Letters B **716** (2012) 30 .

[27] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[28] M. Kobayashi and T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.

[29] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and search for $B^0 \to \mu^+\mu^-$ decays at the LHCb experiment*, Phys. Rev. Lett. **111** (2013) 101805, `arXiv:1307.5024`.

[30] CMS and LHCb collaborations, V. Khachatryan *et al.*, *Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data*, Nature **522** (2015) 68, `arXiv:1411.4413`.

[31] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP violation parameter A_Γ in $D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$ decays*, Phys. Rev. Lett. **118** (2017) 261803, `arXiv:1702.06490`.

[32] LHCb collaboration, R. Aaij *et al.*, *Measurement of the difference of time-integrated CP asymmetries in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays*, Phys. Rev. Lett. **116** (2016) 191601, `arXiv:1602.03160`.

[33] LHCb collaboration, *Update of the LHCb combination of the CKM angle γ using $B \to DK$ decays*, LHCb-CONF-2018-002, 2018.

[34] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $B_s^0 \to D_s^\mp K^\pm$ decays*, JHEP **03** (2018) 059, `arXiv:1712.07428`.

[35] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001, and 2019 update.

[36] 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** (1964) 138.

[37] M. L. Perl *et al.*, *Evidence for anomalous lepton production in $e^+ - e^-$ annihilation*, Phys. Rev. Lett. **35** (1975) 1489.

[38] S. W. Herb *et al.*, *Observation of a dimuon resonance at 9.5 gev in 400-gev proton-nucleus collisions*, Phys. Rev. Lett. **39** (1977) 252.

[39] H. A. et al. *Observation of b0-b0 mixing*, Physics Letters B **192** (1987) 245 .

[40] CDF Collaboration, F. e. a. Abe, *Observation of top quark production in $\overline{p}p$ collisions with the collider detector at fermilab*, Phys. Rev. Lett. **74** (1995) 2626.

[41] D0 Collaboration, S. e. a. Abachi, *Search for high mass top quark production in $p\overline{p}$ collisions at $\sqrt{s} =$ 1.8 tev*, Phys. Rev. Lett. **74** (1995) 2422.

[42] S. Bifani, S. Descotes-Genon, A. R. Vidal, and M.-H. Schune, *Review of lepton universality tests in b decays*, Journal of Physics G: Nuclear and Particle Physics **46** (2018) 023001.

[43] Physics Reports **427** (2006) 257.

[44] A. Pich, *Precision tau physics*, Progress in Particle and Nuclear Physics **75** (2014) 41.

[45] Physics Reports **532** (2013) 119.

[46] T. N. collaboration, *Precision measurement of the ratio of the charged kaon leptonic decay rates*, 2012.

[47] PIENU Collaboration, A. Aguilar-Arevalo *et al.*, *Improved measurement of the $\pi \to e\nu$ branching ratio*, Phys. Rev. Lett. **115** (2015) 071801.

[48] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, Eur. Phys. J. **C77** (2017) 895, `arXiv:1612.07233`, updated results and plots available at `https://hflav.web.cern.ch`.

[49] LHCb collaboration, R. Aaij *et al.*, *Test of lepton flavor universality by the measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays*, Phys. Rev. **D97** (2018) 072013, `arXiv:1711.02505`.

[50] BABAR Collaboration, J. P. e. a. Lees, *Evidence for an excess of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ decays*, Phys. Rev. Lett. **109** (2012) 101802.

[51] The BABAR Collaboration, J. P. e. a. Lees, *Measurement of an excess of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ decays and implications for charged higgs bosons*, Phys. Rev. D **88** (2013) 072012.

[52] Belle Collaboration, M. e. a. Huschle, *Measurement of the branching ratio of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ relative to $\overline{B} \to D^{(*)}\ell^-\overline{\nu}_\ell$ decays with hadronic tagging at belle*, Phys. Rev. D **92** (2015) 072014.

[53] Belle Collaboration, Y. e. a. Sato, *Measurement of the branching ratio of $\overline{b}^0 \to D^{*+}\tau^-\overline{\nu}_\tau$ relative to $\overline{b}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ decays with a semileptonic tagging method*, Phys. Rev. D **94** (2016) 072007.

[54] Belle Collaboration, S. e. a. Hirose, *Measurement of the τ lepton polarization and $r(D^*)$ in the decay $\overline{B} \to D^* \tau^- \overline{\nu}_\tau$*, Phys. Rev. Lett. **118** (2017) 211801.

[55] T. B. Collaboration *et al.*, *Measurement of $\mathcal{R}(d)$ and $\mathcal{R}(d^*)$ with a semileptonic tagging method*, 2019.

[56] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions $\mathcal{B}(\overline{B}^0 \to D^{*+}\tau^-\overline{\nu}_\tau)/\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)$*, Phys. Rev. Lett. **115** (2015) 111803, Publisher's Note ibid. **115** (2015) 159901, arXiv:1506.08614.

[57] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of the $\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)$ and $\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)$ branching fractions using three-prong τ-lepton decays*, Phys. Rev. Lett. **120** (2018) 171802, arXiv:1708.08856.

[58] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions $\mathcal{B}(B_c^+ \to J/\psi\tau^+\nu_\tau)/\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu)$*, Phys. Rev. Lett. **120** (2018) 121801, arXiv:1711.05623.

[59] G. Hiller and F. Krüger, *More model-independent analysis of $\vec{b}$ s processes*, Phys. Rev. D **69** (2004) 074020.

[60] BABAR Collaboration, J. P. e. a. Lees, *Measurement of branching fractions and rate asymmetries in the rare decays $b \to K^{(*)}\ell^+\ell^-$*, Phys. Rev. D **86** (2012) 032012.

[61] Belle Collaboration, J.-T. e. a. Wei, *Measurement of the differential branching fraction and forward-backward asymmetry for $b \to K^{(*)}l^+l^-$*, Phys. Rev. Lett. **103** (2009) 171801.

[62] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*, JHEP **08** (2017) 055, arXiv:1705.05802.

[63] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays*, Phys. Rev. Lett. **113** (2014) 151601, arXiv:1406.6482.

[64] LHCb collaboration, R. Aaij *et al.*, *Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays*, Phys. Rev. Lett. **122** (2019) 191801, arXiv:1903.09252.

[65] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $\Lambda_b^0 \to pK^-\ell^+\ell^-$ decays*, JHEP **05** (2020) 040, arXiv:1912.08139.

[66] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using $3\,fb^{-1}$ of integrated luminosity*, JHEP **02** (2016) 104, arXiv:1512.04442.

[67] M. Bordone, N. Gubernari, D. van Dyk, and M. Jung, *Heavy-quark expansion for $\bar{B}_s \to D_s^{(*)}$ form factors and unitarity bounds beyond the $SU(3)_F$ limit*, The European Physical Journal C **80** (2020) .

[68] LHCb collaboration, R. Aaij *et al.*, *Measurement of $|V_{cb}|$ with $B_s^0 \to D_s^{(*)-}\mu^+\nu$ decays*, Phys. Rev. **D100** (2020) 072004, arXiv:2001.03225.

[69] A. Hoecker *et al.*, *TMVA 4 — Toolkit for Multivariate Data Analysis with ROOT. Users Guide.*, arXiv:physics/0703039.

[70] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[71] W. Verkerke and D. Kirkby, *The roofit toolkit for data modeling*, 2003.

[72] N. Dawe *et al.*, *scikit-hep/root_numpy : The interface between ROOT and NumPy.* Zenodo, 2018. doi: 10.5281/zenodo.592881.

[73] A. Poluektov, *Correction of simulated particle identification response in LHCb using transformation of variables*, Tech. Rep. LHCb-INT-2017-007. CERN-LHCb-INT-2017-007, CERN, Geneva, 2017.

[74] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.