SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Department of Physics Giuseppe Occhialini

PhD program in Physics and Astronomy Cycle XXXV

Curriculum in Subnuclear Physics

Search for lepton flavour violating $\tau^+ \to \mu^+\mu^-\mu^+$ decay at LHCb and study on MCP-PMT detector for future LHCb Upgrade

Surname: Capelli Name: Simone

Registration Number: 789559

Tutor: Prof. Marta Calvi

Co-Tutor: Dr. Claudio Gotti

Coordinator: Prof. Stefano Ragazzi

Academic Year 2021/2022

Contents

Abstract

The physics analysis has been the primary focus of my research activity during the PhD. Within the CERN LHCb collaboration, I've performed an analysis of data collected during the LHC Run2 (2016, 2017 and 2018).

The aim of this work is the search for the decay of the τ lepton into three muons ($\tau^+ \to \mu^+\mu^-\mu^+$), a decay that would violate the conservation of charged lepton flavour number (cLFV). The lepton flavour is an accidental symmetry of the Standard Model, and without the oscillations of neutrinos such decay would be prohibited. In the Minimal extended Standard Model the branching ratio $\mathcal{B}$ ($\tau^+ \to \mu^+\mu^-\mu^+$) is expected to be $\mathcal{O}$ (10^{-55}), well below current and foreseen experimental sensitivity.

Theories of physics beyond the Standard Model predict an enhancement of the $\tau^+ \to \mu^+\mu^-\mu^+$ decay within present experimental sensitivity $\mathcal{O}(10^{-10})$. This decay has not been observed to date, only upper limits have been established by B-factories (BaBar, Belle) or by hadron collider experiments (LHCb). The upper limit improvement implies strengthen of the constraints on exotic theories, while an observation of the decay would be a clear signal of New Physics.

The analysis is performed separately for each year, and the data is divided into two subsamples depending on the number of muon candidates triggered by the LHCb muon system. Multivariate models are used to distinguish signal and background to enhance the signal sensitivity, and to define correction for data-simulation agreement. The $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ channel is used as a reference channel to estimate the upper limit on the branching fraction. The expected upper limit is computed with the CLs method and results in

$$\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+) \leq 1.8(2.2) \times 10^{-8} \text{ @ } 90\%(95\%) \text{ C.L.}$$

The original analysis presented in this work is the result of a fruitful collaboration of many people working on the LHCb experiment at the universities of Milano-Bicocca and Heidelberg. The content of the following chapters is the results of independent efforts of the Author or collaboration among the members of the analysis team. In particular, the studies related to the trigger efficiencies, the methods defined to correct Data-MC differences and the two multivariate classifiers for signal and background discrimination on the main data sample are results of individual work of the Author.

The $\tau^+ \to \mu^+\mu^-\mu^+$ is an example of a very rare decay, and the analysis involving such decays will benefit from the increment of statistics that will be collected in the current Run3 and in the following Run4 period of data taking at the upgraded LHCb. The High-Luminosity phase of LHCb, starting with Run5 of the LHC, will provide a further boost to the amount of available data. The LHCb detector will need to undergo a second upgrade, to cope with the $\times 10$ increase of luminosity. Numerous studies and R&D projects are currently working on the development of technologies for the future detectors of LHCb. A part of my PhD project was devoted to work on a candidate photodetector for the upgraded Ring Imaging Cherenkov (RICH). I've characterized the timing performance of a multianode microchannel plate photomultiplier (MCP-PMT) in single photon regime. For the second upgrade it has been proposed to improve particle identification performance exploiting the use of precise timing information to cope with the increased pileup. MCP-based devices show excellent time resolution, but their use is critical due to saturation at rate

above $\sim$100kHz/mm^2. The expected rate that the future devices will have to stand is $\sim$10MHz/mm^2. The Auratek-Square MCP-PMT produced by Photek is 53×53 mm device with 64×64 anodes grouped into 8×8 pixels. The dependence of the time resolution from the bias voltage and the photon rate was assessed. When operating as single photon counter at low photon rate and with a single pixel illuminated it shows a transit time spread (jitter) of $\sim$100 ps FWHM, saturating at high rate, above $\sim$100kHz/mm^2. Lowering the bias voltage between the photocathode and the MCP input or between the MCP slabs can reduce the worsening of the time resolution at high rate. The charge sharing between the neighbouring pixels can degrade the time resolution to $\sim$170 ps FWHM when the entire pixel area is illuminated, and could become a major crosstalk source if not accounted for.

Sommario

La parte principale della mia attività di ricerca svolta durante il dottorato consiste nella analisi di dati raccolti dalla collaborazione LHCb del CERN durante il Run2 (dal 2016 al 2018).

Lo scopo di questo lavoro consiste nella ricerca del decadimento del leptone τ in tre muoni ($\tau^+ \to \mu^+\mu^-\mu^+$) Questo decadimento raro non ancora osservato violerebbe la conservazione del sapore leptonico, una quantità conservata accidentalmente nel Modello Standard (SM). Nell'estensione minimale del Modello Standard che include neutrini massivi, il rateo di decadimento atteso $\mathcal{B}$ ($\tau^+ \to \mu^+\mu^-\mu^+$) è previsto essere dell'ordine $\mathcal{O}$ (10^{-55}), ben al di sotto del livello di sensibilità degli attuali e prossimi futuri esperimenti. Vi sono tuttavia teorie di Fisica oltre il Modello Standard (BSM) per le quali è previsto un aumento del rateo di decadimento di $\tau^+ \to \mu^+\mu^-\mu^+$ fino a valori $\mathcal{O}$ (10^{-10}). A oggi questo processo non è ancora stato osservato direttamente, ne da esperimenti a collisori leptonici (BaBar, Belle) ne a collisori adronici (LHCb), sono stati invece posti limiti superiori, che impongono limiti sempre più stringenti sulle teorie BSM. Una eventuale osservazione di questo decadimento sarebbe un chiaro segnale di nuova Fisica.

L'analisi presentata in questa tesi è il risultato di una proficua collaborazione di molte persone che lavorano per l'esperimento LHCb alle università di Milano-Bicocca e Heidelberg. Il lavoro svolto è frutto di sforzi individuali e di sforzi collettivi tra i membri del gruppo di analisi. In particolare, l'autore di questa tesi è il responsabile principale per gli studi sulle efficienze di trigger, dei metodi e delle implementazioni delle correzioni alle differenze tra dati acquisiti e dati simulati e dei due classificatori multivariati utilizzati per l'identificazione di eventi di segnale e di fondo.

L'analisi è stata svolta separatamente per ogni anno, utilizzando il decadimento $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ come canale di riferimento rispetto a cui calcolare il $\mathcal{B}$ ($\tau^+ \to \mu^+\mu^-\mu^+$) in bin dei classificatori. È stato utilizzato il metodo CLs per il calcolo del limite atteso, che risulta essere

$$\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+) \leq 1.8(2.2) \times 10^{-8} \text{ al } 90\%(95\%) \text{ di C.L.}$$

Gli studi di decadimenti molto rari come $\tau^+ \to \mu^+\mu^-\mu^+$ beneficeranno notevolmente dall'aumento di dati raccolti durante l'attuale Run3 e il prossimo Run4 dal rivelatore recentemente rinnovato. Durante la futura fase di alta luminosità che per l'esperimento LHCb comincerà con il Run5, esso sarà in grado di acquisire una maggiore quantità di dati grazie alla luminosità 10 volte superiore. Per fare questo tuttavia è necessario progettare e sviluppare nuove componenti per rimpiazzare quelle

attuali, che non saranno in grado di sostenere le nuove condizioni di funzionamento. Per una parte del mio progetto di dottorato mi sono occupato della caratterizzazione della risposta temporale di un fotomoltiplicatore candidato per l'aggiornamento del Ring Imaging Cherenkov (RICH). Per il futuro upgrade è stato proposto di sfruttare l'informazione temporale delle tracce per ridurre il livello di pile-up. Fotorivelatori basati su MCP sono caratterizzati da una ottima risoluzione temporale, ma il loro impiego ad alti ratei di fotoni è complicato dalla saturazione a cui vanno incontro. Il rate atteso nella fase di alta luminosità è di circa $\sim$10MHz/mm^2. L'Auratek-Square è uno strumento multianodo a microcanali (MCP-PMT) prodotto da Photek di 53×53 mm con 64×64 anodi raggruppati in 8×8 pixels. Ne è stata caratterizzata la risoluzione temporale in regime di singolo fotone, in funzione della tensione di alimentazione e del rateo di fotoni. Il rivelatore mostra eccellenti performance quando un singolo pixel viene illuminato, mostrando uno sparpagliamento del tempo di transito di $\sim$100 ps FWHM quando esposto a un rateo di fotoni fino a $\sim$100kHz/mm^2. Oltre tale soglia il fotomoltiplicatore satura e la risoluzione temporale peggiora velocemente. È possibile mitigare questo peggioramento riducendo sia la differenza di potenziale presente tra il fotocatodo e l'ingresso del MCP che la differenza di potenziale tra i piani del MCP, lavorando a basso guadagno. La capacità di timing è influenzata anche dal fenomeno di condivisione di carica (*charge sharing*) tra pixel adiacenti, che porta la risoluzione temporale a circa 170 ps FWHM quando l'intera superficie del rivelatore è illuminata e che può risultare una delle sorgenti principali di crosstalk se non viene adeguatamente considerata.

Search for $\tau^+ \to \mu^+\mu^-\mu^+$

1.1 Introduction

Physicists have always been fascinated by the symmetries of the Nature. The principles of symmetries are intrinsically connected with the conservation laws, as stated by the Noether's theorem [1].

1.1.1 The Standard Model

The Standard Model (SM) of particle physics [2–4] is built upon symmetries and gauge invariance, and it describes the fermions, the fundamental particles of the matter and their mutual interactions mediated by bosons.

The SM fermions are listed in Table 1.1 and are divided into six leptons and six quarks. Both are organized into three generations (or families) by a group of two particles, and each fermion has a correspondent antifermion. Particles in different generations are characterized by different masses, but have the same quantum numbers.

field	J	J_3	Y	Q
$\begin{pmatrix}\nu_{eL}\\ e_L\end{pmatrix}, \begin{pmatrix}\nu_{\mu L}\\ \mu_L\end{pmatrix}, \begin{pmatrix}\nu_{\tau L}\\ \tau_L\end{pmatrix},$	$\frac{1}{2}$	$\begin{pmatrix}+\frac{1}{2}\\ -\frac{1}{2}\end{pmatrix}$	$-\frac{1}{2}$	$\begin{pmatrix}0\\ 1\end{pmatrix}$
e_R, μ_R, τ_R	0	0	-1	0
$\begin{pmatrix}u_L\\ d_L\end{pmatrix}, \begin{pmatrix}c_L\\ s_L\end{pmatrix}, \begin{pmatrix}t_L\\ b_L\end{pmatrix},$	$\frac{1}{2}$	$\begin{pmatrix}+\frac{1}{2}\\ -\frac{1}{2}\end{pmatrix}$	$-\frac{1}{2}$	$\begin{pmatrix}0\\ 1\end{pmatrix}$
u_R, c_R, t_R	0	0	$\frac{2}{3}$	$\frac{2}{3}$
d_R, s_R, b_R	0	0	$-\frac{1}{3}$	$-\frac{1}{3}$

TABLE 1.1: Fermions of the Standard Model and their quantum numbers

There are charged leptons and uncharged leptons. The charged leptons are the electron (e), the muon (μ) and the tau (τ), while the uncharged are the electron neutrino (ν_e), the muon neutrino (ν_μ) and the tau neutrino (ν_τ). The neutrinos are traditionally massless in the Standard Model and occur only left-handed, whereas the charged leptons can be both right-handed and left-handed.

All the six quarks are electrically charged, but they carry only a fraction of the fundamental electric charge. The up (u), charm (c) and top (t) carry $+\frac{2}{3}e$ while instead the down(d), the strange(s) and the bottom (b) carry $-\frac{1}{3}e$. In addition, each quark can have three different colour charges of the strong interaction, while instead leptons don't carry any colour charge.

The observable bosons $W^{\pm}$ and Z are the weak force mediators, that are a linear combination of gauge bosons $W_i^{\mu}(i = 1, 2, 3)$ and B^{μ}. These gauge fields are associated to the generators of the weak isospin $SU(2)_L$ (L stands for Left) group and the one associated to the weak hypercharge group $U(1)_Y$. The combination is defined by the weak mixing angle θ_W (Weinberg angle):

$$W_{\mu}^{\pm} = \frac{1}{\sqrt{2}}\left(W_{\mu}^{1} \pm W_m u^2\right)$$
$$\begin{pmatrix} A_{\mu} \\ Z_{\mu} \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_{\mu} \\ W_{\mu}^{3} \end{pmatrix}$$

The photon (γ) is the mediator of the electromagnetic force and is associated with the field A^{μ}. The electric charge Q is linked to a projection of the weak isospin J of the left-handed fermions and the hypercharge Y of the fermions by the Gell-Mann-Nishijima formula

$$Q = J_3 + \frac{Y}{2} \tag{1.1}$$

There are eight mediators of the strong force interaction, according to the theory of Quantum Chromo Dynamics (QCD). These massless gauge bosons are called gluons and associated to the fields $A_{\mu}^{a}(a = 1, ..., 8)$ of the $SU(3)_C$ gauge group. They couple to the colour charge of the quarks. All the gauge fields mentioned above grant local gauge invariance of the SM lagrangian under transformation of the group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$.

Only left-handed fermions (doublets under $SU(2)_L$) can undergo weak process mediated by $W^{\pm}$ bosons (charge current process), which are coupled to J. The right-handed neutrinos are neutral to any interaction: they are singlets with respect to $SU(2)_L$ and without electric charge, thus neutral also with respect to the $U(1)_Y$ group according to Eq. 1.1. For these reasons, the right-handed neutrinos are removed from the SM and as consequence no Yukawa coupling or Majorana mass term allow them to acquire mass from the spontaneous symmetry breaking mechanism. The right-handed lepton fields and right-handed quarks instead are charged with respect to the $U(1)_Y$ group and can undergo a process mediated by Z^0 or γ (neutral current process).

The mass eigenstate of quarks do not coincide with the flavour eigenstate of the weak interaction. The rotation of the fermionic fields introduced to diagonalize the matrices of the Yukawa interactions in order to obtain mass terms results in a non-diagonal complex matrix in the charge current lagrangian, that is called the Cabibbo-Kobayashi-Maskawa matrix.

$$\begin{pmatrix} d \\ s \\ b \end{pmatrix}_{weak} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{sb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{CKM\ matrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{mass} \tag{1.2}$$

This matrix is the only source of CP violation in the Standard Model due to the presence of a complex phase. In the lepton sector instead, the absence of right-handed neutrino fields allows having charged current interactions diagonal on the mass basis.

The interaction with the $W^{\pm}$ bosons are the only way in which leptons change type, but this change happens within a family. Even though the Z^0 couples with different strength to left and right-handed fermions, it conserves the flavour too, as well as the photon do. Therefore, the lepton flavour number is conserved in the electroweak interaction. The lepton flavour number is defined as the sum of the number of lepton and neutrinos, minus the number of anti-leptons and anti-fermions. It should be noted that the conservation of the lepton flavour is accidental: it is not the consequence of a fundamental symmetry, but arises from the absence of ν_R fields.

1.1.2 Extending the Standard Model

The Standard Model was tested intensively from many experiments and independent collaborations of experimental physicists, and has proven capable to describe most observed physical phenomena, with high accuracy up to $\mathcal{O}$ (TeV) scale. Despite these successes, there are still unresolved problems in the SM, and it is considered an effective theory because it doesn't include gravity (the only remnant fundamental force not yet included) or clear candidates or explanation for dark matter or dark energy (which constitute $\sim$ 95% of the universe). The only source of CP violation in the Standard Model, within the CKM matrix, is not sufficient to explain the large matter/anti-matter asymmetry observed [5, 6], and some of its parameters need to be fine-tuned very precisely in order to obtain the finiteness of the mass of the Higgs boson.

Neutrino oscillation The discovery of neutrino oscillation [7–9] implies that they should have a non-zero mass, contrary to Standard Model prediction. The oscillation between different flavours could happen only if the mass differences $(m_{\nu_i} - m_{\nu_j})^2$ are different from zero. Furthermore, the presence of massive neutrinos entails that the lepton flavour eigenstate and mass eigenstate cannot be diagonal in both the neutrino sector and the charged lepton sector simultaneously. The rotation of the mass eigenstate into the flavour eigenstate of the neutrinos is performed by the unitary PMNS matrix (Pontecorvo–Maki–Nakagawa–Sakata) [10].

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \underbrace{\begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix}}_{PMNS\ matrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{1.3}$$

The observation of oscillating neutrinos is a clear signal of flavour physics beyond the Standard Model.

1.1.3 Charged lepton flavour conservation in the SM and beyond

The only distinctions between the three families of leptons and quark in the SM are the mass of the particles and the values of the Yukawa couplings. The *universality* of the interactions raises from the fact that the fermionic fields have the same gauge charges. The diagonalization of the Yukawa terms induces interactions with different weights between the $W^{\pm}$ bosons and the three quark families, allowing Flavour Changing Charged Current interaction at tree level. On the contrary, transitions between different families are prohibited in the neutral interactions (mediated by Z/γ) at tree level. In the massless neutrinos' framework, the interactions between the leptons and the $W^{\pm}$ bosons are diagonal in both mass and flavour eigenstate:

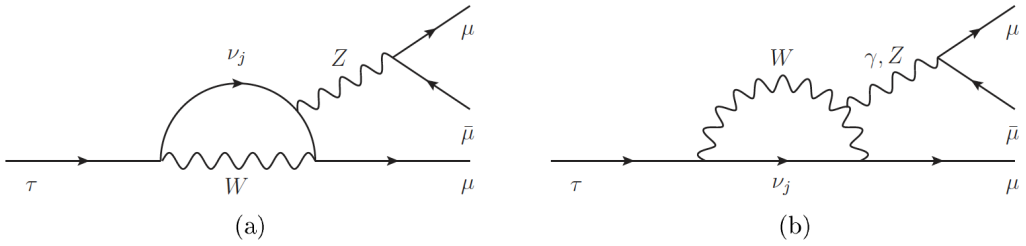


FIGURE 1.1: Feynman diagrams contributing to $\tau^+ \to \mu^+\mu^-\mu^+$ cLFV decay in the Standard Model extended to include the massive neutrinos. The ν_j indicates a neutrino mass eigenstate.

every vertex conserves thus the lepton flavour, and charged lepton flavour violating decays (cLFV) like $\tau^+ \to \mu^+\mu^-\mu^+$ could not happen.

In the minimal extended SM [11] that accounts for neutrino oscillation (known as νSM), there could be vertices that violate the conservation of lepton flavour. The interference between diagrams with different neutrino mass eigenstates could lead to $\nu_\tau \to \nu_\mu$ oscillation, allowing penguin diagrams as in Fig. 1.1 to contribute to the $\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+)$. Such processes are strongly suppressed: the amplitude of each penguin process is proportional to the matrix element $U_{\mu,i}U^*_{\tau,i}$ of the PMNS matrix and summing all over the neutrino families

$$\mathcal{A} = \sum U_{\mu,i}U^*_{\tau,i} + \sum a \cdot U_{\mu,i}U^*_{\tau,i}\frac{m^2_{\nu_i}}{m^2_W}$$

where a is an expansion coefficient for the kinematic dependence. Due to the unitarity of the PMNS matrix, the first term is identically zero, resulting in a $\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+) \sim 10^{-55}$ [12].

There are numerous theories of physics beyond the Standard Model (BSM) that were formulated in order to overcome the shortcomings of the SM. In those new models, the branching fraction of charged lepton flavour violating decays are enhanced by several order of magnitudes. In particular the $\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+)$ is predicted to be of the order of $10^{-10} - 10^{-7}$, definitely more in line with the sensitivity of existing experimental facilities. Charged-lepton flavour changing processes can be allowed at the tree level, by exchanging a new heavy particle, like in the Higgs-mediated decays in supersymmetric seesaw models. Otherwise, the cLFV can occur at loop level with a particle such as right-handed heavy Dirac or Majorana neutrinos, which replace ν_j or γ, Z (Figure 1.1). An overview over some of these BSM theories is reported in Table 1.2, together with the predicted BF for the $\tau^+ \to \mu^+\mu^-\mu^+$ decay.

1.1.4 State of the art

To date, no cLFV has been observed yet. Nonetheless, the improvements of the existing branching fraction limits of charged lepton flavour violating decay help to constrain and exclude the various BSM theories.

Currently, the world's best experimental upper limit to this branching fraction from a single experiment is the one provided by the Belle collaboration, who measured an upper limit of 2.1×10^{-8} @ 90% C.L. using $782\,\text{fb}^{-1}$ of data collected at the KEKB asymmetric-energy e^+e^- collider [19]. The BaBar collaboration also set

Model	$\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+)$
Standard Model + ν oscillation [12]	10^{-55}
SM with right-handed heavy Dirac neutrino [13]	$< 10^{-18}$
SM with right-handed heavy Majorana neutrino [14]	$< 10^{-10}$
left-right SUSY [15]	10^{-10}
SUSY with neutral Higgs [16]	$10^{-10} - 10^{-7}$
SUSY with Higgs triplet [17]	10^{-7}
Non universal $Z^{'}$ (technicolor) [18]	10^{-8}

TABLE 1.2: Predictions for the branching fraction of $\tau^+ \to \mu^+\mu^-\mu^+$ in different BSM models.

a limit to the $\tau^+ \to \mu^+\mu^-\mu^+$ branching fraction of 5.3×10^{-8} at 90% confidence level using 376 fb^{-1} of data collected at the SLAC PEP-II *B-factory*[20]. The search for the $\tau^+ \to \mu^+\mu^-\mu^+$ decay can be performed not only at the B-factories, but also at hadronic collider such as LHC.

The ATLAS collaboration published its sensitivity to this cLFV decay, estimating an upper limit on $\mathcal{B}$ $(\tau^+ \to \mu^+\mu^-\mu^+)$ of 3.76×10^{-7} @ 90% C.L. [21] obtained analysing the data collected in 2012. The CMS collaboration has recently presented the results of their search for $\tau^+ \to \mu^+\mu^-\mu^+$ decay using 33.2 fb^{-1} collected in 2016. The general-purpose experiment set an upper limit on $\mathcal{B}$ $(\tau^+ \to \mu^+\mu^-\mu^+)$ of 8.0×10^{-8} @ 90% C.L. [22].

In this work will be presented the search for $\tau^+ \to \mu^+\mu^-\mu^+$ performed analysing the Run 2 dataset (from 2016 to 2018) collected by the LHCb experiment. The LHCb collaboration already provided a first measurement of $\mathcal{B}$ $(\tau^+ \to \mu^+\mu^-\mu^+)$ in 2012 after analysing Run 1 data [23]. An upper limit of $\mathcal{B}$ $(\tau^+ \to \mu^+\mu^-\mu^+) < 4.6(5.6) \times 10^{-8}$ @ 90%(95%) confidence level was set then.

Comparison with other lepton flavour violating decays

Charged lepton flavour violation can be investigated at LHCb exploiting many decay channels. Among the possible $l \to l'l'l'$ transitions, the $\tau^+ \to \mu^+\mu^-\mu^+$ decay offers a clear signature in the final state provided by the three muons, for which the reconstruction and identification efficiencies are higher than the one for electrons. In addition, the τ reconstructed at LHCb originated mainly from the prompt or secondary decay of b-mesons and c-mesons, produced abundantly in the proton-proton collisions.

Other channels with different final state available are $l \to l'\gamma$, but the amount of background events that can simulate the searched signal is much lower at *B-factories*. The BaBar collaboration [24] has set upper limits for both $\mathcal{B}$ $(\tau \to e\gamma)$ and $\mathcal{B}$ $(\tau \to \mu\gamma)$ as 3.3×10^{-8} and 4.4×10^{-8} @ 90% C.L. respectively. The MEG experiment is dedicated to the search of $\mu \to e\gamma$ at the Paul Scherrer Institut (PSI), and they set the best upper limit so far, being 4.2×10^{-13}@ 90% C.L. [25]. A more detailed list of cLFV decays and the corresponding upper limits can be found in [26].

1.2 Large Hadron Collider

The Large Hadron Collider (LHC) is the largest particle accelerator in the world [27]. It is located at CERN, the European organization for Nuclear Research, in the Geneva area. The accelerator has been built inside the 27 km tunnel that held the Large Electron Positron (LEP) collider. It is operating since 2010, collecting data coming from pp and heavy ions' collisions, with the purpose of precision studies of heavy-quark decays.

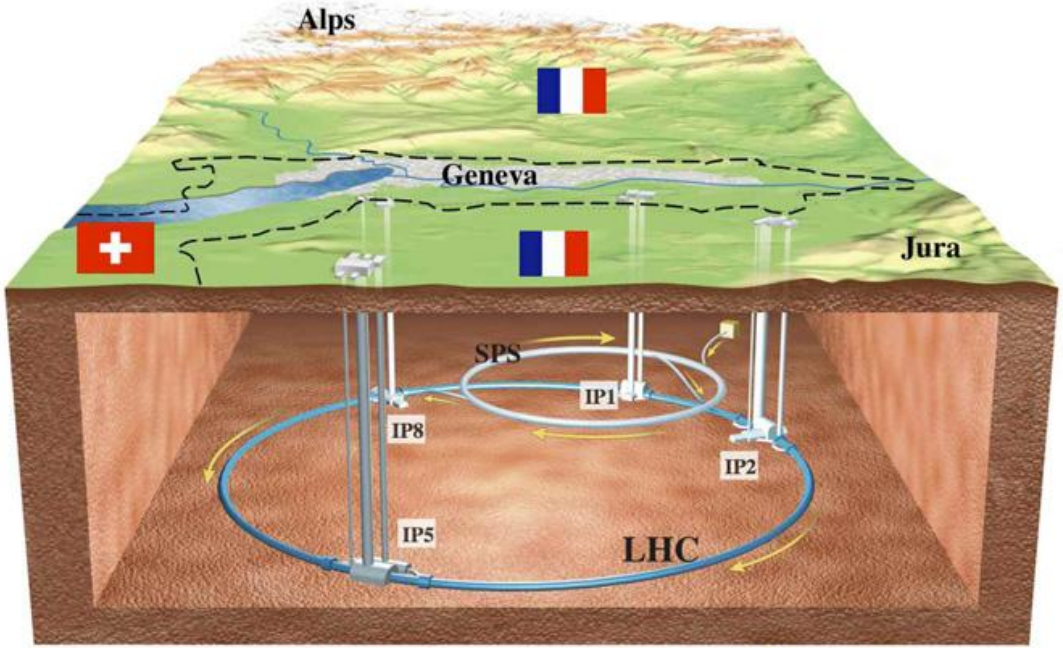


FIGURE 1.2: Schematic representation of the Large Hadron Collider.

Protons are extracted from ionized hydrogen atoms and go through many acceleration stages (Fig. 1.3): from the linear accelerators, to the Proton Synchrotron Booster (PSB), the Proton Synchrotron and Super Proton Synchrotron (SPS), to finally get into in the LHC. This complex structure is designed to stores 2808 proton bunches per ring, each of them made by $\sim 1.1 \times 10^{11}$ protons. The collisions between the bunches happens every 25 ns (40 MHz bunch crossing frequency), and the designed maximum centre of mass energy is $\sqrt{s} = 14$ TeV. The LHC magnet system make use of superconductive NbTi Rutherford cables, cooled down below 2 K by means of liquid helium, in order to operate at fields above 8 T, required to keep the protons on a closed orbit and reduce the transverse beam side.

The LHC design instantaneous luminosity is $\mathcal{L} = 10^{34}$ cm^{-2}s^{-1}, but the number of protons circulating in the rings (beam current) reduces over time, as the protons that collide "exit" the beam, thus at a certain point the beam is deflected out of the beam pipe, towards the LHC beam dump, and a new series of bunches are reinjected. This strategy allows increasing the amount of integrated luminosity over the runs. The luminosity lifetime (the time required to reach $1/e$ of the initial luminosity) is 29 hours.

The CERN facility hosts a large number of experiments, at different stages of the accelerator system. The main one hosted on the LHC ring are briefly described in the following:

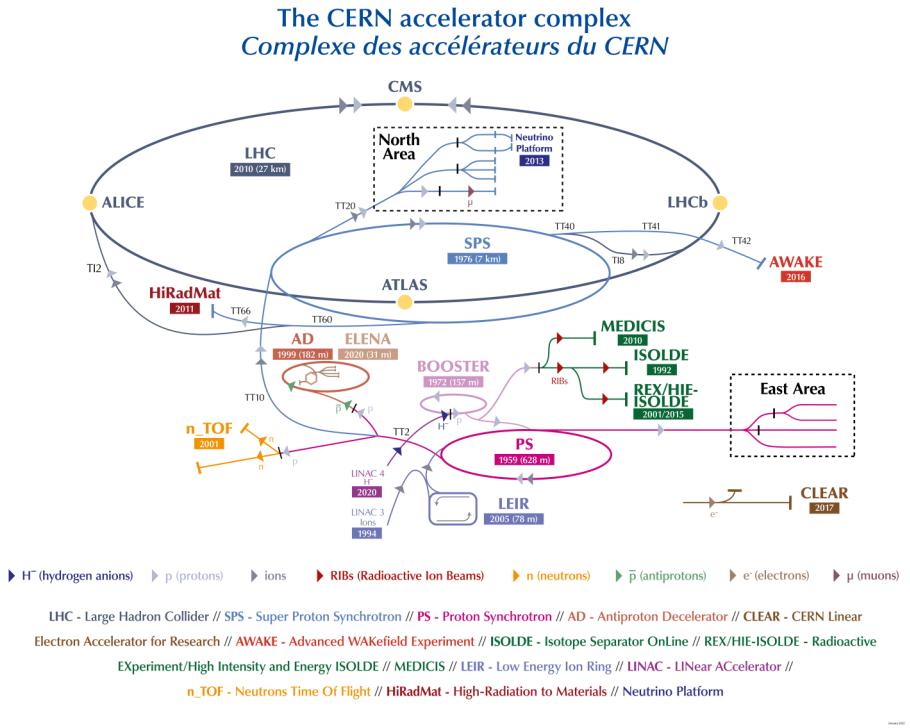


FIGURE 1.3: Summary scheme of the accelerating stages of LHC and the various experiments

- The **ALICE** (A Large Ion Collider Experiment) [28] is a general-purpose detector focused on QCD. It can collide heavy nuclei like Pb or also run proton-nucleus collision in order to study the interaction between the matter and the Quark-Gluon Plasma.

- The **ATLAS** (A Toroidal LHC ApparatuS) [29] and the **CMS** (Compact Muon Solenoid) [30] are the two general purpose detectors, built with the goal to detect the Higgs boson and discover new particles at the TeV scale. In the ATLAS detector there is a thin superconducting solenoid surrounding the inner-detector cavity, and three large superconducting toroids around the electromagnetic and hadronic calorimeters. In the CMS detector instead, a large superconducting solenoid surrounds a silicon pixel and strip tracker, a lead-tungstate scintillating-crystals electromagnetic calorimeter, and a brass-scintillator sampling hadron calorimeter. The four muon stations are mounted in the iron return yoke. Both experiments have a cylindrical structure, composed by one barrel and two end-caps.

- The **LHCb** (LHC beauty) experiment is a single arm forward spectrometer, dedicated to study the physics of heavy-flavour hadrons. The system of sub-detectors will be described in more detail in Sec. 1.2.1.

1.2.1 LHCb detector

The LHCb experiment [31–43] is located at the interaction point 8 (IP8) of the LHC tunnel and collected data from 2010 to 2018 with increasing luminosity up to 4×10^{32} cm^{-2}s^{-1}. It adopted a luminosity-levelling procedure that allows to operate a constant instantaneous luminosity during almost all the duration of the fills: the

proton bunches are displaced vertically at the beginning of the fill, and as the beam current reduces over the time, the displacement is reduced to contrast the diminution of the luminosity. Despite being two orders of magnitude lower than the instantaneous luminosity of the general-purpose detector (Figure 1.4), the LHCb physics yield benefits from the lower and stable number of visible interactions per bunch crossing (pile-up), allowing separate secondary vertices from primary vertices, to keep the same trigger configuration during each fill and reduce the systematic uncertainties due to changes in the detector occupancy [44].

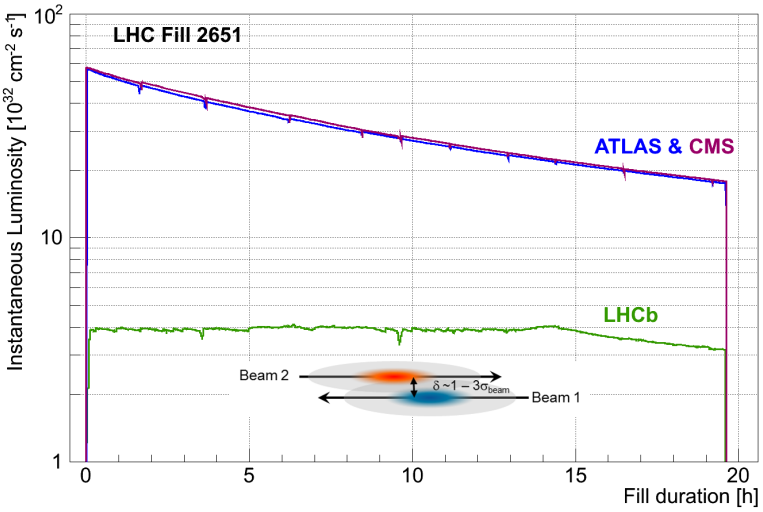


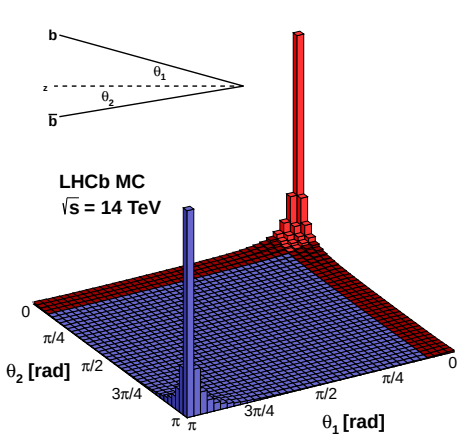
FIGURE 1.4: Comparison between the development of the instantaneous luminosity for LHCb, ATLAS and CMS experiment, during a particularly long LHC fill. The LHCb's luminosity is maintained stable, adjusting the transversal beam overlap.

Production of heavy-flavour hadrons Heavy quarks can be produced as results of strong interactions between partons in pp collisions. The strong interaction conserves the flavour, so when a b-quark is produced, it goes along with a $\bar{b}$. The $b\bar{b}$ pair is usually produced at small angles with respect to the beam axis, and the flight direction is highly correlated (both quark are produced in the forward or in the backward region, as can be seen from Fig. 1.5a). The direction can also be expressed by the pseudorapidity η

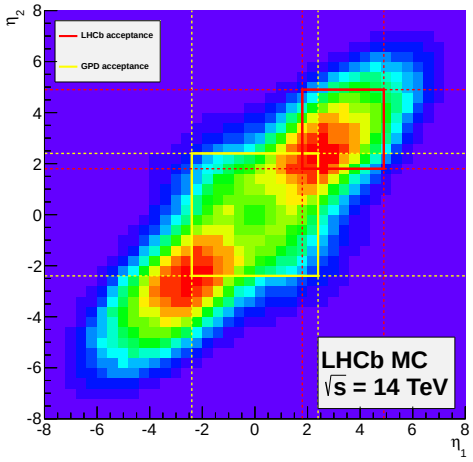
$$\eta = ln\left(tan\frac{\theta}{2}\right)$$

where θ is the angle between the particle flight direction and the beam axis. LHCb is designed to cover the pseudorapidity region $1.8 < \eta < 4.9$, while the others general-purpose detectors cover the region $|\eta| < 2.4$, as shown in Figure 1.5b.

Detector layout The LHCb detector is not symmetrical along the beam axis as the other GPD, but is designed to be a single arm forward spectrometer, with an acceptance very close to the beam axis, from 10 mrad to 300/250 mrad in the horizontal/vertical plane. The detector scheme is shown in Figure 1.6. The coordinate system is defined to have the z axis along the beam pipe, the y axis point upwards and the x axis points toward left (the centre of the LHC ring, if an observer sit on the interaction point and looks into the spectrometer, in the downstream direction).



(A) Distribution of the production angles for b-quarks obtained from simulated data @ $\sqrt{s}$ = 14 TeV [45]



(B) Distribution of b-quarks pseudorapidity obtained from simulation @ $\sqrt{s}$ = 14 TeV. Two boxes indicate the region covered by LHCb and the general-purpose experiments.

In the following sections will be described the system of subdetectors that operated during the Run1 (2011-2012) and Run2 (2015-2018) data taking periods.

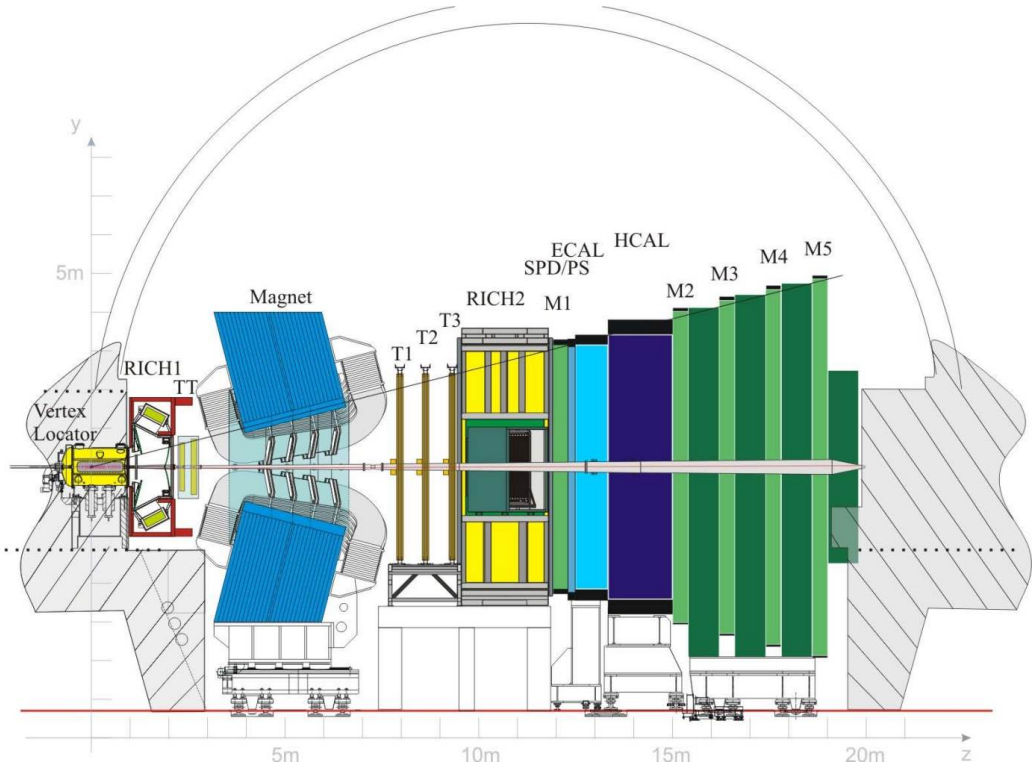


FIGURE 1.6: Scheme of LHCb detector, side view.

Vertexing and Tracking systems

Vertex Locator The detector closer to the beam line is the VErtex LOcator (VELO) [37]. It provides precise measurements of the coordinates of the tracks close to the

primary interaction vertex (PV) of which can provide a measurement with an average resolution of 42 μm on the z direction and 10 μm on the plane perpendicular to the beam. A distinctive feature of the b-hadron decays are displaced vertex, and thanks to precision of the VELO the coordinates are used to reconstruct the production and decay vertices of beauty- and charm- hadrons, providing an impact parameter resolution of 20 μm and a precision on the decay length that ranges from 220 μm to 370 μm.

The vertex locator features a series of silicon stations placed along the beam direction. Each station is composed by a left and right module, in a half-moon shape, and each module is composed by two sensors with a $r - \phi$ geometry (Figure 1.8). The strips of the r sensors are arranged into four segments, tilted by 45°, while the ϕ sensors have two zones, with inner and outer strips. The pitch between the strips varies in the range from 40 to 100 μm across the sensors.

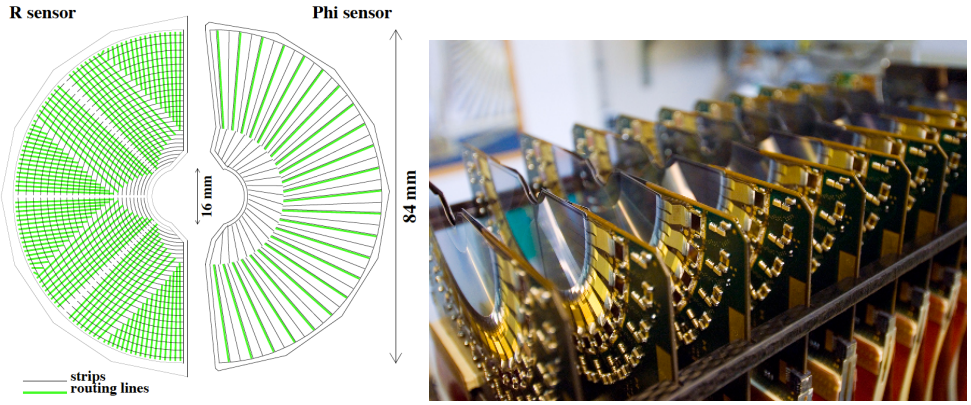


FIGURE 1.7: The left picture shows a scheme of an R and a ϕ sensor. The right picture shows an image of the half-moon shaped modules [46].

The arrangement layout can be seen from Figure 1.7: the blue lines indicate a r-measuring sensor, while the red lines indicate a ϕ-measuring sensor. Two of the r-measuring sensors are placed upstream of the interaction point, with the aim to perform VETO of pile-up events in the first level of the trigger. The VELO measurements are also a vital input to the second level trigger (HLT1), which enriches the b-decay content of the data. More details on the trigger system will be given in Sec. 1.2.1 The minimal radial distance of the module from the beam is 8 mm, but the left and right halves are retracted from the beam axis by 3 cm during LHC injection, in order to reduce the radiation damage. All the stations are mounted in a vacuum vessel, separated by the LHC machine vacuum by an RF foil: 0.3 mm thick corrugated sheets. The foils shield against pickup from the beam and prevent possible gas leaking from the module.

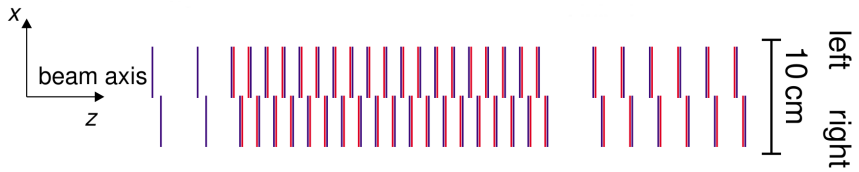


FIGURE 1.8: VELO setup seen from above, showing the overlap between the right and left stations, when the detector is fully closed.

Tracker Turicensis The Tracker Turicensis (TT) is located downstream the VELO, between the RICH1 and the dipole magnet[32, 47]. It is part of the LHCb Silicon Tracker project, together with the Inner Tracker, and its main purposes are the reconstruction of low-momentum particle or long-lived neutral particles that decay outside the VELO, and provide HLT1 with information to assign transverse-momentum information to tracks with large impact parameter. The TT is made by four planar detector layers, each of them equipped with 512 read-out strips of length 10 cm, with a thickness of 500 µm and a pitch of 183 µm. The modules are divided into a top and bottom part, separated by $\sim$ 27 cm along the beam axis, covering an area of approximately $1.5 \times 1.3\,\mathrm{m}^2$. The strips of the first and the last module are oriented along the y axis, while the second and third modules are rotated by $\pm 5^\circ$ respectively, forming an "$x - u - v - x$" configuration that allows to reconstruct the hits in three dimensions. Figure 1.9 shows the layout of the third detector layer, tilted by -5° and the scheme of each module.

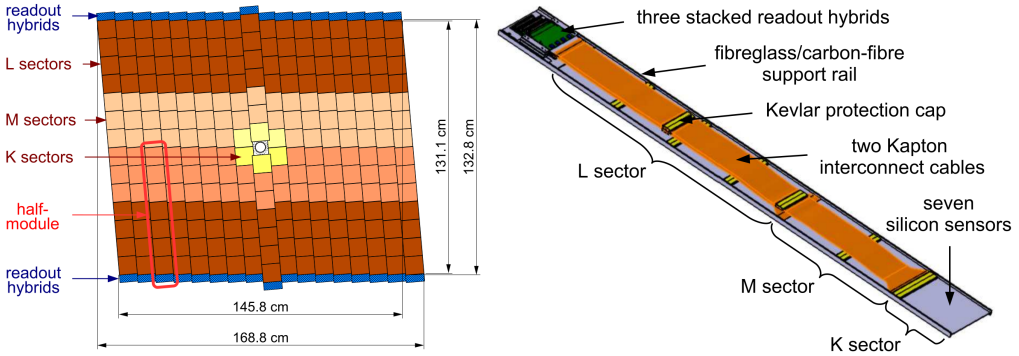


FIGURE 1.9: Layout of the third TT station, with different section of the module indicated by different colours.

Tracking Stations The Tracking Stations, named T1-T2-T3 are placed between the dipole magnet and the RICH2. Unlike the Tracker Turicensis, these tracking stations are built with two different technologies, depending on the distance from the beam pipe: the inner part is instrumented with silicon micro-strips detectors [38] (like the TT), while the outer part is instrumented with straw drift tubes [36].

Each station of the Inner Tracker (IT) consists of four detection layers organized in a "$x - u - v - x$" topology, similarly to the TT, and can be seen in the scheme reported in Figure 1.10.

The Outer Tracker (OT) [48] covers the 98.7% of the active surface of the tracking stations, but it collects only the 80% of the charged particles that are produced close to the interaction point, the remnants enter the smaller IT area. The OT is a straw tube modular detector that cover an area of approximately $5 \times 6\,\mathrm{m}^2$. Each module is designed as a stand-alone device, consisting of 32 or 64 2.4 m long straw tubes, arranged in a two-layers honeycomb geometry, as shown in Figure 1.12. The central anode wires are made of gold-tungsten and have a diameter of 25 µm, while the cathode consists of a 40 µm thick inner foil of electrically conducting carbon doped Kapton-XC (a polyimide film developed by DuPont) and a 25 µm thick outer foil, consisting of Kapton-XC laminated together with a 12.5 µm thick layer of aluminium. The gas mixture chosen to fill the straw tubes, $\mathrm{Ar}(70\%)/\mathrm{CO}_2(28.5\%)/\mathrm{O}_2(1.5\%)$, assures a fast drift time below 50 ns. The position resolution is determined to be $\sim$ 200 µm.

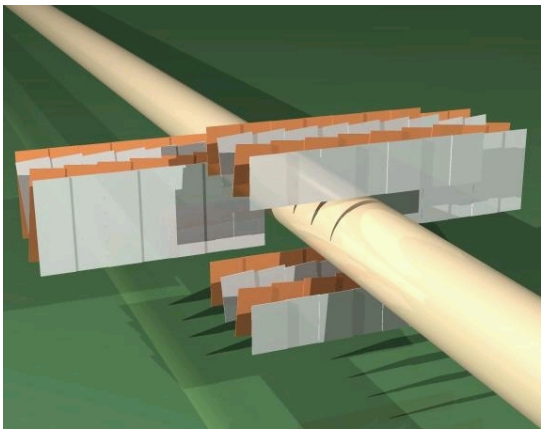


FIGURE 1.10: Isometric view of the layers in one of the T station of the Inner Tracker.

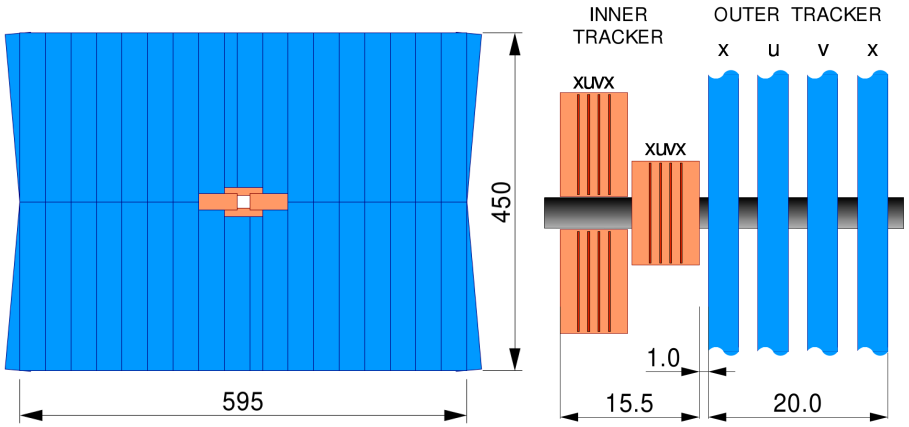


FIGURE 1.11: The left figure shows the front view of a tracking station, while the right figure shows the top view (the lateral dimensions are not to scale). The dimensions are given in cm.

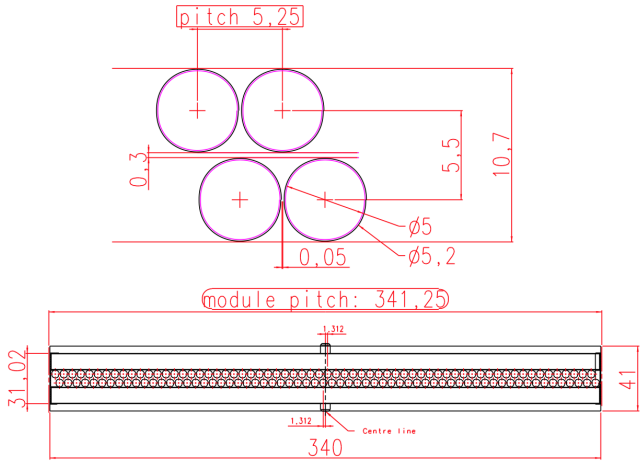


FIGURE 1.12: Sketch of the straw tube layout in the central part of the module. Measures are given in mm.

Dipole Magnet The measurement of the charged particles' momentum can be estimated by the bending of their tracks. A warm dipole magnet [40] is placed among the TT and the T stations, approximately 5 m from the interaction point. It's design,

sketched in Figure 1.13, consists of two trapezoidal coils bent a 45° on the two transverse sides, arranged inside an iron yoke, with a wedge shaped gap in both planes, following the detector acceptance. There are multiple advantages in using a warm magnet instead of a superconducting magnet: Significantly lower costs (the equal expenses for construction and operation will be reached after ten years), rapid construction and lower risks, fast ramping-up of the magnetic field (in sync with the LHC magnets), and most important the possibility to have regular field inversion. On studies of CP violations, the reversal of the magnetic field (done by inverting the electric current in the magnetic coil) reduces the systematic uncertainty due to asymmetries in the detector, since data acquired in the two different configurations (named *MagUp* and *MagDown*) can be mixed together.

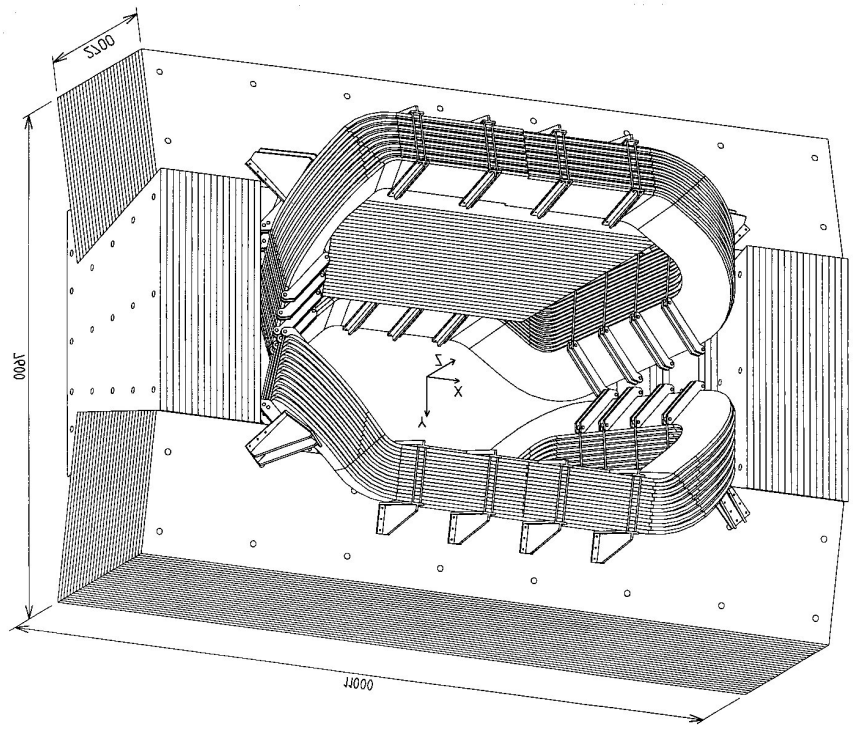


FIGURE 1.13: Prospective view scheme of the LHCb dipole magnet.

The magnetic field is oriented vertically in order to bend the charged particles along the x axis. The direction of the outgoing tracks depends on the input direction and on the particle momentum: The change of direction can be interpreted as a single “kick" done at the magnet centre. The tracks are categorized into different classes, depicted on Figure 1.14, depending on which detector they went through:

- **Long tracks**: The most important tracks for the B-decays, they pass through all the detectors, from the VELO to the T stations.
- **Upstream tracks**: Mainly low momentum tracks, that traverse only the VELO and the TT stations, and eventually are bent outside the LHCb acceptance.
- **Downstream tracks**: Caused by neutral particles decaying into charged particles outside the VELO acceptance, they go through only the TT and T stations.
- **VELO tracks**: Useful for the reconstruction of the primary vertex, are measured only by the VELO because usually they have large angle or the particles are moving backward.

- **T tracks**: Typically produced in secondary interaction, are measured only by the T stations (but are used by the RICH2 for the global pattern recognition).

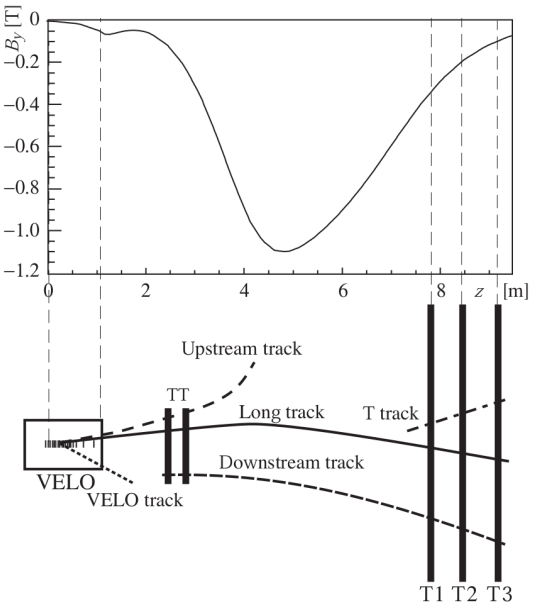


FIGURE 1.14: Illustration of different track types, together with the B_y field component as function of the z coordinate.

The integrated field of 4 Tm allows the tracking system to measure charged particles with a precision of $\sim 0.4\%$ for momenta up to 200 GeV/c. This strong magnetic field also affects the LHCb beam, but the deflection is sufficiently small to keep the bunches within the beam pipe. Nonetheless, this deflection is corrected by additional magnets outside the LHCb cavern.

Particle Identification system

The Particle Identification System (PID) [49] is one of the strengths of the LHCb experiment, a fundamental requirement to study *CP*-violation decays [50], measurements of the unitary angle γ [51] or to investigate rare events like $B_s^0 \to \mu^+\mu^-$ [52]. It is composed by two Ring Imaging Cherenkov (RICH) detectors, used to identify mainly charged tracks of p,π and K; an Electronic Calorimeter (ECAL) and a Hadronic Calorimeter (HCAL) dedicated to the reconstruction of electron, charged hadrons but also of neutral hadrons and photons; five Muon Stations (M1-M5) that are designed to identify muon tracks. In the next paragraphs, each subsystem will be described in details.

Ring Imaging Cherenkov detectors The RICH detectors exploit the light emitted by particles that travel in a medium faster than what light would do. Given a medium with refractive index n, the light propagates in it with a speed $c' = c/n$. If a particle travels with speed $v > c'$ it will emit photons forming a cone around the particle tracks, with angle θ_c defined by

$$v \cdot t \times cos(\theta_c) = \frac{c}{n} \cdot t \implies cos(\theta_c) = \frac{1}{n \cdot \beta} \tag{1.4}$$

In general, from measuring the emission angle and knowing the particle momentum (i.e. from tracking measurements) it is possible to assign a mass hypothesis to the track, thus identify the particle. Figure 1.15 shows how the Cherenkov emission angles depends on particle's momentum differently for each kind of particle.

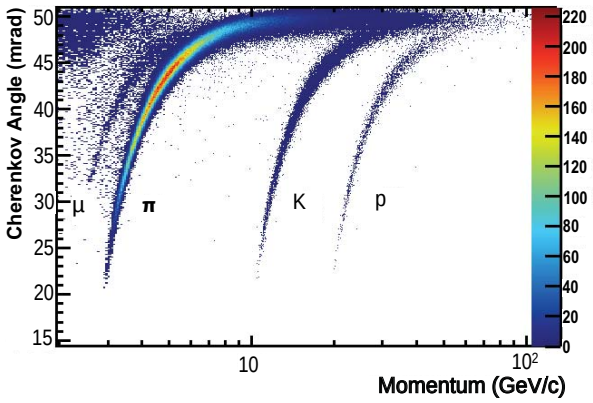


FIGURE 1.15: Dependence of Cherenkov emission angle from the track's momentum, for different types of particles, in the RICH1 gas radiator.

The pattern recognition algorithm used by the RICH system consists in a comparison between the observed hit's pattern on the photodetectors and the theoretical one under a finite set of mass hypothesis, given the knowledge of the detector optics. The direct measure of a ring radius (and the corresponding Cherenkov angle θ_c) is done only on isolated tracks for performance studies. The mass hypotheses of each track in the event are varied to maximize a likelihood function. The estimators for the hadrons' identification are expressed as likelihood ratios

$$\Delta \ln \mathcal{L}_{X\pi} = \ln \mathcal{L}(X) - \ln \mathcal{L}(\pi) = \ln |\mathcal{L}(X)/\mathcal{L}(\pi)| \tag{1.5}$$

where the quantity defined in Eq: 1.5 tends to have positive value for X=K,p and negative value for pions. The RICH system can provide some information useful to discriminate leptons, such as $\Delta\mathcal{L}_{e\pi}$, helping the Calorimeters and Muon systems by multiplying the likelihood obtained by each subsystem. For electrons, muons and hadrons, the combined likelihoods are defined by:

$$\begin{aligned}
\mathcal{L}(e) &= \mathcal{L}^{\mathrm{RICH}}(e)\mathcal{L}^{\mathrm{CALO}}(e)\mathcal{L}^{\mathrm{MUON}}(\text{non } \mu) \\
\mathcal{L}(\mu) &= \mathcal{L}^{\mathrm{RICH}}(\mu)\mathcal{L}^{\mathrm{CALO}}(\text{non } e)\mathcal{L}^{\mathrm{MUON}}(\mu) \\
\mathcal{L}(h) &= \mathcal{L}^{\mathrm{RICH}}(h)\mathcal{L}^{\mathrm{CALO}}(\text{non } e)\mathcal{L}^{\mathrm{MUON}}(\text{non } \mu)
\end{aligned}$$

The first RICH detector (RICH1) is designed to provide particle identification for charged tracks over the range from 1 to 60 GeV/c. It is located between the VELO and the trigger tracker, to identify low momentum tracks that otherwise would have been deflected outside the detector acceptance by the bending magnet. It covers the LHCb acceptance from $\pm$25 mrad to $\pm$300 mrad horizontal ($\pm$250 mrad vertical). The choice of the RICH1's location is part of a reoptimized design with respect to the initial Technical Design Report (TDR) which aimed to reduce the material budget and improve the performance of the HLT1 trigger [32, 41, 53].

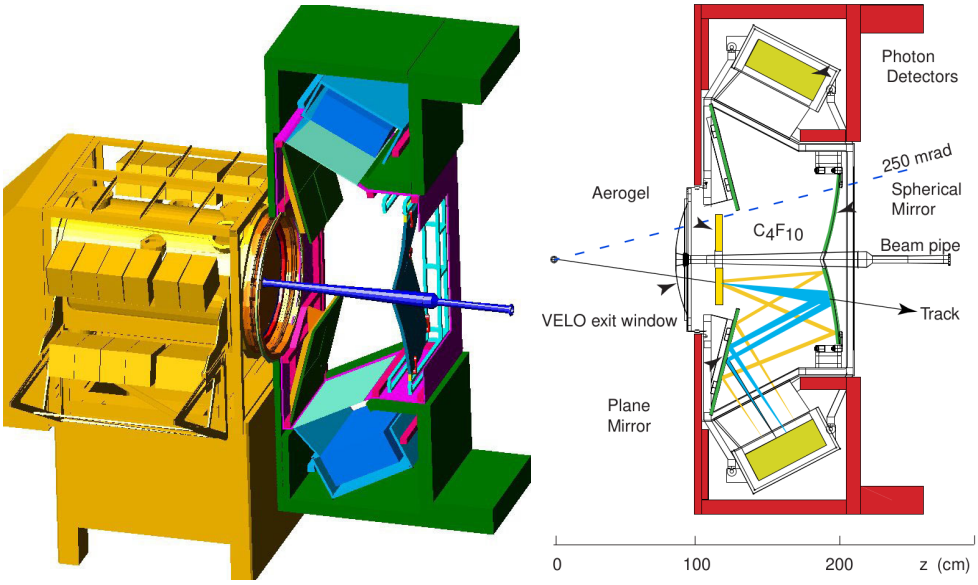


FIGURE 1.16: *left*: 3D model of the RICH1 detector and the VELO tank (shown in yellow). *right*: Layout of the detector and its component: the radiators (aerogel,C_4F_{10}), the spherical and flat mirrors and the photomultipliers planes. The components are surrounded by the magnetic shield boxes (shown in green(red) in the left(right) figure).

The radiators used to produce Cherenkov photons are silica aerogel (n=1.03, removed after Run1) and the perfluoro-n-butane gas (C_4F_{10}, n=1.0014). The first one provides kaon identification above 2 GeV/c and π-K separation up to 10 GeV/c, while the second cover the momentum range from 10 GeV/c to 60 GeV/c. The produced photons are focused with spherical mirrors (with a radius of curvature of 240 cm) on flat mirrors (with dimensions 380 $\times$ 347.5 mm), which deflect them onto the photodetectors planes. The photodetectors chosen for the RICH1 are pixel Hybrid Photon Detectors (HPDs): these vacuum photon detector accelerate photoelectrons generated from the conversion of Cherenkov photons in a photocathode, toward a reverse-biased silicon detector. The output signal is generated from the production of $\sim$5000 electron-hole pairs (one pair for each 3.6 eV deposited by the photoelectron) at the nominal bias voltage of 20kV. The HPD's layout is reported in Fig. 1.17: it has an area of 2.5 $\times$ 2.5 mm^2 at the photocathode, while the silicon detector is made of 1024 squared pixels of 500 $\times$ 500 µm.

The HPDs can work in magnetic fields up to 3mT, above that thresholds the B-field interferes with the focusing geometry of the device. Magnetic shield boxes are necessary to reduce the fringe magnetic field of 60mT generated by the LHCb dipole magnet by a factor of 20, and are shown in Fig. 1.16.

The second RICH station (RICH2) is located at approximately 10 m from the interaction point, between the last tracking station T3 and the first muon station. This detector covers the smaller acceptance of $\pm$120 mrad horizontal ($\pm$100 mrad vertical), necessary to distinguish pions from kaons at energies from 15 to 100 GeV/c. The gaseous radiator used is the carbon tetraflouride (CF_4, n=1.0005), which results in a lower yield of Cherenkov photons for a given radiator length (170 cm). According to simulations, the average number of detected photoelectrons in the rings images for aerogel, C_4F_{10} and CF_4 are 6.6, 32.7 and 18.4 respectively [41]. The HPDs mounted on RICH2 are outside the LHCb acceptance, to minimize the material budget, and the Cherenkov photon are focused and reflected on the photodetector plane

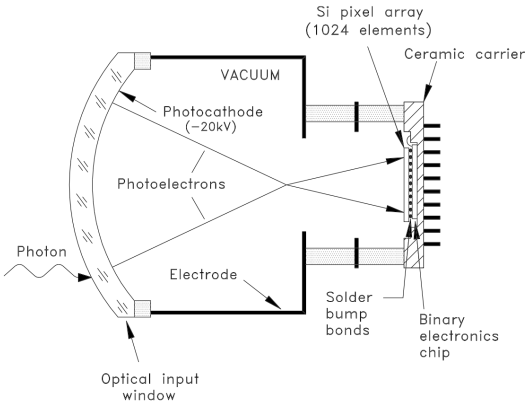


FIGURE 1.17: Layout of the Hybrid Photon Detector, on which the photoelectron trajectories are indicated.

using spherical and flat mirrors similarly to RICH1, as can be seen from Fig. 1.18. In RICH2 however, the photodetectors are positioned on the left and right side, instead of being on top and bottom side.

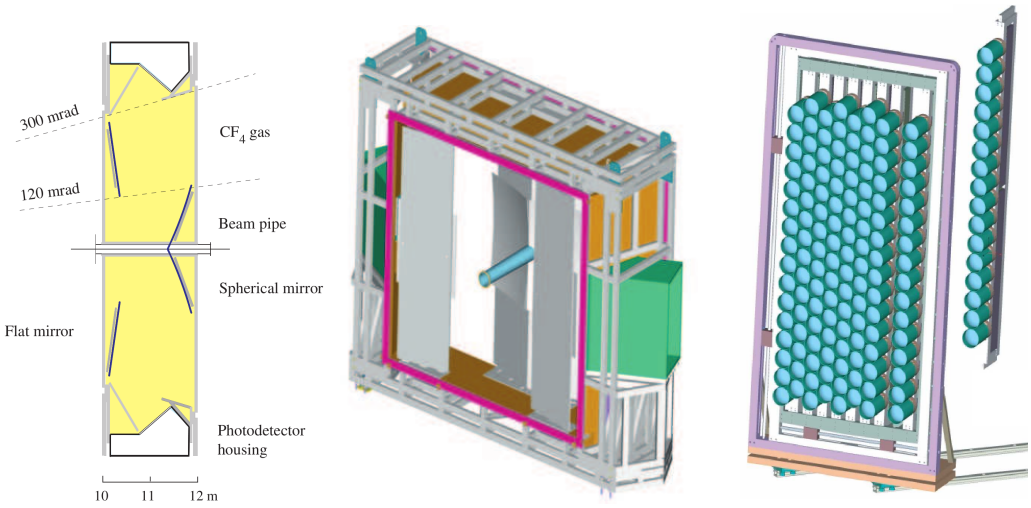


FIGURE 1.18: On the left side the schematic layout of the RICH2 system is shown, while in the central figure there is a tridimensional view. On the right side there is the mounting scheme of a column of HPDs and full detector plane.

Calorimeters The Calorimeter system [39, 54] is a key element for the LHCb, that allows to measure the energies and positions of electrons, photons (performed by the Electromagnetic Calorimeter, ECAL) and hadrons (performed by the Hadronic Calorimeter, HCAL). The collected information is employed by the first level of the Trigger System (described in Sec.1.2.1), and thus it must provide such information as quickly as possible.

The ECAL detector is composed of 2 mm lead sheets interspersed with 4 mm thick scintillator plates, arranged in a sampling structure. The light emitted in the scintillators is collected by wavelength shifting fibres (WLS) and readout by Hamamatsu R7899-20 photomultipliers. In order to fully contain the high energy electromagnetic showers, the thickness of the ECAL is defined to correspond to 25 radiation lengths (X_0). The shashlik structure of the ECAL cell is visible in Fig. 1.19

The ECAL cell's energy resolution measured with test-beam electrons is parametrized as

$$\frac{\sigma(E)}{E} = \frac{(9.0 \pm 0.5)\%}{\sqrt{E}} \oplus (0.8 \pm 0.2)\% \oplus \frac{0.003}{E \sin\theta} \tag{1.6}$$

Where E is the particle energy in GeV and θ the angle between the beam axis and a line from the LHCb interaction point and the centre of the cell. The second constant term is due to systematics and the third one is due to electronic noise.

The HCAL detector is composed of alternating tiles of iron (16 mm thick) and scintillators (4 mm thick) read by WLS fibres and the same photodetectors that are used in the ECAL. This amount of material corresponds to $\sim 5.6 X_0$, sufficient to satisfy the requirements of the first level of Trigger System. Figure 1.19 shows the layout of the HCAL cell, and its energy resolution (measured with test-beam pions) is parametrized as

$$\frac{\sigma(E)}{E} = \frac{(65 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\% \tag{1.7}$$

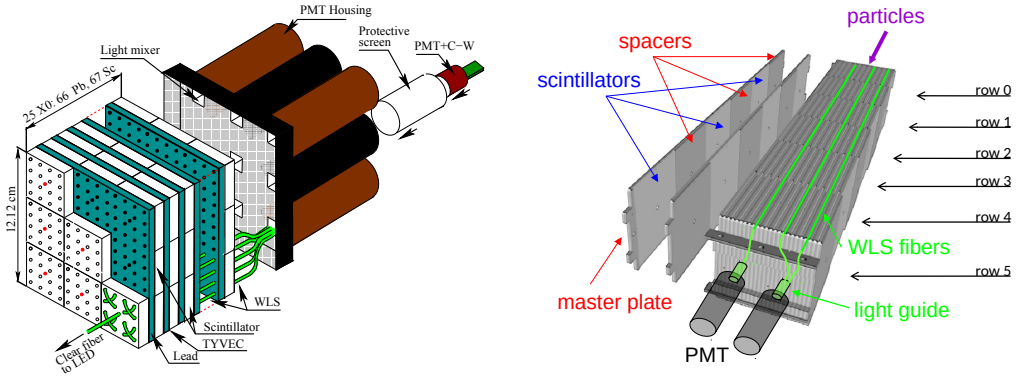


FIGURE 1.19: ECAL cell scheme (left) and HCAL cell scheme (right).

In front of the ECAL there are walls of Scintillator Pads Detector (SPD) and PreShower (PS), with a $2.5 X_0$ of lead foil. When a charged particle go through the SPD the light produced is collected by WLS fibre coil and readout by Hamamatsu 5900 M64 multianode photomultipliers (Ma-PMT). A picture of a SPD cell is shown in Fig. 1.20

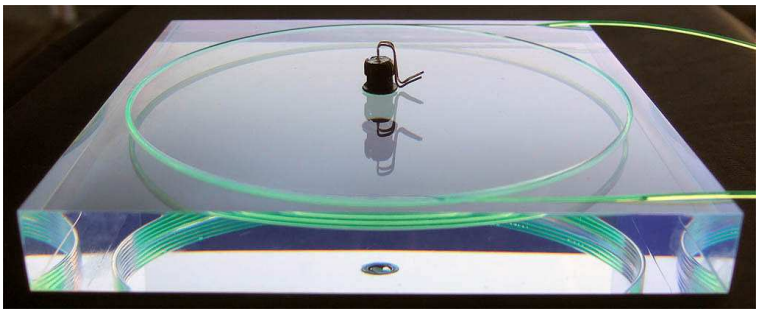


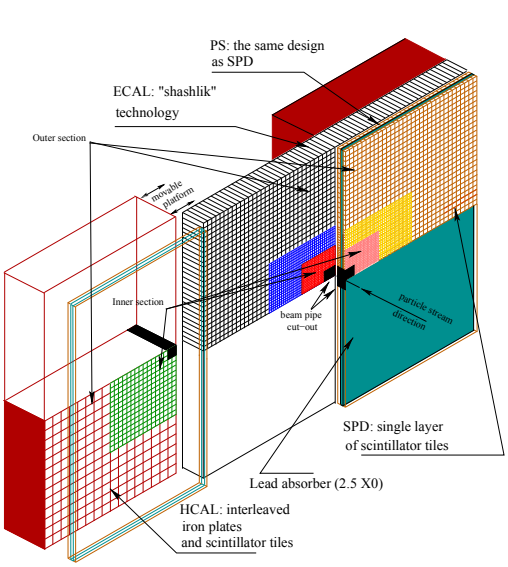
FIGURE 1.20: Picture of a SPD cell, showing the grooved WLS fibre coil.

If the amount of energy deposited in the cell is greater than a configurable threshold, the SPD cell will deliver binary information useful to distinguish the subsequent electromagnetic shower in the corresponding PS cell from the one generated by neutral particles. 25 photoelectrons are expected to be produced on average in response

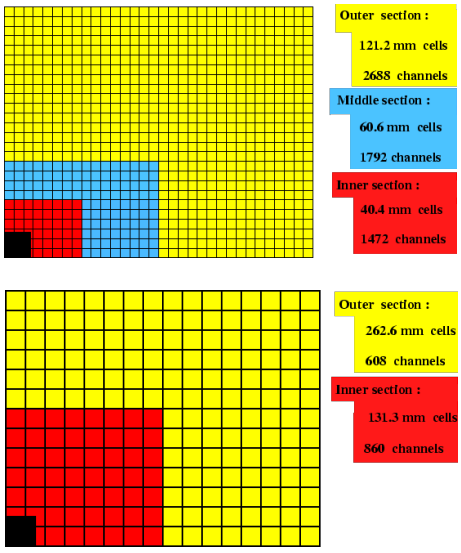
to a minimum ionizing particle (MIP). The PS information is used to separate electrons, photons, and pions.

The whole Calorimeter system covers the LHCb acceptance up to 250 mrad (vertical) and 300 mrad (horizontal), for a total area of $\sim$50 mm^2, located at about 12.5 m from the interaction point. The schematic layout of the four subdetectors is shown in Fig. 1.21a.

All the detectors are segmented into three different zones, with cells size increasing as the distance from the beam pipe increase, to reduce the higher occupancy expected in the central region. In addition, the first $\pm$30 mrad in both direction are not instrumented, to avoid severe radiation damage. Figure 1.21b show the details of the cells' segmentation.



(A) Layout of the calorimeter system, composed by Scintillator Pad Detector, Pre Shower, Electromagnetic Calorimeter and Hadronic Calorimeter



(B) Segmentation of the ECAL's cells on the top figure and of the HCAL on the bottom figure. The cells' size of the PS and SPD are 1.5% smaller than the ECAL cells.

Muon stations The Muon System [35, 49, 53], dedicated to the reconstruction and identification of the muons, is paramount for LHCb detector. Indirect evidence of New Physics can be searched by studying very rare decay such as $B^0 \to \mu^+\mu^-$ or $B_s^0 \to \mu^+\mu^-$, but also other precision measurements that can be done ad LHCb have muons in the final state.

The Muon System is composed by five muon tracking stations, the first one (M1) is placed downstream the RICH2 and before the calorimeters, while the others (M2, M3, M4 and M5) are positioned after the calorimeters, and are interspersed with iron walls, that together with the ECAL and HCAL provide a total absorber-thickness of 20 nuclear interaction-lengths. The total area of the detector is 453 m^2, and the angular acceptance goes from 20 mrad to 306 mrad in the horizontal plane and from 16 mrad to 258 mrad in the vertical plane. The schematic layout of the Muon System is shown in Fig. 1.22, where the two half (side C and side A) are clearly visible. Each quadrant (or logical-pad, named R1 to R4) is treated independently during the extrapolation of the tracks hits, because of the projective structure of the muon stations. The dimension of each reflects the different flux of particles that decreases with the distance from the beam pipe.

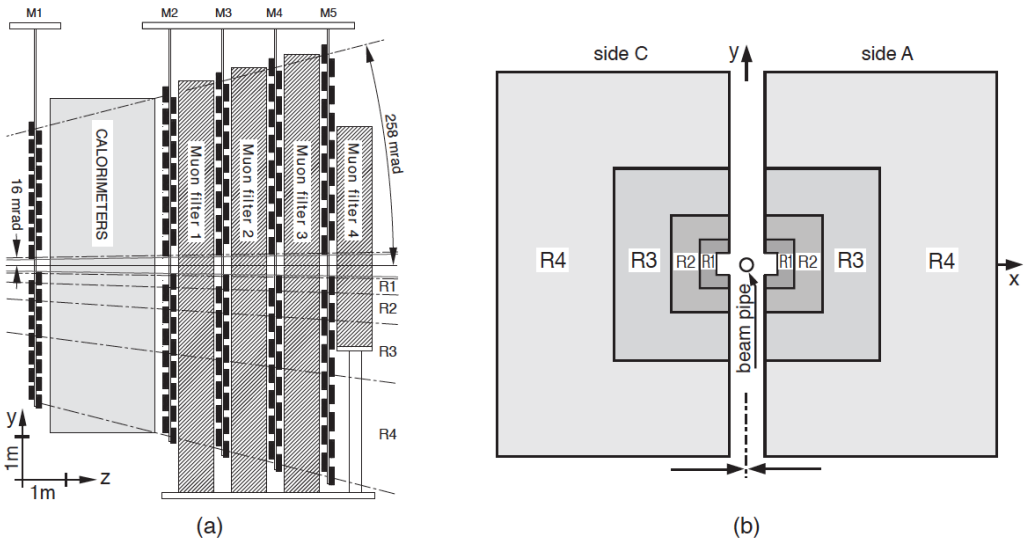


FIGURE 1.22: *left*: Side view of the Muon System, showing the position of the five muon stations and the muon filters (iron wall). *right*: Each station is segmented into four regions (Rs) to cope with the different particle flux observed from the central part (closest to the beam axis) to the detector border.

The M1 station, the most upstream one, is equipped with Gas Electron Multiplier (GEM) detectors, for a total of 12 chambers made of two layers of detectors arrange in a triple layout (Figure 1.23). The electrons produced in the drift gap between the cathode and the first GEM foils by ionization are multiplied by the fields through the three GEM foils, and finally when they drift toward the anode give rise to an induced current signal on the pads. To achieve a time resolution better than 3 ns a gas mixture of $Ar/CO_2/CF_4$ (45:15:40) is used.

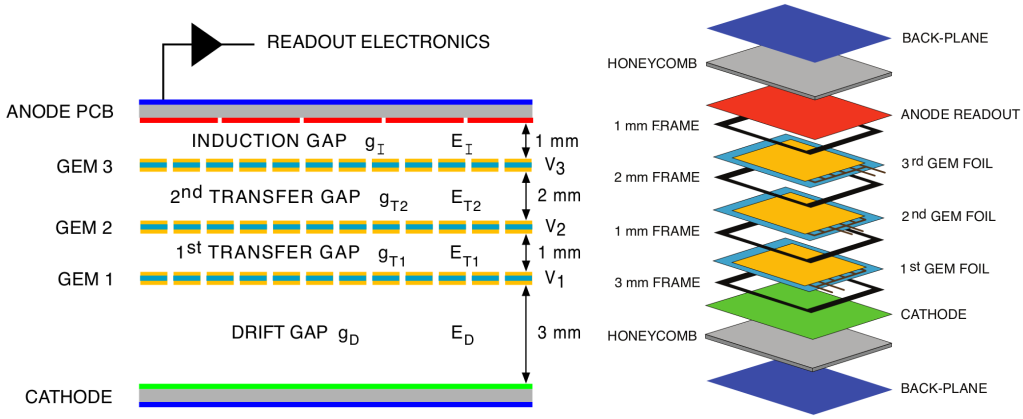


FIGURE 1.23: Schematic (left) and exploded (right) view of a triple-GEM detector.

The stations from M2 to M5 are equipped with 1368 Multi-Wire Proportional Chambers (MWPCs), where the wire planes have a 2 mm spacing and there is a gas gap of 5 mm, as can be seen from Fig. 1.24. The wires are made of Gold-plated Tungsten and have a diameter of 30 µm. The gas mixture used (40%Ar, 55% CO_2, 5% CF_4) allows achieving a time resolution of 5 ns and a gain of $\sim 10^5$ @ 2.65kV.

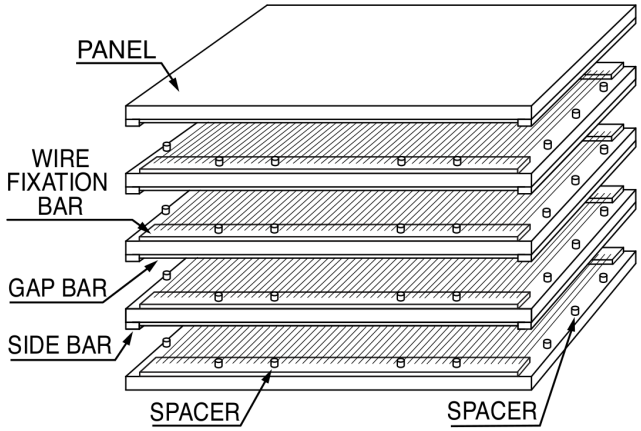


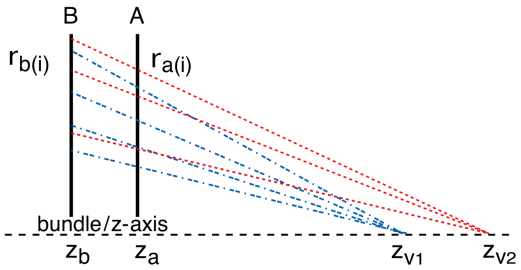
FIGURE 1.24: Exploded layout of a Multi Wire Proportional Chamber.

The information collected from M1 and M2 are used to improve the measurement of p_T in the trigger by means of look-up tables. Tracks are only identified as muon if they reach the M3 station. This condition requires them to have momentum greater than 3 GeV/c. Starting from a logical-pad hit in M3, straight line extrapolations are made to the other stations into search windows called field of interest (FoI). If a hit is present inside the FoI it is assigned to that track. The number of hits in the muons station M2-M5 required to flag a track as muon depends on its momentum: up to M3 for $p < 6$ GeV/c, up to M4 for $p < 10$ GeV/c and up to M5 for $p > 10$ GeV/c.

Trigger system

Given the LHC bunch crossing rate of 40 MHz and the design luminosity of 2×10^{32} cm^{-2}s^{-1}, the rate of $b\bar{b}$-pairs expected is about 100 kHz, but only the 15% of them will have all the decay products of the B meson falling into the detector acceptance. The Trigger System of the LHCb detector [31, 33, 53] was designed to achieve the highest efficiency for the events selected in the offline analysis and reject as much as possible the uninteresting events.

Hardware level: L0 The purpose of the Level-0 (L0) of the trigger system is to preserve the interesting events making decision within 4 μs, reducing the rate from 40 MHz to 1 MHz within which the whole detector can be read out. The Trigger System provides information to a Decision Unit (DU) collected from three different subsystems: pile-up system, Calorimeters Trigger and Muon Trigger as depicted in Fig. 1.26. The pile-up system, composed by two sensors of the VELO (labelled A and B in the scheme of Fig. 1.25), counts the number of primary pp interaction, distinguishing between single and multiple visible interactions. The position of the track origin z_v is estimated by the radii of track hits r_a and r_b according to the relation



$$z_v = \frac{\frac{r_b}{r_a} \cdot z_a - z_b}{\frac{r_b}{r_a} - 1}$$

FIGURE 1.25: Basic scheme illustrating the detection of vertices by the pile-up system.

The particles produced by B meson decays are characterized by large transverse momentum (p_T) and transverse energy (E_T) because of the large mass of the mother particle. The Calorimeter Trigger reconstructs the clusters associated to hadron, electron or photon with the highest E_T. The clusters are formed summing the E_T of 2×2 cells in the calorimeters. The total E_T energy deposited in the HCAL and ECAL is also computed, to perform rejection of empty events, while the charged tracks' multiplicity is estimated by the counting of SPD cells with a hit. The Muon Trigger reconstruct the two muons with the highest p_T for each of the four quadrant of the muon detector.

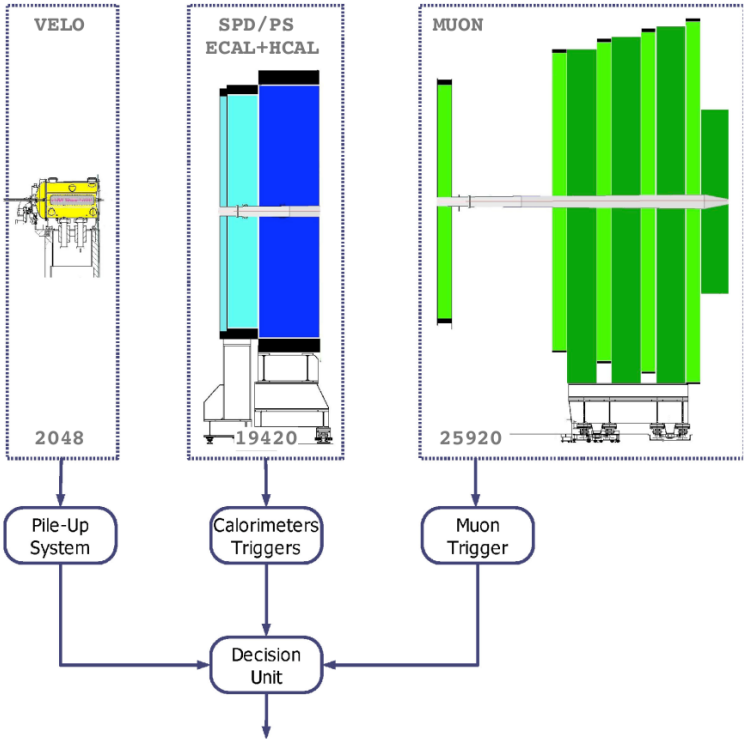


FIGURE 1.26: Overview of the L0 trigger, where the numbers in the dashed boxes represents the number of channels read by each subsystem within 25 ns.

Software level: Hlt The candidates found by the L0 are fed to the following High Level Trigger (HLT), composed by C++ applications that run on the CPU of the Event Filter Farm (EFF).

The HLT reduces the event rate from 1 MHz to 12.5 kHz exploiting the full event data, and the scheme of the trigger flow is reported in Fig. 1.27 The candidates are divided into different alleys, and subjected to different further selection done by two stages, HLT1 and HLT2.

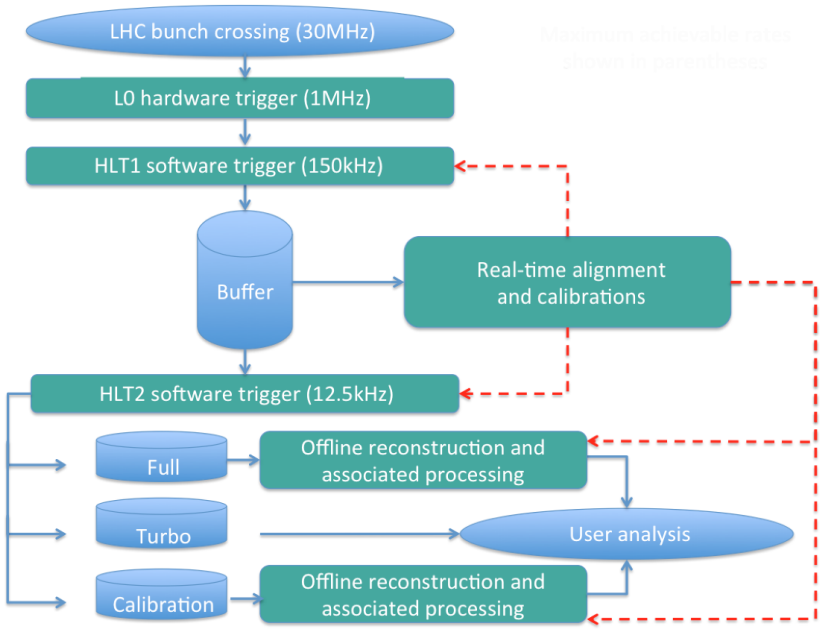


FIGURE 1.27: Scheme of the trigger flow during Run2, illustrating the data process model and the maximum achievable rates. The blue lines represent the data flow, and the red lines the propagation of alignment constants [55].

The aim of HLT1 is to use a portion of the event to confirm the reconstruction of L0 and further reduce the rate in order to allow HLT2 performing full pattern recognition on the remaining events. For neutral candidates like γ and π^0 the HLT1 level check that no other charged particles can match the reconstructed objects, while for the charged particle alleys it exploits the information provided by the VELO and T-stations.

After the HLT1 the events rate reaches hundreds of kHz, and it is sufficiently low to perform a full track reconstruction, similar to what can be done offline. It is possible to use the time between data taking fills, technical stops etc. to process the HLT2 data to have the same offline-data quality already at the trigger level. Buffering the HLT1 data to local storage space and processing it during LHC downtime allows not only to have more time to process each single event, but also to perform calibration and subdetectors alignment in real time using data from HLT1 in dedicated data samples. Together with the Calibration and Full streams, in 2015 LHCb introduced the *Turbo* stream, in which the candidates reconstructed by the trigger lines are directly saved to disk and available to analysts. The different set of trigger conditions applied by the trigger system to the events are also stored, as a unique Trigger Configuration Key (TCK).

TISTOS In order to perform studies on the trigger performance, together with the event and the TCK other summary information are stored, such as all the hits of the tracks and the vertexes that triggered the event. The hits in the subdetectors are identified by a unique number, named LHCbID. It is possible to classify triggered

events into three categories, by comparing the online object that trigger a certain line with the offline object of interest, through the comparison of the set of LHCbID associated to it.

The classification of events is based on the portion of the event that fired a certain trigger [56]:

- Triggered On Signal (TOS): events for which the presence of the signal is sufficient to generate a positive trigger decision.
- Triggered Independent of Signal (TIS): the "rest" of the event is sufficient to generate a positive trigger decision, where the rest of the event is defined through an operational procedure consisting in removing the signal and all detector hits belonging to it.
- Triggered On Both (TOB): these are events that are neither TIS nor TOS; neither the presence of the signal alone nor the rest of the event alone are sufficient to generate a positive trigger decision, but rather both are necessary

It should be noted that an event can be simultaneously TIS and TOS (TISTOS) if both the signal and the rest of the events are sufficient to fire the trigger. This classification permits to evaluate the trigger efficiency directly on the acquired data: given a sample of events selected by a defined trigger configuration, its efficiency can be expressed as

$$\epsilon_{Trig} = \frac{N_{Trig|Sel}}{N_{sel}} = \frac{N_{Trig|Sel}}{N_{TIS|sel}} \times \frac{N_{TIS|Sel}}{N_{sel}} = \frac{N_{Trig|Sel}}{N_{TIS|sel}} \times \epsilon_{TIS} \tag{1.8}$$

If the TIS efficiency of any subsample of the triggered events is the same as that of the whole sample of selected events, it can be measured within the TOS subsample.

$$\epsilon_{TIS} = \epsilon_{TIS|TOS} = \frac{N_{TISTOS}}{N_{TOS}} \tag{1.9}$$

The trigger efficiency can be thus evaluated with quantities that can be measured directly from data:

$$\epsilon_{Trig} = \frac{N_{Trig|Sel}}{N_{TIS|sel}} \times \frac{N_{TISTOS}}{N_{TOS}} \tag{1.10}$$

1.3 Analysis Introduction

This analysis on the search for the lepton flavour violating decay $\tau^+ \to \mu^+\mu^-\mu^+$ is performed with the Run2 data collected at 13 TeV, separately for each year from 2016 to 2018. This will be referred to as *signal channel*, and charge conjugation will be implied henceforth. The number of expected events for $X \to YZ$ can be express as the product of the luminosity ($\mathcal{L}$) times the cross section (σ), times the branching fraction ($\mathcal{B}$), times the efficiency (ϵ):

$$N^{ev}(X \to YZ) = \mathcal{L} \cdot \sigma(pp \to X)\mathcal{B}(X \to YZ)\epsilon(X \to YZ) \tag{1.11}$$

Since the number of τ produced in the sample of analysed data is not known, the branching ratio of the signal channel is normalized to the branching ratio of the *reference channel* $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$, characterized by a similar topology (two muons

in the final state) and kinematics. The branching fraction for the signal channel can therefore be expressed as

$$\begin{aligned}\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+) &= \mathcal{B}(D_s^+ \to \phi(\mu^+\mu^-)\pi^+) \times \frac{f_{D_s}^{\tau}}{\mathcal{B}(D_s^+ \to \tau^+\bar{\nu}_\tau)} \times \frac{\epsilon_{D_s}}{\epsilon_\tau} \times \frac{N_{sig}}{N_{ref}} \\ &= \alpha \times N_{sig}\end{aligned} \tag{1.12}$$

where:

- $f_{D_s}^{\tau} = \frac{\sigma(pp\to D_s\to\tau)}{\sigma(pp\to X\to\tau)}$ represents the fraction of τ produced via D_s decays.
- $\epsilon_{D_s}/\epsilon_\tau$ is the ratio between the overall efficiency evaluated on the reference channel and the efficiency evaluated on the signal channel.
- N_{sig} and N_{ref} are the number of observed events in the signal and reference channel, respectively.
- α is a normalization factor that combines some of the aforementioned terms, and represents the single event sensitivity.

In Sec. 1.4 is described the preselection and the basic cuts and vetoes applied to both signal and reference channel to remove badly reconstructed events and some peaking backgrounds. The reference channel $D_s \to \phi(\mu\mu)\pi$ is used to determine corrections for differences between data and simulation, that are described in Sec. 1.5. The study of the selection of the trigger lines is given in Sec. 1.4.1, while the separation between the signal and background candidates of $\tau^+ \to \mu^+\mu^-\mu^+$ made by multivariate classifiers is presented in Sec. 1.6. One classifier is based on the event topology and the other uses particle identification variables. Partially reconstructed backgrounds and misidentified backgrounds are discussed in Sec. 1.7.

The data samples after the trigger cuts are further divided into two disjoint subsamples according to how many particles in the final state that are identified as muons. In the larger subsample, all the three tracks are identified as muons (3μ sample), while in the other subsample only two out of three tracks are identified as muons (2μ sample). Further details on the selection are given in Sec. 1.4

The upper limit on the branching fraction (in case of the observation of no events) is evaluated on the 3μ sample using the *CLs* method [57] in bins of the aforementioned multivariate classifiers and in bins of the three muons invariant mass ($m_{\mu\mu\mu}$). There are three regions in which $m_{\mu\mu\mu}$ is divided into:

- Signal region: $|m_{\mu\mu\mu} - m_\tau| < 20\,\text{MeV}$ where $m_\tau = 1776.86\,\text{MeV}/c^2$ is the PDG value of the τ mass. In this region, the signal selection efficiency and the expected number of background candidates are evaluated. This region is blinded in data until the end of the analysis.
- Inner sideband region: $20\,\text{MeV}/c^2 < |m_{\mu\mu\mu} - m_\tau| < 30\,\text{MeV}/c^2$. This region is used in the optimization of the trigger selection, in the training phase of the topological classifier and in the optimization of the binning scheme.
- Outer sideband region: $|m_{\mu\mu\mu} - m_\tau| > 30\,\text{MeV}/c^2$. This region is used to fit the background and obtain an expectation of the background yield in the signal region.

Since the 2μ sample is statistically limited, its contribution to the limit evaluation is included as an additional bin. Details of the limit setting procedure are given in Sec. 1.9. The Run2 limit is finally combined with the Run1 limit.

1.3.1 Datasets description

The largest part of the τ leptons produced at LHCb comes from decays of b and c hadrons. [58] The $Z/\gamma^* \to \tau\tau$ and the $W \to \tau\nu_\tau$ production cross-section are three orders of magnitude smaller and thus can be neglected [59, 60]. The charmed mesons D_s and D^+ can either be produced in the primary pp interactions (*prompt* decays), or originate from the decay of B mesons (non-prompt production, or *secondary* decays).

The five sub-channels that contribute the most to $\sigma(p\overline{p}) \to \tau X$ are listed in Table 1.3. To get the fractional contribution to the τ production of each sub-channel its total branching fraction is multiplied by the cross-section in 4π. The LHCb measurements of the prompt charm production cross-section reported in [58] have been performed within the kinematic range $1 < p\,(\mathrm{GeV}/c) < 8$ and $2 < y < 4.5$ and need to be multiplied by a conversion factor to obtain the cross-section in 4π. These conversion factors are determined, both for D_s^+ and D^+ production, with Pythia[61], as the ratio between the generation efficiency and the generator level efficiency of the inclusive production of prompt D_s^+ and D^+, respectively 4.91 ± 0.03 and 4.86 ± 0.02. The two sub-channels that contributes to the D_s production are listed in Table 1.4.

All the branching fractions reported in the fourth and fifth columns can be found in the PDG [62], with the exception of $\mathcal{B}$ ($b \to D^+$) which is still unobserved but can be extrapolated multiplying the B.F. for $b \to D_s$ by $(|V_{cd}|^2/|V_{cs}|^2)$. $b \to D^+_{(s)}$ is an abbreviation for all the channels with $b \to D^+_{(s)}$+X. The only way a D^+ or a D_s^+ can decay into a τ^+ is via the process $D^+_{(s)} \to \tau^+\nu_\tau$, with associated branching fractions $(1.20 \pm 0.27) \times 10^{-3}$ and $(5.48 \pm 0.23) \times 10^{-2}$, respectively. The simulated data samples are produced separately for each sub-channel with an arbitrary luminosity, so when the events are combined they need to be properly weighted to agree with the LHCb measurements of the charm and beauty cross-section at 13 TeV and the known values of the branching fractions. The last column of Tables 1.3 contains the fractional contribution of each sub-channel to the τ production in 4π solid angle ($\mathrm{Calc}_{4\pi}$) at 13 TeV, and the same procedure is applied to the D_s production in the case of reference channel $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ in Table 1.4.

However, only a fraction of generated τ produced in a specific sub-channel passes the cuts at the generator level and decay within the LHCb acceptance. The $\mathrm{Calc}_{4\pi}$ of each sub-channel need to be weighted by ϵ_{CUT} (which depends on the kinematic of a specific sub-channel and can be obtained dividing the generator level cut efficiency, $\epsilon_{\mathrm{GEN|CUT}}$, namely the fraction of generated τ produced in a specific channel of interest, that pass the cuts at the generator level and decay within the LHCb acceptance, by ϵ_{GEN}, the efficiency for a τ to be produced in a specific sub-channel) to determine the corrected contribution (f_{Gauss}) of each sub-channel where the τ decay within the LHCb acceptance:

$$f^i_{\mathrm{Gauss}} = \frac{\epsilon^i_{\mathrm{CUT}} \times \mathrm{Calc}^i_{4\pi}}{\sum_i \epsilon^i_{\mathrm{CUT}} \times \mathrm{Calc}^i_{4\pi}} \tag{1.13}$$

The weight associated to each sub-channel events to have the proper fraction of sub-channels mixing is finally defined as:

$$w^i_{fraction} = \frac{N^i_{\text{exp}}}{N^i_{\text{prod}}} = \frac{N^*_{\text{prod}}}{f^*_{\text{Gauss}}} \frac{f^i_{\text{Gauss}}}{N^i_{\text{prod}}} \tag{1.14}$$

where N^*_{prod} is the number of events generated for the sub-channel with the minimum value of $N_{\text{prod}}/f^i_{\text{Gauss}}$ and f^*_{Gauss} the relative value of f^i_{Gauss}. The same procedure is applied to both signal channel and reference channel.

MC truth-matching In addition to the reconstructed information, in the simulated sample is also possible to know the truth information about the generation of the particles in the decay chain. The `Background Category` tool [63] is used to classify composite particles as either signal or some kind of background. The classification is done on the properties of the final state products, producing a value (`BKGCAT`) that can be used to select the desired categories of event. The signal categories chosen for the signal and reference channels' samples are:

- `BKGCAT=50`: all the final state tracks are associated to a MC truth object with the correct particle ID (no misidentification), but the parent decay has not been fully reconstructed. Moreover, the true common parent must have a mass at most 100 MeV above the mass of what we are trying to reconstruct. This category is included to consider soft photons emitted by the decaying particle.
- `BKGCAT=10`: the decay is correctly and fully reconstructed, but an intermediate resonance or particle is misidentified.
- `BKGCAT=0`: the decay is correctly and fully reconstructed (pure signal).

TABLE 1.3: Fractional contribution of the different sub-channels to the τ production in 4π solid angle at 13 TeV. The indicated channels are an abbreviation for all the channels with τ +X in the final state.

Channel	σ^{meas} **(μb)**	$\sigma^{4\pi}$ **(μb)**	$\mathcal{B}_D$ **(%)**	$\mathcal{B}_\tau$ **(%)**	$\sigma^{4\pi} \times \mathcal{B}_{tot}$ **(μb)**	**Calc$_{4\pi}$ 13 TeV (%)**
$D_s^+ \to \tau$	353 ± 76	1732 ± 373	-	5.48 ± 0.23	94.9 ± 20.8	69.09 ± 3.74
$b \to D_s^- \to \tau$		$2 \times (495 \pm 52)$	14.7 ± 2.1	5.48 ± 0.23	13.5 ± 2.5	9.80 ± 0.53
$b \to D_s^+ \to \tau$			10.1 ± 3.1			
$D^- \to \tau$	834 ± 78	4054 ± 379	-	0.12 ± 0.027	4.9 ± 1.2	3.54 ± 0.19
$b \to D^- \to \tau$		$2 \times (495 \pm 52)$	22.7 ± 1.6	0.12 ± 0.027	0.3 ± 0.1	0.20 ± 0.01
$b \to D^+ \to \tau$			0.5 ± 0.5			
$\bar{b} \to \tau$		$2 \times (495 \pm 52)$	-	2.41 ± 0.23	23.9 ± 3.4	17.37 ± 0.94
Total					137.4 ± 21.3	100

TABLE 1.4: Fractional contribution of the different sub-channels to the D_s production in 4π solid angle at 13 TeV. $b \to D_s^+$ is an abbreviation for all the channels with $b \to D_s^+$+X.

Channel	σ^{meas} **(μb)**	$\sigma^{4\pi}$ **(μb)**	$\mathcal{B}_D$ **(%)**	$\sigma^{4\pi} \times \mathcal{B}_D$ **(μb)**	**Calc$_{4\pi}$ 13 TeV (%)**
prompt D_s^+	353 ± 76	1732 ± 373	-	1732 ± 373	87.58 ± 2.70
$b \to D_s^-$		$2 \times (495 \pm 52)$	14.7 ± 2.1	246 ± 45	12.42 ± 0.38
$b \to D_s^+$			10.1 ± 3.1		
Total				1977 ± 376	100

1.4 Selection

At LHCb the detectors' hits are transformed into objects such as tracks and clusters before being stored. These data are not directly accessible to users, but are filtered through a set of selections called *stripping*. The filtered files are grouped into streams which contains similar selections, in order to save disk space and speed up the access to data. The set of stripping cuts and further fiducial cuts (referred to as offline cuts) applied to the signal channel $\tau^+ \to \mu^+\mu^-\mu^+$ and to the reference channel $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ are reported in Table 1.5

Among the offline cuts, a veto is applied in the signal channel $\tau^+ \to \mu^+\mu^-\mu^+$ against $\phi \to \mu^+\mu^-$ decays in order to reject events from $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ decays where the pion is misreconstructed as a muon, which represents a background source in the signal sample. This is done by requiring the difference between the invariant mass of the two possible oppositely charged muon pairs and the ϕ mass known value to be larger than 20 MeV/c^2. On the contrary, the candidates are required to have the same difference smaller than 20 MeV/c^2 in the reference channel. The invariant mass of the two combinations of opposite-sign muons $\mu^+\mu^-$ is also required to be larger than 450 MeV/c^2 in order to reject events from the background channel $D_s^+ \to \eta(\mu^+\mu^-\gamma)\mu^+\nu_\mu$. Moreover, same-sign dimuon mass is required to be larger than 250 MeV/c^2 in order to reject cloned tracks. Further details on different background sources are given in Sec. 1.7. The PDG [62] values of τ^+ and D_s^+ mass are used as central values for all mass window cuts.

A set of fiducial cuts is also applied to remove poorly reconstructed tracks; the angle θ indicates here the orientation of the momentum vector of the τ^+ (D_s^+) candidate with respect to the line connecting the primary and its decay vertex.

TABLE 1.5: Details of the stripping lines and some offline cuts applied on signal and reference channels.

Stage	Variable	$\tau^+ \to \mu^+\mu^-\mu^+$	$D_s \to \phi(\mu\mu)\pi$
$\mu^\pm$, $\pi^\pm$	p_T	>300 MeV/c	
	Track χ^2/ndf	<3	
	IP χ^2/ndf	>9	
	track ghost probability	<0.45	
μ pairs	$m_{\mu^+\mu^-} - m_\phi$	>20 MeV/c^2	<20 MeV/c^2
	$m_{\mu^+\mu^-}$	>450 MeV/c^2	-
	$m_{\mu^+\mu^+}$	>250 MeV/c^2	-
$\tau^\pm$ and D_s^+	$m_{\mu\mu\mu} - m_\tau$	-	<50 MeV/c^2
	Vertex χ^2	<15	
	IP χ^2	<225	
	$cos(\theta)$	>0.99	
	$c\tau$	>100 μm	
		no PV refitting	
	decay time	>-0.01 ns & <0.025 ns	
		PV refitting	

2μ and 3μ sample

In the previous analysis performed on Run1 data [23] all the muons candidates in the signal channel were required to satisfy the `isMuon` condition: a boolean decision which represents the consistency of a track with the muon hypothesis. If a track has hits in a minimum number of muon stations (including at least M3) within a certain field of interest, it receives the attribute `isMuon`. Table 1.6 summarize the requirements on the muon station hits depending on the momentum range.

TABLE 1.6: Overview of the required stations with hits within the FoI depending on the momentum of the track.

p [GeV/c]	**Required stations**
$p < 3$	*always false*
$3 < p < 6$	M2 & M3
$6 < p < 10$	M2 & M3 & (M4 || M5)
$p > 10$	M2 & M3 & M4 & M5

The stripping line used in the signal channel requires that at least two tracks satisfy the `isMuon` condition, and thus there are events in which one of the three tracks has `!isMuon` condition satisfied: this sample of events (labelled 2μ sample) is now included in the analysis to increase the sample statistics, together with the sample where all the three tracks pass the `isMuon` requirement (referred to as 3μ sample). These two are disjoint samples and will be treated separately in the following steps of the analysis.

In Fig. 1.28 is shown the kinematic distribution of the tracks in 2μ and 3μ sample of 2018. The minimum momentum of a muon to be detected by the muon station is 3 GeV/c. As illustrated, the muons in the 2μ sample are mainly produced at low momentum, with $\sim$ 5% of the events with $p < 3$ GeV/c. Figure 1.29 shows the angular distribution in ϕ and η of the track with opposite charge with respect to the τ, separately for the 3μ and 2μ MC sample of 2018. The different distribution of the events is reasonably related to angular coverage and efficiencies of the Muon Stations.

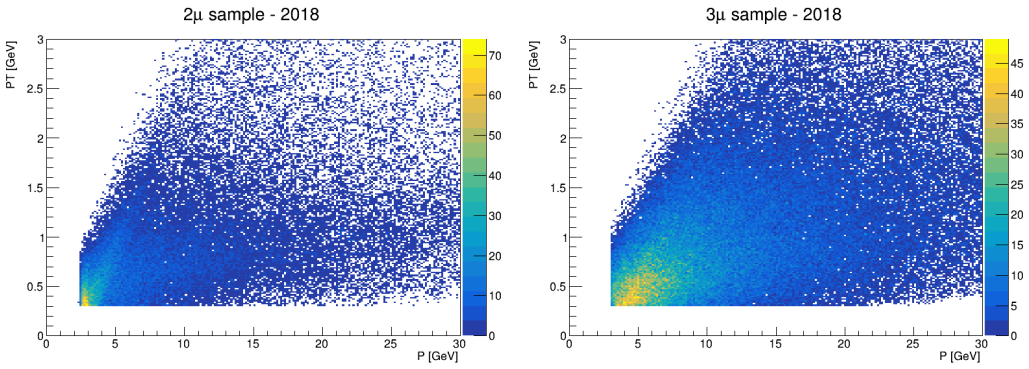


FIGURE 1.28: Distribution of momentum and transverse momentum of a track in the 2μ sample (left) and 3μ sample (right), for 2018 MC.

Fig. 1.30 shows instead the comparison between the momentum distribution of muons who satisfy the `isMuon` condition and of those tracks that have not been detected by the muon stations. As can be seen in the same figure, the transverse momentum distribution of the track with `!isMuon` and the pion from $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$

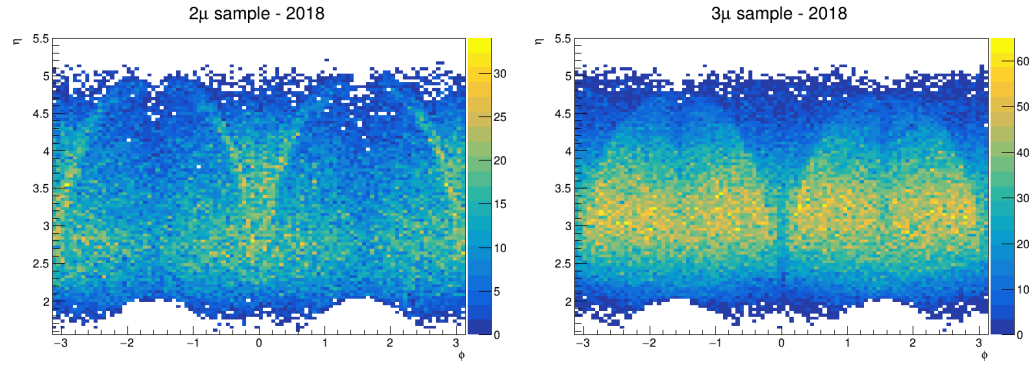


FIGURE 1.29: Distribution of ϕ (angle in the XY plane) and η (pseudorapidity) of a track in the 2μ sample (left) and 3μ sample (right), for 2018 MC.

shows no threshold.

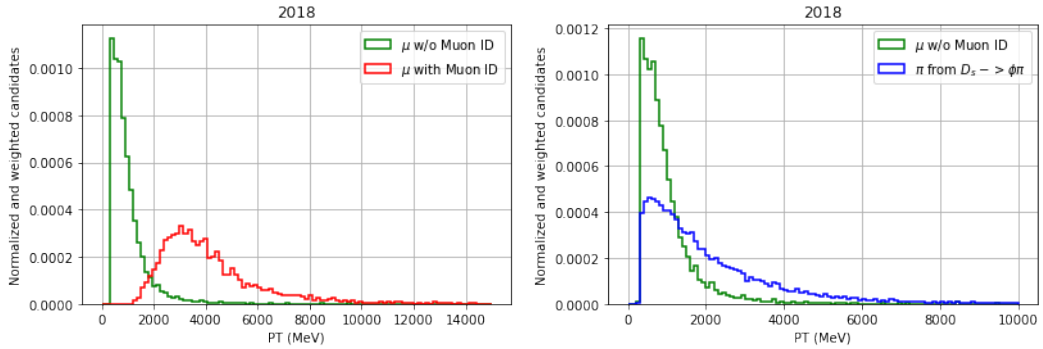


FIGURE 1.30: (left) Comparison between the transverse momentum distribution of the track with `!isMuon` and the track with `isMuon` in the 2μ sample. (right) Same comparison between the track with `!isMuon` and the pion from $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$. Both plots are made from 2018 MC. The momentum distributions for the other years are quite similar except for the muon ID imposing a harder momentum to the muon ($p_T > 2\,\text{GeV}$) for the 2016 sample.

Despite no information from the muon stations is available for these *bad* muons, the RICH detectors' response can be still used to select signal events in case of the 2μ sample. Data acquired by the RICH detectors are processed using a global likelihood approach, and the PID information per track is then stored in the form of differences between the log-likelihood value for a given particle type hypothesis and a pion hypothesis for that track [64]. These differences are named RichDLL[i], where 'i' stands for k (kaon), p (proton), mu (muon), e (electron). A cut on the RichDLLmu variable will be applied in case of the 2μ sample to distinguish between background and signal events (see Sec. 1.6.4).

1.4.1 Trigger selection

The optimization of the set of trigger lines to select signal channel is performed on simulated sample. The trigger efficiency can be evaluated only on the events selected

by the stripping ("|Sel"). The efficiency of the trigger selection (collection of chosen line for each trigger level) is calculated as:

$$\epsilon_{\text{MC,true}} = \frac{\sum w_{\text{fraction}}}{\sum w^{\text{triggered}}_{\text{fraction}}} \tag{1.15}$$

where $w_{fraction}$ is the weight associated to MC events defined in Sec. 1.3.1 to properly mix the events from different tau sources, and $w^{triggered}_{fraction}$ is the subset of weights associated to events that pass the trigger selection. Only up to two lines in `OR` for each trigger level are considered, in order to balance the maximization of the trigger efficiency and the minimization of the complexity of the selection. The addition of a third line provides negligible improvement to the efficiency. Starting from the first level L0 the single line or the pair of lines that triggered the highest number of events are selected. This procedure is repeated for the high-level triggers HLT1 and HLT2 following a bottom-up procedure. The list of the tested combinations and the corresponding trigger efficiencies evaluated using the 3μ sample of the 2018 dataset is given in Table. 1.7.

Selection	TOS lines	$\epsilon^{\text{single}}_{\text{MC,true}}(\%)$	$\epsilon^{\text{two}}_{\text{MC,true}}(\%)$
$L0\|Sel$	L0Hadron	1.79 ± 0.03	56.80 ± 0.12
	L0Muon	56.38 ± 0.12	56.38 ± 0.12
	L0DiMuon	35.20 ± 0.12	59.30 ± 0.12
	L0Electron	0.30 ± 0.01	56.43 ± 0.12
	L0Photon	0.08 ± 0.01	0.08 ± 0.01
$Hlt1\&L0\|Sel$	Hlt1TrackMVA	39.08 ± 0.12	56.87 ± 0.12
	Hlt1TrackMuon	48.49 ± 0.12	57.56 ± 0.12
	Hlt1TrackMuonMVA	39.08 ± 0.12	56.87 ± 0.12
	Hlt1SingleMuonHighPT	2.90 ± 0.04	55.59 ± 0.12
	Hlt1DiMuonLowMass	55.56 ± 0.12	55.56 ± 0.12
	Hlt1TwoTrackMVA	47.55 ± 0.12	47.55 ± 0.12
$Hlt2\&Hlt1\&L0\|Sel$	Hlt2SingleMuonHighPT	0.43 ± 0.02	57.55 ± 0.12
	Hlt2SingleMuonRare	0.60 ± 0.02	57.55 ± 0.12
	Hlt2SingleMuonVHighPT	0.06 ± 0.01	57.55 ± 0.12
	Hlt2Topo2Body	1.21 ± 0.03	57.55 ± 0.12
	Hlt2TopoMu2Body	12.77 ± 0.08	57.55 ± 0.12
	Hlt2TopoMuMu2Body	23.35 ± 0.11	57.55 ± 0.12
	Hlt2Topo3Body	3.62 ± 0.05	57.55 ± 0.12
	Hlt2TopoMu3Body	7.31 ± 0.06	57.55 ± 0.12
	Hlt2TopoMuMu3Body	7.46 ± 0.07	57.55 ± 0.12
	Hlt2DiMuonSoft	0.85 ± 0.02	57.55 ± 0.12
	Hlt2DiMuonDetached	45.03 ± 0.12	57.55 ± 0.12
	Hlt2TriMuonDetached	4.90 ± 0.05	57.55 ± 0.12
	Hlt2TriMuonTau23Mu	57.55 ± 0.12	57.55 ± 0.12

TABLE 1.7: Weighted trigger efficiencies on truth-matched MC signal sample of 3μ, given the preselection and offline cuts for the 2018 sample. The first column contains the efficiency obtained with each single line. The second column contains the efficiency obtained with the OR between the most efficient single line and another one. The HLT efficiencies take into account the previous trigger levels.

All the lines reported in the table are Triggered on Signal (see the definition of TOS in Sec. 1.2.1). There are no significant improvements in the trigger efficiency coming from the inclusion of TIS lines: already from the addition of L0 TIS lines from Table 1.8 it can be seen that the small improvement on signal efficiency entail a great increment of the background retention rate (evaluated on the signal data in a high mass region of the outer sidebands, as the fraction of events that satisfy the selected trigger lines).

TIS lines	$\epsilon_{MC,true}(\%)$	$\epsilon_{Data}(\%)$
w/o TIS line	57.55 $\pm$0.12	36.01 $\pm$0.03
L0Global	66.62 $\pm$0.11	84.47 $\pm$0.02
L0Muon	60.21 $\pm$0.12	48.34 $\pm$0.03
L0DiMuon	58.21 $\pm$0.12	53.31 $\pm$0.03
L0Hadron	62.33 $\pm$0.12	47.30 $\pm$0.03
L0Electron	59.78 $\pm$0.12	41.05 $\pm$0.03
L0Photon	58.31 $\pm$0.12	37.05 $\pm$0.03

TABLE 1.8: The first row contains the efficiencies obtained with the best combination of lines from Table 1.7. The other lines show the total trigger efficiency in MC and the retention rate in data (2018) obtained with the addition of the corresponding TIS line at the L0 requirements (L0Muon TOS and L0DiMuon TOS).

The same procedure described above is followed to determine the trigger selection for the 2μ sample of the signal channel and for the reference channel. The chosen L0 and HLT1 lines are the same for 3μ, 2μ samples and for the reference channel sample (to minimize the differences in the next steps of the analysis), while for the HLT2 stage the choice depends on the number of reconstructed muons in the final state. A dedicated line was introduced in Run2 to select 3μ candidates, while for the 2μ sample and the reference channel the `OR` of two lines is used. The selected lines are shown in Table 1.9 while the corresponding trigger efficiencies on the MC samples are reported in Table 1.10. It should be noted that in 2016 the Hlt2DiMuonDetached line was not available and thus is replaced with different topological line (Hlt2TopoMuMu3Body).

TABLE 1.9: Trigger configuration for signal and reference channel. For each stage of trigger, the specified lines are selected in OR. Events are selected when selected at each separate stage. (nb *Not available in 2016, it is replaced by Hlt2TopoMuMu3Body_TOS)

Stage	3μ Sample	2μ Sample	Reference Sample
L0	L0Muon_TOS L0DiMuon_TOS		
HLT1	Hlt1TrackMuon_TOS Hlt1DiMuonLowMass_TOS		
HLT2	Hlt2TriMuonTau23Mu_TOS	Hlt2TopoMuMu2Body_TOS Hlt2DiMuonDetached_TOS*	

TABLE 1.10: The trigger efficiencies calculated upon the truth-matched events that pass the offline selection $\epsilon_{\text{trig|sel}}$ for each year of data taking and for each different stage of the trigger. All values are in percent.

Year	Trigger Level	Signal 3μ	Signal 2μ	$D_s^+ \to \phi(\mu^+\mu^-)\pi^+$
2016	L0	60.028 ± 0.121	43.141 ± 0.191	46.621 ± 0.126
	L0&HLT1	58.256 ± 0.122	37.805 ± 0.187	41.494 ± 0.125
	L0&HLT1&HLT2	58.236 ± 0.122	8.174 ± 0.105	14.785 ± 0.090
2017	L0	69.047 ± 0.114	51.821 ± 0.192	55.240 ± 0.123
	L0&Hlt1	66.696 ± 0.117	44.961 ± 0.192	48.616 ± 0.123
	L0&Hlt1&Hlt2	66.681 ± 0.117	22.396 ± 0.160	31.765 ± 0.115
2018	L0	59.298 ± 0.122	42.845 ± 0.191	46.572 ± 0.131
	L0&Hlt1	57.564 ± 0.123	37.580 ± 0.187	41.482 ± 0.129
	L0&Hlt1&Hlt2	57.546 ± 0.123	19.054 ± 0.152	27.587 ± 0.117

1.5 Calibration Channel

As mentioned in Sec. 1.3 the decay channel $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ is used as reference to determine the B.R. of the $\tau^+ \to \mu^+\mu^-\mu^+$ decay. The efficiencies that enter in Eq. 1.12 are evaluated on simulated samples, but need to be corrected for Data/MC differences. This correction is evaluated using the reference channel.

The `sPlot` technique [65] allows extracting the signal-only component in a data sample by unfolding the contributions of signal and background in a given variable (the invariant mass of D_s^+ in this case). It is a statistical technique that from the sole knowledge of the *probability density functions* (PDFs) permits to compute a per-event weights (named *sWeights*) that keeps all signal events while get rid of all background events when plotting the distributions of other variables uncorrelated with the fitting variable.

1.5.1 sWeight procedure

The signal component of $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ candidates is described by a Johnson's SU function, defined as:

$$f(x;\mu,\lambda,\gamma,\delta) = \frac{\delta}{\lambda\sqrt{2\pi}} \frac{1}{\sqrt{1+\left(\frac{x-\mu}{\lambda}\right)^2}} exp\left[-\frac{1}{2}\left(\gamma + \delta sinh^{-1}\left(\frac{x-\mu}{\lambda}\right)\right)^2\right] \quad (1.16)$$

The shape parameters γ and δ are determined from a fit on MC and are then fixed in the fit on data, while the mean and width (μ and λ) are allowed to float. The combinatorial background is modelled with an exponential decreasing function. The fit is performed on events that passed the trigger selection (both in data and MC), and the number of estimated signal and background events are extracted. The results of the fits for the three years of are shown in Fig. 1.31

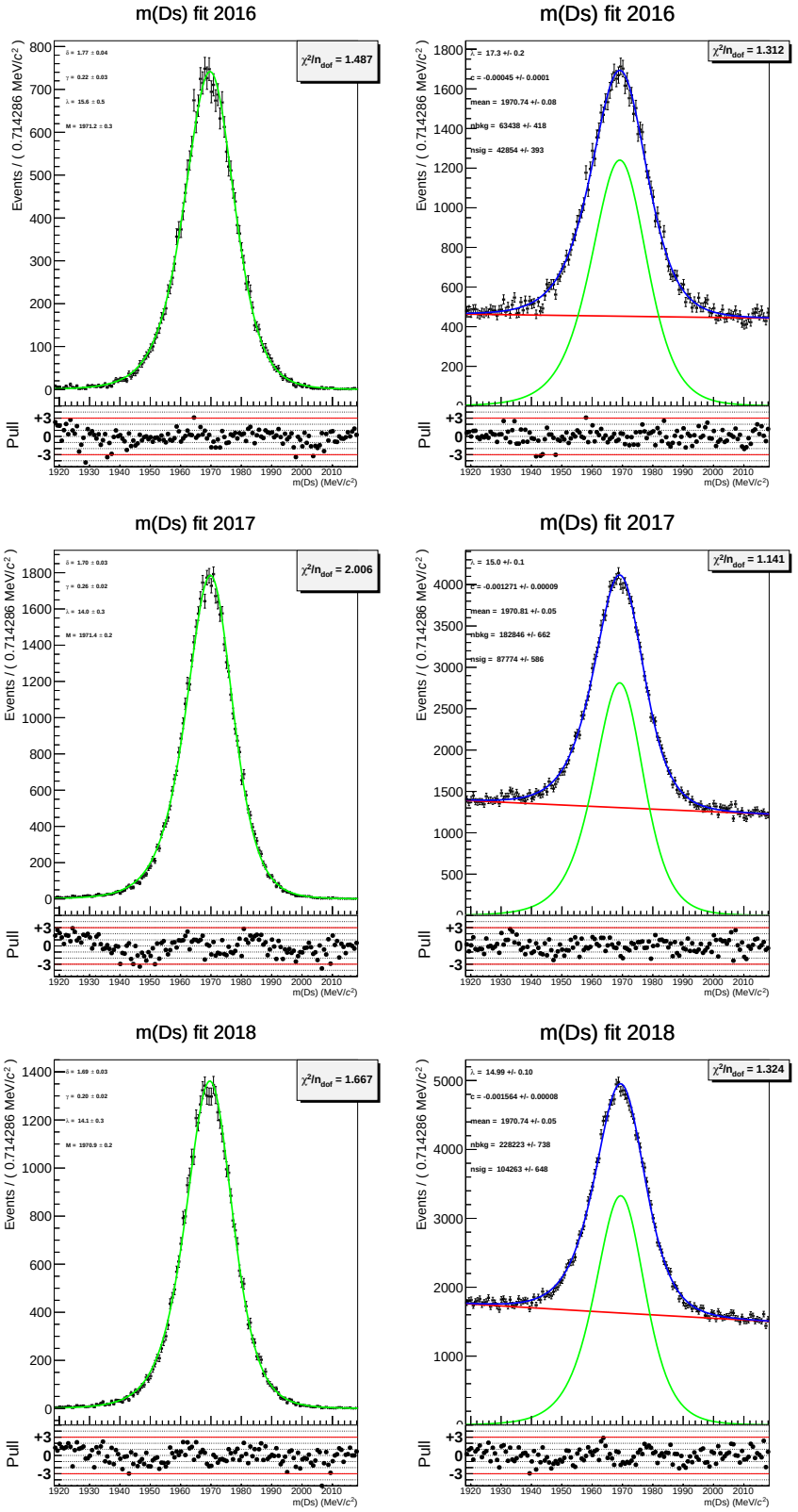


FIGURE 1.31: (Left) Fit to the invariant mass distribution of the reference channel MC using a Johnson's SU function. (Right) Fit to the invariant mass distribution of the reference channel data to extract the sWeights.

1.5.2 Data/MC correction

Some variables in simulated data show differences with the corresponding one obtained from real data. The *reweighting* is a technique that makes distribution of one or several variables identical in an original sample and a target sample, by assign different weights to the events in the original sample. The *hep_ml* package [66] provides a reweighting tool based on decision trees to reduce the Data/MC disagreement. The algorithms are implemented as estimators, fitting and reweighting stages are split: fitted reweighters (one for each year of data taking) are then applied on both $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ and $\tau^+ \to \mu^+\mu^-\mu^+$ MC samples. A gradient-boost based classifier named GBReweighter is used to assign a corrective weight to MC events. The GBReweighter model is composed by many decision trees that are trained individually to greedily maximize a symmetrized χ^2, to find regions (leafs) where the amount of MC events ($w_{leaf,\,MC}$) is greater than the amount of Real Data ($w_{leaf,\,RD}$):

$$\chi^2 = \sum_{leaf} \frac{(w_{leaf,\,MC} - w_{leaf,\,RD})^2}{w_{leaf,\,MC} + w_{leaf,\,RD}} \tag{1.17}$$

During each iteration of the training process, the following tree model tries to adjust the discrepancies not covered by the previous model, in a gradient boosting way. Since the GBReweighter is both trained and applied on the $D_s \to \phi(\mu\mu)\pi$ MC sample, a `k-fold` technique is used to have unbiased predictions: multiple models are trained on different portions of the sample and predicts only those part of data which is not used during its training. When k-folded models are applied on the signal channel $\tau^+ \to \mu^+\mu^-\mu^+$ the average of all models output is taken, since all of them will produce unbiased results. The variables used as input features are listed in Table1.11.

Variable	Definition
ln(D_s_P)	Logarithm of the D_s momentum
ln(D_s_PT)	Logarithm of the D_s transverse momentum
nTracks	Number of tracks in the event
D_s_ENDVERTEX_CHI2	χ^2 fit of the decay vertex
ln(D_s_IPCHI2_OWNPV)	Logarithm of the χ^2 of the impact parameter of the primary vertex of the D_s track
arccos(D_s_DIRA_OWNPV)	Angle between the D_s momentum and the vector from the primary to the decaying vertex

TABLE 1.11: Input variables of the GBReweighter model, trained to reduce the Data/MC difference in the reference channel $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ and in the signal channel $\tau^+ \to \mu^+\mu^-\mu^+$.

Hyperparameters tuning The aim of reweighters is to get identical multidimensional distributions, not only that each variable has identical 1D distribution in the two samples. It is so because multivariate models will be used in the Signal/Background classification task (which are described in Sec. 1.6), and such models are capable to exploit the correlations between the features distributions.

For this reason, the tuning of some hyperparameters of the GBReweighter model was done with the following process: Different sets of hyperparameters are tested

sequentially, and for each of them it was evaluated the effect on the disagreement between the MC and data distribution in each features of Table 1.11. For those sets where the overall effect were to reduce the discrepancies, another multivariate classifier is constructed and trained to distinguish between the reweighted MC and the real data. The optimal set of hyperparameters is chosen among the ones that gives the best overlapping of the one dimensional distributions and the smallest possible accuracy of the additional classifier (which means that the classifier was incapable to distinguish between simulated and real data). The Kolomogorov-Smirnov (KS) distance is the chosen method to quantify the agreement between one-dimensional distributions. The test multivariate classifier that compares the reweighted MC and the real data has 300 estimators with a maximum depth of 6 each, and a learning rate (η) of 0.3.

The detailed process is reported below:

1. The datasets of MC and real data events are combined and split in a train:test sample with proportion 80%:20%, and

2. All possible combinations in a fixed hyperparameters space are inserted in a list and tested sequentially, by training a GBReweighter with such parameters and predict a new set of weights $w_{Data/MC}$.

3. The KS distances computed for each feature are averaged to quantify the overall effect and this value compared with the one obtained without the MC reweighting: if the ratio between the new value and the previous one exceeds 1 (thus if the effect of the weights is to **not** reduce the discrepancies of the one-dimensional distributions) this set of hyperparameters if discarded.

4. If the average KS distance ratio is instead lower than 1, the other multivariate classifier is trained to distinguish the sWeighted data from the MC weighted with $w_{Data/MC}$, and the balanced accuracy (the efficiency to identify each class averaged between the classes) obtained with such model is evaluated on the test dataset

5. The final set of hyperparameters is the one that provides the balanced accuracy closer as much as possible to 50% (equivalent to a random classification of the events) and a KS distance as close as possible close to zero (equivalent to a perfect distributions' overlap)

The set of hyperparameters chosen for each year are listed in Table 1.12. The learning rate and the loss regularization parameters were fixed to 0.2 and 50 respectively during the grid search. However, minimal variations in the GBReweighter output have been observed by exchanging the sets of the hyperparameters between the years.

Figure 1.32 shows how the reweighted MC distributions (red line) are much closer to the sWeighted data (gray distributions), while in Figure 1.33 are shown the ROC curves obtained with a further multivariate classifier trained to distinguish between simulated and real data. The training of the model is performed using only the *fraction* weights ($w_{fraction}$), and then the ROC curves are computed with and without the corrective weights ($w_{Data/MC}$). It can be seen the "confusion" of the classifier, which performances become similar to the one obtained by a random guess of the event class, with Area Under the Curve (AUC) equal to 50%.

TABLE 1.12: Optimized GBReweighter hyperparameters, for each year of data taking.

Variable	2016	2017	2018
n_estimators	40	40	30
learning_rate	0.2	0.2	0.2
max_depth	2	4	3
min_samples_leaf	400	400	200
loss_regularization	50	50	50

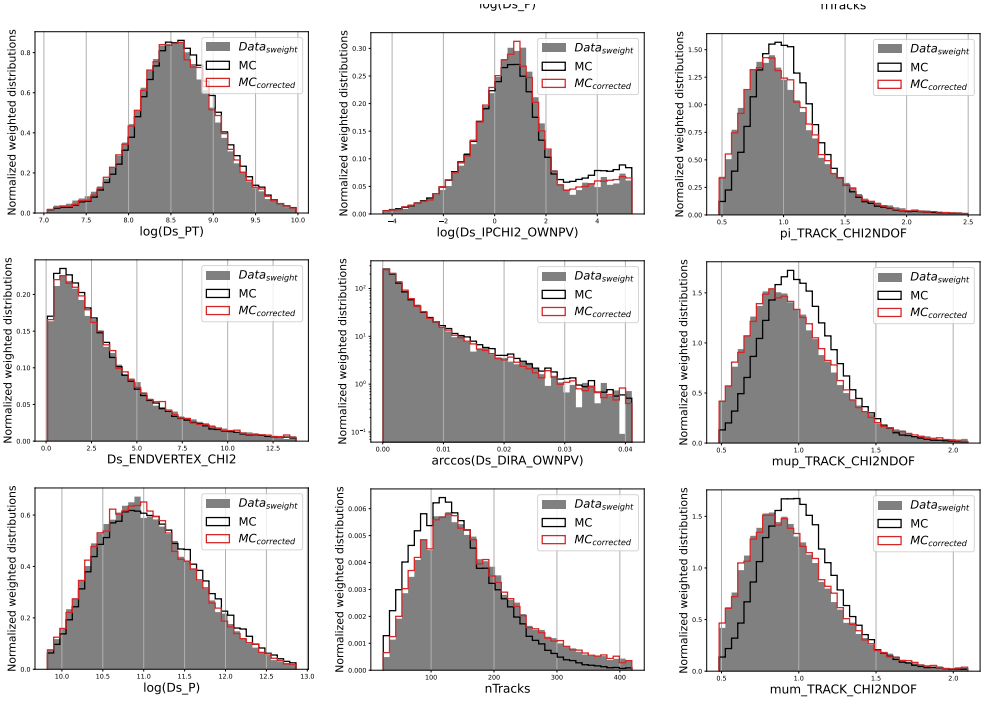


FIGURE 1.32: Example of distributions of sWeighted data and simulated samples in the $D_s \to \phi(\mu\mu)\pi$ channel, with and without the Data/MC correcting weights $w_{Data/MC}$.

1.5.3 TISTOS Method

The trigger efficiencies are computed using the signal and reference channel MC sample. The TISTOS method [56] will be used (Section 1.8.2) to evaluate differences between the trigger efficiency evaluated on simulation and data. The (conditional) trigger efficiency for a selection of TOS lines on MC can be computed directly from the definition because the number of selected events (N_{sel}) is known:

$$\epsilon_{MC,\,TOS|SEL} = \frac{N_{TOS|SEL}}{N_{SEL}} \tag{1.18}$$

As the detector only records events passing the trigger, the number of events in the detector acceptance (N_{Acc}) is not directly observable, and thus the fraction of trigger accepted events that contain a signal candidate within the acceptance ($\epsilon_{Trig|Acc}$) is not known. Is it possible though to exploit the trigger categories TIS, TOS (described in Sec. 1.2.1) to determine the trigger TOS efficiency within a TIS subsample,

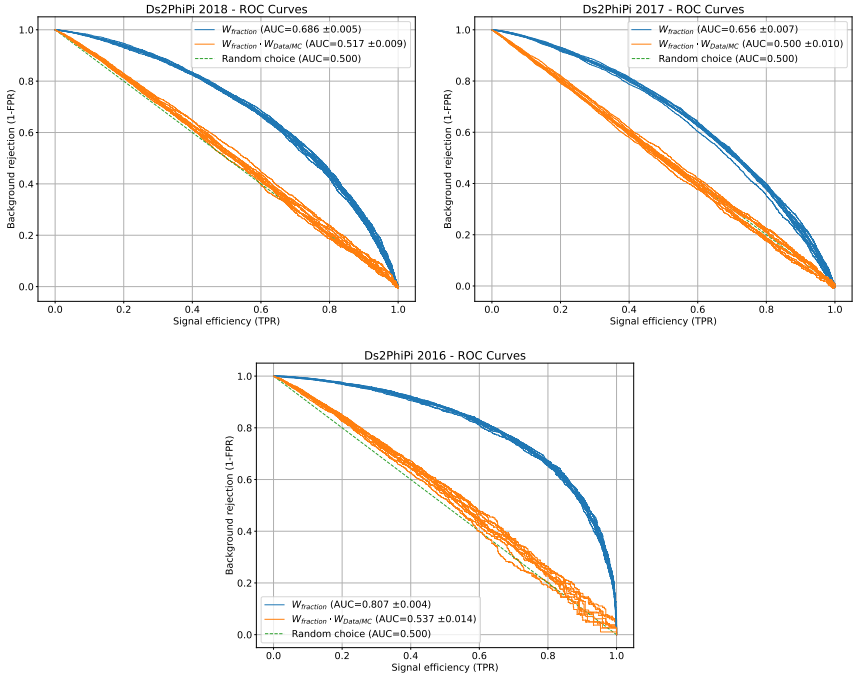


FIGURE 1.33: ROC curves obtained by a multivariate classifier trained to distinguish simulated and real data, for different years of data taking. The training-testing procedure was repeated 10 times, and the ROC curves are computed with and without the corrective weights. The presence of $w_{Data/MC}$ effectively reduce the capability to distinguish real and simulated data.

and provided that the TIS efficiency of any subsample of the triggered events is the same as that of the whole sample of selected events:

$$\frac{N_{\text{TISTOS}}}{N_{\text{TIS}}} = \epsilon_{\text{TOS|TIS}} = \epsilon_{\text{TOS, Data}} \tag{1.19}$$

This assumption implies that the signal candidate were completely uncorrelated with the rest of the events, but as an example B mesons are usually produced correlated with another b-hadron, thus the “rest of the event" is not likely to be independent as far as momentum spectra are concerned (trigger selection is usually based on p_{T} and IP cuts), but in small regions of the phase space this correlation can be considered small.

This correlation can be studied on simulated events, where the trigger efficiency can be computed both with the definition (N_{SEL} is known) and with the TISTOS method. If the overall true TIS efficiency is independent of the chosen TOS subsample, the joint probability of being TIS and TOS ($P(TIS \cap TOS)$) would be the product of the probabilities:

$$\frac{N^{TISTOS}}{N_{sel}} = P(TIS \cap TOS) = P(TIS) \times P(TOS) = \frac{N^{TIS}}{N_{sel}} \times \frac{N^{TOS}}{N_{sel}} \tag{1.20}$$

Since this is only approximately true, it can be expected that the ratio (ρ) between the TOS efficiency evaluated within the TIS sample and the true MC efficiency

should be ~ 1.

$$\rho = \underbrace{\frac{N^{TISTOS}}{N^{TIS}}}_{\epsilon_{MC,tistos}} / \underbrace{\frac{N^{TOS}}{N_{sel}}}_{\epsilon_{MC,standard}} = \frac{N_{sel} \times N^{TISTOS}}{N^{TIS} \times N^{TOS}} \sim 1 \tag{1.21}$$

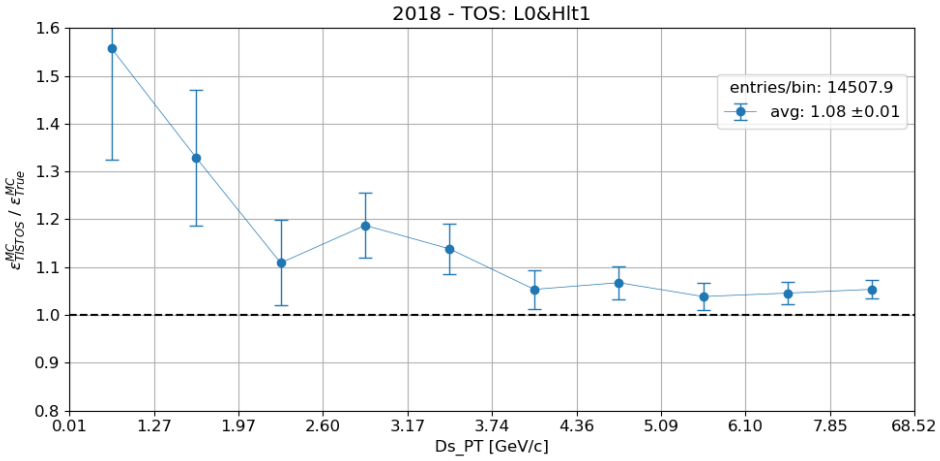


FIGURE 1.34: TISTOS efficiency over true trigger efficiency computed in the reference channel MC up to L0&HLt1, in different bins of D_s p_T.

Figure 1.34 shows the efficiencies ratio ρ evaluated in evenly populated bins of D_s transverse momentum in simulated sample. The numbers of events in each category is obtained by weighting the events by $w_{fraction}$. As can be seen, as the value of p_T increases, the discrepancy between the trigger efficiency computed with the TISTOS method and the true MC efficiency becomes smaller but not zero. This implies that the assumption of no correlation between the signal candidate and the rest of the event of the TISTOS method is not entirely true and thus a small bias is expected in the evaluation of the TISTOS efficiency.

The TIS subsample if formed by the preselected events that triggered the following combination of TIS lines:

$$\begin{aligned}(&L0Hadron|L0Muon|L0Photon|L0Electron)\&\\ &(Hlt1SingleMuonHighPT|Hlt1DiMuonLowMass|Hlt1TwoTrackMVA|\\ &\quad Hlt1TrackMuon|Hlt1TrackMVA|Hlt1TrackMuonMVA)\end{aligned} \tag{1.22}$$

1.6 Signal & Background discrimination

1.6.1 Backgrounds overview

A candidate can be considered as *background* for this analysis if it is characterized by three muonic reconstructed tracks from a common vertex. The three tracks can be produced by actual muons coming from a decay with unreconstructed neutral particles (in this case the event is hereafter classified as "3μ background") or they can be produced by other hadrons like p, π, K misidentified as muons (in this case the event is classified as "mis-ID background". In addition, the three tracks can be

originated by different vertexes but mistakenly associated to a common one and fake the signal. Such event is classified as "combinatorial background".

The most contributing 3μ background channel is the $D_s^+ \to \eta(\mu^+\mu^-\gamma)\mu^+\nu_\mu$, where the photon and the neutrino are not reconstructed. This decay is characterized by a large branching fraction: $(7.50 \pm 1.83) \times 10^{-6}$. The offline selection includes a cut on opposite-sign dimuon invariant mass to reject this background ($m_{\mu^+\mu^-} < 450\,\mathrm{MeV}/c^2$).

The mis-ID backgrounds comes principally from the c meson decays, since the decays of a b hadron into three muons without other charged particles in the final state would result in candidates with an invariant mass outside the signal and sidebands regions, due to the larger mass difference with respect to the τ mass. The particle identification system provides information to reject fake muon tracks, variables like $ProbNN_\mu$ and $ProbNN_K$, which are used to quantify the probability of a track to be an actual muon or kaon respectively. They are outputs of neural networks based on the PID and tracking information. The physics decays whose contribution to the misidentification backgrounds could be significant in this analysis are listed in Table 1.13. The approximate number of events produced in the LHCb acceptance in the three years of data taking (which correspond to $6\,\mathrm{fb}^{-1}$) is reported to compare qualitatively each contribution, since no further selection, trigger, or PID cuts are applied. Those mis-ID backgrounds are rejected by a multivariate classifier based on PID information that will be described in Section 1.6.4. In-depth studies of the dominant background channels are described in Sec. 1.7.

The remnant combinatorial background is rejected using multivariate classifiers that exploit kinematic and particle identification information to distinguish uninteresting events from signal $\tau^+ \to \mu^+\mu^-\mu^+$ candidates.

Decay	$\mathcal{B}\,[10^{-2}]$	$\sigma_{4\pi}(pp \to c\bar{c} + b\bar{b} \to ...)[\mu b]$	$\mathcal{N}[10^9]$
$D^+ \to K^-\pi^+\pi^+$	9.38 $\pm$0.16	5044 $\pm$131	568
$D^+ \to K^-K^+\pi+$	0.968 $\pm$0.018	5044 $\pm$131	59
$D^+ \to \pi^+\pi^+\pi^-$	0.327 $\pm$0.018	5044 $\pm$131	20
$D^+ \to K^-\pi^+\pi^+\pi^0$	6.25 $\pm$0.18	5044 $\pm$131	378
$D^+ \to \pi^+\pi^+\pi^-\pi^0$	1.16 $\pm$0.08	5044 $\pm$131	70
$D^+ \to K^-\pi^+\mu + \nu$	3.65 $\pm$0.34	5044 $\pm$131	221
$D^+ \to K^-K^+\pi^+\pi^0$	0.662 $\pm$0.032	5044 $\pm$131	40
$D_s^+ \to K^-K^+\pi^+$	5.39 $\pm$0.15	2722 $\pm$387	176
$D_s^+ \to K^+\pi^+\pi^-$	0.65 $\pm$0.04	2722 $\pm$387	21
$D_s^+ \to \pi^-\pi^+\pi^+$	1.08 $\pm$0.04	2722 $\pm$387	35
$D_s^+ \to K^-K^+\pi^+\pi^0$	6.2 $\pm$0.6	2722 $\pm$387	203
$D^{*+} \to D^0(K^-\pi^+\pi^0)\pi^+$	9.75 $\pm$0.35	5044 $\pm$131	590
$\tau^+ \to \pi^+\pi^-\pi^+\nu_\tau$	9.31 $\pm$0.05	137.4 $\pm$21.3	15

TABLE 1.13: List of the physics decays contributing to the misidentified backgrounds with their relative branching fraction. The number of events produced in about $6\,\mathrm{fb}^{-1}$ in the LHCb acceptance shown in the fifth column is estimated using the LHCb measured inclusive 4π cross-section (summing over charm and beauty contributions), the PDG branching fraction and a 20% acceptance efficiency, without any further selection.

1.6.2 Classifier with Kinematic information

The analysis strategy to identify signal events is defined separately for 3μ and for 2μ sample because of differences in the kinematic and geometric properties of the track without the `isMuon` requirement, and due to the unavailability of some information for such tracks ($ProbNN_\mu$ variable is defined only for tracks that satisfy the `isMuon` condition). Multivariate classification techniques use machine learning for the construction of a *classifier*, a tool that separate signal from background events by performing cuts on selected *features* of the event, variables that exhibit a different distribution in signal and background. The two different classifiers that make use of kinematic and geometric properties of the candidates are described in this section, while another classifier that exploits PID variables will be described in Sec. 1.6.4

Two different classifiers are developed: The one for the 3μ sample is an implementation of the eXtreme Gradient Boosting (XGBoost) [67] algorithm (which output is labelled XGB), while for the 2μ sample the chosen algorithm is a Gradient Boosted Decision Tree (which output is labelled BDTG), available in the TMVA toolkit [68]. The training of those models uses simulated events as signal proxy, and real data from the inner sidebands as background proxy. To make use of all the available events in the dataset, the training is performed using the `k-folding` technique: simulated and real data are combined in a single dataset that is shuffled and divided into k consecutive parts (folds), each fold is then used once as a test set while the $k-1$ remaining folds form the training set (a k value of 5 was used in this analysis, that correspond to 80% of training events and 20% of test events). During the training phase, the model learn to classify the events by iterative optimization of the thresholds that are used in the decision trees to split events in different groups (or *leafs*) [69]. The output of the models represent the probability of the event of being signal-like. The variables used as features of the models (for both XGB and BDTG) are listed in Table 1.14 and the corresponding distributions for signal and background are shown in Figure 1.35. A detailed description of the isolation variables `LONGMAX` and `VELOMAX` is given in Sec. 1.6.2.

The classifiers are defined by a set of hyperparameters that characterize the structure of the base estimators or the learning process:

- `Ntrees`: number of trees in the forest;
- `MaxDepth`: maximum depth of the decision tree allowed;
- `MinNodeSize`: minimum percentage of training events required in a leaf node;
- `nCuts`: number of grid points in the variable range used in finding the optimal cut in the node splitting.
- `LearningRate`: controls how much to change the model each time its weights are updated in response to the estimated error.

The search for the best set of parameters is performed separately for each year using a grid search over the space parameters and `k-folding`. In a binary classifier, its output is compared with a threshold and the event is classified as background(signal) if the output is lower(greater) than the threshold. Given the true class of the events, there are four different categories on which events are divided into:

- True signal event (P) can be correctly identified (True Positive, TP) or mistaken for a background event (False Negative, FN)

Variable	Definition
tau_DIRA_OWNPV	Cosine of the angle between the τ momentum and the vector from the primary to the decaying vertex
tau_ENDVERTEX_CHI2	χ^2 fit of the decay vertex
tau_TAU	Particle decay time
μ_TRACK_CHI2NDOF	χ^2 of the track fit divided by the degrees of freedom, for each of the daughter particles
IPCHI2_OWNPV	Difference in the vertex-χ^2 of the PV reconstructed with and without the particle, for τ and daughters particles
PT	Transverse momentum of τ and daughters particles.
tau_ConeMultMuon	Number of tracks within a cone built around a muon track
tau_ConePtAsymMuon	Asymmetry between the transverse momentum of a muon and the sum of the transverse momenta of all the particles within a cone built around that muon
VELOMAX	Average of tau_TRKISOBDTVELO variables computed for each pair of daughter particles
LONGMAX	Average of tau_TRKISOBDTLONG variables computed for each pair of daughter particles

TABLE 1.14: Features of the events used by the XGBoost model.

- True background event (P) can be correctly identified (True Negative, TN) or mistaken for a signal event (False Positive, FP)

Different metrics can be defined to characterize the performance of a classifier to recognize signal or background events, the most commons are shown in Table 1.15. The recall score is also known as *signal efficiency* (ϵ_S) while the specificity is also called *background rejection* ($1 - \epsilon_B$). The Receiver Operating Characteristic (ROC) curve is obtained from the values of signal efficiency and the background rejection at different thresholds, and its area (AUC) can be used to evaluate the classifier (the greater value, the better the performance).

Metric	Definition
Accuracy	(TP+TN)/(P+N)
Recall	TP/P
Specificity	TN/(TN+FP)
Precision	TP/(TP+FP)
AUC	Area under ROC curve

TABLE 1.15: Table with metrics used to evaluate the classifier performance.

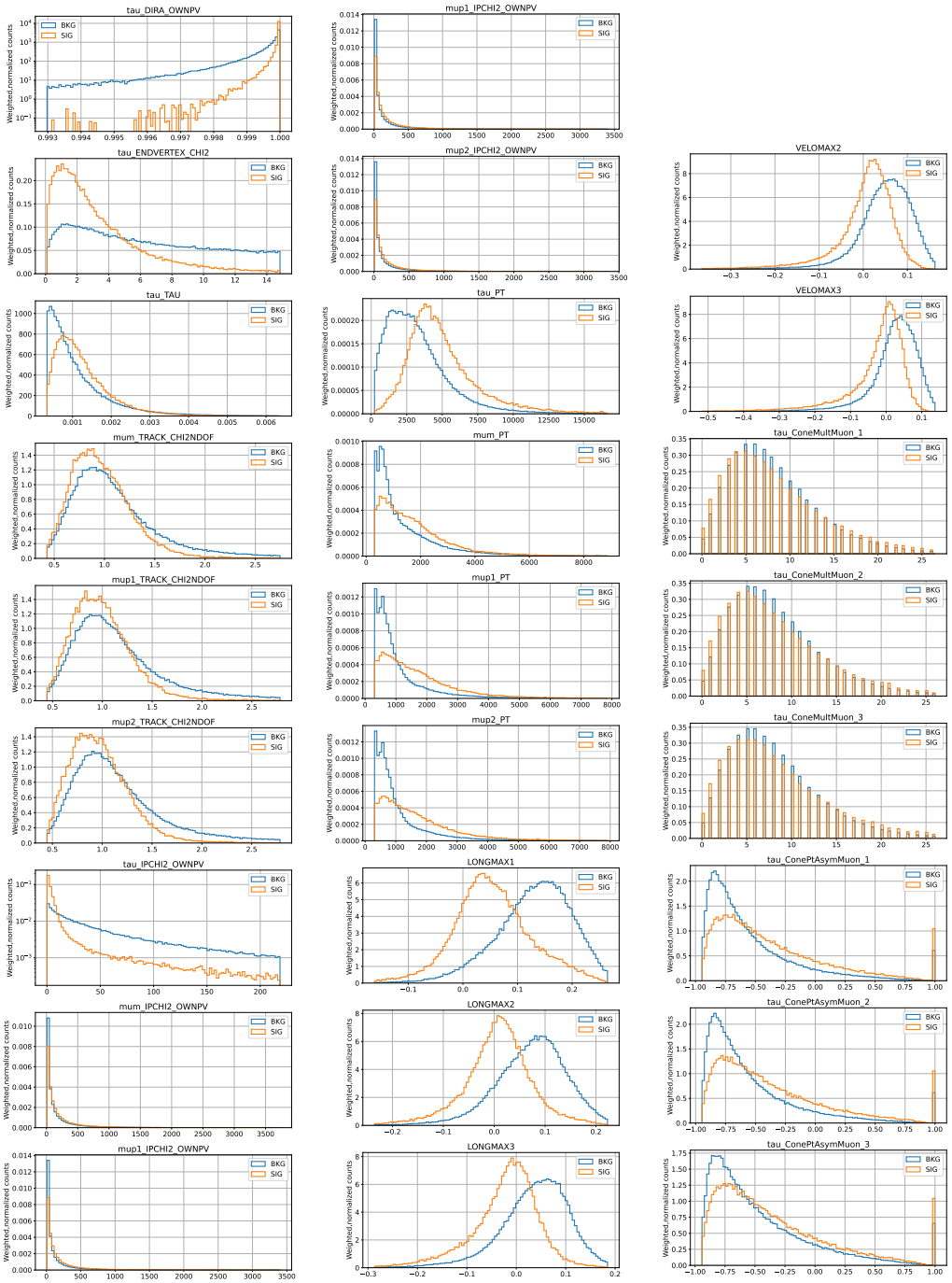


FIGURE 1.35: Distributions of the input features of the XGBoost classifier, for the signal(MC) and background(data inner sidebands) sample. The histograms are normalized to unit area.

The performance obtained with the XGBoost classifier and the BDTG classifier are shown respectively in Table 1.16 and Table 1.17.

During the training phase and in the evaluation of the performance of the models, the MC events were weighted by the product $w_{train} = w_{fraction} \times w_{Data/MC} \times w_{balance}$, a weight obtain by multiplication of three different sets of weights, while the background events have unitary weight:

- $w_{fraction}$: It considers the different τ production subchannels, as defined in Section 1.3.1. It is the same for all the events of a certain subchannel and polarity.

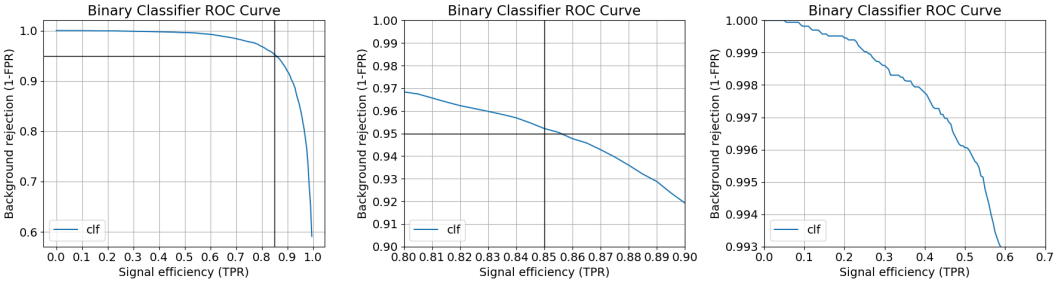


FIGURE 1.36: (Left) ROC curve of the XGB binary classifier trained on the 2018 datasample. (Center, Right) Close-up of the ROC curve in different range of signal efficiency or background retention rate.

scores [%]	2018	2017	2016
Accuracy	91.09 $\pm$0.16	90.87 $\pm$0.07	92.30 $\pm$0.12
Recall (ϵ_S)	87.72 $\pm$0.15	89.83 $\pm$0.13	88.90 $\pm$0.35
Specificity (1-ϵ_B)	93.20 $\pm$0.20	91.73 $\pm$0.12	94.27 $\pm$0.11
Precision	88.94 $\pm$0.30	89.99 $\pm$0.12	94.27 $\pm$0.11
AUC	97.01 $\pm$0.08	97.04 $\pm$0.05	97.69 $\pm$0.05
ϵ_S @ 1-ϵ_B = 95%	94.67 $\pm$0.18	94.69 $\pm$0.10	95.97 $\pm$0.11
1-ϵ_B @ ϵ_S = 85%	84.24 $\pm$0.32	84.30 $\pm$0.29	87.60 $\pm$0.34

TABLE 1.16: Classifier average performance obtained on the 3μ sample with the chosen set of hyper-parameters quantified by different metrics. The last two scores are the signal efficiency obtained when requiring a background rejection of 95% and the background rejection obtained when requiring a signal efficiency of 85%. The test dataset is 10% of the 3μ sample, and the training-testing task is repeated 5 times to evaluate uncertainty on the scores.

Score	**2016**	**2017**	**2018**
AUC (%)	90.16 ± 0.58	93.58 ± 0.24	93.94 ± 0.23
ϵ_S @ $1 - \epsilon_B = 30\%$	88.18 ± 2.01	94.40 ± 0.52	94.72 ± 0.18

TABLE 1.17: Performance of the BDTG model on the 2μ sample in terms of the AUC and signal efficiency at a background rejection of 30%.

- $w_{Data/MC}$: It reduces the difference between simulated and real events, thus enhancing the differences of *signal* and *background*. It is defined on an event-by-event basis (details in Section 1.5.2).

- $w_{balance}$: It balances the model training, correcting for different sizes of the MC and data samples. It is identical for each MC event and is defined as the ratio between the number of background events and the weighted sum of MC events ($\sum_i w^i_{fraction} \cdot w^i_{Data/MC}$)

Isolation Variables

Among the input variables of the classifiers, large discrimination power is given by the *isolation* variables. In the distinction between signal and combinatorial background, the muons' isolation play a key role. The isolation variables described in this section were developed for the $B_{(s)} \to \mu^+\mu^-$ analysis [70] but can be used in this analysis too. With muon *isolation* it is indicated the property of the muon to be relatively isolated from the other tracks of the event, in comparison with background channels where one or more daughter particles can be produced close to it.

These variables are the output of a global BDT classifier[1] used to distinguish between non-isolating tracks in $b\bar{b} \to \mu^+\mu^- X$ and isolating tracks in $B_s \to \mu^+\mu^-$ MC events. Two different outputs can be produced, depending on whether long or velo tracks are fed as input to the BDT (see track's definition in Section 1.2.1).

The BDT returns an isolation score value for each pair of tracks in the final state, namely the track whose isolation is computed for and the track it is computed against. Since in case of our signal channel three are the muons in the final state, three isolation variables are computed for each of the two categories. To summarize the content of the isolation variables, they are combined into three `LONGMAX` variables and three `VELOMAX` variables:

$$\text{LONGMAX}i = \frac{1}{3} \times \left\{ \begin{array}{l} tau\_TRKISOBDTLONG\_D1MAXi\_01+ \\ tau\_TRKISOBDTLONG\_D1MAXi\_12+ \\ tau\_TRKISOBDTLONG\_D2MAXi\_02 \end{array} \right\} \text{for } i = 1,2,3$$

$$\text{VELOMAX}i = \frac{1}{3} \times \left\{ \begin{array}{l} tau\_TRKISOBDTVELO\_D1MAXi\_01+ \\ tau\_TRKISOBDTVELO\_D1MAXi\_12+ \\ tau\_TRKISOBDTVELO\_D2MAXi\_02 \end{array} \right\} \text{for } i = 1,2,3$$

Where $D1$ and $D2$ represents the first and the second daughter (between the pair of tracks under analysis) around which the cone is computed, and the suffix *_jk* represents which combination of the three particles is being considered.

Calibration of the Classifiers

In order to check possible differences between the performance of the XGBOOST classifier on data and MC, even though the $w^{sig}_{Data/MC}$ weights defined in Sec 1.5.2 are used during the model training, the calibration channel is used. These differences can be due to remaining small Data/MC differences in some features used by the model, as can be seen from the plots in Figure 1.37.

The input features of the S/B classifier are defined for events with three muon tracks. Since there are only two muons in the reference channel, the input features are slightly modified. The definition of isolation variables (LONGMAXi and VELOMAXi) changed into:

$$\text{LONGMAX}i = \frac{1}{2} \times \left\{ \begin{array}{l} Ds\_TRKISOBDTLONG\_D1MAXi+ \\ Ds\_TRKISOBDTLONG\_D2MAXi \end{array} \right\} \text{for } i = 1,2,3$$

[1]The BDT is developed using the *Adaboost* boosting algorithm and the TMVA toolkit.

$$\text{VELOMAX}i = \frac{1}{2} \times \left\{ \begin{array}{l} Ds\_TRKISOBDTVELO\_D1MAXi+ \\ Ds\_TRKISOBDTVELO\_D2MAXi \end{array} \right\} \text{for } i = 1,2,3$$

Regarding the variables of single particles in the signal channel, each one is simply replaced with one of the particles in the reference channel.

To estimate the signal efficiency of the S/B classifier on Data, the signal efficiency evaluated on MC $\epsilon^{sig}_{S,MC}$ is corrected for residual Data/MC differences estimated on the reference channel using histograms ratio. The ratio between the binned distributions of the XGBoost output on the reference channel Data and MC, shown in Figure 1.38, is used to compute correction factors ($r^{ref}_{XGB\_corr}$). These factors are used as a per-event corrective weight ($w_{XGB-corr}$) which depends on the XGB output associated to that event. The corrected distributions of the XGB output for signal and reference channel are shown in Fig 1.39. The small rise of the XGB distribution for outer sidebands when approaching 1 is due to some mis-ID contributions in the left outer sideband, as can be seen from Figure 1.44. Those events will be removed by another classifier, described in Section 1.6.4.

The BDTG classifier's output is also corrected for residual Data/MC differences, analogously to the XGBoost classifier: the features of the pion in the $D_s \to \phi(\mu\mu)\pi$ channel replaces the features of the muon with `!isMuon`, while the two good reconstructed muons are matched with the two muons in the reference channel. The ratio of the BDTG output distributions on Data and MC is again used to define a correction weight.

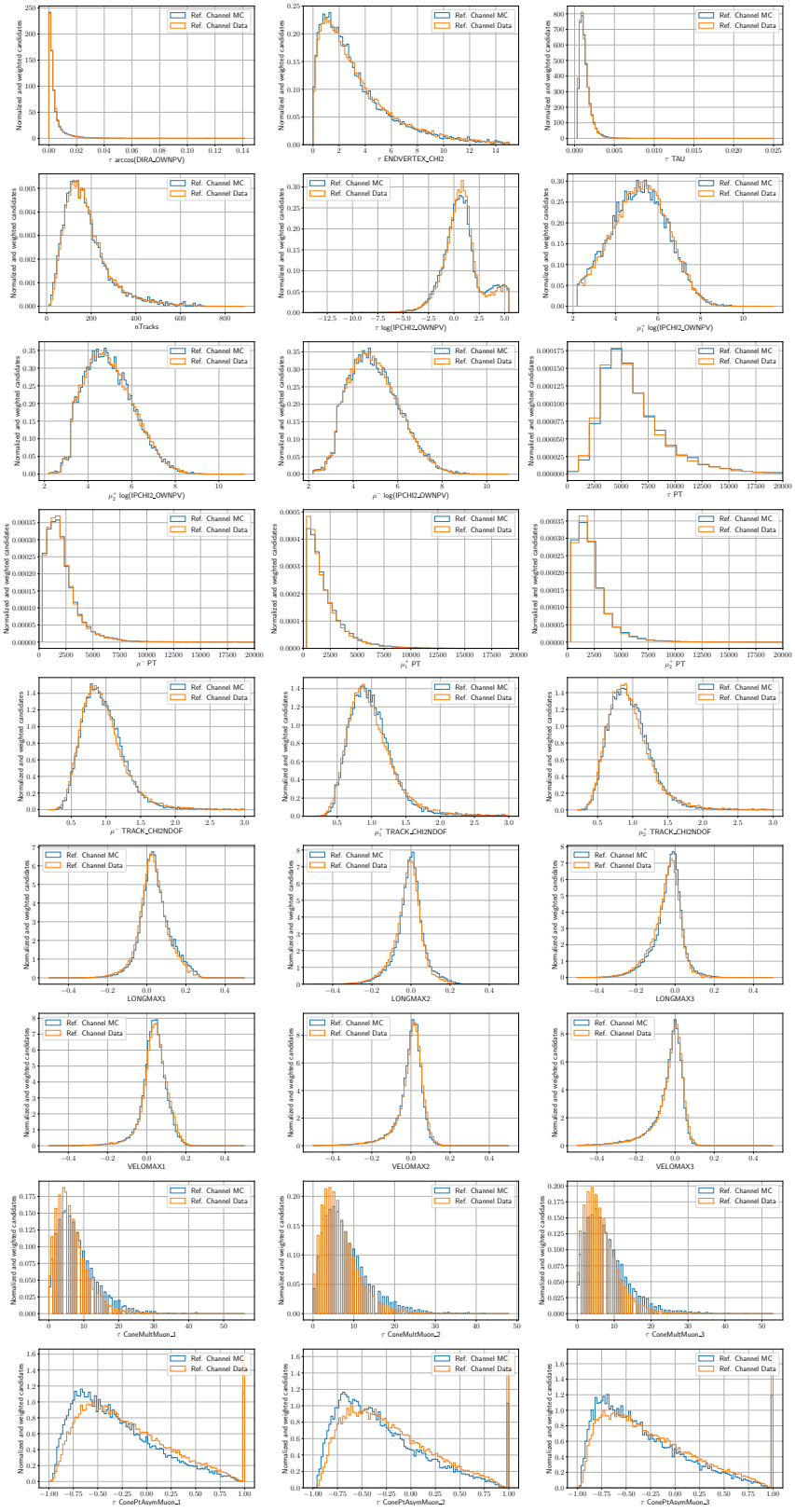


FIGURE 1.37: Distributions of the input features of the XGBoost classifier model reference channel MC and sWeighted Data for 2018. The features of the particles D_s, π, μ^+, μ^- are matched with the ones of $\tau, \mu_1^+, \mu^-, \mu_2^+$.The variables of the pion are displayed together with the one of μ_1.

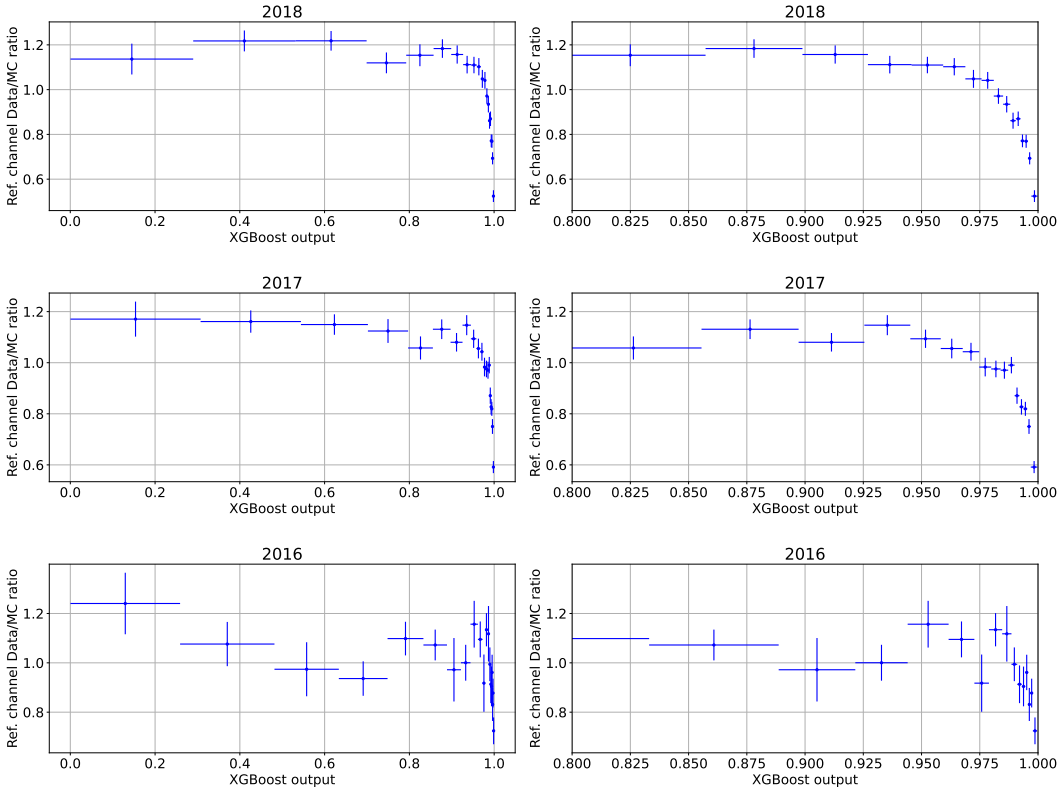


FIGURE 1.38: (Left) Bin-by-bin ratio between the Data and MC distributions, used to reweight the signal sample MC and the reference channel MC itself. (Right) Zoom in of the region at high values of XGB output. The bins edges are computed to have uniform content in the simulated sample.

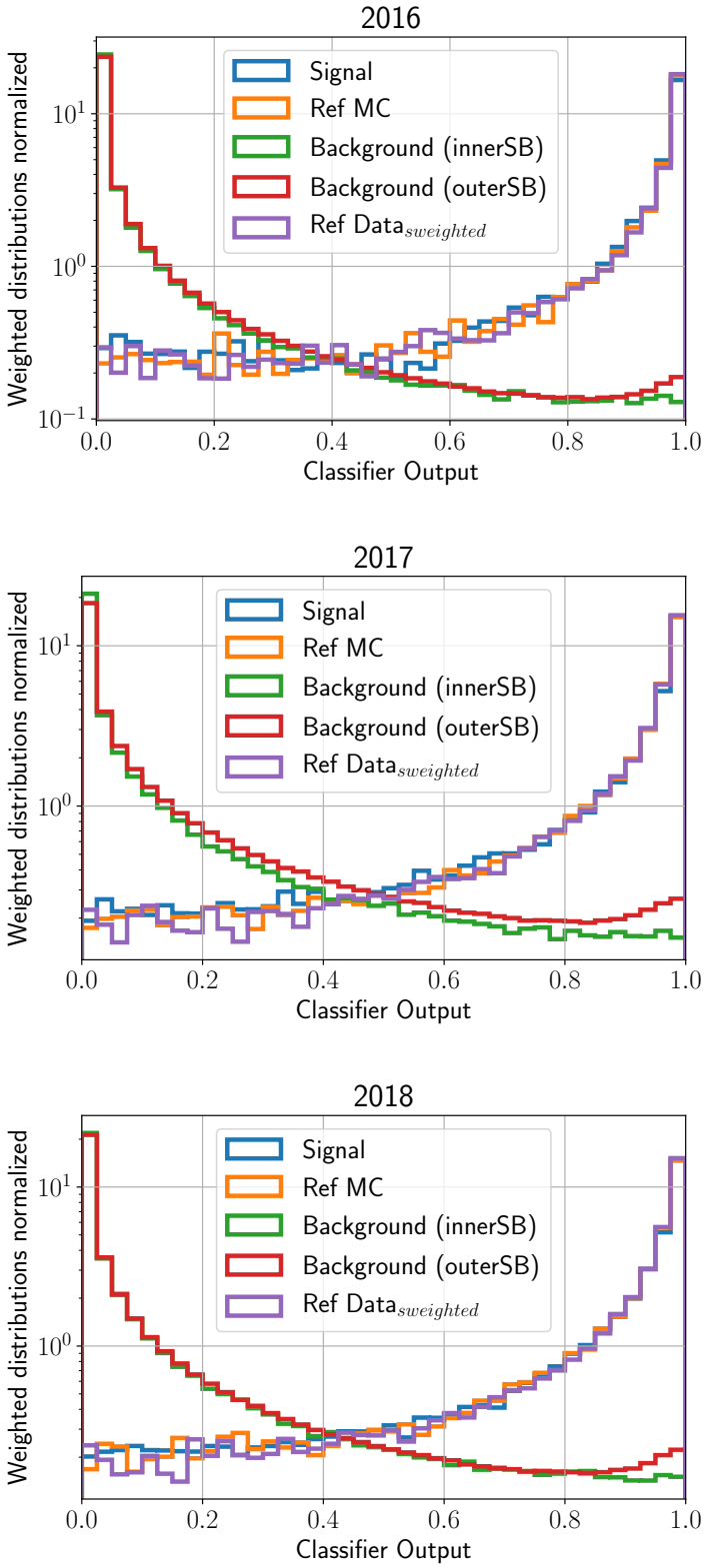


FIGURE 1.39: Classifier output's distributions for different samples, normalized to unit area, for different years. Simulated samples are weighted using $w_{fraction}$, $w_{Data/MC}$, $w_{XGBcorr}$, and PID correction weights defined in Sec.1.3.1, Sec. 1.5.2 and Sec. 1.6.2

1.6.3 Particle Identification

The variable $ProbNN_\mu$ is the main variable used for the muon identification in the analysis, together with the `IsMuon` flag. Usually, Data and MC differences in particle identification variables are corrected using the PIDCalib package [71]. The strategy used in this analysis consists in constructing a 4-dimensional map to extract a per-event weight to correct the PID variables in the simulated samples, similarly to what is done in PIDCalib but using the reference channel $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ instead of its standard samples (with the aim of using samples with more signal-like muons). The variables used to compute the 4-dimensional map are:

- $min(ProbNN_\mu)$: minimum $ProbNN_\mu$ value among the muons in the decay chain,
- p_μ: momentum of the muon with the lowest value of $ProbNN_\mu$,
- η_μ: pseudorapidity of the muon with the lowest value of $ProbNN_\mu$,
- $nTracks$: track multiplicity in the event.

The binning scheme is defined to have 10 equally populated (uniform) bins in $min(ProbNN_\mu)$, and next for each $min(ProbNN_\mu)$ region the remnant 3 variables are binned with a variable number of uniform bins to consider correlations with $min(ProbNN_\mu)$. The 4D map is created for both data and MC sample, and from the ratio of the bins' content a corrective weight (w_{PID}) can be extracted for each candidate. To account for possible differences in each data taking period, the maps are computed separately for each year. The map of the simulated sample is filled with events weighted for the production rates and for the Data/MC differences (as described in Section 1.3.1 and Section 1.5.2), while the sWeights used for the data map to suppress the combinatorial background are computed separately for each bins of $min(ProbNN_\mu)$ (to take into account correlation between the distribution of D_s^+ mass and the PID variables). The performed fit procedure is similar to the one described in Section 1.5.1, but with a slight change to the PDF that describe the signal component (while the background component is still described by a simple negative exponential): a Gaussian function is added to the Johnson's SU function, resulting in

$$\mathcal{P}^{sig}(m) = (1 - f_{tail})\mathcal{G}(m, \mu, \sigma) + f_{tail} \cdot \mathcal{J}(m, \mu, r\sigma, \gamma, \delta) \tag{1.23}$$

where f_{tail} is the relative fraction between the two functions, which share the same mean μ, while the width of $\mathcal{P}$ is scaled by a factor r.

The validation of this reweighting procedure is done by comparing the distributions of few kinematic variables in data and simulation subsample, where $ProbNN_\mu > 0.4$: D_s momentum p_{D_s}, the D_s transverse momentum p_{T,D_s}, $nTracks$, $min(ProbNN_\mu)$, $\mu^\pm$ momentum $p_{\mu^\pm}$, $\mu^\pm$ transverse momentum $p_{T,\mu^\pm}$ and $\mu^\pm$ pseudorapidity $\eta_{\mu^\pm}$. The effect of the additional set of weights w_{PID} is to reduce the mismatch between data and simulation in all the variables in Figure 1.40.

PID correction in $\tau^+ \to \mu^+\mu^-\mu^+$ channel The 4D-maps defined to correct Data/MC differences in the reference channel are also used to reweight $\tau^+ \to \mu^+\mu^-\mu^+$ events. The procedure is straight-forward for the 3μ sample because all the muons passed the `IsMuon` requirement. The value to access the map are taken from the one with the lowest value of $ProbNN_\mu$, but the weight w_{PID} correct the distributions of each of

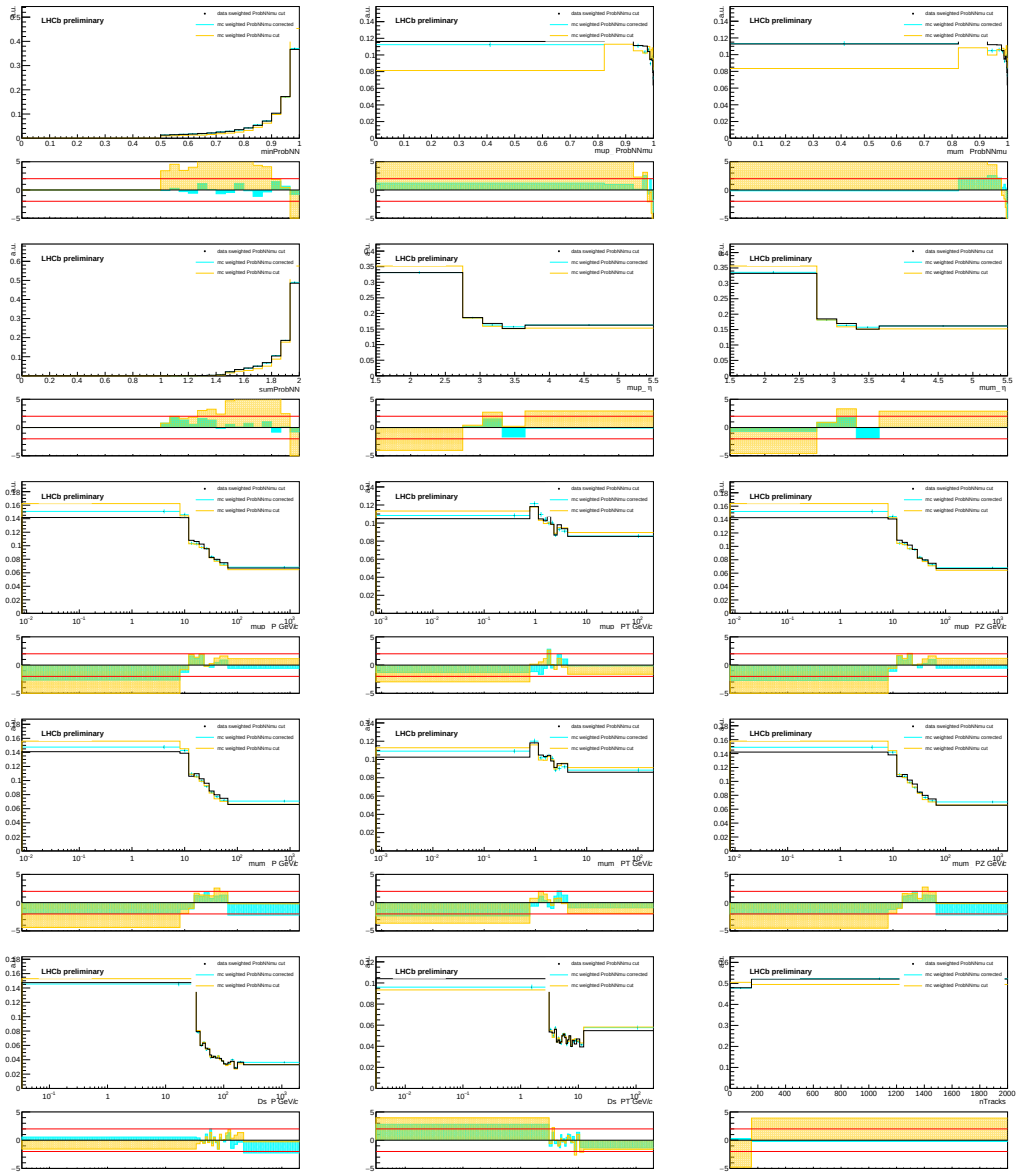


FIGURE 1.40: Comparison of the 2018 distributions of: $min(ProbNN_\mu)$, $ProbNN_{\mu_{\mu^+}}$ and $ProbNN_{\mu_{\mu m}}$ (first row), $sum(ProbNN_\mu)$, η_{μ^+} and $\eta_{\mu m}$ (second row), p_{μ^+}, p_{T,μ^+} and p_{Z,μ^+} (third row), $p_{\mu m}$, $p_{T,\mu m}$ and $p_{Z,\mu m}$ (fourth row), p_{D_s}, p_{T,D_s} and $nTracks$ (fifth row). In each plot the distribution of data and simulated sample before and after the reweighting (top) and the related pulls (bottom) are depicted.

the three particles. In the 2μ sample, the variable $RichDLL_\mu$ replace the missing variable $ProbNN_\mu$ and therefore it needs to be corrected too. A new set of 4-dimensional maps are created to correct Data/MC differences and the aforementioned strategy is repeated, replacing the $ProbNN_\mu$ bins with bins of $RichDLL_\mu$ and considering for each event only the muon with lowest $min(ProbNN_\mu)$. In fact, as shown in Figure 1.41 the distribution of $RichDLL_\mu$ in the 2μ signal channel for the muon without the `IsMuon` flag is similar to the one obtained considering the one of the two other muons with the minimum value of $ProbNN_\mu$.

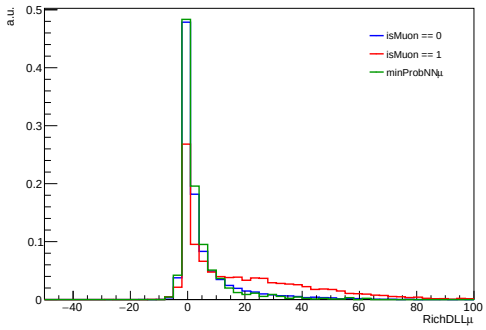


FIGURE 1.41: Example of the distribution of $RichDLL_\mu$ variable for the muon with the lowest value of $ProbNN_\mu$ in the 2μ sample of the signal channel, compared to the distribution obtained from the muon with and without the `IsMuon` flag.

1.6.4 Classifier with PID information

The XGBoost multivariate classifiers described in Sec. 1.6.2 has proven to reach excellent performance in discriminating the signal $\tau^+ \to \mu^+\mu^-\mu^+$ events from the combinatorial background. In this section is described a similar strategy followed to construct a new classifier, called $SuperProbNN$ and based on the same XGBoost algorithm, which uses the PID variables listed in Table 1.18 instead of kinematic variables to discriminate signal from background. The output of this classifier will be used together with the previous XGBoost model's output (XGB) to define a binned region where the branching ratio limit for $\tau^+ \to \mu^+\mu^-\mu^+$ is computed. The multivariate model is built with 300 decision trees, with a maximum depth of 4 and minimum sample leaf of 300, trained with a learning rate of 0.3 over simulated $\tau^+ \to \mu^+\mu^-\mu^+$ sample as signal proxy, and on the right inner sideband of data as background proxy. The left data inner sideband is excluded from the training sample because sensible to contamination of mis-ID background from the $D^+ \to K^-\pi^+\pi^+$ channel. This contribution is mainly rejected by the multivariate classifier, the small remnant fraction of events can be removed by a dedicated veto, as will be described in Section 1.7.2.

Variable	Definition
$ProbNN_\mu$'s	Probability of the tracks to be identified as a muon
$ProbNN_K$'s	Probability of the tracks to be identified as a muon
η's	Track's pseudorapidity

TABLE 1.18: Features of the events used by the SuperProbNN model.

A `k-folding` technique is applied for the $SuperProbNN$ classifier (with $k = 5$), to use the whole dataset without introducing any bias. The training is performed separately for each year of data taking. The distribution of the input features is shown in Fig. 1.42, where data are sWeighted and the simulated events are weighted with the product of $w_{fraction}$, $w_{Data/MC}$ and w_{PID}. The discriminating power of the SuperProbNN classifier is visible from the distribution of its output ($SuperProbNN$) for signal MC and data sidebands. The value assigned to outer sidebands' events is obtained by averaging the response of five models trained with the `k-folding`

technique. Table 1.19 reports the scores obtained by the SuperProbNN model for different years of data taking.

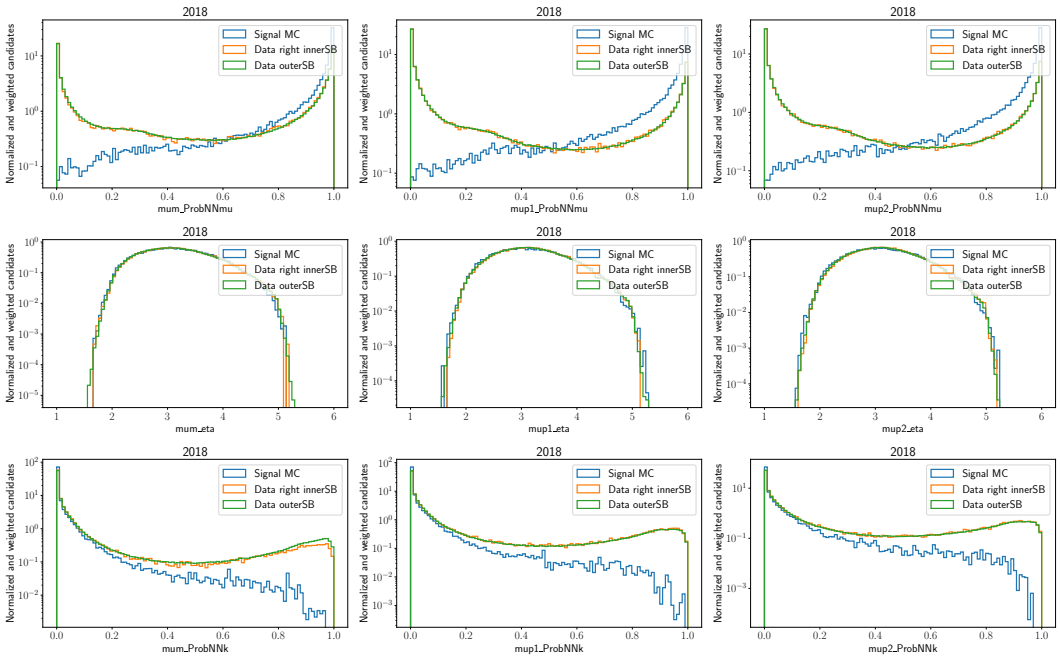


FIGURE 1.42: Distributions of the input features for the SuperProbNN model in signal MC, right inner Data sideband and outer Data sidebands, for the 2018.

scores [%]	2018	2017	2016
Accuracy	88.86 ±0.38	89.17 ±0.38	86.75 ±1.25
Recall (ϵ_S)	82.53 ±0.69	84.67 ±0.62	77.65 ±2.00
Specificity (1-ϵ_B)	96.77 ±0.12	96.35 ±0.13	97.35 ±0.03
Precision	96.96 ±0.11	97.37 ±0.09	97.16 ±0.07
AUC	97.97 ±0.08	97.92 ±0.04	97.81 ±0.20
ϵ_S @ 1-ϵ_B = 95%	89.04 ±0.39	89.31 ±0.18	88.59 ±0.57
1-ϵ_B @ ϵ_S = 85%	96.21 ±0.14	96.22 ±0.14	95.94 ±0.32

TABLE 1.19: SuperProbNN Classifier average performance obtained with the chosen set of hyper-parameters quantified by different metrics. The last two scores are the signal efficiency obtained when requiring a background rejection of 95% and the background rejection obtained when requiring a signal efficiency of 85%. The results are averaged over the predictions made on the 5 folds (train:test ratio is 80%:20%)

Combined Performance

The size of signal-enriched sample on which the limit will be calculated depends on the cuts that are applied on the discriminating variables XGB and superProbNN. Lower levels of background can be achieved by tight cuts at the cost of reject potential signal candidates. Another strategy to recover these candidates consists of

applying minimal cuts on the discriminating variables, and calculate the limit in various bins of such variables. The correlation between XGB and superProbNN should be minimized in order to get the most benefit from the 2D binning. As can be seen from Figure 1.43, there is no significant correlation between the output of the two multivariate classifiers in the MC or data sample. In addition to the outputs of the multivariate classifiers, the invariant mass of the signal candidates is considered: the mass signal window is defined in Sec. 1.3.1 to be $\pm 20\,\mathrm{MeV}/c^2$ around the known τ mass ($1776.86 \pm 0.12\,\mathrm{MeV}/c^2$). Since the limit will be computed also in mass bins, the invariant mass of the candidate should not correlate with XGB or superProbNN. No visible correlation can be observed either from the distributions of the kinematic classifier as function of the invariant mass (Figure 1.44) or from the distributions of the PID classifier (Figure 1.45).

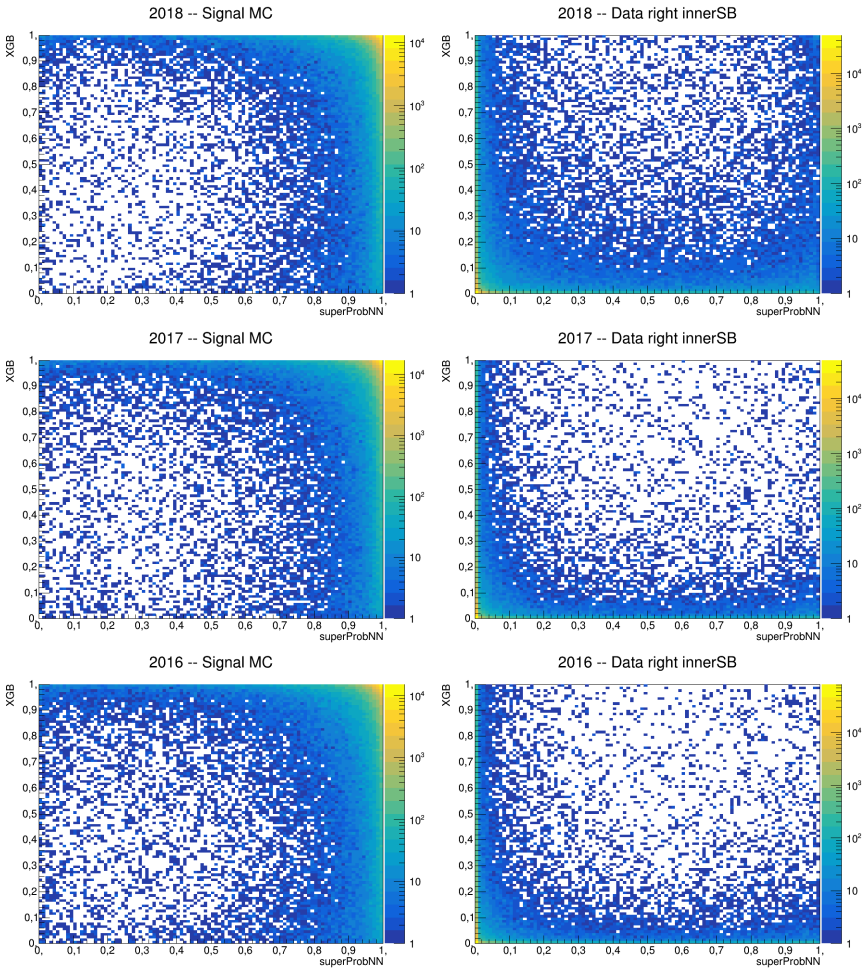


FIGURE 1.43: Correlation between the XGB classifier output and the superProbNN output, for signal MC (left) and for Data right inner sidebands (right).

The choice of the minimal cuts' thresholds and the finer binning is made to optimize the signal sensitivity α (defined in Eq. 1.12). The minimum value to classify as signal a candidate with the XGB classifier is 0.8, while for the SuperProbNN classifier the value is 0.88. Further details on the definition of the binning scheme can be found in Section 1.9. How the values of the thresholds can affect the signal efficiency and the number of expected background can be observed from the left plots on Figure 1.46. The number of expected events is estimated by integrating in the

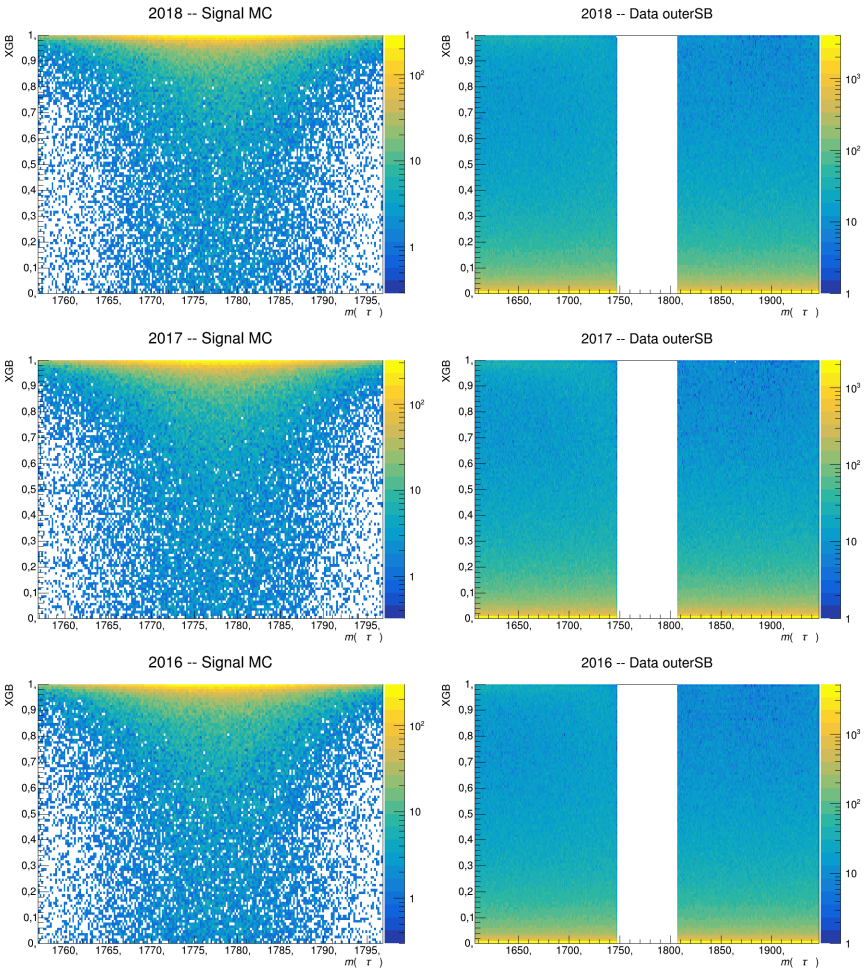


FIGURE 1.44: Correlation between the XGB classifier output and the invariant mass distribution in the signal channel, for signal MC (left) and for Data outer sidebands (right).

signal window the exponential function used to fit the data outer sidebands, populated by events that passed the minimal cut $XGB > 0.8$, which has an efficiency of 74.03±0.23% for 2018, 75.87 ±0.23% for 2017 and 75.93 ±0.50% for 2016. The same cut is applied on the simulated sample, and the signal efficiency shown in the plots is computed within the subsample that survived the cut. If the analysis of the 3μ sample had been performed without binning, the optimal value to define a threshold on a discriminant variable for a counting experiment would be given by the optimization of the Punzi Figure of Merit (FoM) [72].

$$FoM = \frac{\epsilon_S}{\frac{3}{2} + \sqrt{N_{bkg}}} \tag{1.24}$$

Where ϵ_S is the signal efficiency extracted from MC, while N_{bkg} is the number of the expected background events in the signal region (obtained from a fit to the data outer sidebands). The maximum value is obtained with the requirement $superPNN > 0.965$, given the cut on the XGBoost classifier $XGB > 0.8$. A similar strategy is applied on the 2μ sample, where the statistics is lower and its contribution to the limit is evaluated without binning (with a single bin).

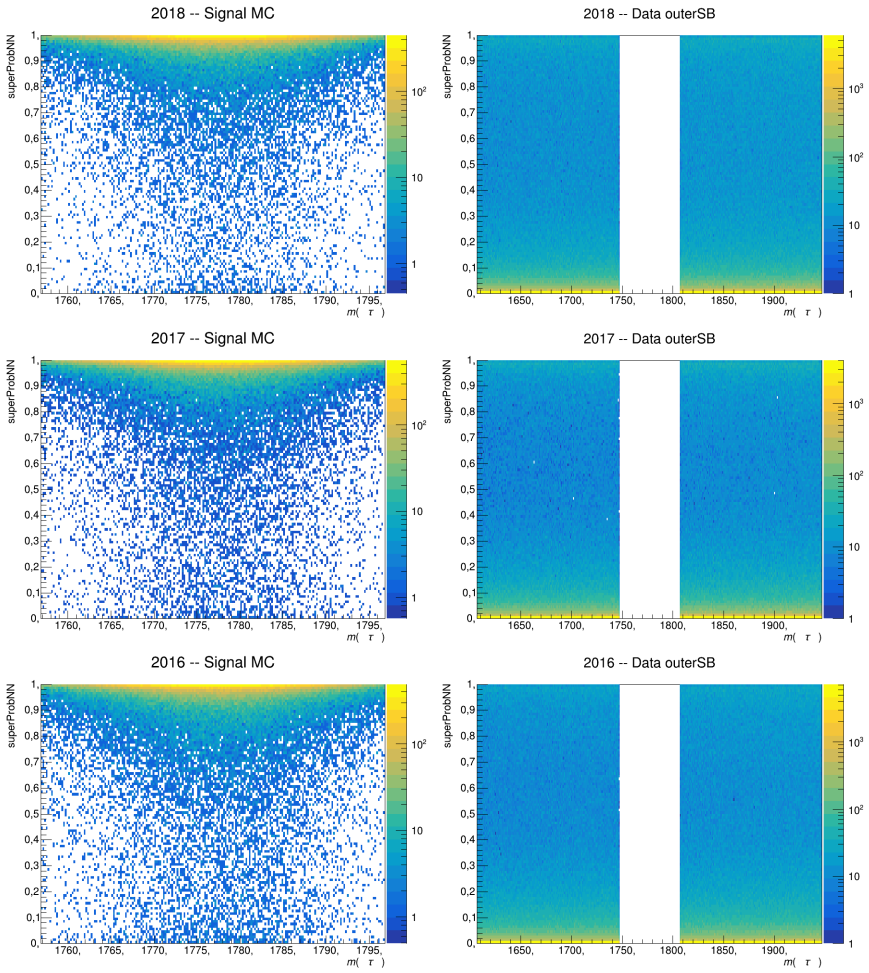


FIGURE 1.45: Correlation between the superProbNN classifier output and the invariant mass distribution in the signal channel, for signal MC (left) and for Data outer sidebands (right).

Selection of the 2μ sample The selection of the signal candidates is performed with a set of cuts on the PID information provided by $ProbNN_\mu$ for the tracks correctly identified as muons, by $RichDLL_\mu$ for the other track not identified and on the kinematic multivariate classifier output (BDTG). The optimization of each cut is done separately for the three variables, as no correlation between them is observed, as shown is Figure 1.47. The best cuts chosen for each data taking year are listed in Table 1.20

Year	**BDTG**	$RichDLL_\mu$	**sum**($ProbNN_\mu$)
2016	> 0.8	> 30	> 1.55
2017	> 0.9	> 30	> 1.75
2018	> 0.8	> 35	> 1.75

TABLE 1.20: Best BDTG and PID cuts for the 2μ sample.

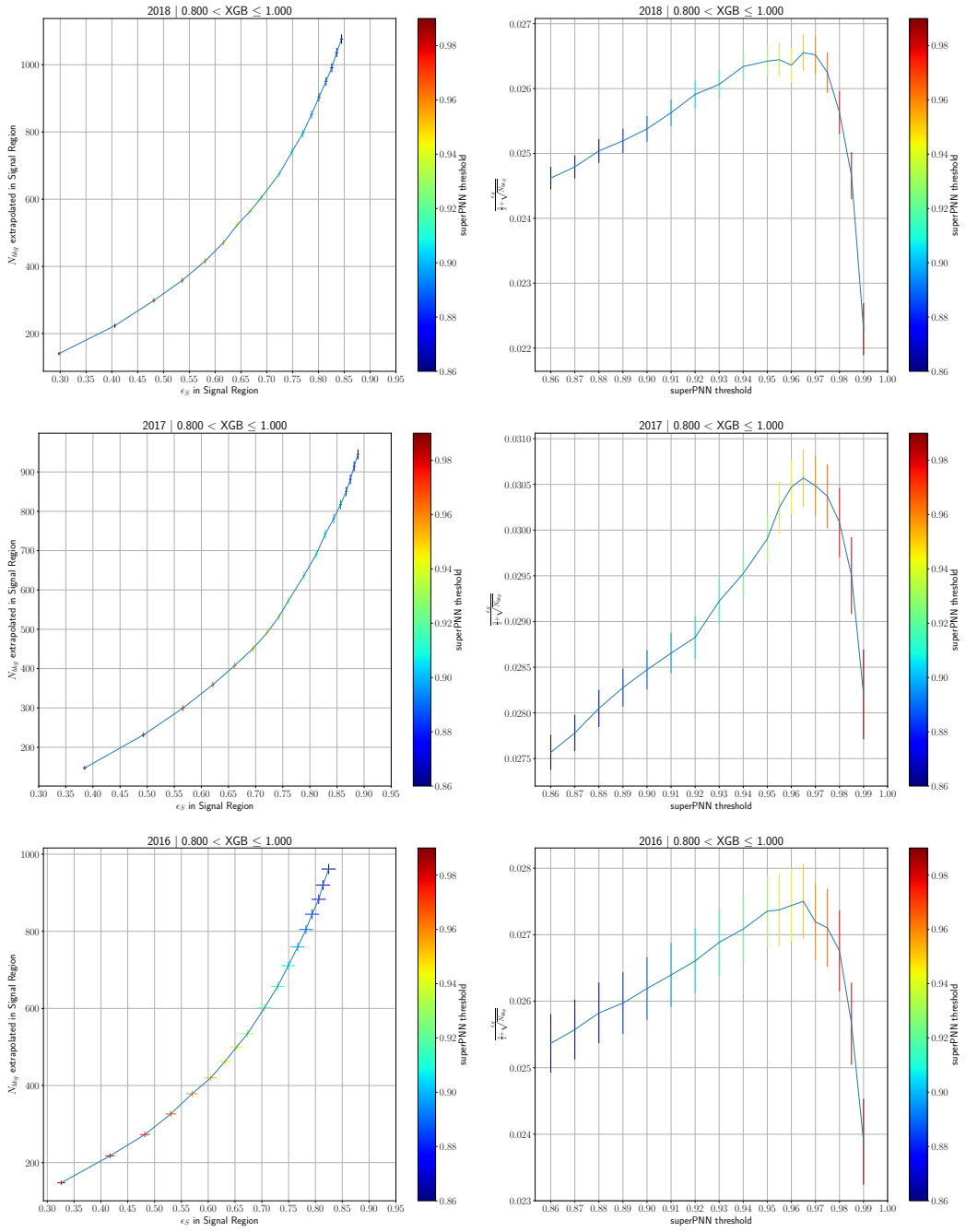


FIGURE 1.46: (Left) Expected number of background events extracted from a fit to data outer sidebands and the corresponding signal efficiency evaluated on $\tau^+ \to \mu^+\mu^-\mu^+$ MC sample at different thresholds of superProbNN classifier. The signal efficiency is evaluated on the events in the signal window that pass the minimal cut on XGB. (Right) Punzi figure of merit evaluated at different thresholds of superProbNN variable.

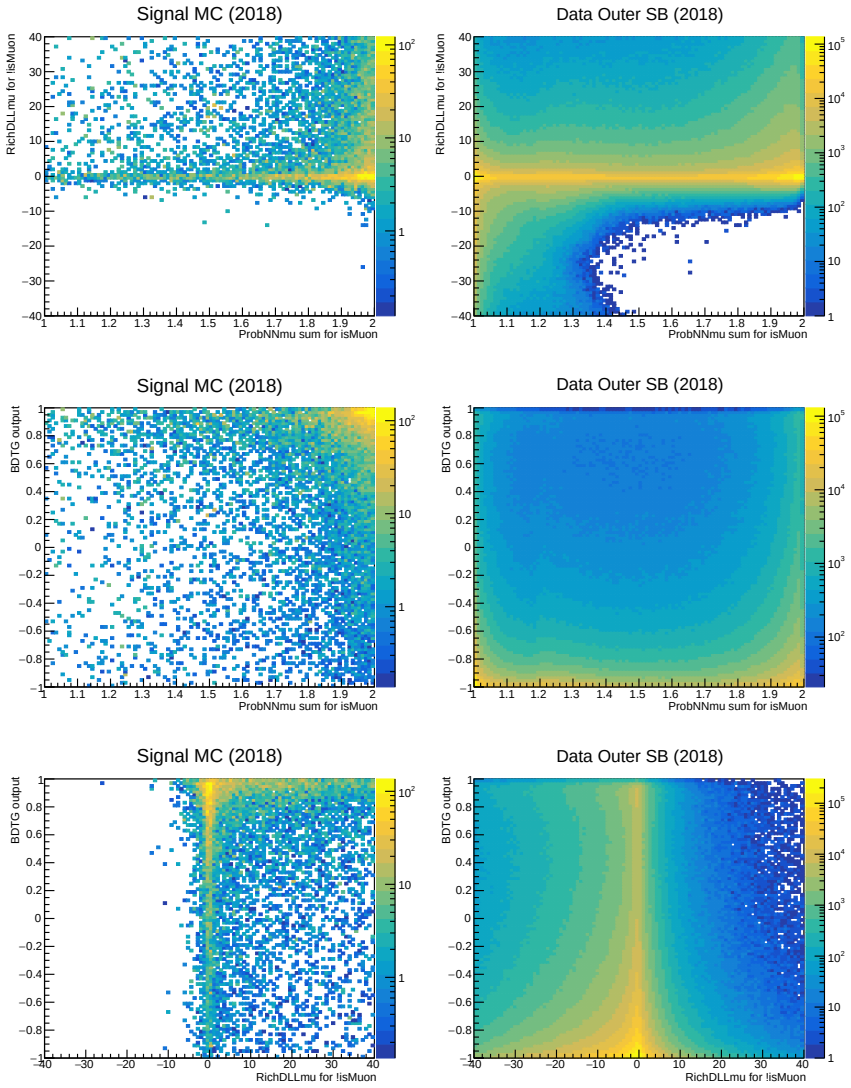


FIGURE 1.47: Correlation between the $RichDLL_\mu$ of the *bad* track, the sum over the $ProbNN_\mu$ of the *good* tracks and the BDTG output for 2018 MC and outer sidebands data.

1.7 Backgrounds

As anticipated in Sec. 1.6.1 in this section are described the main background sources in the search for the $\tau^+ \to \mu^+\mu^-\mu^+$ decay.

1.7.1 Physical background

To study of the contribution of $D_s^+ \to \eta(\mu^+\mu^-\gamma)\mu^+\nu_\mu$ decay to the number of remaining events in the signal region, about 3M of simulated events have been produced for the three years of data taking, of which $\sim 2.7\%$ of them pass the same stripping cuts used to select the $\tau^+ \to \mu^+\mu^-\mu^+$ events. A truth-matching based on the `Background Category` tool and parent particle identification number (ID) is applied on the simulated data to obtained pure sample of $D_s^+ \to \eta\mu^+\nu_\mu$. The background category 40 selects events where the mother particle is not fully reconstructed and whose mass value is more than 100 MeV/c^2 lower than the expected mass value. In

the MonteCarlo events, each generated particle has a numeric ID that characterize them uniquely. When the mother ID of the oppositely charged muons (the ID associated to the decaying particle that produced them) matched the η ID (± 221) their grandmother ID is required to match the D_s ID (± 431).

The same set of offline, IsMuon and trigger cuts described in Section 1.4 is applied to this background channel. The offline selection includes a veto on the opposite-sign dimuon invariant mass to specifically remove this background channel: $m_{\mu^+\mu^-} > 450\,\text{MeV}/c^2$. The threshold value has been optimized in the Run1 analysis. The cut preserves $\sim$ 84% of the $\tau^+ \to \mu^+\mu^-\mu^+$ signal events, while it removes $\sim$ 92% of the background events. The effect of the cuts on the invariant mass distribution of the three muons in the $D_s^+ \to \eta\mu^+\nu_\mu$ MC can be seen in Figure 1.48. The red lines in Dalitz plot of the two possible opposite-sign dimuon combinations indicates the veto.

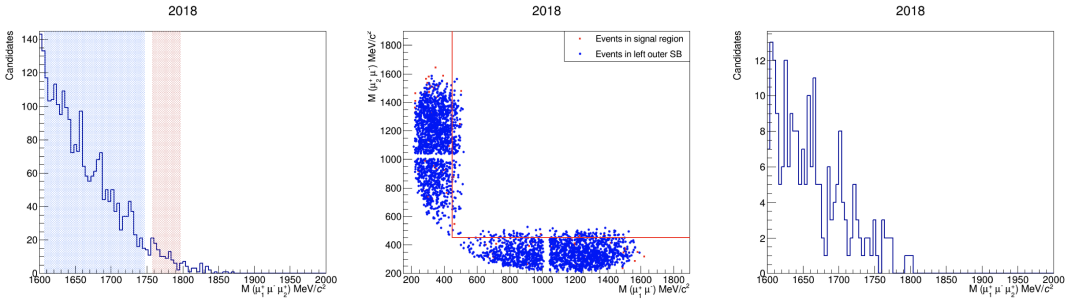


FIGURE 1.48: (Left) invariant mass distribution of the three muons in the $D_s^+ \to \eta(\mu^+\mu^-\gamma)\mu^+\nu_\mu$ MC, after the offline cuts but before the veto. (Center) Dalitz plot of the two opposite-sign dimuon pairs showing the bottom-left region excluded by the $m_{\mu^+\mu^-} > 450\,\text{MeV}/c^2$ veto. (Right) Same distribution of the events after the offline cuts, once the veto is applied.

The number of expected $D_s \to \eta\mu\nu_\mu$ in the signal region after the minimal cuts on the XGB and superProbNN of the multivariate classifiers is estimated using the $D_s \to \phi(\mu\mu)\pi$ as normalization channel: given the efficiencies and the branching fractions reported in Table 1.21 that combines the three years of data taking, the total number of expected events that fall within the signal region is 191$\pm$69.

	$D_s \to \phi(\mu\mu)\pi$	$D_s \to \eta(\mu\mu)\mu\nu_\mu$
ϵ	$(2.20 \pm 0.06) \times 10^{-3}$	$(2.46 \pm 0.61) \times 10^{-6}$
B.F.	$(1.03 \pm 0.10) \times 10^{-5}$	$(7.5 \pm 1.8) \times 10^{-3}$
N	235189$\pm$960	191$\pm$69

TABLE 1.21: Estimate of the number (N) of $D_s^+ \to \eta(\mu^+\mu^-\gamma)\mu^+\nu_\mu$ background events expected in the signal region, obtained by normalization with respect to the number of events of $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ in the signal region. It considers the different efficiency (evaluate on the simulated sample) and the different branching fraction.

1.7.2 mis-ID backgrounds

A decay with three charged hadrons in the final state (plus neutral particles) can be misidentified as signal if the tracks are wrongly identified as muons. Table 1.13

summarize the main contributions in the search for the $\tau^+ \to \mu^+\mu^-\mu^+$, with dominant contributions coming from the $D^+ \to K^-\pi^+\pi^+$ and $D^{*+} \to D^0\pi^+$ (with $D^0 \to K^-\pi^+\pi^0$). However, since the D^{*+} mass can only be partially reconstructed due to the missing π^0, this second contribution will fall outside the 3μ mass range. Replacing the muon mass hypothesis with $K\pi\pi$, $KK\pi$ or $\pi\pi\pi$ in the 3μ data sample is possible to get a qualitative estimation of this kind of backgrounds. As can be seen from Figure 1.49 only the $K\pi\pi$ mass hypothesis exhibits a large peak.

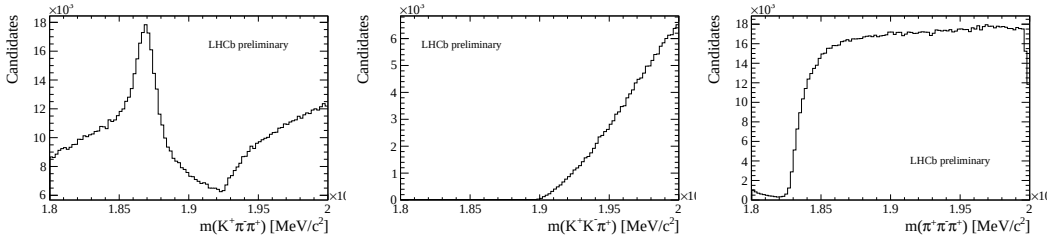


FIGURE 1.49: Invariant mass distribution in the 2018 data outer sidebands for $K\pi\pi$ (left), $KK\pi$ (center) and $\pi\pi\pi$ (right) mass hypothesis.

TABLE 1.22: Efficiency of the veto and global $ProbNN_K$ requirement evaluated on the signal and the $D^+ \to K^-\pi^+\pi^+$ simulated sample. The signal efficiencies are evaluated with respect to the number of candidates surviving the XGB requirement.

Requirement	$\tau^+ \to \mu^+\mu^-\mu^+$ [%]	$D^+ \to K^-\pi^+\pi^+$ [%]
Veto D^+	99.30 ± 0.44	4.80 ± 0.35
$ProbNN_K$ cut	95.09 ± 0.43	12.79 ± 0.87
Veto & $ProbNN_K$	94.43 ± 0.55	0.61 ± 0.06
$sum(ProbNN_\mu)$	72.89 ± 0.36	$< 0.0001 \pm 0.0007$
$min(ProbNN_\mu)$	97.80 ± 0.59	$< 0.0001 \pm 0.0007$
Total	67.36 ± 0.44	$< 0.0001 \pm 0.0007$

A simulated and data sample of $D^+ \to K^-\pi^+\pi^+$ are used to study the suppression of this charmed background mode. A veto in a window around the D^+ invariant mass from 1800 to 1900 MeV/c^2 based on $ProbNN_K$ is found to be particularly effective: requiring $ProbNN_K < 0.05$ for the daughter with opposite charge with respect to the mother reduces the $\tau^+ \to \mu^+\mu^-\mu^+$ signal efficiency by a few percent, while removes most of the $D_s \to K^-\pi^+\pi^+$ events. Other additional cuts were studied to find the optimal selection, and their efficiencies on the simulated samples are indicated in Table 1.22. The efficiencies are evaluated on the subsample of events that pass the minimal cut on XGB ($XGB > 0.8$). The global cut on $ProbNN_K$ consists in requiring $ProbNN_K < 0.4$ for all the particles in the final state; $sum(ProbNN_\mu) > 2.4$ and $min(ProbNN_\mu) > 0.5$ are found to be good choices to suppress the $D^+ \to K^-\pi^+\pi^+$ background without rejecting much of the $\tau^+ \to \mu^+\mu^-\mu^+$ signal candidates.

The SuperProbNN classifier described in Section 1.6.4 is based on the $ProbNN_\mu$ and $ProbNN_K$ of the three particle in the final state, and thus it is expected that it should be able to replicate the rejection power of the cuts in Table 1.22. Removing the candidates with a superProbNN output lower than 0.88 rejects a similar amount of background in the signal data outer sidebands, as can be seen by comparing the

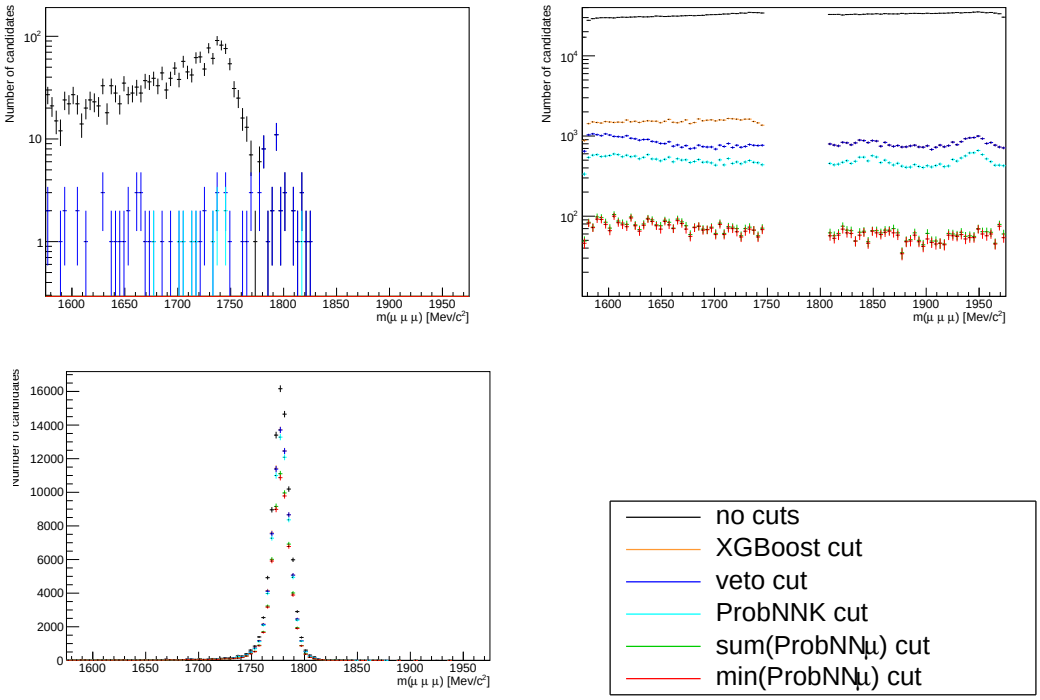


FIGURE 1.50: Invariant mass distribution under the $\mu^-\mu^+\mu^-$ hypothesis for the simulated $D^+ \to K^-\pi^+\pi^+$ (left), data (right) and simulated signal (bottom) sample. Three different requirements are applied: XGB classifier > 0.8, veto cut and global $ProbNN_K$ cut.

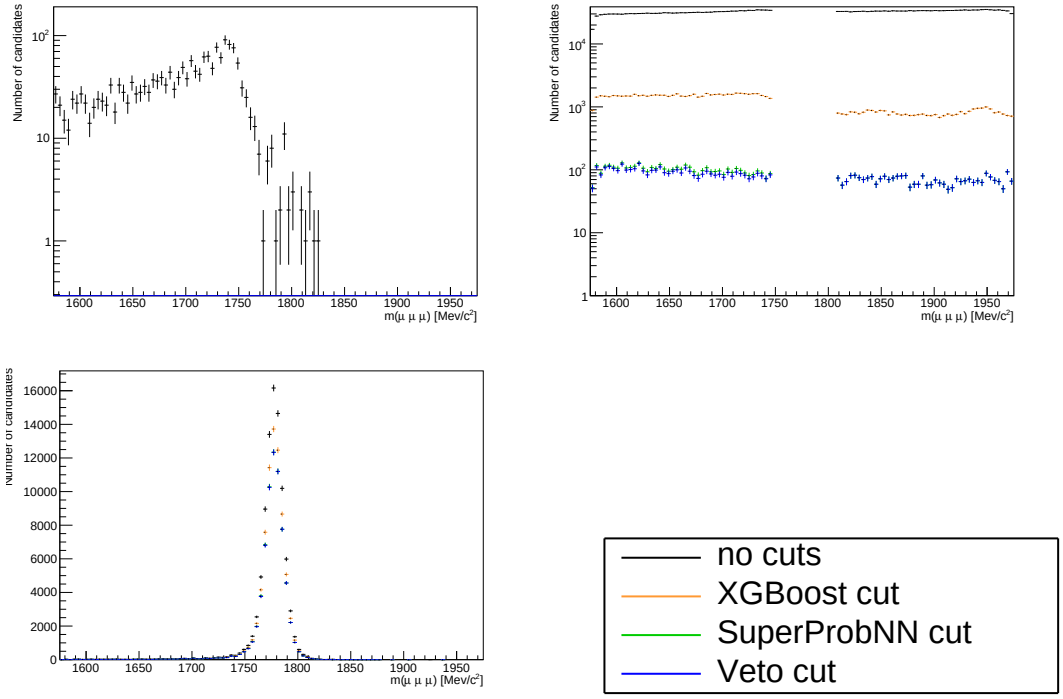


FIGURE 1.51: Invariant mass distribution under the $\mu^-\mu^+\mu^-$ hypothesis for the simulated $D^+ \to K^-\pi^+\pi^+$ (left), data (right) and simulated signal (bottom) sample. Three progressive requirements are applied: XGB classifier > 0.8, superProbNN > 0.88 and the $D^+ \to K^-\pi^+\pi^+$ veto cut.

right plots from Figure 1.50 and Figure 1.51, but results in a higher signal efficiency (it retains $\sim$ 80% of the events that pass the minimal XGB cut). The performance of both strategy on the $D^+ \to K^-\pi^+\pi^+$ MC sample are similar. For these reasons, the requirement of $SuperProbNN > 0.88$ was chosen to remove the charmed mis-ID components, with the addition of the veto described previously to remove a small residual peak in the signal data sample.

In the right data outer sidebands of Figure 1.51 are visible two small peaks due to the mis-ID contribution of $D^+ \to \pi^+\pi^+\pi^-$ and $D_s \to \pi^-\pi^+\pi^+$.

$D^+ \to K^-\pi^+\pi^+$ contamination in signal region

The contamination of $D^+ \to K^-\pi^+\pi^+$ mode in signal region [1756.8, 1796.8] MeV/c^2 is estimated extrapolating the yield from a fit to the 3μ invariant mass distribution using the two outer and inner sideband regions: [$1547 < m(3\mu) < 1757$ & $1797 < m(3\mu) < 2007$] MeV/c^2. In data sample, in addition to the selection described in Section 1.4, in order to reduce the amount of combinatorial background the optimal requirement on XGB classifier is applied, removing all the candidates with $XGB < 0.8$.

When applying the XGB requirements three peak become clearly visible, as shown in Figure 1.52: a $D^+ \to K^-\pi^+\pi^+$ peak in the left sideband and the $D^+ \to \pi^+\pi^+\pi^-$ and $D_s \to \pi^-\pi^+\pi^+$ peaks on the right sideband. The components playing a role in the fit are: the $D^+ \to K^-\pi^+\pi^+$ mode, described with a double-sided Crystal Ball function (see Equation 1.25), the $D_s \to \pi^-\pi^+\pi^+$ and $D^+ \to \pi^+\pi^+\pi^-$ decays, described by two Johnson functions (see Equation 1.16), and the combinatorial background, described as a simple exponential. The Crystal Ball function describing the $D^+ \to K^-\pi^+\pi^+$ background is defined as:

$$C(m,\mu,\sigma,\alpha_{L,R},n_{L,R}) = \begin{cases} e^{-\frac{\alpha_L^2}{2}}\left(\frac{n_L}{\alpha_L}\right)^{n_L}\left(\frac{n_L}{\alpha_L} - \alpha_L - \frac{x-\mu}{\sigma}\right)^{-n_L} & \frac{x-\mu}{\sigma} < -\alpha_L \\ \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]\sigma & \frac{x-\mu}{\sigma} < \alpha_R \\ e^{-\frac{\alpha_R^2}{2}}\left(\frac{n_R}{\alpha_R}\right)^{n_R}\left(\frac{n_R}{\alpha_R} - \alpha_R + \frac{x-\mu}{\sigma}\right)^{-n_R} & \frac{x-\mu}{\sigma} > \alpha_R \end{cases} \quad (1.25)$$

where L, R stand for the left and the right side of the function, μ is the mean, σ is the resolution and $\alpha_{L,R}$ and $n_{L,R}$ are the tail parameters.

The mass parameters that describe the tails shape are fixed during the fit on Run2 data sample to the values obtained from simulation. The central value and the width of the distributions are floating. The results of the fit on MC shown in Figure 1.53 are validated separately for the $D^+ \to K^-\pi^+\pi^+$ and for $D^+_{(s)} \to \pi^+\pi^-\pi^+$, by means of a veto on $ProbNN_K$ (described in Section 1.7.2). The fit on data for the three years are reported in Figure 1.54. The small peak around 1.8 GeV/c^2 could be related to the $D_s \to K^-\pi^+\pi^+$ mode that is expected to be about 25 times smaller than the $D^+ \to K^-\pi^+\pi^+$ mode, as reported in Table 1.13. Once that the full signal selection is applied, its residual contribution should be completely negligible. Another small increase in data distribution is observed around 1860 and 1960 Mev/c^2, probably due to $D^+ \to \pi^+\pi^+\pi^-$ and $D_s \to \pi^-\pi^+\pi^+$ candidates surviving the veto requirement. Due to their very low yields, however, the fit performance is not significantly affected even if these two components are not included in the total p.d.f..

Once all the shape parameters have been fixed, the fit on the data sample can be performed without any requirement on $ProbNN_K$. The final yield (n_X^{sig}), expected in the signal region, for the three charmed modes surviving both the XGB

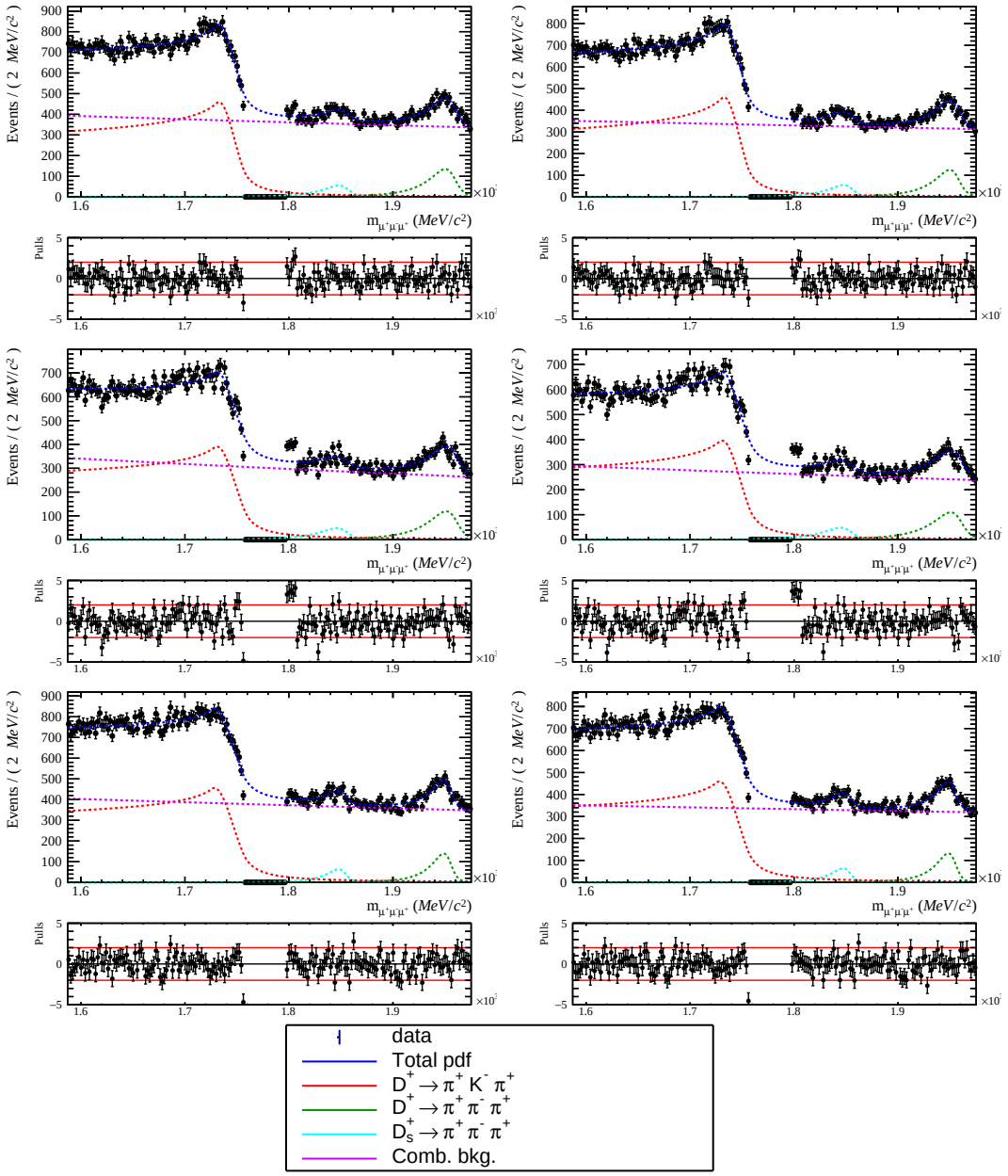


FIGURE 1.52: Data distribution under the 3μ mass hypothesis applying the *XGB* requirement only (left) and adding the reverse *SuperProbNN* requirement and the veto around the D^+ mass (right). The results of the fits are superimposed, separately for the three data taking periods: 2016 (top), 2017 (middle), 2018 (bottom). Legend: total pdf (blue), $D^+ \to K^-\pi^+\pi^+$ (red), $D^+ \to \pi^+\pi^+\pi^-$ (cyan), $D_s \to \pi^-\pi^+\pi^+$ (green), comb. bkg (violet).

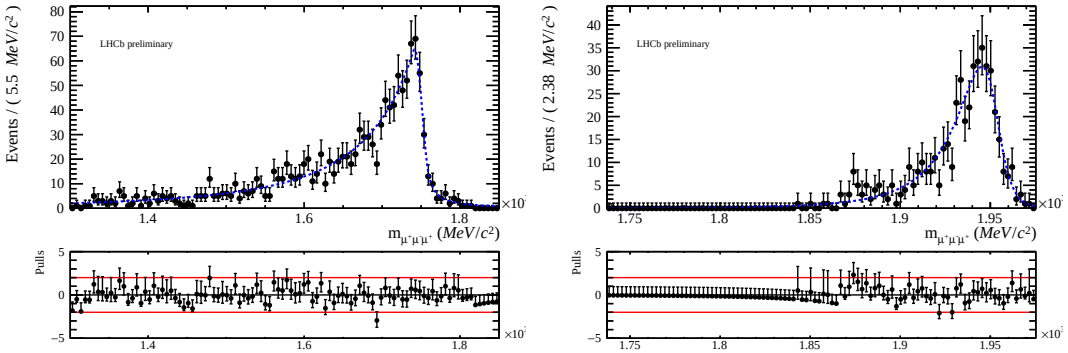


FIGURE 1.53: Distribution of the $D^+ \to K^-\pi^+\pi^+$ (left) and $D_s \to \pi^-\pi^+\pi^+$ (right) modes under the 3μ mass hypothesis in fully simulated Run2 sample, the results of the fit is superimposed.

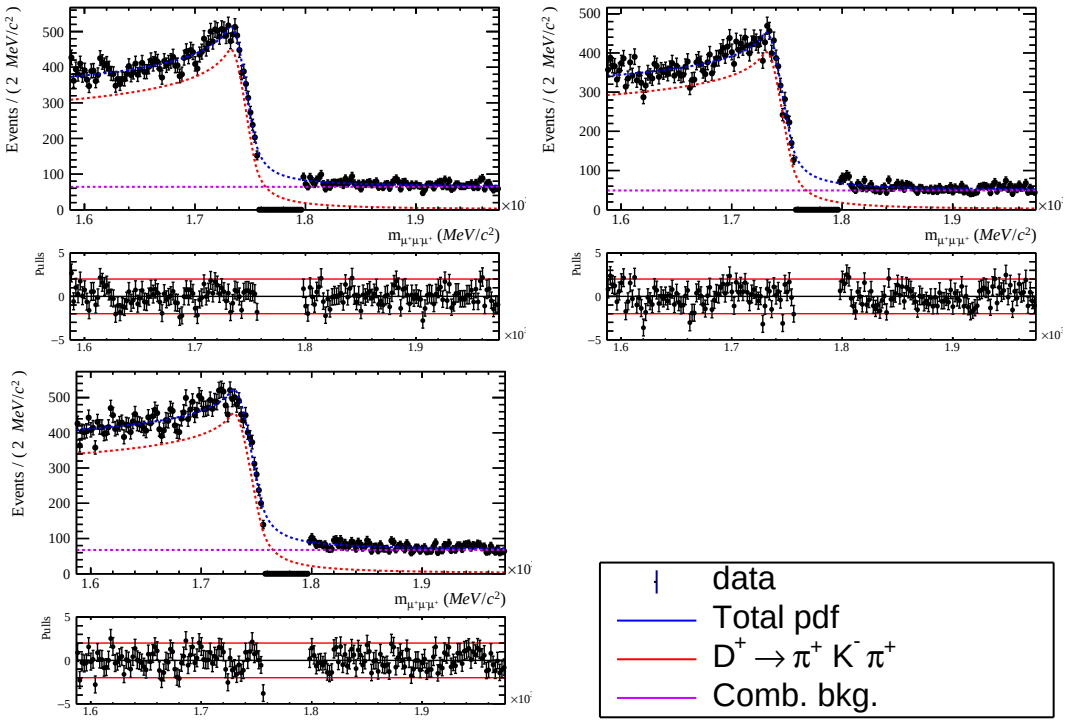


FIGURE 1.54: Data distribution under the 3μ mass hypothesis applying the *XGB* selection, the $ProbNNk > 0.05$ requirement on the opposite charge final state track and without any requirement on $SuperProbNN$; the results of the fit is superimposed. Three data taking periods are fitted independently: 2016 (top left), 2017 (top right), 2018 (bottom). Legend: total pdf (blue), $D^+ \to K^-\pi^+\pi^+$ (red), comb. bkg (violet).

and $SuperProbNN$ cut are determined as the difference between the yield obtained without the cuts and the yield obtained with the reverse of the cuts. The values are reported in Table 1.23, and their uncertainty is estimated using Equation 1.26 (which considers the correlation between the results of the fits).

$$\sigma_X^{sig} = \sqrt{(\sigma_{X,XGB}^{sig})^2 + (\sigma_{X,XGB\&PID}^{sig})^2 - 2\rho\sigma_{X,XGB}^{sig}\sigma_{X,XGB\&PID}^{sig}}\,, \tag{1.26}$$

$\sigma_{X,XGB}^{sig}$ is the statistical uncertainty related to the fit performed using only the XGB requirement, $\sigma_{X,XGB\&PID}^{sig}$ is the statistical uncertainty related to the fit performed using both the XGB and $SuperProbNN$ & veto requirements, and ρ is the correlation between the two fits results assumed to be equal to the fraction of total data candidates in common in the two fit selections: $\rho = \frac{n_{XGB\&PID}^{total}}{n_{XGB}^{total}}$.

TABLE 1.23: Number of candidates for the $D^+ \to K^-\pi^+\pi^+$, $D^+ \to \pi^+\pi^+\pi^-$, $D_s \to \pi^-\pi^+\pi^+$ and combinatorial background surviving the XGB, $SuperProbNN$ and veto requirements determined as the difference of the yields expected in the signal region. The statistical uncertainties are estimated taking into account the correlation between the two fits.

Parameter	**2016**	**2017**	**2018**
$n_{D^+\to K^-\pi^+\pi^+}$	28 ± 29	6 ± 21	7 ± 19
$n_{D^+\to \pi^+\pi^+\pi^-}$	3 ± 3	2 ± 1	1 ± 1
$n_{D_s\to \pi^-\pi^+\pi^+}$	11 ± 8	8 ± 8	2 ± 3
$n_{comb.bkg}^{sig}$	649 ± 491	719 ± 375	818 ± 497

The yields are found to be compatible with 0 within two standard deviations, so their contribution is neglected in the final limit estimation. The last row of Table 1.26 reports the number of combinatorial events expected for the three years in the signal region. Compared to them, the mis-ID backgrounds contribution is totally negligible. It is important to check that this statement holds in each bin of the binning scheme defined in Section 1.9.1. The maps containing the number of $D^+ \to K^-\pi^+\pi^+$ candidates expected in the signal region and in the outer and inner sidebands in bins of XGB and $SuperProbNN$ for the 2018 data taking years are reported in Figure 1.56. The same distributions for the combinatorial background are reported in Figure 1.57 (with a further distinction between the left and the right sidebands). The further binning of the signal region will be defined in Section 1.9.2.

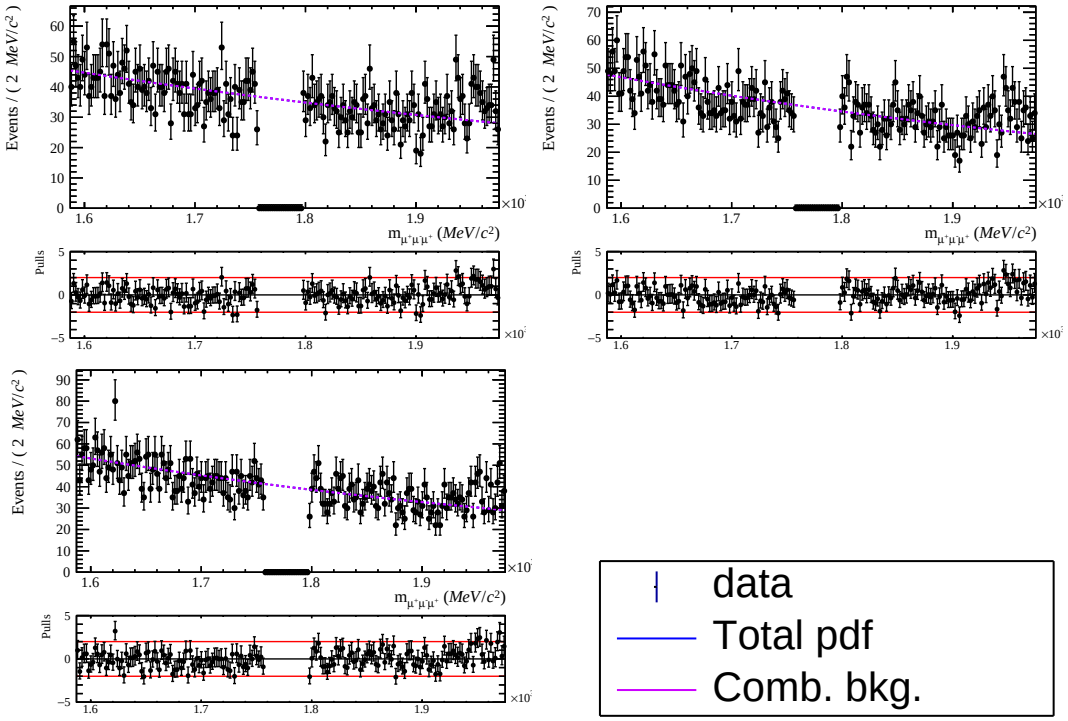


FIGURE 1.55: Data distribution under the 3μ mass hypothesis applying both the *XGB*, *SuperProbNN* and veto requirements; the results of the fit is superimposed. Three data taking periods are fitted independently: 2016 (left), 2017 (middle), 2018 (right). Legend: comb. bkg (violet).

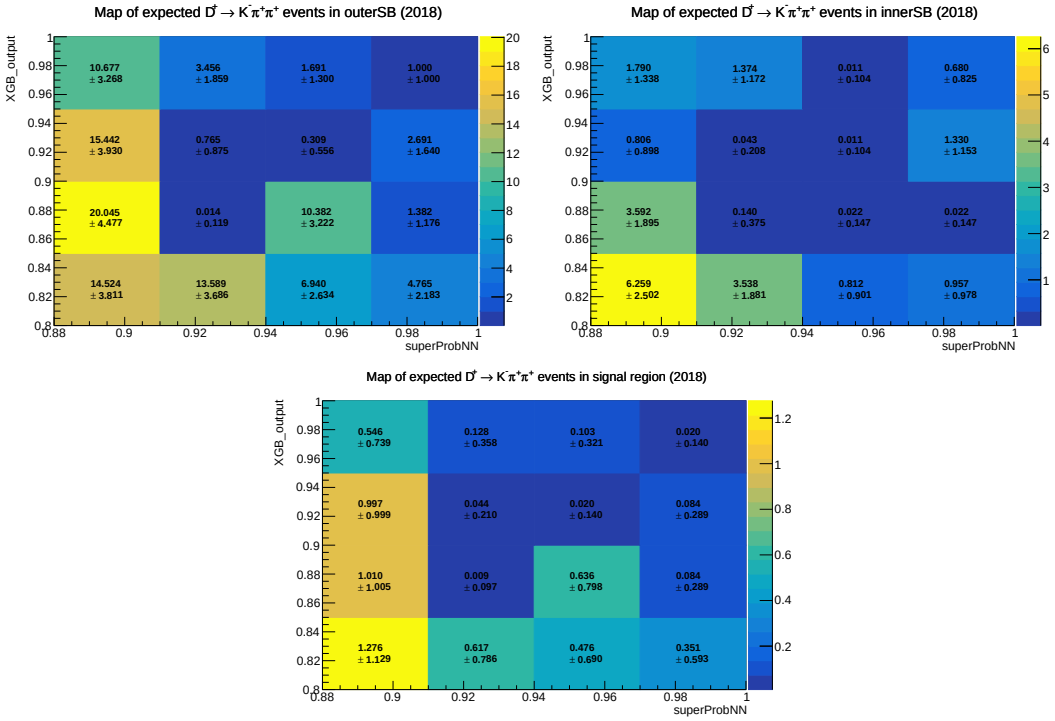


FIGURE 1.56: Maps containing the number of $D^+ \to K^-\pi^+\pi^+$ candidates expected in the outer (top left) and inner (top right) sidebands and in the signal region (bottom) in bins of *XGB* and *SuperProbNN* for the 2018 data taking year.

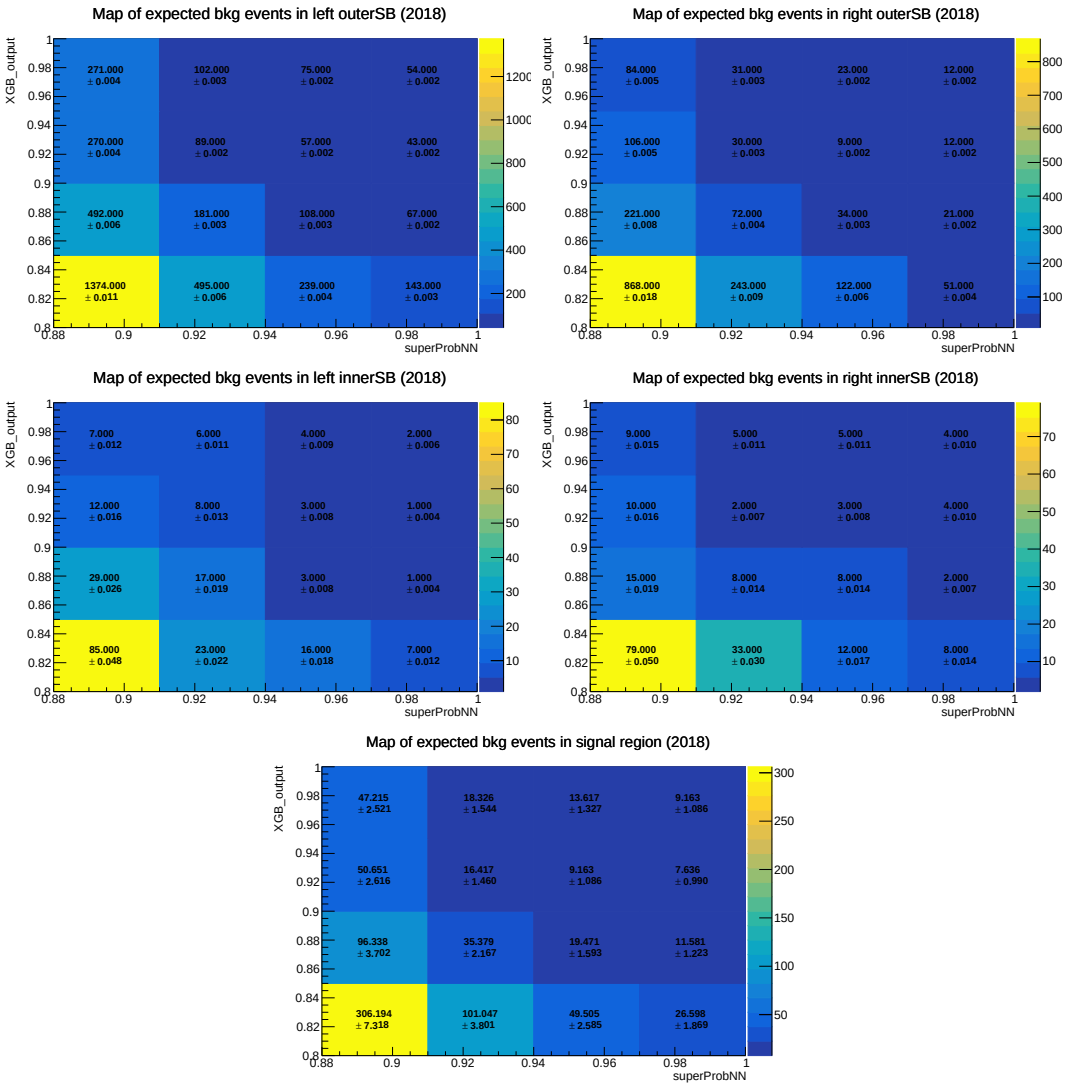


FIGURE 1.57: Maps containing the number of combinatorial candidates expected in the signal region and in the outer and inner sidebands in bins of XGB and $SuperProbNN$ for the 2018 data taking year. The maps are related to the candidates populating the left (top left) and right (top right) outer sidebands, the left (middle left) and right (middle right) inner sidebands and the signal region (bottom).

1.8 Normalization

In this section are described the factors that enter the Equation 1.12 of the branching fraction (reported below for ease of reading), the efficiencies, and their corrections:

$$\begin{aligned}\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+) &= \mathcal{B}(D_s^+ \to \phi(\mu^+\mu^-)\pi^+) \times \frac{f_{D_s}^{\tau}}{\mathcal{B}(D_s^+ \to \tau^+\bar{\nu}_\tau)} \times \frac{\epsilon_{D_s}}{\epsilon_\tau} \times \frac{N_{sig}}{N_{cal}} \\ &= \alpha \times N_{sig}\end{aligned}$$

The branching fraction of the reference channel $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ is computed from the $D_s^+ \to \phi(\to K^+K^-)\pi^+$ without the non-resonant part of the D_s^+ since its contribution in data is negligible ($\mathcal{B}(D_s^+ \to \mu^+\mu^-\pi^+) < 1.8 \times 10^{-7}$, [62]). The values of the branching fractions are listed in Table 1.24.

Channel	$\mathcal{B}$
$D_s^+ \to \phi(\to K^+K^-)\pi^+$	$(2.24 \pm 0.08) \times 10^{-2}$
$\phi \to K^+K^-$	$(49.2 \pm 0.5) \times 10^{-2}$
$\phi \to \mu^+\mu^-$	$(2.86 \pm 0.19) \times 10^{-4}$
$D_s^+ \to \tau^+\bar{\nu}_\tau$	$(5.48 \pm 0.23) \times 10^{-2}$
$D_s^+ \to \phi(\mu^+\mu^-)\pi^+$	$(1.302 \pm 0.099) \times 10^{-5}$

TABLE 1.24: Branching ratios used in the evaluation of $\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+)$.

1.8.1 Efficiencies

The factor $\epsilon_{D_s}/\epsilon_\tau$ in Equation 1.12 represents the ratio between the overall efficiency evaluated on the reference channel and the one evaluated on the signal channel. Expanding the terms at each step of the analysis it results in:

$$\frac{\epsilon_{D_s}}{\epsilon_\tau} = \frac{\epsilon_{mix,D_s}^{GEN} \times \epsilon_{D_s}^{RR} \times \epsilon_{D_s}^{REC} \times \epsilon_{D_s}^{SEL} \times \epsilon_{D_s}^{TRIG}}{\epsilon_{mix,\tau}^{GEN} \times \epsilon_\tau^{RR} \times \epsilon_\tau^{REC} \times \epsilon_\tau^{SEL} \times \epsilon_\tau^{TRIG}} \cdot \frac{1}{\epsilon_\tau^{Clf\&PID}} \tag{1.27}$$

The first part of the overall efficiency contains, for both the reference and signal channel, the acceptance or generation efficiency ϵ_{mix}^{GEN}, the efficiency of the cuts applied at the filtering level or retention rate ϵ^{RR}, the reconstruction efficiency ϵ^{REC} which also includes the stripping selection efficiency, the offline selection efficiency ϵ^{SEL} and the trigger selection efficiency ϵ^{TRIG}. The offline selection efficiency also includes the isMuon selection in the signal channel. In the signal channel there is a further selection based on multivariate classifiers and on PID (described in Section 1.6), for both the 2μ and the 3μ sample, and therefore an additional term $\epsilon^{Clf\&PID}$ is included in the determination of the overall efficiency. The values of these efficiencies are listed in Table 1.25, separately for each decay channel and year of data taking.

1.8.2 Efficiencies corrections

The efficiencies defined in the previous section are evaluated on simulated data of signal $\tau^+ \to \mu^+\mu^-\mu^+$ and reference $D_s \to \phi(\mu\mu)\pi$ channels, so their ratio $\epsilon_{D_s}/\epsilon_\tau$ need to be corrected for Data/MC differences. The corrections of PID efficiencies have already been calculated in Section 1.6.3.

Efficiency	Channel	2016 (%)	2017 (%)	2018 (%)
ϵ^{GEN}_{mix}	$\tau \to \mu\mu\mu$		10.72 ± 0.43	
	$D_s \to \phi(\mu\mu)\pi$		12.04 ± 0.33	
ϵ^{RR}	$\tau \to \mu\mu\mu$	57.17 ± 0.04	57.09 ± 0.04	57.04 ± 0.04
	$D_s \to \phi(\mu\mu)\pi$	47.53 ± 0.03	47.62 ± 0.03	47.50 ± 0.03
ϵ^{REC}	$\tau \to \mu\mu\mu$	39.62 ± 0.05	39.54 ± 0.05	39.66 ± 0.05
	$D_s \to \phi(\mu\mu)\pi$	16.94 ± 0.04	17.05 ± 0.03	17.09 ± 0.03
ϵ^{SEL}	$\tau \to \mu\mu\mu$ (3μ)	53.60 ± 0.09	53.67 ± 0.09	53.61 ± 0.09
	$\tau \to \mu\mu\mu$ (2μ)	20.50 ± 0.08	20.50 ± 0.08	20.47 ± 0.08
	$D_s \to \phi(\mu\mu)\pi$	91.12 ± 0.07	91.18 ± 0.07	91.11 ± 0.07
ϵ^{TRIG}	$\tau \to \mu\mu\mu$ (3μ)	58.40 ± 0.12	66.78 ± 0.11	57.67 ± 0.12
	$\tau \to \mu\mu\mu$ (2μ)	08.17 ± 0.11	22.40 ± 0.16	19.05 ± 0.15
	$D_s \to \phi(\mu\mu)\pi$	14.79 ± 0.09	31.76 ± 0.11	27.59 ± 0.12
$\epsilon^{Clf\&PID}$	$\tau \to \mu\mu\mu$ (3μ)	60.72 ± 0.14	66.45 ± 0.14	63.18 ± 0.15
	$\tau \to \mu\mu\mu$ (2μ)	3.41 ± 0.18	7.50 ± 0.36	9.20 ± 0.16
ϵ^{tot}	$\tau \to \mu\mu\mu$ (3μ)	0.46 ± 0.02	0.58 ± 0.02	0.47 ± 0.02
	$\tau \to \mu\mu\mu$ (2μ)	0.0014 ± 0.0001	0.0083 ± 0.0005	0.0087 ± 0.0004
	$D_s \to \phi(\mu\mu)\pi$	0.131 ± 0.004	0.283 ± 0.008	0.246 ± 0.007

TABLE 1.25: Summary table of the efficiencies that enter the evaluation of the $\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+)$, for both signal and reference channel. When a single value is given, it is intended to be equal for the three years of data taking. The last row contains the overall efficiencies.

Correction to $\phi(1020)$ mass cut efficiency Among the offline cuts listed in Table 1.5 there is one regarding the mass of the intermediate resonance $\phi(1020)$: The signal candidates' pairs of opposite-sign dimuon are required to satisfy $|m_{\mu^+\mu^-} - m_\phi| > 20\,\text{MeV}/c^2$, while the requirement is inverted for the reference channel ($|m_{\mu^+\mu^-} - m_\phi| < 20\,\text{MeV}/c^2$). In the MC simulation, the $\phi(1020)$ is generated for masses within $\pm 15\Gamma$ window [955.7, 1083.2] MeV/c^2, where Γ is the decay width of $\phi(1020)$. Since there is no such truncation in the data, a Data/MC correcting factor c_ϕ need to be evaluated. To measure the portion of excluded events in the tails outside the $\phi(1020)$ peak, its distribution is approximated with a non-relativistic Breit-Wigner (BW) pdf, shown in Figure 1.58. The pdf is defined in the phase space available in the $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ decay, from $2m_\mu$ up to $m_{D_s^+} - m_\pi$. In the truncated MC, the fraction of excluded area outside the $\pm 20\,\text{MeV}/c^2$ window (depicted in green) corresponds to 4.7% of the total area (green and orange area). In the data, the same area corresponds to 6.6% of the whole pdf in the available phase space. Thus, the $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ selection efficiency, as determined on MC, would need to be corrected by 1.066/1.047, i.e. a factor of 1.018 to which a conservative systematic uncertainty of 50% of its value is assigned. This lead to an increment of the $\phi(1020)$ mass cut efficiency of (+1.8 $\pm$0.9)%.

Track reconstruction efficiency correction The track reconstruction efficiency at LHCb is estimated with a tag-and-probe approach [73]. This method uses the two-particle decay $J/\psi \to \mu^+\mu^-$, where one of the decay product is reconstructed as a long track (the "tag"), while the other is only partially reconstructed (the "probe"). The track reconstruction efficiency for long tracks is obtained as the fraction of probe tracks matched to a long track. The $J/\psi \to \mu^+\mu^-$ decay channel is chosen because the

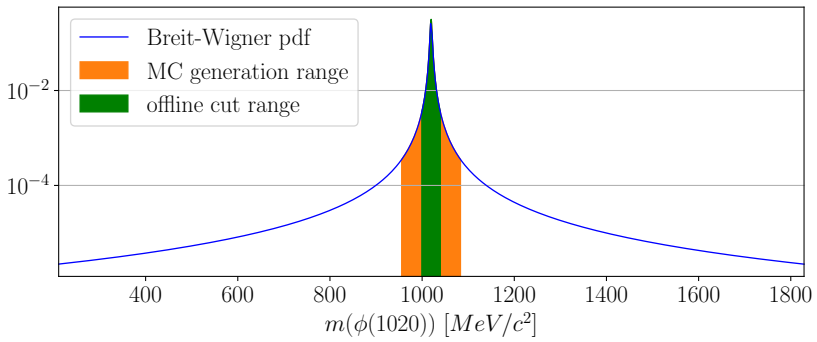


FIGURE 1.58: Non-relativistic Breit-Wigner mass distribution of $\phi(1020)$. The green area covers the offline cut range, while the orange area shows the portion excluded by the cut used during the generation of MC sample. The axis limit corresponds to the feasible range in Data, from $2m_\mu$ up to $m_{D_s^+} - m_\pi$.

daughter particles have PID information from the muon system that can be exploited in the reconstruction of the probe track.

The track reconstruction efficiency depends reasonably on the kinematics, and thus the corrections are evaluated in bins of track momentum ($5 < p < 200$) GeV/c and pseudo-rapidity ($1.9 < \eta < 4.9$) from maps generated from the TrackCalib tool, separately for different years. The track reconstruction correction is computed for both the signal and the reference channel, and their ratio (R_{Track}) is used to compensate Data/MC differences. Table 1.26 lists their calculated values for the three years of data taking.

Channel	**2016**	**2017**	**2018**
C^{τ}_{Track}	$0.9579 \pm 0.0106 \pm 0.0139$	$0.9974 \pm 0.0080 \pm 0.0069$	$0.9761 \pm 0.0074 \pm 0.0069$
$C^{D_s}_{Track}$	$0.9618 \pm 0.0114 \pm 0.0197$	$0.9972 \pm 0.0078 \pm 0.0156$	$0.9776 \pm 0.0074 \pm 0.0156$
R_{Track}	$1.0041 \pm 0.0051 \pm 0.0141$	$0.9999 \pm 0.0047 \pm 0.0140$	$1.0015 \pm 0.0050 \pm 0.0140$

TABLE 1.26: Correction to the track reconstruction efficiency for the signal and the reference channel for the different data taking years, and the resultant Data/MC correcting factor R_{Track}. The first is the statistical uncertainty while the second is the systematic uncertainty.

Most of the signal events are characterized by a momentum lower than 5 GeV/c, as can be seen from Figure 1.59. For those events in the region from 3 to 5 GeV/c it has been decided to assign the correction of the closest bin in the efficiency maps. The correction to the track reconstruction efficiency for a given channel is calculated, as in [74], with the formula:

$$C^{ch} = \prod_{tr} C^{ch}_{tr} = \prod_{tr} \left\{ \sum_i (f_i^{MC,tr} C_i) \beta_{RW} \beta^{tr}_{mat} \right\} \quad \text{where:} \begin{cases} tr = \mu_1^+, \mu^-, \mu_2^+ & \text{if ch}=\tau^+ \\ tr = \mu^+, \mu^-, \pi^+ & \text{if ch}=D_s^+ \end{cases} \tag{1.28}$$

where $f_i^{MC,tr}$ is the fraction of reconstructed and selected events in bin i of the distribution of track tr, C_i is the content of the correction maps in bin i. The three correction factors C^{ch}_{tr}, one for each particle in the final state, are multiplied together to obtain the overall correction factor: $C^{D_s} = C^{D_s}_{\mu^+} \times C^{D_s}_{\mu^-} \times C^{D_s}_{\pi^+}$, and $C^{\tau} = C^{\tau}_{\mu_1^+} \times C^{\tau}_{\mu^-} \times C^{\tau}_{\mu_2^+}$

for the signal and reference channel respectively. The factor β_{RW} provides an uncertainty per track on the correction resulting from different data/simulation reweighting techniques, while the factor β^{tr}_{mat} takes into account the increase of the tracking uncertainty for kaon and pion tracks compared to muons due to the hadronic interactions of the former with the detector material, respectively. The values for 2018 are shown below (all results to be compatible with unity for 2017 and 2016 as well).

$$\beta_{RW} = 1 \pm 0.004$$
$$\beta^{K}_{mat} = 1 \pm 0.011$$
$$\beta^{\pi}_{mat} = 1 \pm 0.014$$

The uncertainties evaluated on the correction factors in Table 1.26 also include the MC statistics and the outliers' uncertainty, in addition to the factor β_{RW} and the hadron uncertainty. The outliers' uncertainty is error introduced by the percentage of events whose track momentum and pseudo-rapidity do not fall within the ranges where the correction maps are defined, and is evaluated by the change in C^{ch}_{tr} when those events are not corrected ($C_i = 1$).

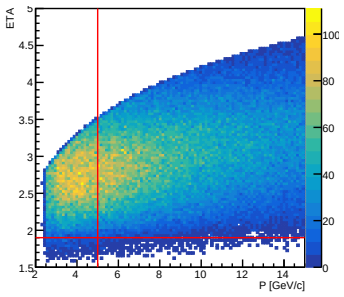


FIGURE 1.59: Pseudo-rapidity versus momentum distribution for a track in the final state for 2018 signal MC. The red lines indicate the lower limits of the tracking efficiency maps in terms of pseudo-rapidity and momentum.

Trigger efficiency correction The trigger efficiency ratio that enters the calculation of the $\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+)$ is estimated using simulated samples, thus needs to be corrected for Data/MC differences. Possible differences are expected only in L0 and Hlt1 triggers, since HLT2 is a fully software trigger, running on offline reconstructed variables, producing very similar results in data and simulation. In this section, the term $\tilde{\epsilon}$ indicates the trigger efficiency computed up to L0&HLT1 with respect to the selection cuts, while ϵ includes also Hlt2 selection. The ratio of trigger efficiencies between the reference and the signal channels evaluated on MC is corrected by a double ratio factor $R_{D_s}(D_s, \tau)$ to account for Data/MC differences

$$\frac{\epsilon^{D_s}_{\text{Data}}}{\epsilon^{\tau}_{\text{Data}}} = \frac{\epsilon^{D_s}_{\text{MC,TOS}}}{\epsilon^{\tau}_{\text{MC,TOS}}} \cdot R(D_s, \tau) \tag{1.29}$$

The double ratio is defined as

$$R(D_s,\tau) = \frac{\frac{\tilde{\epsilon}^{D_s}_{\text{Data,TISTOS}}}{\tilde{\epsilon}^{D_s}_{\text{MC,TISTOS}}}}{\frac{\tilde{\epsilon}^{\tau}_{\text{Data,TISTOS}}}{\tilde{\epsilon}^{\tau}_{\text{MC,TISTOS}}}} = \frac{\frac{\tilde{\epsilon}^{D_s}_{\text{Data,TISTOS}}}{\tilde{\epsilon}^{D_s}_{\text{MC,TISTOS}}}}{\frac{\sum_i f_i^{\tau}\tilde{\epsilon}^{D_s,i}_{\text{Data,TISTOS}}}{\sum_i f_i^{\tau}\tilde{\epsilon}^{D_s,i}_{\text{MC,TISTOS}}}} \tag{1.30}$$

Where $\tilde{\epsilon}^{D_s}_{\text{Data,TISTOS}}$ and $\tilde{\epsilon}^{D_s}_{\text{MC,TISTOS}}$ are the trigger efficiency of L0&Hlt1 obtained using the TISTOS method. It is implemented differently in real data and MC due to the different type of weight associated to the events. In the first case the sWeights (W_{sig-sw}) described in Sec. 1.5.1 are used to extract statistically the signal-only component in data, while the $w_{fraction}$ (described in Sec. 1.3.1) weights correctly the prompt and secondary events in the simulated sample. The TISTOS trigger efficiency on preselected Data ($\epsilon_{\text{Data,TISTOS}}$) is than defined as:

$$\epsilon_{\text{Data,TISTOS}} = \frac{\sum w^{\text{TIS\&TOS}\,|\,\text{SEL}}_{\text{sig\_sw}}}{\sum w^{\text{TIS}\,|\,\text{Sel}}_{\text{sig\_sw}}} \tag{1.31}$$

where TIS | Sel represents the preselected events that triggered the selection of TIS lines in Eq.1.22, while TIS&TOS | Sel represents the preselected events that also triggered the TOS lines from Table 1.9.

The MC trigger efficiencies are defined as:

$$\epsilon_{MC,TOS} = \frac{\sum w^{fraction}_{TOS|Sel}}{\sum w^{fraction}_{Sel}} \quad \epsilon_{MC,TISTOS} = \frac{\sum w^{fraction}_{TOS\&TIS|Sel}}{\sum w^{fraction}_{TIS|Sel}} \tag{1.32}$$

Where *Sel* represents the events that pass the offline cuts listed in Table 1.5, and $TIS|Sel$, $TOS|Sel$ and $TIS\&TOS|Sel$ represents the events that satisfy also the corresponding trigger requirements.

The TISTOS efficiencies in the signal channel are obtained by the corresponding one in the reference channel by convolving them with the phase space distribution of the signal events: The convolution of the MC signal sample with these efficiencies is done in bins of transverse momentum (p_{T}) and χ^2 of impact parameter (IPχ^2), and f_i^{τ} correspond to fraction of signal events within each bin. The values of p_{T} and χ^2 of impact parameter associated to each event come from the muon with greatest p_{T}. In the 2μ signal sample, only the good reconstructed muon tracks are compared to determine the one with greatest p_{T}.

The binning scheme consists of uniform bins that cover the range 10^2 to 10^5 MeV/c of p_{T} and from 1 to 10^5 of IPχ^2. In the computation of the bins edges, the MC sample of signal channel (2μ and 3μ samples) and the MC sample of reference channel are combined to cover properly the whole binned space. The number of bins in each dimension is optimized to satisfy the TISTOS assumptions, and the default number of bins chosen (7,5) is a compromise between enough statistics per bin and $\rho \sim 1$ (even though some correlation is still present). Figure 1.60 shows the ρ ratio between the $\tilde{\epsilon}_{MC,\ TISTOS}$ and $\tilde{\epsilon}_{MC,\ true}$ efficiencies for the signal and reference channel computed as integral over the bins: $\tilde{\epsilon}^x = \sum_i f_i^x \tilde{\epsilon}^{x,i}$. The percentage distributions of the events that passed the offline cuts in the binned maps are shown in Figure 1.61. Figure 1.62 shows the ϵ_{TISTOS} up to L0&Hlt1 computed in each bin of p_{T} and IPχ^2 for the signal channel $\tau^+ \to \mu^+\mu^-\mu^+$ (2μ and 3μ samples) and for the reference channel $D_s \to \phi(\mu\mu)\pi$ (MC and sweighted data).

The Table 1.27 shows the Data/MC correction factors defined at Eq. 1.30, computed for the two signal samples (3μ and 2μ). The systematic uncertainty on the

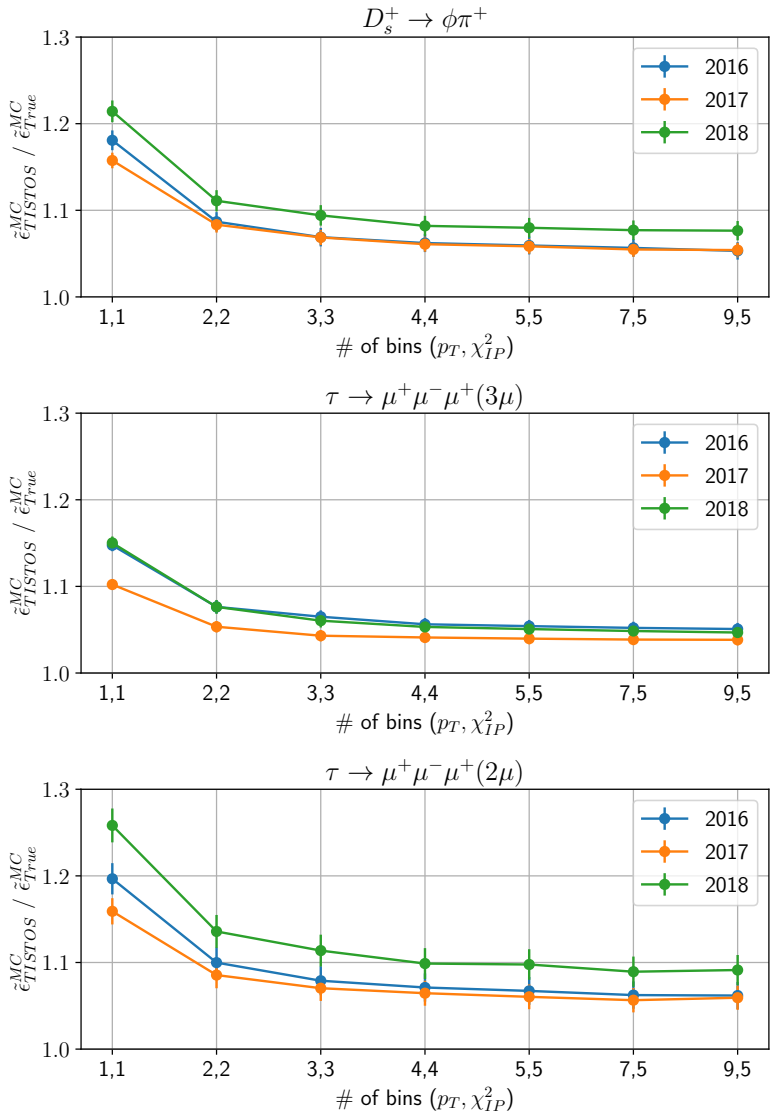


FIGURE 1.60: TISTOS efficiency over true trigger efficiency computed in the simulated sample up to L0&HLt1, as function of the number of uniform bins of p_T and IPχ^2 used.

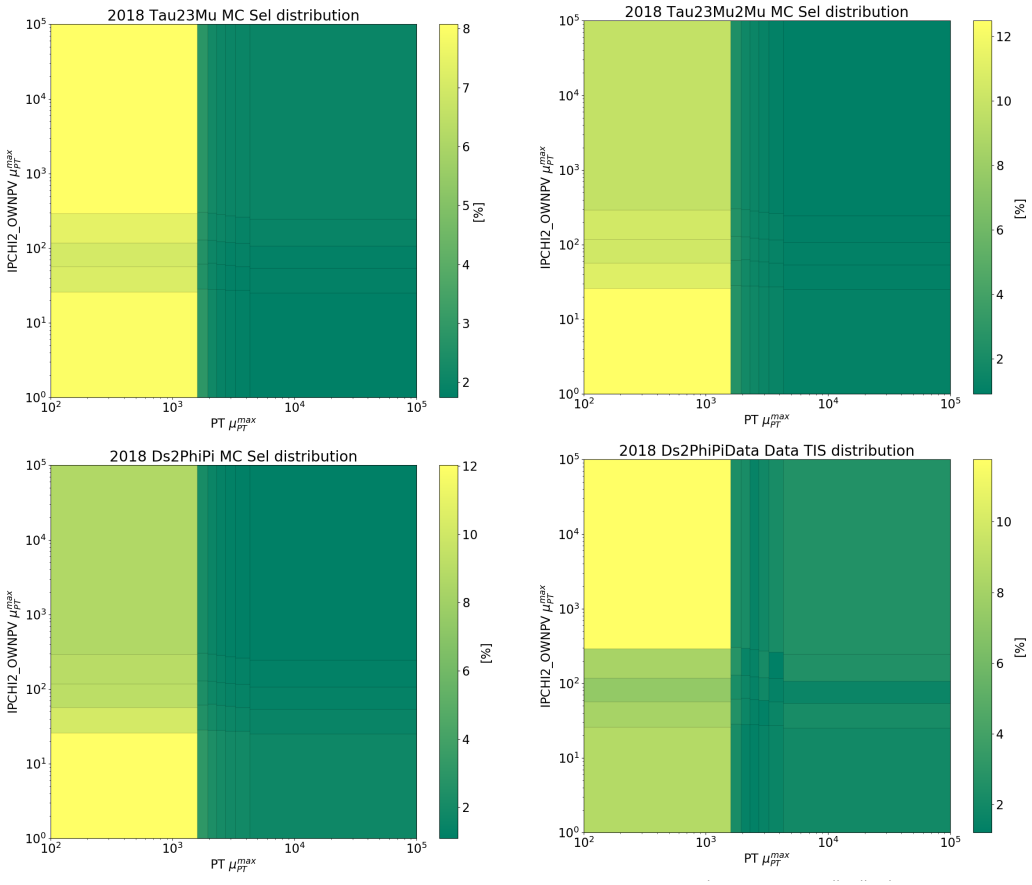


FIGURE 1.61: (Top) Distributions of the signal MC (2μ and 3μ samples) and (Bottom) distributions of reference channel (data and MC), with the default binning (7,5) in p_{T} and IPχ^2 for the events that passed the selection and offline cuts. The content of each bin enter as f_i^x in the evaluation of the integrated efficiencies.

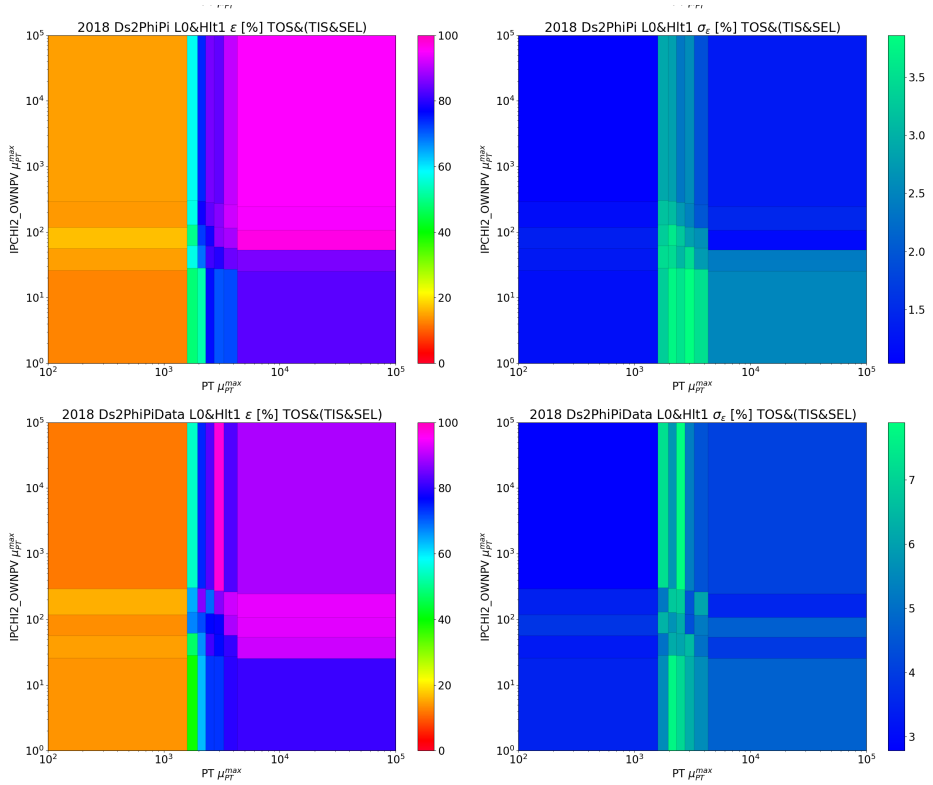


FIGURE 1.62: (Left) Efficiencies up to L0&HLT1 for 2018, computed with TISTOS method in the reference channel $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ (top: simulated data, bottom: real data), with the default binning (7,5) in p_T and IPχ^2. (Right) Binned maps with the corresponding error on the efficiencies.

Data/MC factor $R(D_s, \tau)$ that comes from the choice of the number of p_T and IPχ^2 bins is the maximum difference observed, varying the number of bins by one. All the efficiencies correction's factors are summarized in Table 1.28.

	2016	**2017**	**2018**
$R(D_s, \tau_{3\mu})$	$1.0372 \pm 0.0432 \pm 0.0032$	$0.9644 \pm 0.0300 \pm 0.0079$	$0.9755 \pm 0.0321 \pm 0.0075$
$R(D_s, \tau_{2\mu})$	$0.9934 \pm 0.0477 \pm 0.0060$	$0.9615 \pm 0.0332 \pm 0.0122$	$0.9787 \pm 0.0360 \pm 0.0097$

TABLE 1.27: Data/MC correction factor (calculated considering L0&Hlt1 efficiencies) for the trigger efficiencies ratio, for different years and for both the 2μ and the 3μ signal sample. The first one is the statistical uncertainty while the second one is the systematic coming from the choice of a binning scheme.

	2016	**2017**	**2018**
C_ϕ		1.018 ± 0.009	
R_{Track}	1.0041 ± 0.0149	0.9999 ± 0.0148	1.0015 ± 0.0149
$R(D_s, \tau_{3\mu})$	1.0372 ± 0.0433	0.9644 ± 0.0310	0.9755 ± 0.0330
$R(D_s, \tau_{2\mu})$	0.9934 ± 0.0481	0.9615 ± 0.0354	0.9787 ± 0.0373

TABLE 1.28: Summary of the efficiencies correction's factors. The reported uncertainties combine statistical and systematics uncertainties.

1.9 Limit Setting

The expected limit on the $\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+)$ is calculated with the CLs method [57]. The goal of a search is to either exclude the existence of a signal in its absence or to confirm the existence of a true signal while the probabilities of falsely excluding a true signal or falsely discovering a non-existent signal are below a certain threshold. In a counting experiment, this can be done by defining a *test-statistics* λ that depends on an observable of the experiment (i.e. the number of events that satisfy a certain selection criteria). Since a certain level of background can always be expected, the two hypothesis that are being compared are a Signal+Background and a Background-only hypotheses. The *confidence* on these hypotheses a given by the probability that the test statistics is equal or lower than the observed λ_{obs} under each hypothesis:

$$CL_{s+b} = P_{s+b}(\lambda < \lambda_{obs}) \qquad CL_b = P_b(\lambda < \lambda_{obs}) \tag{1.33}$$

The CLs can be thus defined as the normalization of the confidence of the Signal+Background hypothesis over the Background-only hypothesis:

$$CL_s = \frac{CL_{s+b}}{CL_b} \tag{1.34}$$

The signal hypothesis will be considered rejected at a confidence level CL (usually 90% or 95%) when $1 - CL_s \leq CL$.

1.9.1 Binning definition

In Section 1.6 were discussed the multivariate classifiers which outputs are used to define a signal region within which count the number of observed events to calculate the limit. This region can be binned, allowing for lower cut on the discriminant variable (higher background level) to recover potential signal candidates, and the limit being calculated in each bin. The binning scheme is determined following a technique described in [75] based on the optimization criteria

$$\Delta LQ = 2\ln Q_{\mathrm{S+B}} - 2\ln Q_{\mathrm{B}}, \tag{1.35}$$

where Q is a test-statistics defined as the product of the ratio of probability of observing events under the Signal+Background hypothesis and the probability of observing the same events under the Background-only hypothesis. The authors demonstrate that the separation power ΔLQ can be described by the media of the $Q_{S+B(B)}$ distributions, that can be approximated as:

$$\tilde{Q}^{\mathrm{med}}_{\mathrm{SB}} = \prod \frac{\mathcal{P}(s_i + b_i; s_i + b_i)}{\mathcal{P}(s_i + b_i; b_i)} \qquad \tilde{Q}^{\mathrm{med}}_{\mathrm{B}} = \prod \frac{\mathcal{P}(b_i; s_i + b_i)}{\mathcal{P}(b_i; b_i)}. \tag{1.36}$$

with $\mathcal{P}(x, \mu)$ being a Poisson distribution of mean μ calculated in x, and $s_i + b_i$ (b_i) the expected number of candidates in the Signal+Background (Background) hypothesis. The optimal binning is determined by maximizing the ΔLQ quantity for the expected number of candidates in each bin for the two hypotheses. The signal region defined by the output of the two multivariate classifiers for the 3μ signal sample, ($0.8 < XGB < 1.0$ and $0.88 < superProbNN < 1.0$), is divided into a variable number of bins (spanning from 2x2 to 7x7) with equal amount of expected signal candidates. The binning scheme chosen (4x4) is reported in Table 1.29, and the score and signal sensitivity in each bin are shown in Figure 1.63.

TABLE 1.29: Binning scheme that optimises the signal event sensitivity to $\tau^+ \to \mu^+\mu^-\mu^+$ decays.

SuperProbNN	XGB
[0.800, 0.949]	[0.880, 0.973, 0.988, 0.994, 1.000]
[0.949, 0.983]	[0.880, 0.975, 0.989, 0.995, 1.000]
[0.983, 0.993]	[0.880, 0.979, 0.991, 0.995, 1.000]
[0.993, 1.000]	[0.880, 0.980, 0.990, 0.995, 1.000]

Since the 2μ sample is statistically limited compared to the 3μ sample, as previously described in Section 1.6.4 its contribution to the limit is evaluated in a single bin.

1.9.2 Signal invariant mass distribution

In addition to XGB and superProbNN, the invariant mass of the signal candidate is used to define a further binning scheme. The signal window of $\pm 20\,\mathrm{MeV}/c^2$ is divided into 8 bins of $5\,\mathrm{MeV}/c^2$ and for each of them is evaluated the fraction of events which falls within. Since the classifiers' output do not correlate with the invariant mass, the fractions are not evaluated in each XGB and superProbNN bins.

As mentioned in Sec. 1.5.1 a Johnson's SU function is used to model the invariant mass spectra of the reference channel. The same function can describe the signal channel; As shown in Fig. 1.64 both distributions have similar resolutions and shape.

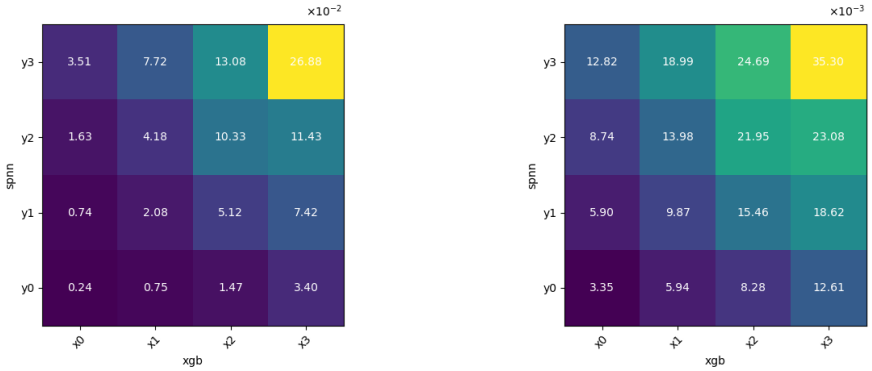


FIGURE 1.63: Individual bin score (left) and signal sensitivity (right) for the binning scheme resulted from the optimization. In the plots xi and yi represent the i-th bin in $SuperProbNN$ and XGB, respectively, as defined in Table 1.29.

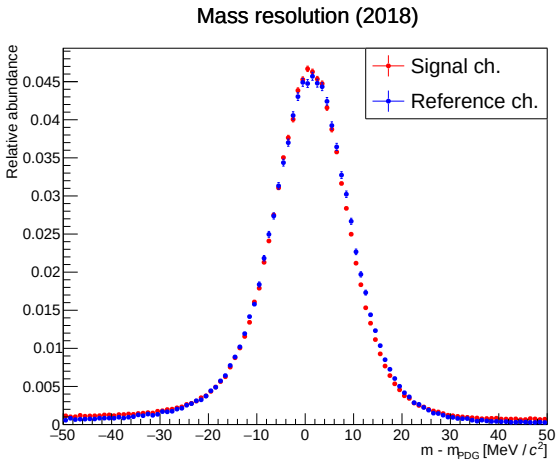


FIGURE 1.64: Comparison between the invariant mass distributions of the signal and the reference channel from 2018 MC. Both distributions are subtracted from the τ and D_s PDG mass value respectively in the signal and reference channel, to facilitate the comparison of the invariant mass resolution in the two channels.

As other efficiencies described in Section 1.8.2 also the fraction of the pdf in the mass bins needs to be corrected for Data/MC differences. The extrapolated signal pdf shown in Figure 1.65 is obtained by scaling its mean and width by correction factors c_M and c_λ evaluated from a fit to the reference channel invariant mass distribution in data and simulation:

$$c_M = \frac{M^{ref}_{Data}}{M^{ref}_{MC}} \qquad c_\lambda = \frac{\lambda^{ref}_{Data}}{\lambda^{ref}_{MC}} \tag{1.37}$$

Where M^{ref} and λ^{ref} represents the mean and width of the mass distribution of events in the reference channel. The values for the correction factors for different years are reported in Table 1.30. From the calibration's fit to the invariant mass distribution are also extracted the number of events observed in this channel (N_{cal}) used in the evaluation of the limit. These numbers are found to be 42946 ± 394 (2016), 88478 ± 594 (2017), and 103762 ± 643 (2018).

	2016	2017	2018
c_M	0.99990 ± 0.00015	0.99968 ± 0.00006	0.99989 ± 0.00008
c_λ	1.170 ± 0.037	1.092 ± 0.014	1.087 ± 0.032

TABLE 1.30: Correction factors for the M and λ parameters of the fit to signal MC invariant mass distribution.

The fraction of extrapolated signal pdf in the eight mass bins are shown in Table 1.31.

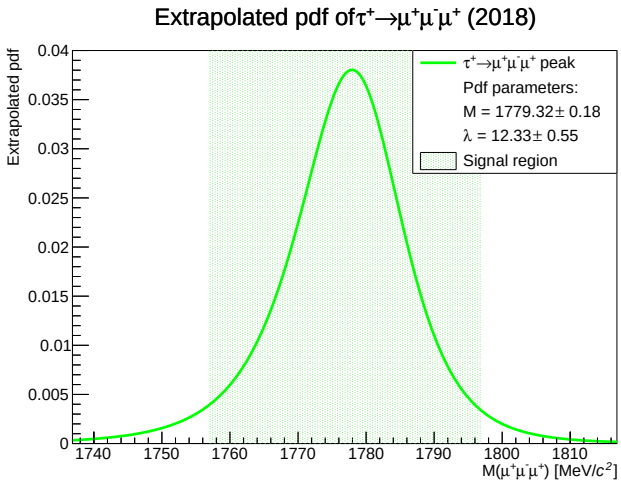


FIGURE 1.65: Extrapolated $\tau^+ \to \mu^+\mu^-\mu^+$ mass shape obtained from the fit to the MC invariant mass in 2018 sample. The green area indicates the signal window of $\pm 20\,\text{MeV}/c^2$.

Mass bin	2016	2017	2018
1	0.040 ± 0.001	0.038 ± 0.001	0.0348 ± 0.0005
2	0.076 ± 0.001	0.075 ± 0.001	0.069 ± 0.001
3	0.135 ± 0.002	0.137 ± 0.001	0.131 ± 0.001
4	0.197 ± 0.002	0.209 ± 0.002	0.208 ± 0.001
5	0.210 ± 0.002	0.220 ± 0.001	0.228 ± 0.001
6	0.153 ± 0.002	0.150 ± 0.001	0.158 ± 0.001
7	0.083 ± 0.001	0.074 ± 0.001	0.078 ± 0.001
8	0.038 ± 0.001	0.033 ± 0.001	0.0336 ± 0.005
tot	0.932 ± 0.004	0.936 ± 0.003	0.940 ± 0.006

TABLE 1.31: Fraction of events falling in each mass bin for the three data taking years, evaluated as the integral of the extrapolated signal pdf after the trigger cuts in each mass bin.

Since the outputs of the kinematic and PID classifiers (XGB and superProbNN) are uncorrelated with the invariant mass of the signal candidate, the fractions listed in 1.31 are fixed for each bin of the signal region defined by the minimal cut on the output of the classifiers.

1.9.3 Extrapolated limit from Run1

The expected upper limit (UL) for the Run1 dataset was estimated to be $5.0(6.1) \times 10^{-8}$ at 90%(95%) [23]. It is possible to extrapolate an expected limit for the Run2 dataset by considering the increment of luminosity recorded and higher cross-section of the production of τ at 13 TeV.

As shown in Section 1.3, the branching fraction is proportional to the ratio of the number of events observed in the signal and reference channel: $\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+) \propto N_{sig}/N_{cal}$.

In the case of no signal events are observed, the upper limit on the number of events in the signal channel is equal to the uncertainty on the number of background events (roughly $\sqrt{B_{exp}}$). Since both $\sqrt{B_{exp}}$ and N_{cal} are proportional to the product of the integrated luminosity (L) and the production cross-section of the predominant mother particle of τ, the expected upper limit is proportional to:

$$\mathrm{UL}_{exp} \propto \frac{\sqrt{L \times \sigma_{bkg}}}{L \times \sigma_{D_s}} \tag{1.38}$$

From the studies in Section 1.7 we can assume that the cross-section (σ_{bkg}) of the main background source (the combinatorial background) and the cross of the reference channel (σ_{D_s}) scale with the same factor, the expected upper limits results in:

$$\mathrm{UL}_{exp} \propto \frac{1}{\sqrt{L \times \sigma_{D_s}}} \tag{1.39}$$

It is possible though to write the extrapolated limit for Run2 as follows:

$$\mathrm{UL}_{extrap,Run2} = \mathrm{UL}_{exp,Run1} \frac{\sqrt{L_{2011} \times \sigma_{D_s,2011} + L_{2012} \times \sigma_{D_s,2012}}}{\sqrt{L_{Run2} \times \sigma_{D_s,Run2}}} \tag{1.40}$$

where the integrated luminosities and cross sections of the corresponding data taking periods are listed in Table 1.32 [58, 76].

TABLE 1.32: Integrated luminosity and prompt charm production cross section for each data taking period.

	2011	2012	Run 2
L (fb^{-1})	1	2	5.4
σ_{D_s} (μb)	197 ± 31	225 ± 35	353 ± 77

The prompt charm production cross section for 2012 is obtained by scaling the 2011's for the ratio between the corresponding centre-of-mass energies (8/7). The extrapolated upper limit is therefore evaluated as

$$\mathrm{UL}_{extrap,Run2} = 2.9(3.6) \times 10^{-8} \text{ @ } 90\%(95\%) \text{ C.L.}$$

1.10 Expected limit for Run2

The final values for the single event sensitivity α are shown in Table 1.33 separately for the 2μ and for the 3μ sample. The single terms entering the definition of α are also shown. The efficiency ratio $\epsilon_{D_s}/\epsilon_\tau$ is already corrected for the correction factors described in Section 1.8.2.

	2016	**2017**	**2018**
$\mathcal{B}(D_s^- \to \phi(\mu^+\mu^-)\pi^-)$		$(1.302 \pm 0.099) \times 10^{-5}$	
$f_{D_s}^{\tau}$ (%)		78.89 ± 3.78	
$\mathcal{B}(D_s^- \to \tau^-\bar{\nu}_\tau)$		$(5.48 \pm 0.23) \times 10^{-2}$	
N_{cal}	42946.23 ± 394.71	88478.34 ± 594.19	103762.50 ± 643.13
$\epsilon_{D_s}/\epsilon_{\tau,3\mu}$	0.300 ± 0.020	0.482 ± 0.029	0.515 ± 0.032
$\epsilon_{D_s}/\epsilon_{\tau,2\mu}$	0.970 ± 0.100	0.326 ± 0.029	0.281 ± 0.022
$\alpha_{3\mu}$	$(1.31 \pm 0.16) \times 10^{-9}$	$(1.02 \pm 0.12) \times 10^{-9}$	$(9.32 \pm 1.09) \times 10^{-10}$
$\alpha_{2\mu}$	$(4.17 \pm 0.56) \times 10^{-7}$	$(7.05 \pm 0.90) \times 10^{-8}$	$(5.09 \pm 0.61) \times 10^{-8}$

TABLE 1.33: Summary of the factors entering the normalization factor α in the 2μ and in the 3μ sample. The reported uncertainties include both statistical and systematics uncertainties. The ratios of efficiencies $\epsilon_{D_s}/\epsilon_{\tau}$ include the correction factors defined in Sec. 1.8.2.

The systematics uncertainties that are currently being considered come from the single event sensitivity α (listed in Table 1.34), from the background shape in the reference channel mass fits to the sidebands, and from the uncertainty on the number of background events expected in the signal region.

	relative uncertainty [%]		
	2016	**2017**	**2018**
$\alpha_{3\mu}$	11.96	11.63	11.68
$\alpha_{2\mu}$	13.35	12.73	11.96
$f_{D_s}^{\tau}$		4.79	
$\mathcal{B}(D_s^- \to \phi(\mu^+\mu^-)\pi^-)$		7.60	
$\mathcal{B}(D_s^- \to \tau^-\bar{\nu}_\tau)$		4.20	
N_{cal}	0.92	0.67	0.62
$\epsilon_{D_s}/\epsilon_{\tau,3\mu}$	4.92	4.90	4.91
$\epsilon_{D_s}/\epsilon_{\tau,2\mu}$	7.31	6.90	5.26
C_ϕ		0.88	
R_{Track}	0.51(stat) 1.40(syst)	0.47(stat) 1.40(syst)	0.50(stat) 1.40(syst)
$R(D_s, \tau_{3\mu})$	4.17(stat) 0.31(syst)	3.11(stat) 0.82(syst)	3.29(stat) 0.77(syst)
$R(D_s, \tau_{2\mu})$	4.80(stat) 0.60(syst)	3.45(stat) 1.27(syst)	3.68(stat) 0.99(syst)

TABLE 1.34: Relative uncertainties of the single event sensitivity α, followed by the relative uncertainties of each term that enters in their evaluation, for the three years of data taking. The ratios of efficiencies $\epsilon_{D_s}/\epsilon_{\tau}$ don't include the correction factors defined in Sec. 1.8.2.

To consider the effect of the systematics on α, its value is randomly fluctuated by a fraction (δ) of its combined statistical+systematic uncertainty (σ_α). The fluctuated value of α is thus used to compute the number of expected signal events given the $\mathcal{B}$ under test:

$$N_{sig} = \frac{\mathcal{B}}{\alpha \pm (\delta \cdot \sigma_\alpha)} \tag{1.41}$$

The effect of the systematic uncertainty on the background shape is estimated through two additional binned maps. These maps contain the number of expected background events in the signal region estimated by the exponential function used

to describe the background component, which slope was changed by $\pm 1\sigma$. The maps are then used to fluctuate the binned pdf used to get the number of background events in each bin of the signal region. Further systematic uncertainties are now under scrutiny, but these are thought to be smaller with respect to the ones quoted in the text.

The CLs method described 1.9 is applied iteratively to scan a region of $\mathcal{B}$ values and compute the confidence level at which they can be excluded. The number of background events and the signal efficiency in the signal region are stored separately for each year in binned maps. The values of α's are rescaled for the total signal efficiency obtained by the sum of the efficiencies in each bin, separately for each year, to obtain an overall value used to divide the $\mathcal{B}$ under test and get the number of expected signal events (since the analysis is still blinded). These quantities are used to compute the observed test statistics Q_{exp}, where the observed number of events corresponds thus to the expected number of background events. Q_{exp} is taken from the median of the background only hypothesis distribution. The distributions of the test statistic for the background only hypothesis (H_0) and the signal plus background hypothesis (H_1) are evaluated with toys. The expected value of CLs is finally obtained as:

$$CLs_{exp} = \frac{\int_{Q>Q_{exp}} H_1}{\int_{Q<Q_{exp}} H_0}$$

For both the 2μ and the 3μ sample, the maps of the three years are combined to determine the expected CLs as function of the$\mathcal{B}$ ($\tau^+ \to \mu^+\mu^-\mu^+$). Figure 1.66 shows the expected upper limits at 90% and 95% of confidence level, for the 3μ sample with and without the inclusion of the uncertainties described in Section 1.10:

$$\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+) \leq 1.9(2.3) \times 10^{-8} \text{ @ } 90\%(95\%) \text{ C.L.} \quad \text{without systematics}$$
$$\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+) \leq 1.8(2.2) \times 10^{-8} \text{ @ } 90\%(95\%) \text{ C.L.} \quad \text{with systematics}$$

The 2μ and the 3μ sample can be combined to obtain the expected upper limit for the Run2, but the contribution of the 2μ sample results to be negligible.

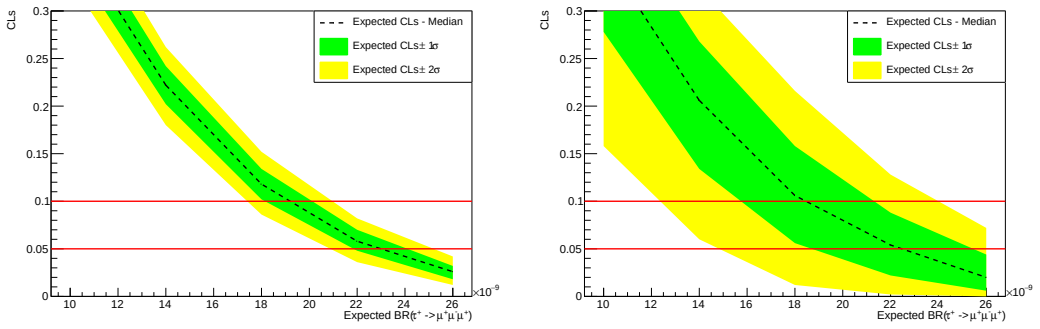


FIGURE 1.66: CLs as a function of the assumed branching fraction for the 3μ sample without including uncertainties (left) and with the statistical and systematics uncertainties included (right).

1.10.1 Model dependence

The LHCb simulated sample of $\tau^+ \to \mu^+\mu^-\mu^+$ decays is generated for simplicity without any assumption on the resonant structure of the decay (as the amplitude

structure of the decay may depend on beyond the Standard Model effects). By modifying the phase space distribution of the signal decays, the global efficiency of the decay may vary since the reconstruction and selection efficiency is not uniform along the phase space. This effect is taken into account by considering various amplitude models and reweighing the simulated sample after reconstruction according to each of them to study the variation of the global efficiency.

The alternative models are defined by using an effective-field theory approach as described in [77] in which 6-dimensional chirality operators are combined together with 6-dimensional radiative operators for producing the decay amplitudes. The chirality operators are

$$\begin{aligned} O_1 &= \left(\overline{L}\gamma_\mu L\right)\left(\overline{L}\gamma^\mu L\right) \\ O_2 &= \left(\overline{L}\tau^\alpha\gamma_\mu L\right)\left(\overline{L}\tau^\alpha\gamma^\mu L\right) \\ O_3 &= \left(\overline{R}\gamma_\mu R\right)\left(\overline{R}\gamma^\mu R\right) \\ O_4 &= \left(\overline{R}\gamma_\mu R\right)\left(\overline{L}\gamma^\mu L\right) \end{aligned} \tag{1.42}$$

with

$$L = \begin{pmatrix} \nu_L \\ l_L \end{pmatrix} \quad R = \begin{pmatrix} 0 \\ l_R \end{pmatrix}. \tag{1.43}$$

Operators other than $O_{1\text{–}4}$ are not considered since they do not contain helicity-conserving currents. The radiative operators are

$$\begin{aligned} R_1 &= g'\left(\overline{L}H\sigma_{\mu\nu}R\right)B^{\mu\nu} \\ R_2 &= g(\overline{L}\tau^\alpha H\sigma_{\mu\nu}R)W^{\mu\nu,\alpha}. \end{aligned} \tag{1.44}$$

The following amplitude structures for the various processes are obtained by combination of the operators above:

- 4-fermion $LL \to LL$ (same for $RR \to RR$):

$$\frac{\mathrm{d}\Gamma_V^{(LL)(LL)}}{\mathrm{d}m_{23}^2\mathrm{d}m_{12}^2} = \frac{|g_V^{(L_\mu L^\tau)(L_\mu L^\mu)}|^2}{\Lambda^4}\frac{(m_\tau^2 - m_\mu^2)^2 - (2m_{12}^2 - m_\tau^2 - 3m_\mu^2)^2}{256\pi^3 m_\tau^3} \tag{1.45}$$

- 4-fermion $LL \to RR$ (same for $RR \to LL$ with $g_V^{(L_\mu L^\tau)(R_\mu R^\mu)} \to g_V^{(R_\mu R^\tau)(L_\mu L^\mu)}$):

$$\begin{aligned} \frac{\mathrm{d}\Gamma_V^{(LL)(RR)}}{\mathrm{d}m_{23}^2\mathrm{d}m_{12}^2} = \frac{|g_V^{(L_\mu L^\tau)(R_\mu R^\mu)}|^2}{\Lambda^4}\Bigg[&\frac{(m_\tau^2 - m_\mu^2)^2 - 4m_\mu^2(m_\tau^2 + m_\mu^2 - m_{12}^2)^2}{512\pi^3 m_\tau^3} \\ &- \frac{(2m_{13}^2 - m_\tau^2 - 3m_\mu^2)^2 + (2m_{23}^2 - m_\tau^2 - 3m_\mu^2)^2}{1024\pi^3 m_\tau^3}\Bigg] \end{aligned} \tag{1.46}$$

- Radiative:

$$\frac{\mathrm{d}\Gamma_{\mathrm{rad}}^{(LR)}}{\mathrm{d}m_{23}^2\mathrm{d}m_{12}^2} = \alpha_{\mathrm{em}}^2 \frac{|g_{\mathrm{rad}}^{(L_\mu R^\tau)}|^2}{\Lambda^4}\left[\frac{m_\mu^2(m_\tau^2-m_\mu^2)^2}{128\pi^3 m_\tau^3}\left(\frac{1}{m_{13}^4}+\frac{1}{m_{23}^4}\right)\right.$$
$$+\frac{m_\mu^2(m_\tau^4-3m_\tau^2m_\mu^2+2m_\mu^4)^2}{128\pi^3 m_\tau^3 m_{13}^2 m_{23}^2}+\frac{(2m_{12}^2-3m_\mu^2)^2}{128\pi^3 m_\tau^3}$$
$$\left.+\frac{(m_{13}^2+m_{23}^2)\left(m_{12}^4+m_{13}^4+m_{23}^4-6m_\mu^2(m_\tau^2+m_\mu^2)\right)}{256\pi^3 m_\tau^3 m_{13}^2 m_{23}^2}\right] \quad (1.47)$$

- Interference of Radiative and 4-fermion $LL\to LL$ (same for $RR\to RR$):

$$\frac{\mathrm{d}\Gamma_{\mathrm{mix}}^{(LL)(LL)}}{\mathrm{d}m_{23}^2\mathrm{d}m_{12}^2} = \alpha_{\mathrm{em}}^2 \frac{2v\mathrm{Re}(g_V^{(L_\mu L^\tau)(L_\mu L^\mu)} g_{\mathrm{rad}}^{*(L_\mu R^\tau)})}{\Lambda^4}\times$$
$$\left[\frac{m_{12}^2-3m_\mu^2}{64\pi^3 m_\tau^2}+\frac{m_\mu^2(m_\tau^2-m_\mu^2(m_{13}^2+m_{23}^2}{128\pi^3 m_\tau^2 m_{13}^2 m_{23}^2}\right] \quad (1.48)$$

- Interference of Radiative and 4-fermion $LL\to RR$ (same for $RR\to LL$):

$$\frac{\mathrm{d}\Gamma_{\mathrm{mix}}^{(LL)(RR)}}{\mathrm{d}m_{23}^2\mathrm{d}m_{12}^2} = \alpha_{\mathrm{em}}^2 \frac{2v\mathrm{Re}(g_V^{(L_\mu L^\tau)(R_\mu R^\mu)} g_{\mathrm{rad}}^{*(L_\mu R^\tau)})}{\Lambda^4}\times$$
$$\left[\frac{m_\tau^2-m_{12}^2-3m_\mu^2}{256\pi^3 m_\tau^2}+\frac{m_\mu^2(m_\tau^2-m_\mu^2(m_{13}^2+m_{23}^2}{256\pi^3 m_\tau^2 m_{13}^2 m_{23}^2}\right] \quad (1.49)$$

and the relevant distribution of the events in the phase space is shown in Figure 1.67.

The weights to redistribute the events are calculated after normalizing the amplitude distributions, therefore being concerned of the phase-space dependence only. For each amplitude model, the weight is calculated by dividing the Dalitz plot of the amplitude model over one of the phase-space generated events. To avoid border effects, squared Dalitz plots are considered. The weights distributions are shown in Figure 1.68.

Once the weights distributions are determined, the effect of the reweighing is calculated by integrating them after the selection cuts ($SuperProbNN > 0.88$ and $XGB > 0.8$) and after correcting for data/MC difference and particle identification. The ratios of the integrals for each amplitude model with respect to the phase-space are shown in Table 1.35, where the last column (*Total*) merges the results of each simulation sample after correcting for the fraction of the relevant production process (as taken from Table 1.3).

The assumption on the amplitude model introduces an efficiency variation ranging from -20% in the case of the purely radiative process to $+10\%$ for the 4-fermion $LL\to LL$.

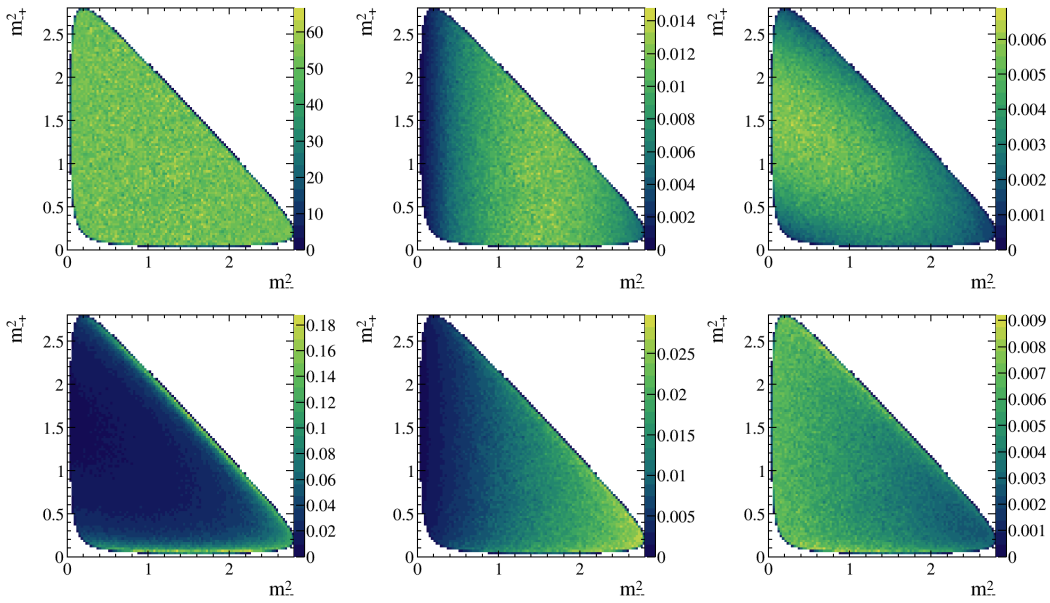


FIGURE 1.67: Dalitz plots of the distribution of the events generated in the phase space (top-left) and reweighted by the amplitude of the processes calculated with effective-field theory: 4-fermion $LL \to LL$ (top-center), 4-fermion $LL \to RR$ (top-right), radiative (bottom-left), interference of radiative and 4-fermion $LL \to LL$ (bottom-center), interference of radiative and 4-fermion $LL \to RR$ (bottom-right). The Dalitz plot variables are $m^2_{--} = m^2(\mu^-\mu^-)$ and $m^2_{-+} = m^2(\mu^-\mu^+)$.

TABLE 1.35: Ratio of the sum of weights between the various amplitude models and the phasespace in the simulated samples after applying the selection. The final column shows the combination of the results of the simulated samples after correcting for the fraction of the relevant production process.

Model	$D^+ \to \tau^+\nu$		$D_s \to \tau^+\nu$		$B \to \tau X$	Total
	from B	Prompt	from B	Prompt		
4-fermion $LL \to LL$	1.09	1.11	1.08	1.10	1.09	1.10
4-fermion $LL \to RR$	1.05	1.02	1.04	1.04	1.04	1.04
Radiative	0.80	0.86	0.79	0.81	0.79	0.81
Rad+4f $LL \to LL$	1.02	1.06	1.01	1.04	1.02	1.04
Rad+4f $LL \to RR$	0.95	0.94	0.96	0.94	0.95	0.95

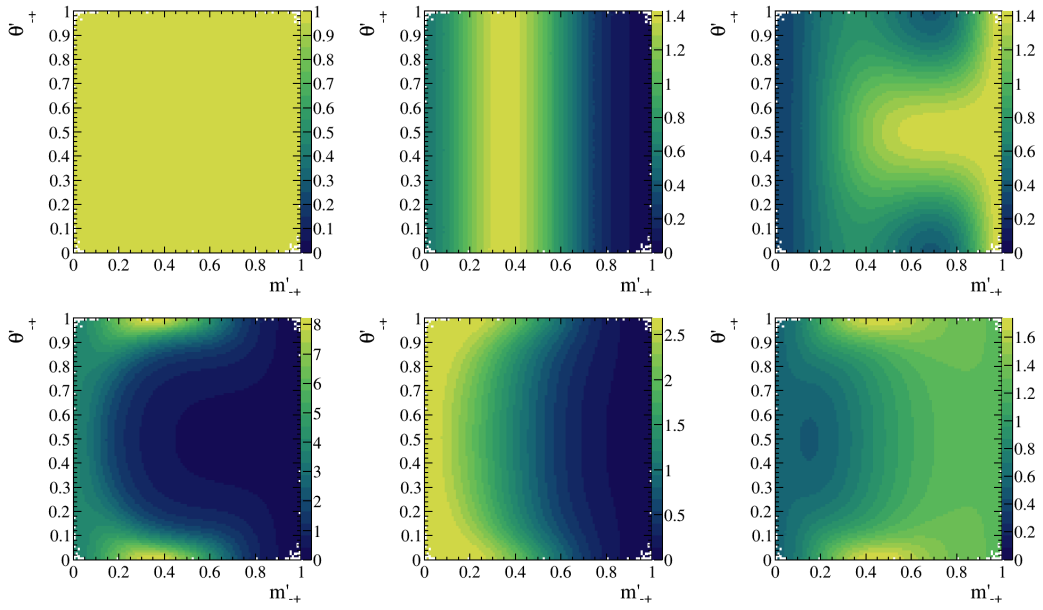


FIGURE 1.68: Squared Dalitz plots showing the distribution of the weights obtained by calculating the bin-by-bin ratio of the normalized amplitudes calculated with effective-field theory and the phase space. They are shown as follows: the phase space as a cross-check (top-left), 4-fermion $LL \to LL$ (top-center), 4-fermion $LL \to RR$ (top-right), radiative (bottom-left), interference of radiative and 4-fermion $LL \to LL$ (bottom-center), interference of radiative and 4-fermion $LL \to RR$ (bottom-right). The squared Dalitz plot variables are $m'_{-+} = \frac{1}{\pi} \arccos\left(2\frac{m(\mu^-\mu^+) - m(\mu^-\mu^+)_{\min}}{m(\mu^-\mu^+)_{\max} - m(\mu^-\mu^+)_{\min}} - 1\right)$ and $\theta'_{-+} = \frac{1}{\pi}\theta_{\mu^-\mu^+}$, where $\theta_{\mu^-\mu^+}$ is the helicity angle of the $\mu^-\mu^+$ system, namely the angle between the μ^- momentum in the $\mu^-\mu^+$ rest frame and the opposite of the momentum of the $\mu^-\mu^+$ system in the τ^- rest frame.

1.11 Conclusion

In this chapter, the analysis of the Run2 dataset in search for the lepton flavour violating decay $\tau^+ \to \mu^+\mu^-\mu^+$ has been presented. The data were collected at 13 TeV during 2016, 2017 and 2018, and correspond to an integrated luminosity of 5.4 fb^{-1}.

The analysis is still blind at the moment of the writing of the thesis, thus only the expected upper limit for the branching ratio of $\tau^+ \to \mu^+\mu^-\mu^+$ can be evaluated. The upper limits for the 3μ sample, obtained combining the three years, results in

$$\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+) \leq 1.8(2.2) \times 10^{-8} \text{ @ } 90\%(95\%) \text{ C.L.}$$

This limit sets a lower bound than the one found by the Belle experiment, and it will decrease even more when combined with the upper limit found with the Run1 dataset. The LHCb experiment will continue the operations to collect more valuable data and further improve the search for new Physics.

Study of the timing performance of multianode MCP-PMT in single photon regime

2.1 Introduction

In this chapter the study of the time resolution of a multianode microchannel plate photomultiplier produced by Photek[1] will be presented. The purpose of this study lies in the framework of R&D projects for the future upgrade of the LHCb RICH detector. In Section 1.2.1 the former LHCb detector's subsystems were described. Within the next decade, lots of researches will be conducted to finalize the shape of future enhancements of the LHCb experiment. The Author's work given in Section 2.3 is preceded by an overall introduction on the upgrades of LHCb in Section 2.1.1 and an overview of MCP photomultipliers in Section 2.2.

2.1.1 LHCb Upgrade

On the 5$^{\text{th}}$ of July 2022 LHCb resumed data-taking operations (Run3) after a break that lasted more than three years. During this second Long Shutdown (LS2) the LHCb detector previously described in Sec. 1.2.1 was almost completely renovated. The main feature of the upgraded LHCb detector [42] is the capability of being fully readout at 40 MHz, synchronously with the LHC bunch crossing rate. This is made possible by the use of a very flexible software-based trigger, which allows the experiment to increment significantly its sensitivity for flavour physics by increasing the amount of data acquired compared to what was collected during Run1 and Run2.

Figure 2.1 shows a comparison of the trigger sequence and throughput rate between the Run2 and Run3.

The upgraded LHCb detector operates at an instantaneous luminosity of $\mathcal{L} = 2 \times 10^{33}$ cm^{-2}s^{-1}, with $\nu = 7.6$ nominal visible interactions per crossing, resulting in a total sample of $\sim$50 fb^{-1} collected before the end of 2032, during Run3 and Run4. The higher luminosity, combined with the increased energy of the collisions ($\sqrt{s}$ =13.6 TeV) will improve the precision of the measurements of *CP*-violation observables, rare decays and fundamental parameters of the CKM matrix.

The abundance of the B and D mesons that will be collected by the upgraded LHCb detector won't be matched by any other experiments in the same temporal windows of operation; study of $\overline{B}^0_s$ decays and *CP*-violation will be a field dominated by LHCb. In addition to the heavy flavour physics, other topics will provide interesting opportunities to exploit the uniqueness of LHCb: improvement of the effective electroweak mixing angle, measure of W mass, search for long-living particles and even QCD studies complementary to the ones of ATLAS and CMS, making LHCb a full-fledged general purpose detector.

[1] `https://www.photek.com/`

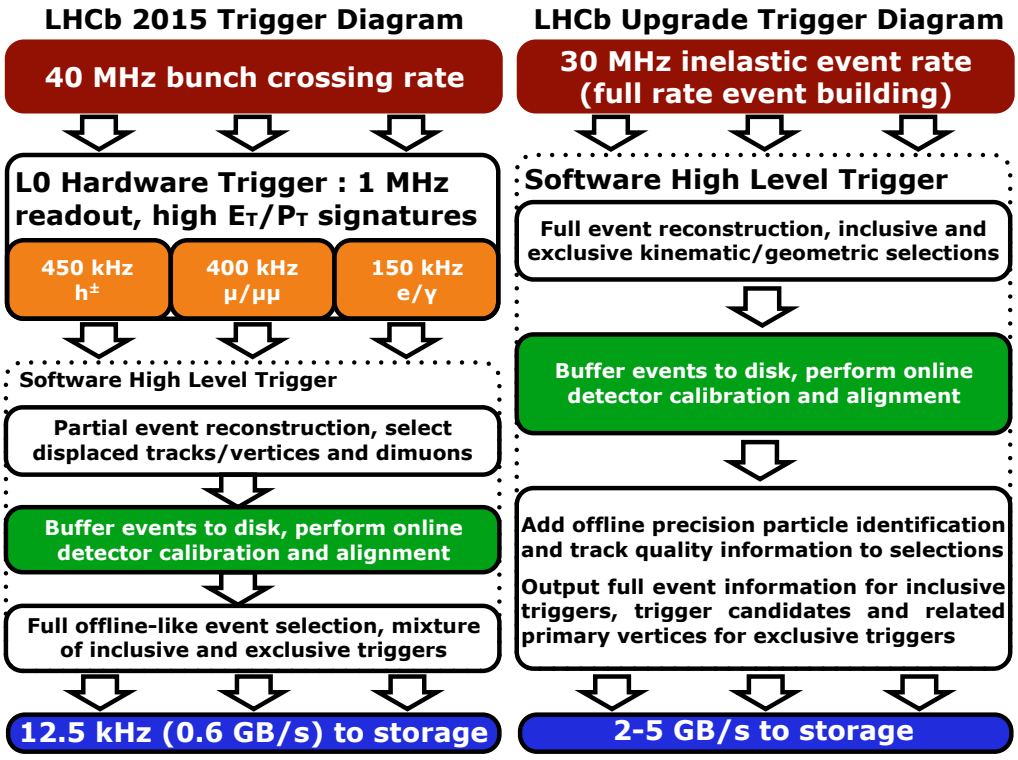


FIGURE 2.1: Illustration of the trigger scheme for Run2 and the update trigger scheme for Run3, with approximate throughput rates of each subsystem.[78]

Upgraded subsystems

The whole LHCb Data Acquisition (DAQ) system, all the readout electronics and almost all the detectors (except for calorimeters and the M2-M5 muon stations) are completely renewed. The level-0 hardware trigger, that reduced the inelastic collision rate of 30 MHz to the readout rate of 1 MHz, was removed in favour of a full software trigger [31]. The new readout system consists of the event builder, the Timing and Fast Control (TFC) distribution, the Experiment Control System (ECS) and the Event Filter Farm (EFF) which will host the full software trigger. The architecture of the trigger system is summarized in Fig. 2.3.

It is composed by a Low Level Trigger (LLT) that identify clusters with high E_T in the calorimeters or tracks with high p_T in the muon stations, but unlike the previous system, there are no trigger decisions sent to the front-end electronics (the readout can be thus defined triggerless). The High level trigger is still composed by two stages, the first one runs a partial reconstruction of the event while the second one run inclusive and exclusive selections, fit to tracks with kalman filter and particle identification obtained from the RICH information, and it requires $\sim$20 ms for processing an event. In the next paragraphs will be briefly described the updated subdetectors of LHCb, that will face the increased occupancies and rates.

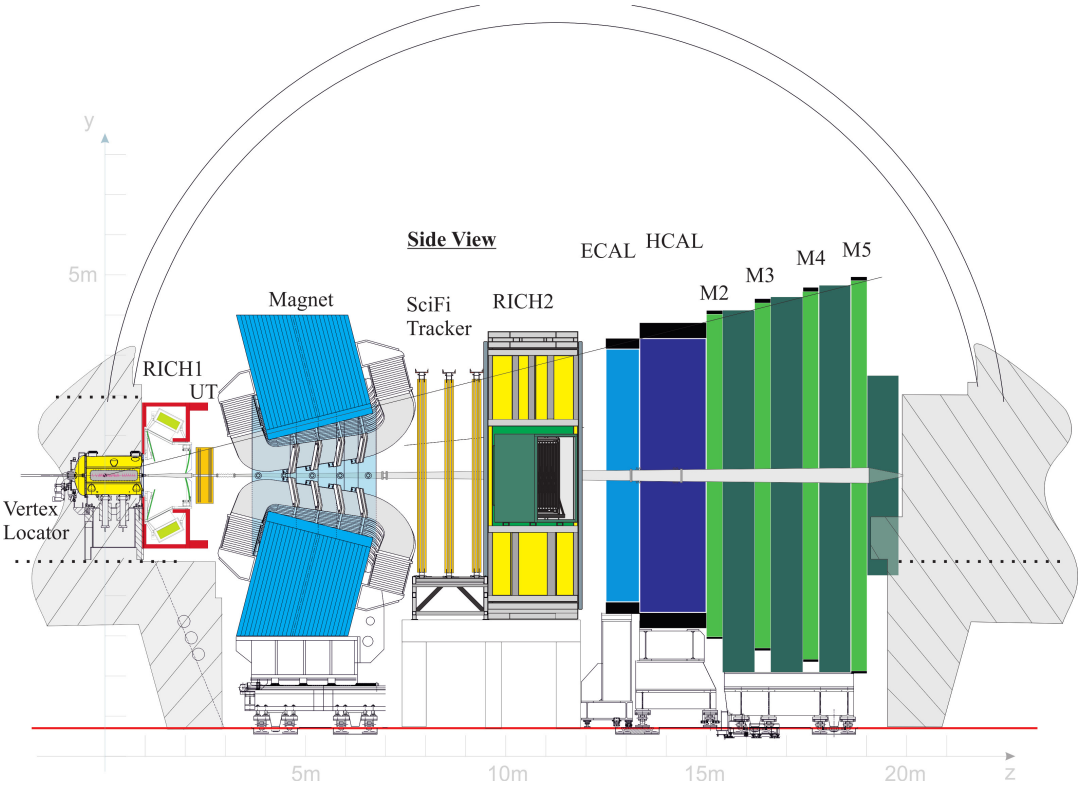


FIGURE 2.2: Layout of the upgraded LHCb detector.

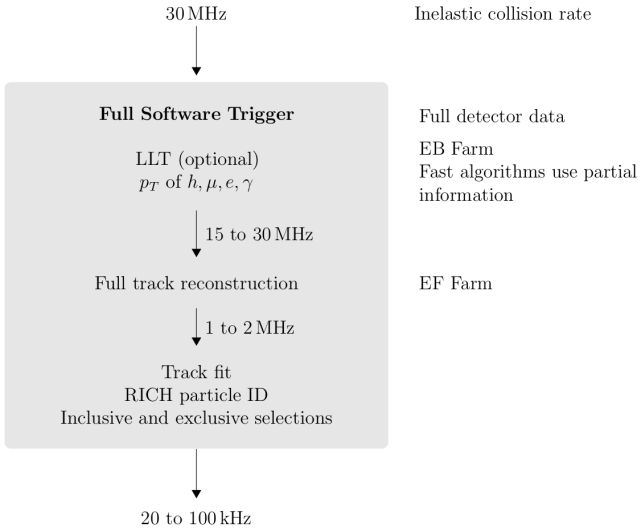


FIGURE 2.3: Summary of the upgraded trigger scheme of LHCb.

VErtex LOcator In the upgraded VELO [79] the silicon r- and ϕ- sensors are replaced by hybrid pixel sensors cooled by evaporative CO_2, readout by the VeloPix, a custom radiation hard ASIC. The layout of the new VELO module and the positions of the layers are shown in Figure 2.4.

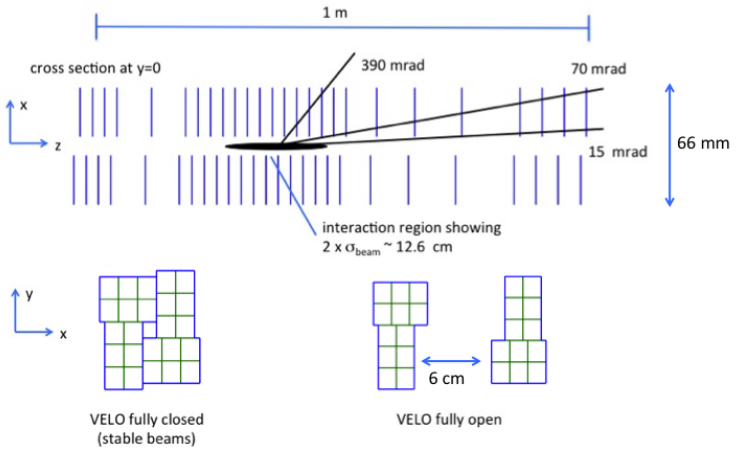


FIGURE 2.4: Layout scheme of the new VELO detector.

Tracker system The updated Tracker system [47] consists of an Upstream Tracker (UT), placed before the dipole magnet, and a Scintillating Fibre Tracker (SFT) placed downstream the dipole magnet.

The UT replaced the Tracker Turicensis. It has a similar geometry, but the sensors that constitute the four planes are thinner, with a larger coverage and capable of processing the signal directly, exploiting a dedicated ASIC (SALT). Figure 2.5 illustrate the layout of the Upstream Tracker. One of the main characteristics of the UT detector is the capability to provide hits for all the tracks in the LHCb detector, in order to reduce the mismatch between the segments of tracks reconstructed by the VELO and the downstream tracker and thus reduce the number of ghost tracks.

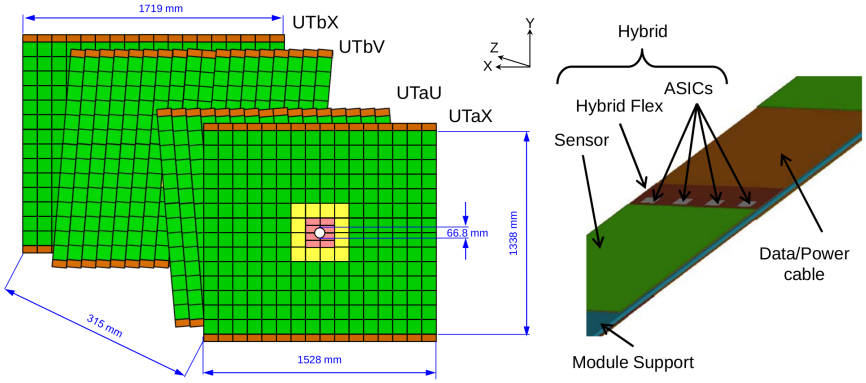


FIGURE 2.5: *left*: Layout of a module of the Upstream Tracker, illustrating the four layers organized in a "$x - u - v - x$" scheme. *right*: Basic module of the UT showing the stave structure.

The downstream tracking stations T1,T2, and T3 are replaced by a new SciFi Tracker, composed by 2.5 m fibers with a 125 μm radius and readout by Silicon Photomultipliers (SiPM). When a charged particle traverse a scintillating fibre it produces photons in its core, and then they travel thanks to total internal reflection toward the end of the fibre where the SiPM are placed. The total reflection is given by two different cladding layers with smaller refraction indices (n=1.49 and n=1.42) than the one of the scintillating fibre (n=1.60).

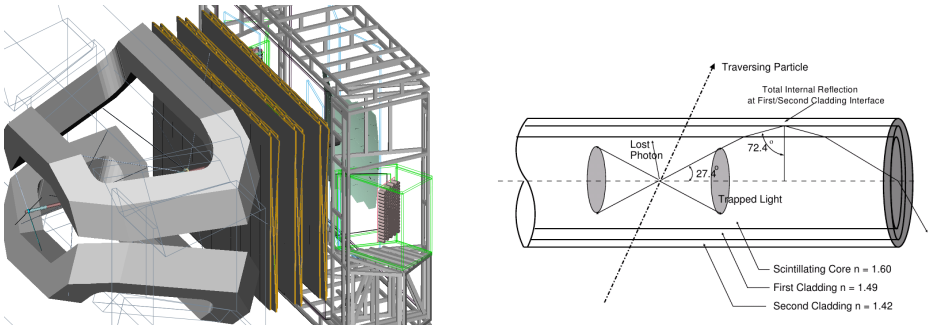


FIGURE 2.6: *left*: 3D Illustration of the three modules of the SciFi detector, each composed by four layers. The two inner layers are tilted by $\pm 5^\circ$ with respect to the outer ones.

Ring Imaging Cherenkov detectors The HPDs and the readout electronics are replaced by multianode photomultipliers (MaPMTs) produced by Hamamatsu[2]. The whole RICH1 and the central part of RICH2 are instrumented with 1" R11265 MaPMTs ($2.8 \times 2.8\,\text{mm}^2$ pixels arranged in a 8×8 matrix), while for the outer part of RICH2 where a lower photon rate is expected 2" R13742 MaPMTs (based on H12700) with bigger pixels ($5.6 \times 5.6\,\text{mm}^2$) are used [80]. The readout system is provided by an 8-channel digital ASIC named CLARO [81]. It is designed to have a fast recovery time (less than 25 ns) and to be rad-hard. The CLAROs are mounted on Front End Boards (FEBs) placed directly behind MaPMTs, in a compact structure named Elementary Cell (EC) visible in Fig. 2.7. The actual photoelectron rate faced by the MaPMTs is $\sim$10MHz/pixel, and given the pixel size of about $10\,\text{mm}^2$ this corresponds to $1\text{MHz}/\text{mm}^2$.

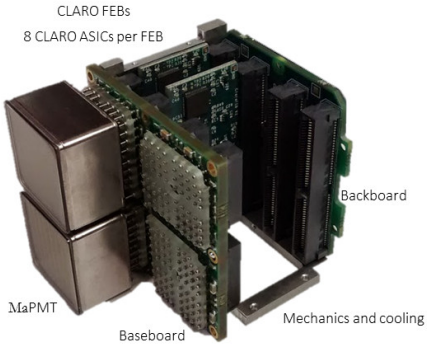


FIGURE 2.7: Isometric view of a half-instrumented Elementary Cell of the Upgraded RICH detector.

Calorimeters The redesign of the trigger system lead to the removal of scintillating pad and preshower detectors from the Calorimeters System, since the main purpose of them was to provide information to the L0 trigger, and an upgrade of the electronic system to cope with the 40 MHz readout. The ECAL and HCAL modules are kept, as well the photodetectors, but the PMTs operate at a reduced gain ($\times 5$ less than Run2 and balanced by an increased gain in the electronics) to fight off the ageing [49].

[2]Hamamatsu photonics: `http://www.hamamatsu.com`

Muon stations All the muon stations equipped with Multi-wire proportional chambers that are placed downstream the HCAL (M2-M5) are maintained and protected from the increased particle flux by additional shielding around the beam pipe. The M1 station placed between the RICH2 and the ECAL is removed consequently to the removal of the L0 trigger, similarly to what happened to the PS and SPD in the Calorimeters System. The off-detector readout electronics were updated to operate at the rate of 40 MHz, while the front-end electronics already satisfied this requirement [49].

2.1.2 HL-LHC

The High-Luminosity Large Hadron Collider (HL-LHC) project will push the limits of the actual collider to reach a peak luminosity of 5×10^{34} cm^{-2}s^{-1} and deliver a total of 3000 fb^{-1} in the 12 years of operations [82] at a center-of-mass energy of 14 TeV. It will start after the LS3 scheduled between the 2026 and 2029. During these years, the ATLAS and CMS collaborations will upgrade their detectors to cope with the more challenging environment [83, 84].

The importance of Timing

The LHCb upgrade for the HL-LHC (Upgrade II [85]) will take place later, during the LS4 scheduled for 2033. From Run5 and onward, LHCb will operate at an instantaneous levelled luminosity of 1.5×10^{34} cm^{-2}s^{-1}, with an average of $\sim$ 42 pile-up collisions, to be compared with the nominal $\nu = 7.6$ faced by the actual detector. It will be a challenge for the software trigger system and for the subdetectors to maintain the same performance of Upgrade I in such higher pile-up and occupancy conditions. The identification of primary vertex (PV) and the right match with the decay products is extremely important to reconstruct correctly the heavy flavour decay chains. It has been demonstrated that the mis-association levels can be reduced from $\sim$ 20% to about 5% exploiting the use of precision timing [86]. The addition of temporal information will spread the spatial density of PVs to a level compatible with the real time pattern recognition developed for the first upgrade, together with an increment of the subdetectors granularity to reduce the occupancy.

Figure 2.8 illustrate the effects of the use of temporal information to select the right PV for an event simulated in the Upgrade II conditions. The primary vertices are represented by a 2D Gaussian with central value and widths associated to spatial(x axis) and temporal (y axis) resolution. Without the time information, the correct PV “A" closet to (0,0) is discarded in favour of “B", which is only spatially closer to the origin of the axis. There are studies that demonstrate the positive impact of timing on the CPU resources spent for the reconstruction of tracks, due to the reduction of combinatorics at an early stage.

RICH The fast timing can also reduce the combinatorial background seen by the RICH detectors in high-multiplicity events: The arrival time of the photons from any PV can reach RICH1 and RICH2 within of $\sim$ 5 ps and $\sim$ 50 ps respectively, thanks to the focusing optics. Time-stamping each photon could also allow the RICH detectors to complement the VELO in the determination of the primary vertex time. The PID algorithms can predict the time of arrival of the photons with $\sigma \sim 10$ ps exploiting the tracking information, thus the aim of Upgrade II is to instrument the RICH with photon detectors and readout electronics with a time resolution better than 100 ps.

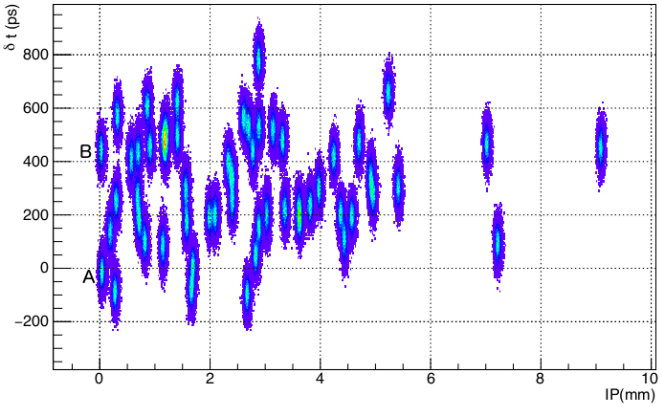


FIGURE 2.8: Primary vertices to be matched with an event containing a $B^0 \to \pi^+\pi^-$ candidate, simulated under the Upgrade II conditions. PVs are represented with 2D Gaussian distribution with spatial and temporal uncertainties. Without the temporal information, all the PV would have been aligned on the same y-axis, leading to the wrong choice of vertex "B" instead of vertex "A".

The MaPMTs installed for the Upgrade I will need to be replaced in the second upgrade because don't satisfy the pixels size, time resolution, maximum anode current and sensitivity in the green wavelength region required. The two main technologies that are currently envisaged as alternative to MaPMT are MCP and SiPM. Microchannel place detectors can achieve time resolution of ~ 30 ps and exhibit low dark-count rate (DCR), while the SiPMs usually have high DCR, especially when exposed to irradiation up to $10^{13} n_{eq}/cm^2$ [87], that will be treated by annealing during detector downtime and the installation of additional neutron shielding and cryogenic cooling (which can occupy the place of the magnetic shielding, since the SiPM devices are insensitive to stray magnetic fields). In addition to a good time resolution, SiPMs have high photon-detection efficiency at longer wavelength than MCPs and operate at lower voltage, reducing the ageing due to the integrated collected charge. Extension of MCP lifetime can be achieved through atomic layer deposition (ALD) processes [88, 89] but some studies demonstrates that the process can affect the load capacity and the recovery time of MCP-PMTs [90].

TORCH The precise measurement of Time-of-Flight (ToF) of kaons and protons made by the TORCH detector (Timing Of internally Reflected CHerenkov photons) will provide an increased identification capability of such particle by the LHCb detector at low momenta (up to 10 GeV/c for kaons and up to 20 GeV/c for protons, regions below the threshold of the C_4F_{10} gas in RICH detectors) [91, 92]. The TORCH collect the Cherenkov photons emitted by the hadrons traversing a thin quartz plane by means of total internal reflection, which are then focused onto microchannel plate photomultipliers (MCP-PMT), as can be seen from the detector layout shown in Fig. 2.9. The system exhibits a time resolution of few tens of picoseconds per photon.

ECAL Fast timing will be a mandatory requirement for the Upgrade of the Electromagnetic calorimeter, due to the increased number of candidates in the events [85, 86]. The large combinatoric background can be reduced by the correct association of the candidates to the individual pp interactions in the bunch crossing. A dedicated

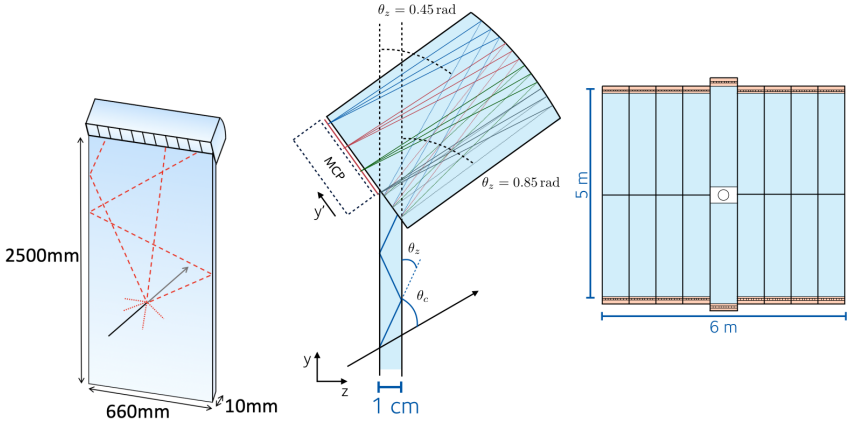


FIGURE 2.9: Schemes of the TORCH detector. From left to right there is the layout of a single module, the focusing block illustrating the paths of the reflected photons and the layout of the whole detector assembled.

timing layer capable to achieve a timing resolution of ~ 20 ps for electrons and photons is under evaluation, with Large Area Picosecond Photo-Detectors (LAPPDs) combined with MCPs as the most promising technology [93–96].

In order to retain the detector performance achieved during Run2, the cell size and the Molière radius of the converter need to be reduced, to sustain operation at high luminosity of 1.5×10^{34} cm^{-2}s^{-1}: the current transverse dimension of the modules are 4×4 mm^2 in the inner region and should be reduced to 2×2 mm^2. The inner region is the one more susceptible to radiation damage (200 mrad of total dose is expected for the innermost modules), and the upgraded modules will need to be very radiation hard. The current modules haven't been replaced during LS2 (only the readout electronics were changed) but will definitely be substituted after Run3. The Upgrade II will recover the performance degradation due to the increased cluster overlap and combinatorics expected during Run3 and Run4. The most promising technology for the instrumentation of the innermost part of ECAL consists of longitudinal crystal fibres that are used both as scintillator and for light transportation, without the need of wavelength shifters [97, 98]. This Spaghetti-Calorimeter (SpaCal) will be able to satisfy the requirement of $>$ 200kGy radiation tolerance for the central region.

RICH LS3 enhancements

The benefits of precision timing to the event reconstruction in the high luminosity environment from Run5 were briefly presented in the previous section.

However, the stop of operations during LS3 will be a good opportunity to improve parts of the LHCb experiment, in order to enhance the performance of the recently updated detector already in Run4 and acquire valuable experience for the next major upgrade in particular in the particle identification. A recent study [99] shown how the use of sub-ns timing information can improve the PID performances of the upgraded RICH detectors, that otherwise would have been deteriorated by the higher detector occupancy and greater event multiplicity.

The photodetectors of the RICH system will not be replaced before the Upgrade II to reduce the costs, only the readout electronics and DAQ will undergo changes

principally to include temporal information. There is the possibility to apply a hardware time-gate at the level of 3-6 ns already in Run3, but the electronics enhancements will allow applying software time-gate of 600 ps (corresponding to a window of $\pm 2\sigma$ of the time resolution of MaPMTs). One of the possible solutions under evaluation is the replacement of the CLARO readout ASIC from the Elementary Cell and the FPGA (Field Programmable Gate Array) on the Front End Board (described in Sec.2.1.1) with a FastIC ASIC combined with a TDC (Time-to-Digital Converter) whose goal is to timestamp photons with an intrinsic time resolution of 25 ps [100]. This device accepts inputs from different kind of photodetectors, such as MaPMTs, SiPMs or MCPs. The MaPMT's time resolution would limit the system's achievable time resolution.

Another solution [101] to enable single photon counting with sub-ns resolution would require the replacement of the FPGA with a radiation hard digital ASIC with GHz sampling rate, and a time over threshold measurement (ToT) to compensate the CLARO time walk. This second cost-saving option leaves intact the Elementary Cell, allowing to make use of its intrinsic time resolution ($\sim$ 200 ps RMS [102]) which will corresponds to an improvement of one order of magnitude with respect to the Run3 conditions.

2.2 Microchannel Plate Photomultiplier

Photomultipliers that use microchannel plates (MCP) as an electron multiplier are named MCP-PMT.

Microchannel plates are made of a two-dimensional array of a great number of glass capillaries (channels) bundled in parallel, where the electron multiplication takes place [103]. A schematic representation of the structure of an MCP and of the electron multiplication is shown in Fig. 2.10.

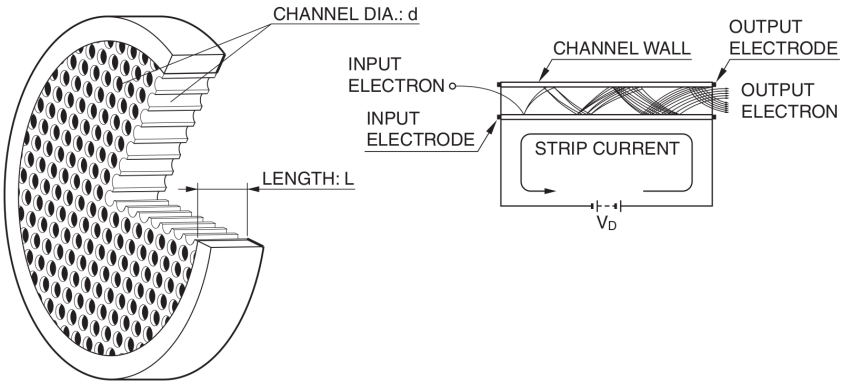


FIGURE 2.10: (Left) Scheme of a microchannel plate of length 'L' and with channels of diameter 'd'. (Right) Representation of the electron multiplication process inside a single MCP channel.

The detection of an incoming photon makes use of the external photoelectric effect: the electrons in the photocathode (made of alkali metals, with low work function) absorbed the photon energy ($\hbar\nu$) and diffused toward the surface of the photocathode. If the electrons have enough energy to overcome the forbidden-band gap, they can be emitted in the vacuum as photoelectrons. Once a photoelectron is emitted, it is accelerated by the electric field generated within the MCP, which is parallel to the channel axis, producing secondary electrons when colliding onto the channel

wall, as can be seen from Figure 2.10. The gain (μ) of an MCP depends on the length (L) and diameter (d) of the channels [103], and can be approximated as:

$$\mu = e^{G \cdot \frac{L}{d}} \tag{2.1}$$

where G is called gain factor (usually has values between 40 and 60), it is a secondary emission characteristic of the channel wall material and depends on the electric field intensity inside the channel. In common photon counting measurements, the gain level is $\mathcal{O}$ (10^6). In order to obtain such high value, usually two MCP slabs are stack together in a Chevron (v-shape) as shown in Fig. 2.11. If only one MCP is used, when the gain reaches $\mathcal{O}$ (10^4) the ions feedback begin to affect the signal output. The ion back-bombardment causes deterioration of the photocathode, thus an ion barrier film is placed on the MCP to stop the accelerated ions to reach it.

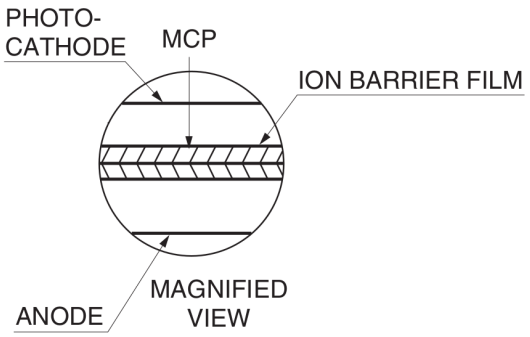


FIGURE 2.11: Layout of an MCP-PMT that illustrate the photocathode, where the photons convert into photoelectrons, the MCP slabs in which happens the electrons' multiplication process and the anode, that collects the electrons' cloud. An ion barrier film on the MCP prevent the drift of the ions back to the photocathode, in order to block delayed showers or damage to it.

Time characteristics The excellent timing capability of MCP-PMT compared to traditional photomultiplier tubes comes from the almost straight flow of the electrons from the photocathode to the anode, following the strong electric field parallel to the channels. The initial velocity and emission angle is very similar for the photoelectrons produced on the whole surface, so their contribution to the transit time (the amount of time between the interaction of the light on the photocathode and the collection of the electrons' cloud at the anode) is negligible. The spread of the transit time (T.T.S.), which measures the time resolution of the device, is kept low by the very short multiplication time.

Saturation characteristics When the output signal of a photodetector is not proportional anymore to the intensity of the incident light, it means that it has reached *saturation*. The saturation of MCP-PMT devices is caused by the non-negligible *dead time* (or recovery time) that is needed to restore the multiplication capability of a channel after an electron shower was created in it.

As the secondary emission process proceeds within an MCP channel, a corresponding amount of positive charges is generated at the MCP end. These charges alter the potential distribution and reduce the intensity of the electric field, suppressing the electron multiplication. It takes some time to neutralize these charges

because the strip current that flows through the channel wall is small, as a consequence of the high resistance of the material (usually between tens and hundreds of Ω/mm). The dead time (or recovery time) is the amount of time needed to restore the initial potential distribution and strength of the electric field. If the time elapsed between two electron multiplication is smaller than the recovery time, the output of the second cascade will be saturated.

Magnetic characteristics The resistance to magnetic field of MCP depends on the direction between the field and the tube axis: the distance between the photocathode and the anode is small compared to classical photomultipliers with discrete dynode structure, and thus the electron transit distance is minimized. MCP-PMT can sustain up to 2 T if the field is oriented alongside the tube axis, whereas only up to $\sim$ 70 mT if it is oriented perpendicular to the tube axis [103].

Ageing and lifetime characteristics One of the main problems for the longevity of vacuum photomultiplier detectors is the degradation of the photocathode. Positive ions can be produced at the last stages of the multiplication process within the MCP pores and flow back to strike the photocathode, reducing the photocathode's quantum efficiency. Tilted MCP slabs and/or thin barrier film are standard solution to prevent a direct path to the photocathode. A different solution is provided by atomic layer deposition (ALD). Studies demonstrate that the deposition of atomic mono-layers on a substrate inside the MCP pores can improve the secondary emission yield (SEY), resulting in higher gain at the same operating voltage (or equivalently, it is possible to operate at a lower bias voltage while maintaining a high gain). The secondary electron energy is reduced, and so is the probability of an ion to be produced in the electron's collisions [104]. Additional consequences of disentangling the MCP mechanical structure and the multiplication process are the possibility to replace high resistivity lead glass with other substrates in the construction of MCP [105], or even convert photons into photoelectrons directly within the MCP pores [96]. ALD coated MCP-PMT can sustain an accumulated charge up to few C/cm^2, more than one order of magnitude greater than standard MCP, without significant effect on the timing performance. Despite this, there are signs of side effects of the ALD process regarding the low current saturation levels and the recovery time of high gain [90, 106]. In particular, it was observed that multianode MCP-PMT exposed to high photon rates to induce saturation for a brief period of time takes significantly more time to recover from the gain loss when ALD coating is present, but more in-depth studies are needed.

2.3 Setup & Analysis

The study presented in this chapter was performed in view of possible employment of MCP-PMT in the future upgrade of the LHCb RICH detector. The specifications that an MCP photomultiplier should have to meet the requirements imposed by the future RICH operating conditions are: an effective pixel size of $1 \times 1\,\mathrm{mm}^2$ and a time resolution better than $\sim 100ps$ RMS per photon, to reduce the photon occupancy per pixel and mitigate the combinatorial backgrounds; dark-count rate (DCR) well below 100kHz/ mm^2, read-out at 40MHz and must be insensitive to residual magnetic fields from the dipole magnet (3mT). They need to be able to operate with a photon rate of $\sim$10MHz/ mm^2 and withstand radiation levels of about $10^{13}\mathrm{n}_{eq}/\mathrm{cm}^2$ (5kGy for a total integrated luminosity of $350\,\mathrm{fb}^{-1}$) [85].

MCP pore diameter	15 μm
MCP pitch	19 μm
MCP pore length:diameter ratio	47:1
Photocathode - MCP gap	1.6 mm
MCP - Anode gap	3.0 mm
active area	59×59 mm^2
anode pitch	0.828 mm

TABLE 2.1: Mechanical specifications of the Auratek MCP-PMT.

A comparison between the timing performance at high photon rate in single photon regime of the current MaPMT installed and other MCP-PMT has been previously performed [102]. It suggests that the deterioration of time resolution due to recovery time experienced at high photon rates could be mitigated by operating at low gain. The following section will describe the instrumentation and analysis procedure used to study the time resolution of the Auratek-Square (model number MaPMT253, produced by Photek) in single photon events.

2.3.1 Setup description

This multianode photomultiplier tube utilizes a dual microchannel plate to provide electrons multiplication with optimal time resolution. The two MCP slabs are arranged in a Chevron (v-shape) configuration to reach a gain of $\mathcal{O}(10^6)$. The mechanical properties of the MCP-PMT are given in Table 2.1.

The setup adopted for the measurements is represented by the block scheme in Figure 2.12 and described in details in this paragraph.

The anode native configuration of 64 $\times$ 64 is readout with a custom PCB that groups the 4096 individual anodes into an 8 $\times$ 8 grid using 4x Samtec TMM-116-01-L-D-SM, 16 channels per connector, resulting in 64 “pixels" of 6.62 mm per side. The connection of MCP to the anode output is made with an Anisotropic Conductive Film (ACF). Each pixel's signals are amplified by LMH6702 current feedback operational amplifier, mounted on readout cards shown in Fig. 2.13 right side, which operate as fast integrator with $C_F \simeq 3$ pF (internal) and $R_F = 1k\Omega$ [107]. The signals at the output of the amplifier had a rise time $t_r = 1.5$ ns and a fall time $t_f \simeq 10$ ns. The measured gain, halved by the 50 Ω termination, calibrated by injecting a known charge through a test capacitor, was 27.5 mV/Me^-. The baseline noise, as seen at the oscilloscope, was 0.25 mV RMS, or σ_Q = 9 ke^- RMS, the main contributor being the current noise at the inverting input of the operational amplifier (18.5 pA/$\sqrt{Hz}$).

The source of pulsed light is a Hamamatsu PLP-10 (C10196 controller and M10306-29 head), which provides laser pulses of 70 ps FWHM and 405 nm wavelength. The amplitude of the laser pulses was kept at a medium setting on the PLP-10 controller (knob set at 9), since in a previous study a larger pulse duration was observed at lower settings, and a smaller secondary pulse delayed by about 200 ps was observed at higher settings [102].

The amplified anode signals are routed to the oscilloscope Rohde & Schwarz RTO1044 (4 GHz, 20 GS/s), digitally low-pass filtered at 300 MHz (to eliminate high frequency environmental disturbances, without reducing the signal bandwidth) and acquired. Pixels not readout are connected to ground. The electronic noise contribution to the measured time resolution of the MCP-PMT, which also includes the laser contribution, can be estimated as $\sigma_t = t_r \sigma_Q / Q$, where σ_Q is the equivalent noise

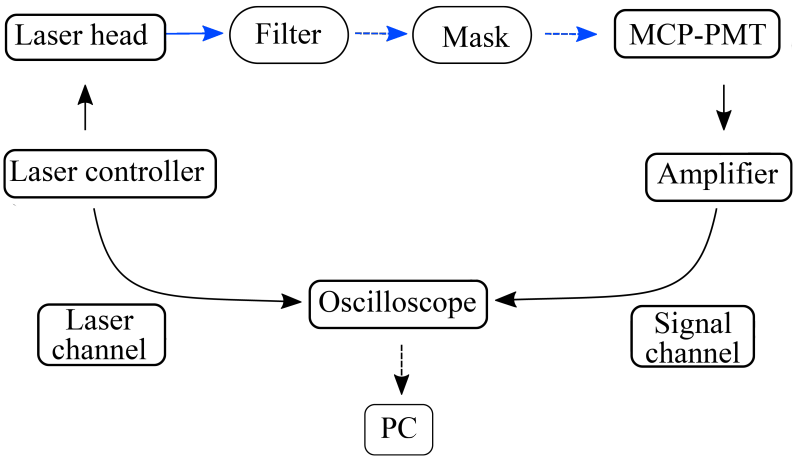


FIGURE 2.12: Block schematic of the test setup. The black arrows represent electrical signals, the blue arrows represent light.

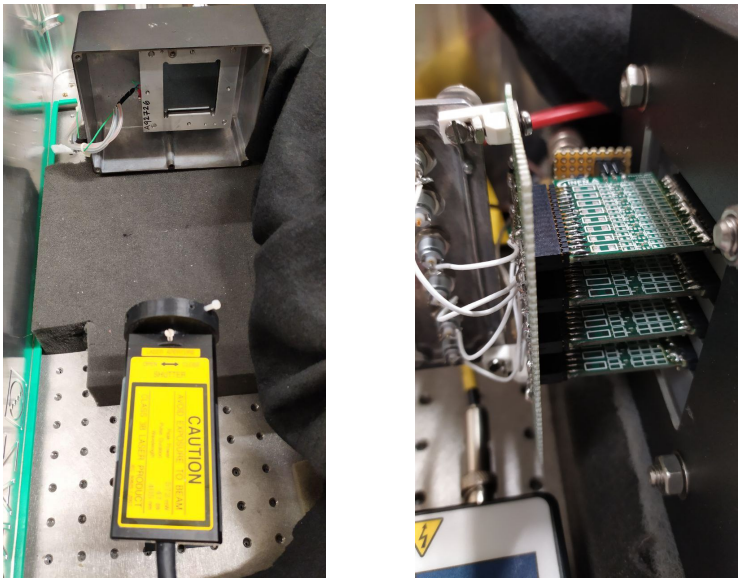


FIGURE 2.13: (Left) Picture of the laser head and MCP-PMT, both placed inside a large shielded enclosure with hermetic latch to isolate from external light. (Right) Readout cards that amplify the signals from the pixels of interest and ground the others.

charge and Q is the charge carried by the signal. It can be considered sufficiently small for signals above 5×10^5 e^-, being < 30 ps RMS.

The oscilloscope was set to trigger on the sequence (within a 200 ns window) of a signal from the photodetector, detected with a configurable threshold set above electronic noise, and "sync out" signal from the laser controller, which is delayed by 20 ns and provides also the time reference for each signal. This configuration avoids the acquisition of many empty signals while working in single photon regime. Figure 2.14 shows a typical event, with the digital "sync-out" on the top part, and the amplified analog signal from the MCP on the bottom.

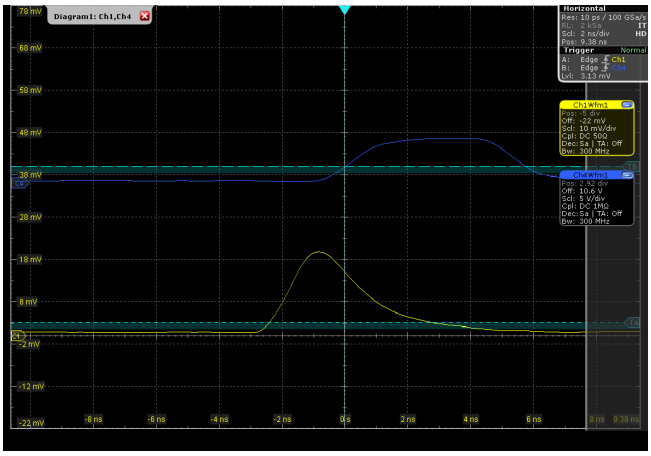


FIGURE 2.14: Screenshot from the oscilloscope that shows the delayed laser sync-out signal (blue) and below the amplified MCP-PMT signal (yellow). Horizontal scale is 2 ns/div.

The photon rate hitting the detector was varied by changing the laser pulse repetition rate ν_L from a few Hz up to 100 MHz. To operate in single photon regime, Thorlabs AR-coated absorptive neutral density filters are placed in front of the laser head, with attenuation factors ranging from 10^3 to 10^5 (NE30A-A to NE50A-A). The filters will be denoted by F3 to F5 in the following, where the number indicates the optical density, or absorbance, of the filter. A plastic mask with 1 mm diameter holes was used to illuminate only a portion of the pixels (the center and/or the corners), when needed.

In a counting experiment, the probability of having n events when the mean is μ is given by the Poisson distribution:

$$P(n,\mu) = \frac{\mu^n e^{-\mu}}{n!} \tag{2.2}$$

The fraction of non-empty events n over the total number of laser pulses n_{TOT} in test acquisition (where the oscilloscope was set to acquire the photodetectors signals synchronously with the laser pulses) corresponds to the probability of having at least one photon reaching the detector:

$$\frac{n}{n_{TOT}} = 1 - P(0) \quad \rightarrow \quad \mu = -log(1 - \frac{n}{n_{TOT}}) \tag{2.3}$$

The filters in front of the photodetector are chosen so that the actual rate of non-empty events ν_P was below $\nu_L/10$, which ensures a sufficient purity of single photon events ($> 95\%$) as can be seen from plots in Fig. 2.15.

2.3.2 Analysis strategy

The digitized signal waveforms are analysed with an offline algorithm to extract the timestamp of the moment where each of the waveforms crosses a threshold set at 15% of its amplitude (defined threshold-crossing time in the following). In Figure 2.16 is visible an example of analysed waveform. Defining a threshold proportional to the waveform amplitude is equivalent to performing constant fraction discrimination, thus avoids spurious contributions to time resolution due to amplitude walk. Typically, each acquisition consists of 10k events. The threshold crossing

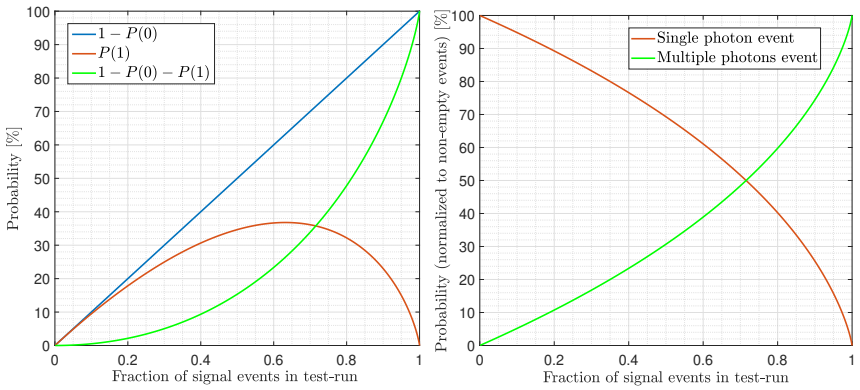


FIGURE 2.15: (Left) Probability of having one or more photons for each event. (Right) Single and multiple photons events probability

timestamps of a given set of waveforms (usually 10k) are collected in a 10 ps binned distribution. The transit time spread of the system (photodetector, amplifier, and laser) is thus evaluated with a fit to the histogram, and an example is shown in Figure 2.17.

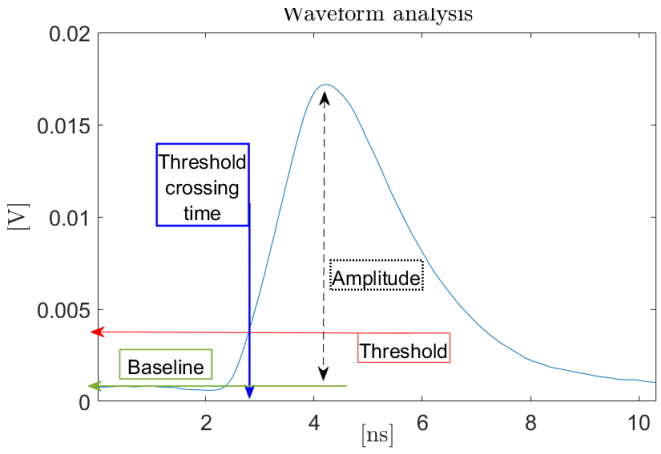


FIGURE 2.16: Sketch of an acquired waveform, where the interesting parameters are depicted. The threshold is set to 15% of the waveform amplitude.

Data are fitted with two Gaussian functions to accurately measure the FWHM of the distribution: the primary peak is followed by a second peak, caused probably by backscattered photoelectrons on the MCP input (some contribution from the laser cannot be excluded). The fit function is reported in Eq. 2.4. The iterative fit procedure follows four steps:

1. A single Gaussian fit is performed on all the binned data, to get an approximate mean (μ_1) and width (σ_1)

2. a second fit is done using a dedicated function f_2 with two Gaussians (Eq. 2.4), excluding points that don't satisfy the condition $|X - \mu_1| < 5 \cdot \sigma_1$. The error associated to each point is the square root of the corresponding bin content. From this second fit the parameters μ and σ are extracted.

3. The range $\mu \pm 5 \cdot \sigma$ is sampled into 1000 points $f_2(\tilde{x})$ to measure f_2 FWHM (indicated by the red line in Fig. 2.17).

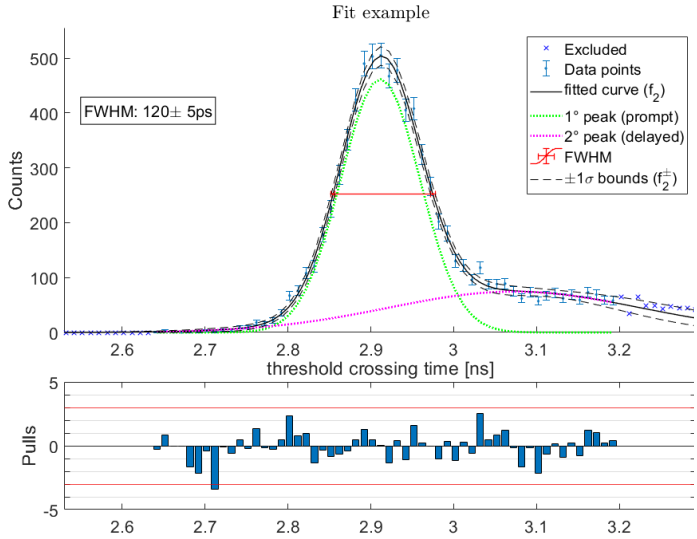


FIGURE 2.17: Example of the fit performed to the histogram of threshold crossing time. The green and purple Gaussian functions are combined to obtain the final fit function f_2, of which the FWHM is depicted in red.

4. $f_2^{\pm}$ are the functions obtained from the f_2 confidence intervals at $\pm 1\sigma$. They are used to computing FWHM$^+$ (which is an overestimate of FWHM) and FWHM$^-$ (which is an underestimate of FWHM), and half of the difference between them is the error associated to the measurement of FWHM.

$$f_2(x) = a \cdot \left(e^{-\frac{(x-\mu)^2}{2\cdot\sigma^2}} + b \cdot e^{-\frac{(x-(\mu+\delta))^2}{2\cdot\sigma_N^2}} \right) \tag{2.4}$$

2.4 Single pixel performance

The timing performance of one pixel was studied as function of the power supply voltage, the photon rate and fraction of illuminated area. The mask shown in Fig.2.18 was mounted in front of the detector, It has 5 holes for each pixel, arranged in a quincunx. For this first set of measurements, all holes are covered with black tape, except those corresponding to the studied pixel, labelled S3_4.

2.4.1 Varying bias voltage

The time resolution of the MCP-PMT was studied as a function of the total voltage difference between the photocathode and the anode (ΔV_{tot}). The ΔV_{tot} is divided into three different ΔV steps: the first between the photocathode and the MCP input (ΔV_{1-2}), the second between the input and the output of the MCP (ΔV_{2-3}) and the last between the MCP output and the anode (ΔV_{3-a}).

For these measurements the Photek SH01 voltage divider was used, which is built with the following resistors: R1 = 1 MΩ, R2 = 14.75 MΩ, R3 = 4.7 MΩ. The voltage division, shown in Fig. 2.19, results in:

$$\Delta V_{1-2} = 0.049 \times \Delta V_{tot} \qquad \Delta V_{2-3} = 0.721 \times \Delta V_{tot} \qquad \Delta V_{3-a} = 0.230 \times \Delta V_{tot}$$

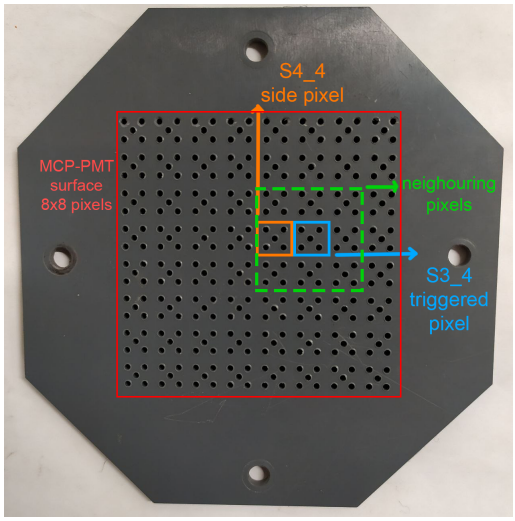


FIGURE 2.18: Plastic mask placed in front of the detector in dedicated measurements, in order to illuminate different areas of certain pixels. The most studied pixel is labelled S3_4, the neighbouring pixels are tested during the charge sharing studies.

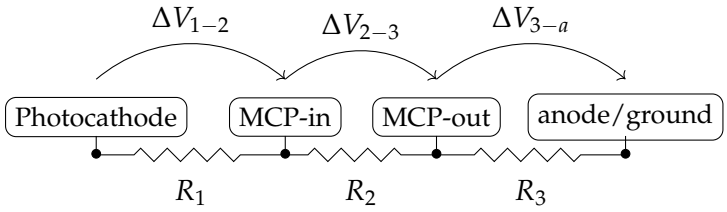


FIGURE 2.19: Schematic layout of the voltage divider. The total voltage drop between the photocathode and the anode is $\Delta V_{tot} = \Delta V_{1-2} + \Delta V_{2-3} + \Delta V_{3-a}$

Figure 2.20 shows the measured time resolution and the histogram of the collected charge at the anode illuminating S3_4. The Gain is defined as the average of the collected charge distribution (roughly corresponding to the peak of the distribution). It drifts towards larger values with the increase of ΔV_{tot}, as can be seen from the second Y axis of the bottom plot in Fig. 2.20. The time resolution is not very sensitive to changes of the total voltage ΔV_{tot}. Some deterioration starts to occur below 3050 V, where the gain reduces below 10^6.

Since the ageing of MCPs depends on the total extracted charge [108], it makes sense to look for a working point that minimizes the gain while preserving other performance parameters such as the time resolution. Various configuration of inter-stage voltages are tested using the Photek SH02 voltage distribution and the Triple Output Digital High Voltage Power Supply unit DPS3-6N6N6N, that allows to set different ΔV for each stage, in different ranges.

The effect on the time resolution of changes in the voltage difference between the photocathode and the input side of the MCP can be seen in Figure 2.21. Increasing ΔV_{1-2} improves the time resolution. However, after few tens of V, the time resolution tends to degrade. Comparing the threshold crossing time distribution at different ΔV_{1-2} it can be observed that the second peak becomes more relevant, resulting in a wider overall distribution, even though each of the two populations has a smaller width (Figure 2.22). This is most likely due to the contribution of delayed signals

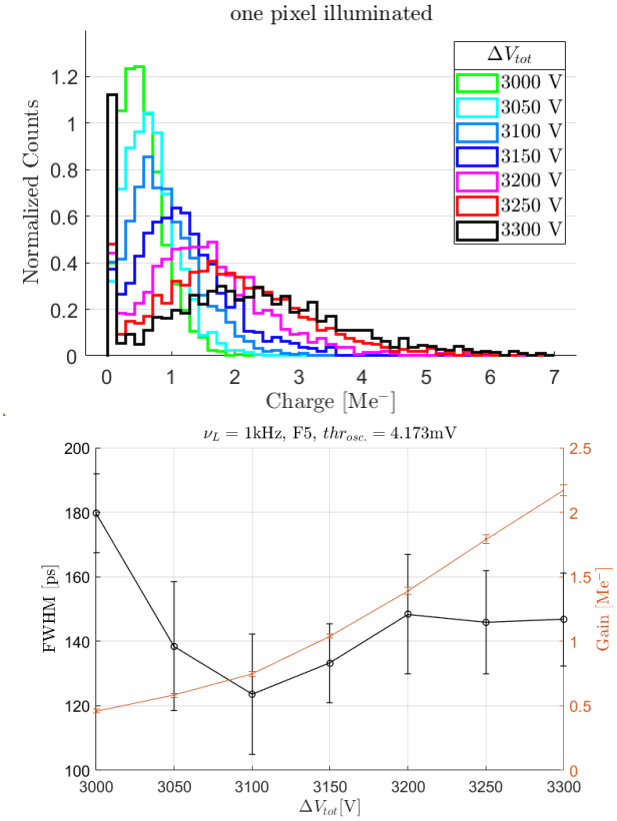


FIGURE 2.20: (Top) Distributions of the collected charge obtained with only one pixel illuminated through the 5 holes in the mask, for different total voltages. (Bottom) Measured FWHM and the corresponding gain at different total voltages.

generated by backscattered photoelectrons, whose arrival time will be closer to the one of prompt signals.

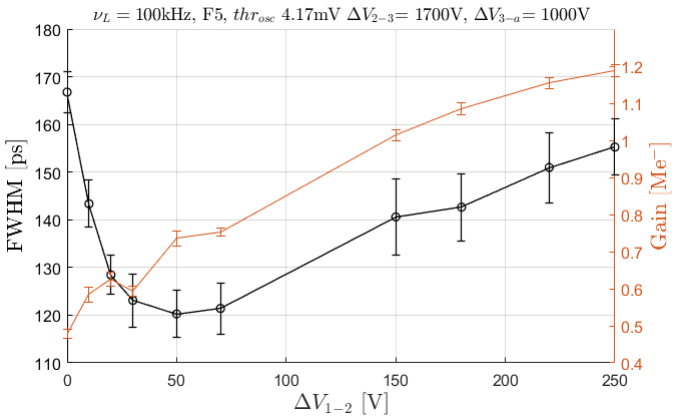


FIGURE 2.21: Time resolution and Gain measured with one pixel illuminated, as function of ΔV_{1-2} values.

Once the electron shower exits the MCP it is directed to the anode through ΔV_{3-a}. Changes on this voltage don't impact the time resolution, as can be seen from Figure 2.23. The $\sim$ 15% increment of the gain can be explained by an increase in the collection efficiency of the electrons. Two different values for ΔV_{1-2} are shown, and

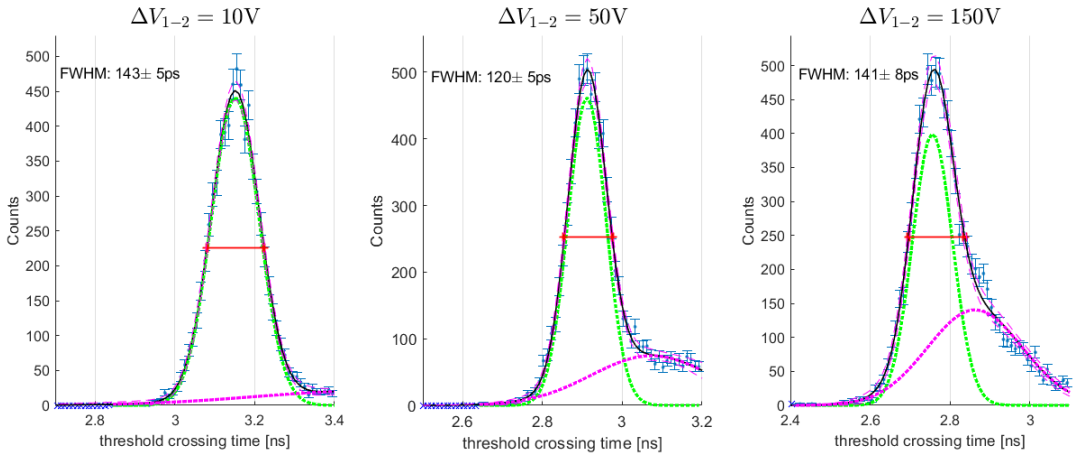


FIGURE 2.22: Effects of ΔV_{1-2} on the distribution of threshold crossing time. The right tail increases with the increment of voltage between the photocathode and the MCP input, resulting in a worse overall FWHM.

their trend is similar to what observed in Fig. 2.21. Increasing from 600 V to 1300 V reduces the average time to collect the electrons by ~ 0.1 ns.

The electron multiplication that happens within the MCP layers is controlled by ΔV_{2-3}. Figure 2.24 shows that the higher the voltage difference the higher the gain, while the time resolution improves up to 130 ps (ΔV_{3-a} =650 V) or 120 ps (ΔV_{3-a} =1300 V), but further increase in voltage above 1700 V doesn't affect the timing. For values of ΔV_{2-3} lower than 1650 V, the gain is lower than $5 \times 10^5 e^-$ and the contribution to the timing due to the amplifier's noise becomes predominant.

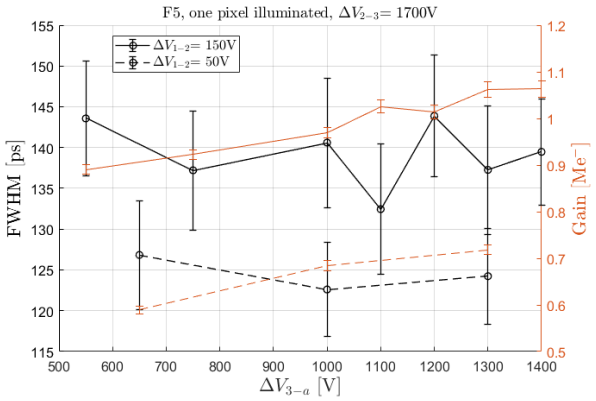


FIGURE 2.23: Time resolution and corresponding gain measured with one pixel illuminated, as function of ΔV_{3-a} values. ΔV_{2-3} was set to 1700V while two different values ΔV_{1-2} are tested.

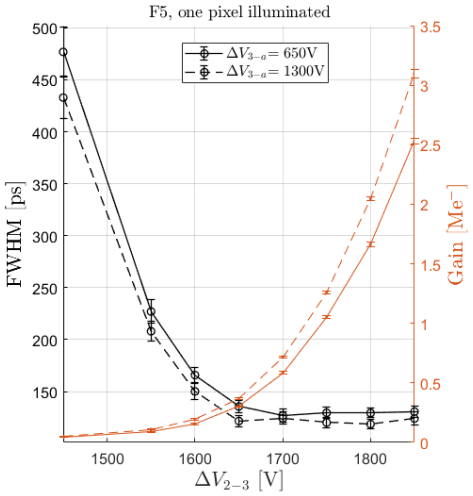


FIGURE 2.24: Time resolution and corresponding gain measured with one pixel illuminated as function of ΔV_{2-3} values. ΔV_{1-2} was set to 50V while two different values ΔV_{3-a} are tested.

2.4.2 Photon rate dependence

The response of the detector when only one pixel is illuminated was studied as function of the effective rate ν_P of incident photons on the photodetector. A time resolution of $\sim$130 ps is achieved in the range $10^{-2} - 10^2$ kHz/mm^2.According to the measurements described in the previous section, the optimal working point seems to be obtained powering the device with a total voltage of 2850V (divided into $\Delta V_{1-2} = 150$ V, $\Delta V_{2-3} = 1700$ V and $\Delta V_{3-a} = 1000$ V).

In order to obtain an effective rate of photons larger than $\sim$100 kHz it is necessary to use filters with lower attenuation factors, labelled F4 ad F3. The distributions of collected charge from Figure 2.25 shows how the gain peak is shifted towards lower values as the photon rate exceeds 100 kHz/mm^2. The saturation of the collected charge is reflected on the degradation of the timing performance.

As mentioned before, the saturation is related to the electron multiplication stage. The contribution to the saturation of each interstage voltage is shown in Figure 2.26. Increasing ΔV_{1-2} or ΔV_{2-3} leads to larger signals in comparison with the reference configuration because in the first case the photoelectron that starts the shower is more energetic, while in the second case more electrons are extracted by secondary emission.

The degradation of the timing performance is similar in both cases, whilst the gain increment is more significant with the latter. In view of this, a lower gain effectively mitigates the degradation of the timing performance observed at high rate, reducing the recovery time needed to recharge the glass channels of the MCP. Nonetheless, at very low voltages the time resolution is degraded again, as visible in Fig. 2.21 and Fig. 2.24. It must be considered that operating at small gain the electric noise contribution to the time resolution is not anymore negligible. The recovery time of the MCP can be influenced negatively by the presence of ALD coating, a procedure that is usually employed to extend the devices' lifetime [90].

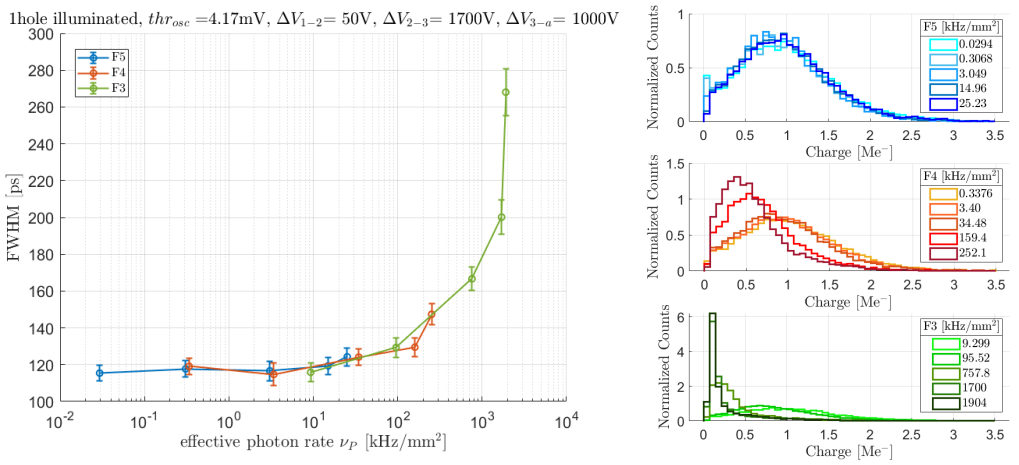


FIGURE 2.25: (Left) FWHM time resolution as function of the effective photon rate ν_P, obtained using different absorptive filters. (Right) Collected charge spectra, for the three filters used. The distortion of the distributions caused by the saturation can be seen clearly for rate above 100kHz/mm^2.

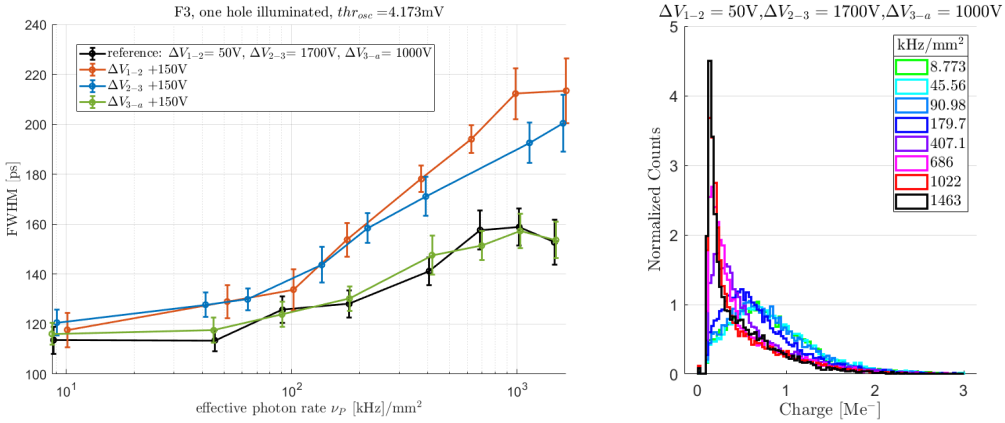


FIGURE 2.26: FWHM time resolution as function of the effective photon rate ν_P. Each ΔV step was increased separately by +150V with respect to the reference configuration, and the corresponding collected charge spectra are shown in the bottom plots. Increasing the ΔV after the multiplication stage doesn't affect the timing performances, while a similar degradation is observed for the other two changes.

2.5 Charge Sharing studies

To study charge sharing effects, the tape was removed from the mask in front of the detector in order to illuminate all the pixels at the same time. In Figure 2.27 the spectra taken in different conditions are compared. The gain peak visible when a single pixel was illuminated, through one or five holes (green histograms), disappears when more pixels are illuminated. A similar distortion of the charge spectra is observed when all pixels or just two neighbouring pixels are illuminated (S3_4 and S4_4, dashed histogram). The reason of this distortion is the charge sharing, as will be described in this section.

The time resolution measured when more than one pixel is illuminated increases up to about 170 ps, as shown in Figure 2.28, even though the bias voltage and all the

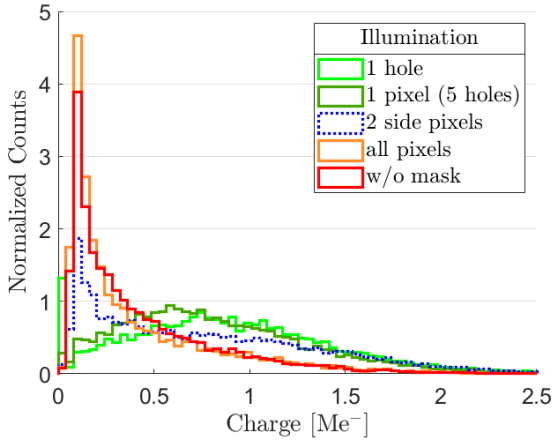


FIGURE 2.27: Comparison of collected charge for different illuminated area of the detector, with the same bias voltage (ΔV_{1-2} = 50 V, ΔV_{2-3} = 1700 V and ΔV_{3-a} = 1000 V). The gain peak is not visible when more than one pixel is illuminated by laser light.

other parameters are the same.

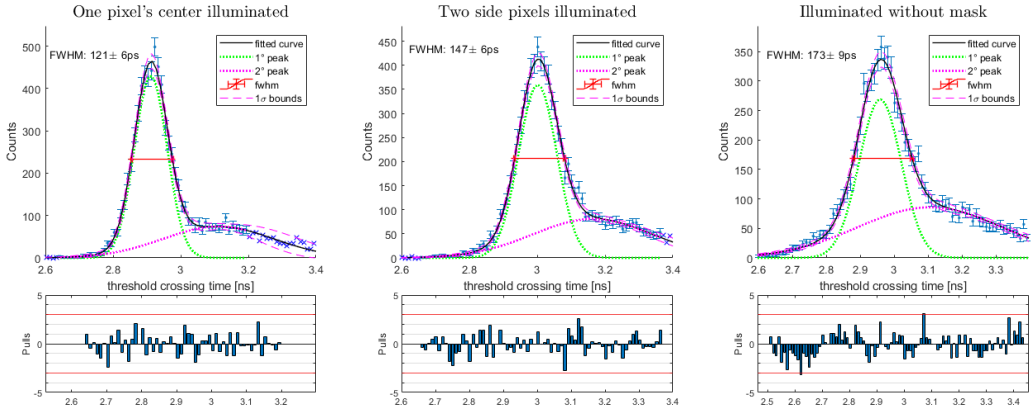


FIGURE 2.28: Fit to threshold crossing time distribution in different configuration of illumination from Fig.2.27. The presence of crosstalk signals lead to a decrement of the timing performance.

In order to understand the causes of the modification to the collected charge, an in-depth study of the charge sharing between the pixels is carried out. The outputs of two neighbouring pixels are acquired simultaneously, but only one (S3_4) was illuminated by the laser through 5 holes in the mask, the holes over S4_4 are covered with black tape, as all the others. The acquisition trigger setup remains unchanged.

The distribution of waveforms' amplitudes are reported in Figure 2.29. The left plot shows three regions where the waveforms amplitudes have a different correlation. This is due to the layout of the holes in the mask that cover the pixels: the closer is the illuminated point to the neighbouring pixel and the greater is the correlation between the amplitude of the acquired signals. The charge sharing happens only between neighbouring pixels, indeed no correlation is observed between non-consecutive pixels, as can be seen from the right plot of Figure 2.29. Illuminating all the pixels except for the first neighbours of the triggered pixel doesn't affect S3_4 nor S4_4.

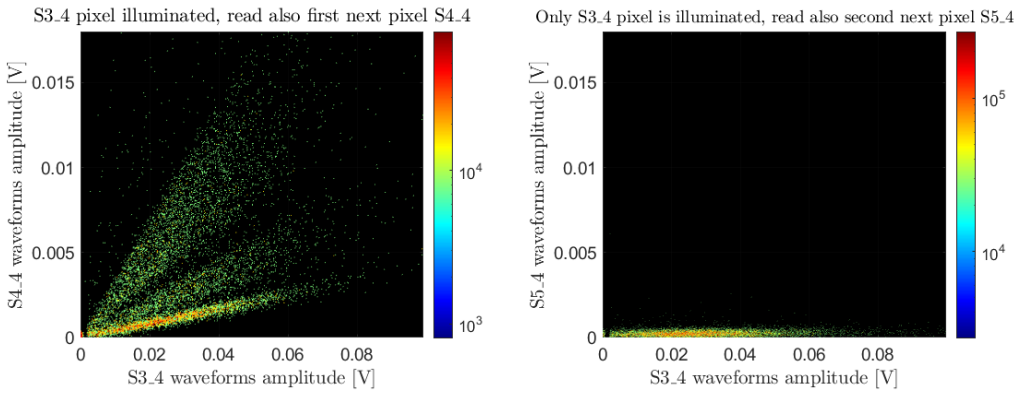


FIGURE 2.29: Waveforms amplitude of two pixels (S3_4 is the pixel on which the oscilloscope trigger the acquisition), in different configuration of illumination. (Left) waveforms acquired from two consecutive pixels. (Right) waveforms acquired from triggered pixel and next side pixel. The induced signal appears only between neighbouring pixels.

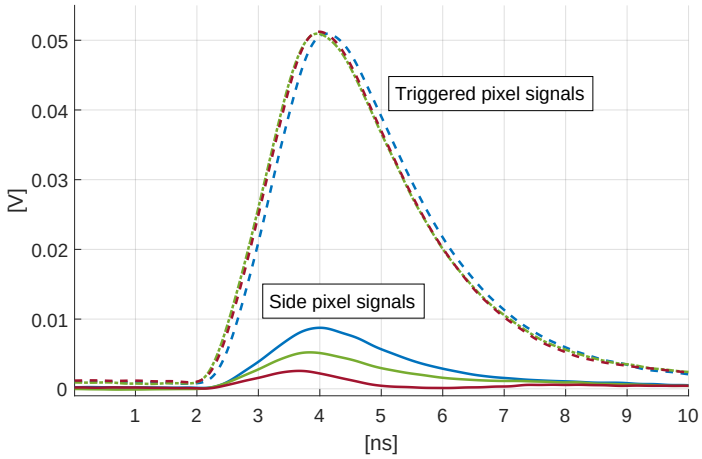


FIGURE 2.30: Sketch of pairs of waveforms acquired from the three regions of Figure 2.29. Given signals with the same amplitude on the Triggered pixel, the signal acquired simultaneously from the Side pixel depends on the photon interaction point: the closer it is to the Side pixel the higher is the amplitude.

When two neighbouring pixels are readout and both illuminated by the laser light, two additional groups of events are visible on the scatterplot of waveforms' amplitudes. The *C/D* events on Figure 2.31 represent events caused by photons incident on the triggered pixel (similar to what it's shown in Fig. 2.29). Group *B* represents events where photons reached the side pixel through the holes closer to the triggered pixel. Group *A* represents events where the photon hit the side pixel far from the triggered pixel, but the induced signal was sufficiently high to trigger the acquisition. It is possible to distinguish two tiny bands within the events of group A, which corresponds to the different distance between the central and external holes on the side pixel from the triggered pixel. For events in group D the two bands are not visible, due to the high density of events in the plot, but they were previously observed in Figure 2.29.

Figure 2.32 shows the amount of charge sharing between pixels S3_4 and S4_4,

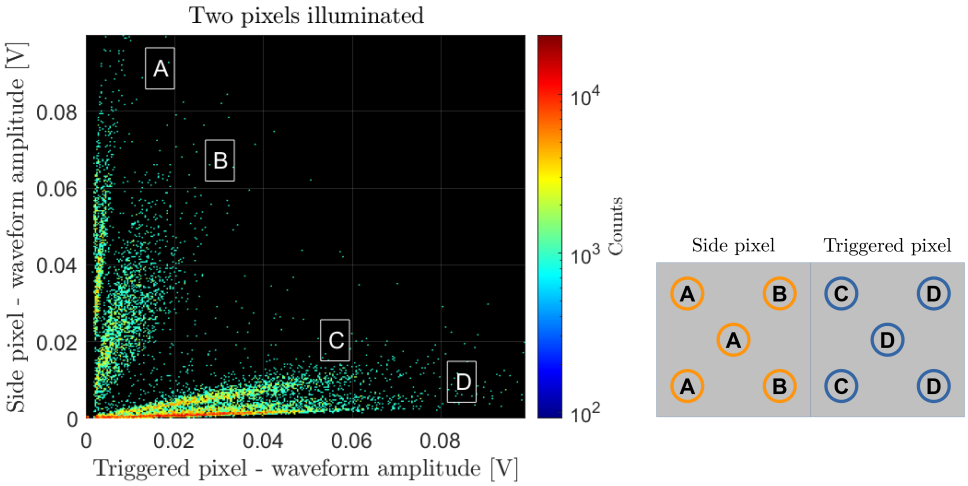


FIGURE 2.31: (Left) 2D distribution of waveforms amplitude of two neighbouring pixels illuminated through 5 holes in the plastic mask mounted in front of the detector. Four different regions can be identified: in A/D photons hit Side/Triggered pixel center and corner holes far away from Triggered/Side pixel, while in B/C photons hit corner holes close to the Triggered/Side pixel. (Right) Scheme illustrating the corresponding holes on the mask of the different regions

measured when the device is powered at 3050V (with ΔV_{1-2} = 50 V, ΔV_{2-3} = 1700 V and ΔV_{3-a} = 1300 V). This quantity is estimated by the slope of a linear fit between the amplitudes of the waveforms acquired from both pixels, when only one hole of the quincunx is illuminated. The fraction of charge sharing decreases with the distance of the illuminated spot from the common border of the pixels (it is ∼7% when the center of the side pixel is illuminated).

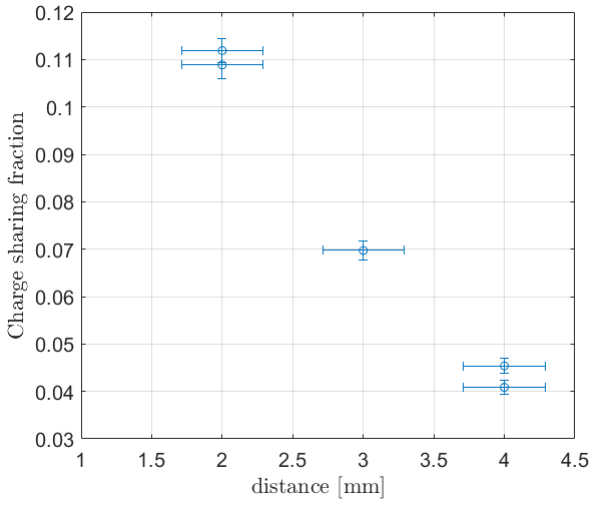


FIGURE 2.32: Fraction of collected charge between neighbouring pixels S3_4 and S4_4, for illuminated spots of the quincunx at different distance from the common border between the pixels.

Figure 2.33 shows the different contributions (*A*, *B*, *C/D*) to the total distribution of waveforms acquired through the triggered pixel S3_4. The gain peak that was visible when only one pixel is illuminated is present, but masked by the charge sharing

events.

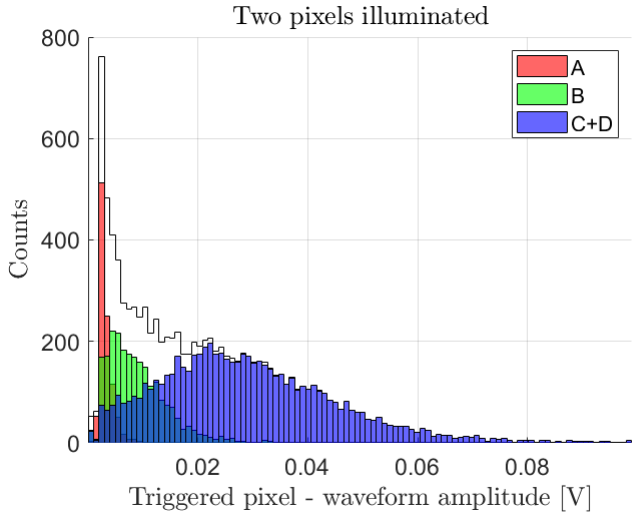


FIGURE 2.33: Distribution of waveforms' amplitude of the triggered pixel (S3_4) obtained illuminating also a side pixel (S4_4). Selecting the labelled region from Fig 2.31 with cuts on waveforms' amplitudes the single photon peak is once again visible (C+D). The contribution of the induced signals from the side pixel (A,B) is clearly appreciable.

2.6 Conclusion

In this chapter the timing performance of multianode MCP photomultiplier Auratek-Square MaPMT253 are presented. The device exhibits an excellent time resolution of ∼120 ps FWHM (∼100 ps FWHM subtracting the laser contribution) at low rate and without charge sharing. When the entire active surface is illuminated, the resolution degrades because of the induced crosstalk from the first neighbouring pixels. The maximum sustainable rate before saturation at nominal gain is just below 100 kHz/mm^2. The main effect of saturation is on gain and efficiency, since the time resolution degrades by just about 30% up to 1 MHz/mm^2.

The LHCb RICH system is currently mounting MaPMT detectors with a time resolution of ∼400 ps FWHM, which is stable up to 10 MHz/mm^2 [102]. Operation of MCP-based photodetectors as single photon counters at MHz/mm^2 and higher rates will likely require using microchannels of smaller diameter and lower resistivity, in order to reduce the hit probability and recharge time of each microchannel. Adjusting the interstage voltages (between the photocathode and the MCP input, or within the MCP) can mitigate the degradation of timing capabilities of the detector. The amplitude of the signals induced on adjacent pixel depends on the distance between the impact point of the photon and the center of neighbouring pixels.

Charge sharing between neighbouring pixels, if not accounted for, constitutes a crosstalk which deteriorates performances to ∼170 ps FWHM (∼ +60%). It is usually possible to exploit charge sharing to improve the detector's spatial resolution beyond the pitch of the anode pads, but in the actual RICH system the readout rate and the throughput are too high to permit more than a simple binary output.

As previously described in Section 2.1.2 the LHCb detector will operate at a nominal instantaneous luminosity of 1.5×10^{34} cm^{-2}s^{-1} during the high luminosity phase, which will be 7.5 times higher than what expected in Run3. Assuming

a photon rate of $\sim$7.5MHz/mm^2 due to increased luminosity, the charge collected by photodetectors operating with a gain factor of 10^6 will be $7.5\times10^{14}e^-$ / mm$^2\cdot s$ (1.2×10^{-4} C/ cm$^2\cdot s$). The estimated lifetime of the Auratek-Square MCP-PMT is about 5 C/mm^2, which would corresponds to $\sim$12 hours of continuous operation. This extremely short duration make the use of MCP-PMT critical in such operating conditions. To consider the use of MCP based devices, the future RICH photomultipliers will need to work at much lower gain ($10^3 - 10^4$). Any improvement in technology that could extend the ageing and rate capability of MCP-based photodetectors will be much sought-after.

Conclusions

The results of my research during the PhD have been presented in this thesis. Chapter 1 contains the results of the analysis performed on data collected during Run2, searching for the decay of $\tau^+ \to \mu^+\mu^-\mu^+$. This process is prohibited in the Standard Model because the lepton flavour is an accidental symmetry. It is allowed in the minimal extended SM, but with extremely low branching ratio. The expected upper limit obtained from the analysis of the 3μ sample in Run2 corresponds to

$$\mathcal{B}(\tau^+ \to \mu^+\mu^-\mu^+) \leq 1.8(2.2) \times 10^{-8} \text{ @ } 90\%(95\%) \text{ C.L.}$$

and it is expected to improve even further when it will be combined with the upper limit from Run1.

Chapter 2 contains the characterization of the time resolution of a multianode MCP-PMT detector (Auratek-Square produced by Photek). In single photon regime, it exhibits a transit time spread of $\sim$100 ps FWHM (42 ps RMS) when only a single pixel is illuminated and the photon rate is below 100kHz/mm^2. The increment of rate and illuminated area deteriorate the timing performance (due to saturation and charge sharing). Decreasing the gain mitigates the deterioration of time resolution, but to contrast the ageing due to the integrated charge it should go down to $\mathcal{O}$ (10^3-10^4). This is the most critical yet interesting challenge to the employment of MCP-based photomultipliers for the future RICH upgrade.

Validation of Binned maps for Trigger efficiencies correction

An additional control channel was used to cross-check the evaluation of the TIS-TOS efficiencies that are used to compute the correction factor R as described in Section 1.8.2. The chosen channel is the $B^+ \to J/\psi(\mu^+\mu^-)K^+$. The muons' tracks in the final state are characterized by a greater transverse momentum, and thus the distribution of the events in the binned IPχ^2 and p_T is different from the signal or reference channel ones, as can be seen from Figure A.1

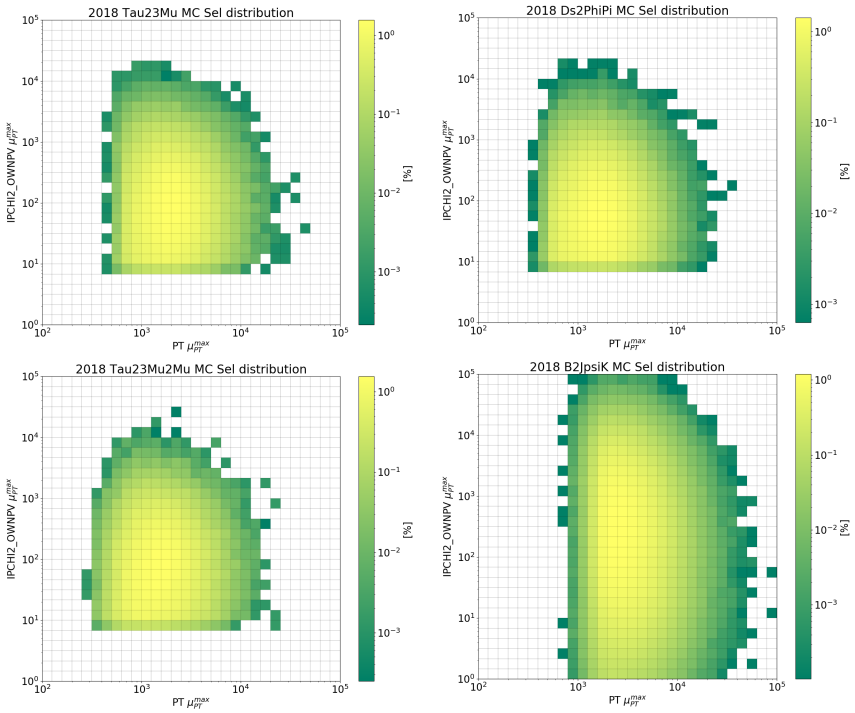


FIGURE A.1: p_T and IPχ^2 distribution of simulated events from $\tau \to \mu\mu\mu$ signal channel 3μ sample (top-left), $\tau \to \mu\mu\mu$ signal channel 2μ sample (bottom-left), $D_s \to \phi(\mu\mu)\pi$ reference channel (top-right) and $B^+ \to J/\psi(\mu^+\mu^-)K^+$ control channel (bottom-right)

For this control channel, the ratio between the trigger efficiency evaluated with the TISTOS method and the true efficiency in the simulated sample is closer to unity already without any binning. Figure A.2 shows this ratio as function of the number of bins of p_T and IPχ^2, and it results more flat when compared to signal or reference channel, especially at smaller number of bins.

The trigger efficiency corrector factor R could be computed convoluting the signal channel MC with the $B^+ \to J/\psi(\mu^+\mu^-)K^+$ channel, but given the differences

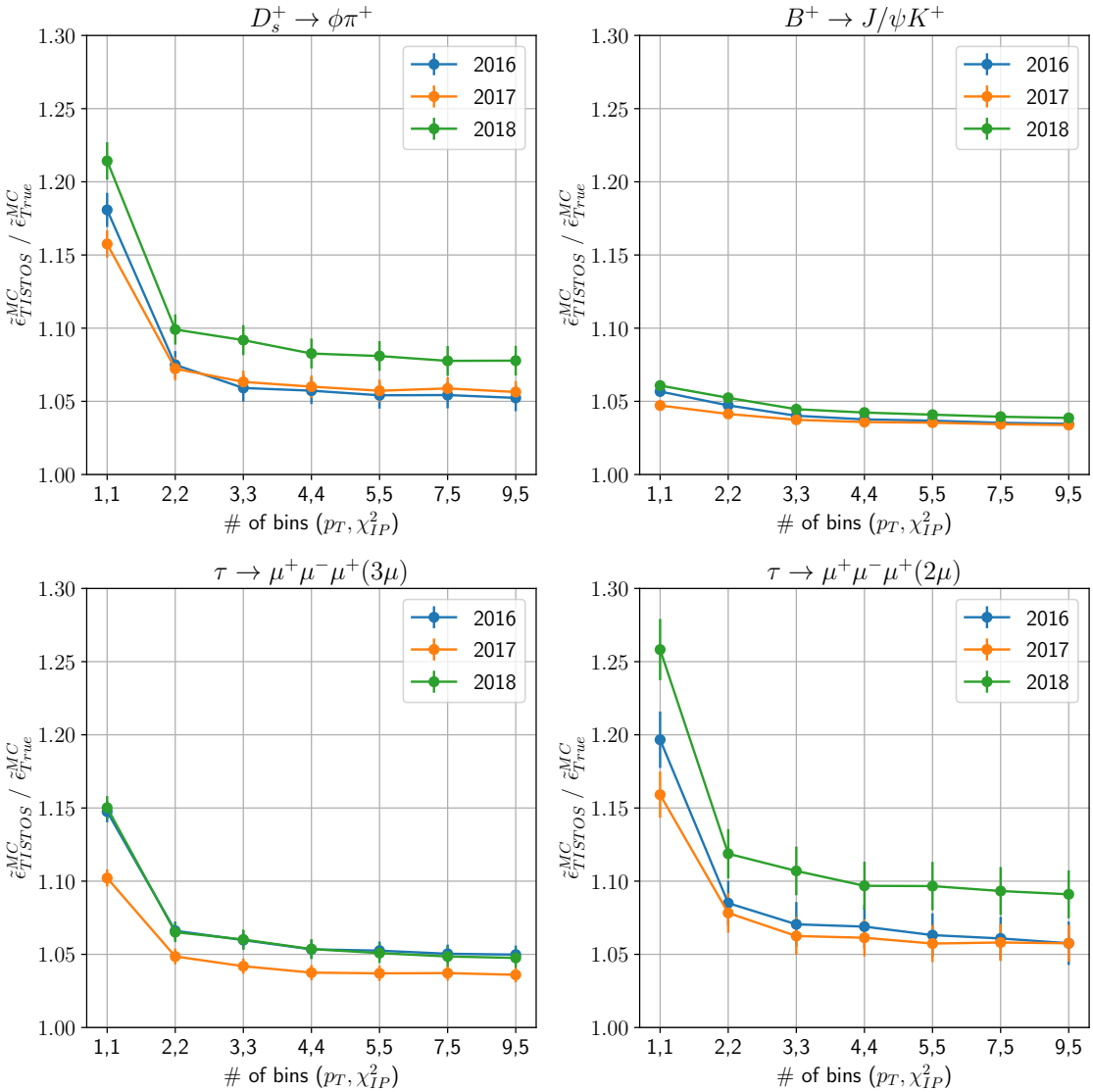


FIGURE A.2: TISTOS efficiency over true trigger efficiency computed in MC up to L0&HLT1, as function of the number of bins used in the TISTOS method. The (1,1) configuration correspond to not divide the region in bins, but consider the whole sample within the edges as in a single bin.

in the kinematic distribution it was chosen to use the reference channel $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$.

Binary XGBoost classifier alternatives

Within the training dataset for the Signal/Background classifier, there are five different sources of τ. The distributions of the kinematic and topological variables used by the model to classify the events shows some differences, that can be exploited by a more complex multiclass model to further separate signal from background.

The signal subsample is divided into three different subsets: *prompt* and *secondary* events (either from D_s^+ or D^+) and *prompt b* events, while the background subsample is composed by events in the middle sidebands around the blind region. The multiclass model is configured to distinguish between four different classes in total during the training phase. In the prediction phase, the three different signal classes are combined, in order to have a binary output. The probability of being signal-like is computed as one minus the probability of being background-like. Table B.1 shows the performance obtained with the multiclass model, compared to the standard binary one. Both models are trained using 5 folds on the 2017 dataset. The first uncertainty is statistical, while the second term is the standard deviation of the metrics obtained by the 5 different models.

3μ Signal channel [%]	**binary**	**multiclass**
Accuracy	90.71 $\pm$0.03 $\pm$0.20	90.66 $\pm$0.08 $\pm$0.18
ϵ_S	90.14 $\pm$0.05 $\pm$0.39	89.84 $\pm$0.14 $\pm$0.31
AUC	96.97 $\pm$0.00 $\pm$0.10	96.96 $\pm$0.00 $\pm$0.11
1-ϵ_B	91.32 $\pm$0.02 $\pm$0.17	91.50 $\pm$0.07 $\pm$0.16
Precision	91.21 $\pm$0.03 $\pm$0.19	91.35 $\pm$0.08 $\pm$0.20
ϵ_S [ϵ_B@5%]	84.24 $\pm$0.00 $\pm$0.49	84.24 $\pm$0.00 $\pm$0.67
1-ϵ_B [ϵ_S@85%]	94.65 $\pm$0.00 $\pm$0.22	94.65 $\pm$0.00 $\pm$0.36

TABLE B.1: Performance comparison between the binary and multiclass XGBoost classifier (2017 dataset). The first uncertainty is statistical, while the second one is the standard deviation of the scores obtained by the different k-fold models.

The multiclass model performs differently on each class, but the final efficiency on the Signal/Background classification task is very close to the one obtained using the standard binary model, whilst taking more computing time in the training phase.

Confusion Matrix [%]	**Prediction**			
True	BKG	Prompt	Secondary	Prompt b
BKG	93.11	04.52	00.63	01.74
Prompt	03.94	94.92	00.45	00.69
Secondary	26.92	38.44	18.94	15.70
Prompt b	45.51	20.21	10.48	23.81

TABLE B.2: Confusion Matrix of the multiclass model, showing the fraction of events correctly reconstructed or mislabelled. The model is trained on the 90% of the 2017 dataset and tested on the 10% left.

Bibliography

[1] Emmy Noether and M A Tavel's English. "Invariante Variationsprobleme". In: *Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math-phys. Klasse* 1.3 (1918).

[2] Sheldon L. Glashow. "Partial-symmetries of weak interactions". In: *Nuclear Physics* 22.4 (Feb. 1961), pp. 579–588. ISSN: 0029-5582. DOI: 10.1016/0029-5582(61)90469-2.

[3] Steven Weinberg. "A Model of Leptons". In: *Physical Review Letters* 19.21 (Nov. 1967), p. 1264. DOI: 10.1103/PhysRevLett.19.1264. URL: https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.19.1264.

[4] Abdus Salam. "Weak and electromagnetic interactions". In: (May 1994), pp. 244–254. DOI: 10.1142/9789812795915{_}0034.

[5] Andrei D Sakharov. "Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe". In: *Soviet Physics Uspekhi* 34.5 (May 1991), p. 392. ISSN: 0038-5670. DOI: 10.1070/PU1991V034N05ABEH002497. URL: https://iopscience.iop.org/article/10.1070/PU1991v034n05ABEH002497.

[6] Patrick Huet and Eric Sather. "Electroweak Baryogenesis and Standard Model CP Violation". In: *Physical Review D* 51.2 (Apr. 1994), pp. 379–394. DOI: 10.1103/physrevd.51.379. URL: https://arxiv.org/abs/hep-ph/9404302v1.

[7] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Physical Review Letters* 81.8 (Aug. 1998), p. 1562. DOI: 10.1103/PhysRevLett.81.1562. URL: https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.81.1562.

[8] SNO Collaboration. "Measurement of the rate of nu_e + d –> p + p + e- interactions produced by 8B solar neutrinos at the Sudbury Neutrino Observatory". In: *Physical Review Letters* 87.7 (June 2001), pp. 71301–1. DOI: 10.1103/physrevlett.87.071301. URL: https://arxiv.org/abs/nucl-ex/0106015v2.

[9] K. Eguchi et al. "First Results from KamLAND: Evidence for Reactor Antineutrino Disappearance". In: *Physical Review Letters* 90.2 (Jan. 2003), p. 021802. DOI: 10.1103/PhysRevLett.90.021802. URL: https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.90.021802.

[10] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. "Remarks on the Unified Model of Elementary Particles". In: *Progress of Theoretical Physics* 28.5 (Nov. 1962), pp. 870–880. ISSN: 0033-068X. DOI: 10.1143/PTP.28.870. URL: https://academic.oup.com/ptp/article/28/5/870/1858382.

[11] Rabindra N Mohapatra and Palash B Pal. "Massive Neutrinos in Physics and Astrophysics". In: World Scientific Lecture Notes in Physics 41 (Mar. 1991). DOI: 10.1142/1246. URL: https://www.worldscientific.com/worldscibooks/10.1142/1246.

[12] Patrick Blackstone, Matteo Fael, and Emilie Passemar. "$\tau \to \mu\mu\mu$ at a rate of one out of 1014 tau decays?" In: *The European Physical Journal C 2020 80:6* 80.6 (June 2020), pp. 1–6. ISSN: 1434-6052. DOI: 10.1140/EPJC/S10052-020-8059-7. URL: https://link.springer.com/article/10.1140/epjc/s10052-020-8059-7.

[13] Kristopher J. Healey, Alexey A. Petrov, and Dmitry Zhuridov. "Erratum: Nonstandard neutrino interactions and transition magnetic moments [Phys. Rev. D 87, 117301 (2013)]". In: *Physical Review D* 89.5 (Mar. 2014), p. 059904. DOI: 10.1103/PhysRevD.89.059904. URL: https://journals.aps.org/prd/abstract/10.1103/PhysRevD.89.059904.

[14] G. Cvetič et al. "Lepton flavor violation in tau decays". In: *Physical Review D* 66.3 (Aug. 2002), p. 034008. DOI: 10.1103/PhysRevD.66.034008. URL: https://journals.aps.org/prd/abstract/10.1103/PhysRevD.66.034008.

[15] S. T. Petcov et al. "Charged lepton flavor violating decays: leading logarithmic approximation versus full RG results". In: *Nuclear Physics B* 676.1-2 (Jan. 2004), pp. 453–480. ISSN: 0550-3213. DOI: 10.1016/J.NUCLPHYSB.2003.10.020.

[16] Athanasios Dedes, John Ellis, and Martti Raidal. "Higgs-mediated Bs,d0$\to \mu\tau$,eτ and $\tau \to 3\mu$,e$\mu\mu$ decays in supersymmetric seesaw models". In: *Physics Letters B* 549.1-2 (Nov. 2002), pp. 159–169. ISSN: 0370-2693. DOI: 10.1016/S0370-2693(02)02900-3.

[17] Ernest Ma. "Theoretical expectations for rare and forbidden tau decays". In: *Nuclear Physics B - Proceedings Supplements* 123 (July 2003), pp. 125–128. ISSN: 0920-5632. DOI: 10.1016/S0920-5632(03)80316-X.

[18] Tai-Fu Feng et al. "Lepton dipole moments and rare decays in the CP-violating MSSM with nonuniversal soft-supersymmetry breaking". In: *Physical Review D* 68.1 (July 2003), p. 016004. DOI: 10.1103/PhysRevD.68.016004. URL: https://journals.aps.org/prd/abstract/10.1103/PhysRevD.68.016004.

[19] K. Hayasaka et al. "Search for lepton-flavor-violating τ decays into three leptons with 719 million produced τ+τ- pairs". In: *Physics Letters B* 687.2-3 (Apr. 2010), pp. 139–143. ISSN: 0370-2693. DOI: 10.1016/J.PHYSLETB.2010.03.037.

[20] B. Aubert et al. "Improved limits on the lepton-flavor violating decays τ-→l-l+l-". In: *Physical Review Letters* 99.25 (Dec. 2007). DOI: 10.1103/PHYSREVLETT.99.251803/FIGURES/1/THUMBNAIL.

[21] G. Aad et al. "Probing lepton flavour violation via neutrinoless $\tau \to 3$ decays with the ATLAS detector". In: *The European Physical Journal C 2016 76:5* 76.5 (Apr. 2016), pp. 1–25. ISSN: 1434-6052. DOI: 10.1140/EPJC/S10052-016-4041-9. URL: https://link.springer.com/article/10.1140/epjc/s10052-016-4041-9.

[22] A. M. Sirunyan et al. "Search for the lepton flavor violating decay $\tau \to 3\mu$ in proton-proton collisions at s $$ \sqrt{\mathrm{s}} $$ = 13 TeV". In: *Journal of High Energy Physics 2021 2021:1* 2021.1 (Jan. 2021), pp. 1–37. ISSN: 1029-8479. DOI: 10.1007/JHEP01(2021)163. URL: https://link.springer.com/article/10.1007/JHEP01(2021)163.

[23] R. Aaij et al. "Search for the lepton flavour violating decay τ - $\rightarrow$ μ - μ + μ -". In: *Journal of High Energy Physics 2015 2015:2* 2015.2 (Feb. 2015), pp. 1–20. ISSN: 1029-8479. DOI: 10.1007/JHEP02(2015)121. URL: https://link.springer.com/article/10.1007/JHEP02(2015)121.

[24] B. Aubert et al. "Searches for lepton flavor violation in the decays $\tau\pm\rightarrow$ e$\pm\gamma$ and $\tau\pm\rightarrow\mu\pm\gamma$". In: *Physical Review Letters* 104.2 (Jan. 2010). DOI: 10.1103/PHYSREVLETT.104.021802/FIGURES/1/THUMBNAIL.

[25] A. M. Baldini et al. "Search for the lepton flavour violating decay μ + $\rightarrow$ e + γ with the full dataset of the MEG experiment". In: *The European Physical Journal C 2016 76:8* 76.8 (Aug. 2016), pp. 1–30. ISSN: 1434-6052. DOI: 10.1140/EPJC/S10052-016-4271-X. URL: https://link.springer.com/article/10.1140/epjc/s10052-016-4271-x.

[26] Marco Ardu and Gianantonio Pezzullo. "Introduction to Charged Lepton Flavor Violation". In: *Universe* 8.6 (May 2022), p. 299. ISSN: 2218-1997. DOI: 10.3390/UNIVERSE8060299. URL: https://www.mdpi.com/2218-1997/8/6/299/htm.

[27] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08001. ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08001. URL: https://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08001.

[28] The ALICE Collaboration et al. "The ALICE experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08002. ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08002. URL: https://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08002.

[29] The ATLAS Collaboration et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08003. ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08003. URL: https://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08003.

[30] The CMS Collaboration et al. "The CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08004. ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08004. URL: https://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08004.

[31] *LHCb Trigger and Online Upgrade Technical Design Report*. May 2014. ISBN: 9789290834021. URL: https://cds.cern.ch/record/1701361.

[32] S Cadeddu et al. *LHCb reoptimized detector design and performance : Technical Design Report*. September. CERN, 2003. ISBN: 9290832096. URL: https://cds.cern.ch/record/630827.

[33] S Cadeddu et al. *LHCb trigger system : Technical Design Report*. Tech. rep. 2003. URL: https://cds.cern.ch/record/630828.

[34] S Cadeddu et al. *LHCb computing : Technical Design Report*. CERN, 2005. ISBN: 9290832487. URL: https://cds.cern.ch/record/835156.

[35] Ya V Pavlenko et al. *LHCb muon system : Technical Design Report*. CERN, 2001. ISBN: 9789290831808. URL: https://cds.cern.ch/record/504326.

[36] Ya V Pavlenko et al. *LHCb outer tracker : Technical Design Report*. CERN, 2001. ISBN: 9789290832003. URL: https://cds.cern.ch/record/519146.

[37] Ya V Pavlenko et al. *LHCb VELO (VErtex LOcator) : Technical Design Report*. CERN, 2001. URL: https://cds.cern.ch/record/504321.

[38] Ya V Pavlenko et al. *LHCb inner tracker : Technical Design Report*. CERN, 2002. URL: https://cds.cern.ch/record/582793.

[39] O Omelaenko et al. *LHCb calorimeters : Technical Design Report*. CERN, 2000. ISBN: 9789290831693. URL: https://cds.cern.ch/record/494264.

[40] O Omelaenko et al. *LHCb magnet : Technical Design Report*. 2000. URL: https://cds.cern.ch/record/424338.

[41] O Omelaenko et al. *LHCb RICH : Technical Design Report*. CERN, 2000. ISBN: 9290831707. URL: https://cds.cern.ch/record/494263.

[42] I Bediaga et al. *Framework TDR for the LHCb Upgrade : Technical Design Report*. Apr. 2012. ISBN: 9789290833741. URL: https://cds.cern.ch/record/1443882.

[43] P R Barbosa-Marinho et al. *LHCb online system, data acquisition and experiment control : Technical Design Report*. CERN, 2001. ISBN: 9290831901. URL: https://cds.cern.ch/record/545306.

[44] The LHCb Collaboration. "LHCb detector performance". In: *International Journal of Modern Physics A* 30.7 (Mar. 2015). DOI: 10.1142/S0217751X15300227. URL: https://doi.org/10.1142/S0217751X15300227.

[45] Christian Elsässer. *bb production angle plots*. URL: https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.htmlhttps://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html.

[46] R Aaij et al. "Performance of the LHCb Vertex Locator". In: *Journal of Instrumentation* 9.09 (Sept. 2014), P09007. ISSN: 1748-0221. DOI: 10.1088/1748-0221/9/09/P09007. URL: https://iopscience.iop.org/article/10.1088/1748-0221/9/09/P09007.

[47] LHCb Collaboration. *LHCb Tracker Upgrade Technical Design Report*. February. Feb. 2014. ISBN: 9789290833970. URL: https://cds.cern.ch/record/1647400.

[48] R. Arink et al. "Performance of the LHCb Outer Tracker". In: *Journal of Instrumentation* 9.01 (Jan. 2014), P01002. ISSN: 1748-0221. DOI: 10.1088/1748-0221/9/01/P01002. URL: https://iopscience.iop.org/article/10.1088/1748-0221/9/01/P01002.

[49] LHCb Collaboration. *LHCb PID Upgrade Technical Design Report*. Nov. 2013. ISBN: 9789290833932. URL: https://cds.cern.ch/record/1624074.

[50] R. Aaij et al. "First observation of CP violation in the decays of Bs0 mesons". In: *Physical Review Letters* 110.22 (May 2013). DOI: 10.1103/PHYSREVLETT.110.221601/FIGURES/1/THUMBNAIL.

[51] R. Aaij et al. "Measurement of the CKM angle γ from a combination of B±→Dh± analyses". In: *Physics Letters B* 726.1-3 (Oct. 2013), pp. 151–163. ISSN: 0370-2693. DOI: 10.1016/J.PHYSLETB.2013.08.020.

[52] R. Aaij et al. "Measurement of the Bs0→μ+μ- branching fraction and search for B0→μ+μ- decays at the LHCb experiment". In: *Physical Review Letters* 111.10 (Sept. 2013). DOI: 10.1103/PHYSREVLETT.111.101805/FIGURES/1/THUMBNAIL.

[53] The LHCb Collaboration et al. "The LHCb Detector at the LHC". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08005. ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08005. URL: https://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08005.

[54] C. Abellán Beteta et al. "Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC". In: (Aug. 2020). URL: http://arxiv.org/abs/2008.11556.

[55] Dordei Francesca. "LHCb detector and trigger performance in Run II". In: (). DOI: 10.1051/epjconf/201716401016.

[56] S Tolk et al. *Data driven trigger efficiency determination at LHCb*. May 2014. URL: https://cds.cern.ch/record/1701134.

[57] A L Read. "Presentation of search results: the CLs technique". In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (Sept. 2002), p. 2693. ISSN: 0954-3899. DOI: 10.1088/0954-3899/28/10/313. URL: https://iopscience.iop.org/article/10.1088/0954-3899/28/10/313.

[58] R. Aaij et al. "Erratum to: Measurements of prompt charm production cross-sections in pp collisions at s = 13 $$ \sqrt{s}=13 $$ TeV". In: *Journal of High Energy Physics 2017 2017:5* 2017.5 (May 2017), pp. 1–29. ISSN: 1029-8479. DOI: 10.1007/JHEP05(2017)074. URL: https://link.springer.com/article/10.1007/JHEP05(2017)074.

[59] A. M. Sirunyan et al. "Measurement of the Z / γ * $\rightarrow$ τ τ cross section in pp collisions at s = 13 TeV and validation of τ lepton analysis techniques". In: *The European Physical Journal C 2018 78:9* 78.9 (Sept. 2018), pp. 1–42. ISSN: 1434-6052. DOI: 10.1140/EPJC/S10052-018-6146-9. URL: https://link.springer.com/article/10.1140/epjc/s10052-018-6146-9.

[60] Georges Aad and others. "Measurement of the W $\tau\nu\_\tau$ cross section in pp collisions at sqrt(s)=7 TeV with the ATLAS experiment". In: *Phys. Lett. B* 706 (2012), pp. 276–294. DOI: 10.1016/j.physletb.2011.11.057.

[61] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. "A brief introduction to PYTHIA 8.1". In: *Computer Physics Communications* 178.11 (June 2008), pp. 852–867. ISSN: 0010-4655. DOI: 10.1016/J.CPC.2008.01.036.

[62] M. Tanabashi et al. "Review of Particle Physics". In: *Physical Review D* 98.3 (Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://journals.aps.org/prd/abstract/10.1103/PhysRevD.98.030001.

[63] V Gligorov. *The BackgroundCategory tool for background classification*. 2013.

[64] A Maevskiy et al. "Fast Data-Driven Simulation of Cherenkov Detectors Using Generative Adversarial Networks". In: *Journal of Physics: Conference Series* 1525.1 (Apr. 2020), p. 012097. ISSN: 1742-6596. DOI: 10.1088/1742-6596/1525/1/012097. URL: https://iopscience.iop.org/article/10.1088/1742-6596/1525/1/012097.

[65] M. Pivk and F. R. Le Diberder. "sPlots: A statistical tool to unfold data distributions". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 555.1-2 (Dec. 2005), pp. 356–369. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2005.08.106.

[66] Alex Rogozhnikov. "Reweighting with Boosted Decision Trees". In: *Journal of Physics: Conference Series* 762.1 (Oct. 2016), p. 012036. ISSN: 1742-6596. DOI: 10.1088/1742-6596/762/1/012036. URL: https://iopscience.iop.org/article/10.1088/1742-6596/762/1/012036.

[67] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (). DOI: 10.1145/2939672. URL: http://dx.doi.org/10.1145/2939672.2939785.

[68] Andreas Höcker et al. "TMVA : The Toolkit for Multivariate Data Analysis eith ROOT". In: (2008). DOI: 10.5170/CERN-2008-001.184. URL: https://cds.cern.ch/record/1099990.

[69] Jerome H. Friedman. "Greedy function approximation: A gradient boosting machine." In: *The Annals of Statistics* 29.5 (Oct. 2001), pp. 1189–1232. ISSN: 0090-5364. DOI: 10.1214/AOS/1013203451. URL: https://projecteuclid.org/journals/annals-of-statistics/volume-29/issue-5/Greedy-function-approximation-A-gradient-boosting-machine/10.1214/aos/1013203451.full.

[70] R. Aaij et al. "Measurement of the Bs0 →μ+μ- Branching Fraction and Effective Lifetime and Search for B0 →μ+μ- Decays". In: *Physical Review Letters* 118.19 (May 2017). DOI: 10.1103/PHYSREVLETT.118.191801/FIGURES/1/THUMBNAIL.

[71] Roel Aaij et al. "Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2". In: *EPJ Techniques and Instrumentation 2019 6:1* 6.1 (Feb. 2019), pp. 1–16. ISSN: 2195-7045. DOI: 10.1140/EPJTI/S40485-019-0050-Z. URL: https://epjtechniquesandinstrumentation.springeropen.com/articles/10.1140/epjti/s40485-019-0050-z.

[72] Giovanni Punzi. "Sensitivity of searches for new signals and its optimization". In: (Aug. 2003). DOI: 10.48550/arxiv.physics/0308063. URL: https://arxiv.org/abs/physics/0308063v2.

[73] The LHCb collaboration. "Measurement of the track reconstruction efficiency at LHCb". In: *Journal of Instrumentation* 10.02 (Feb. 2015), P02007. ISSN: 1748-0221. DOI: 10.1088/1748-0221/10/02/P02007. URL: https://iopscience.iop.org/article/10.1088/1748-0221/10/02/P02007.

[74] Flavio Archilli et al. *Measurement of B0s,d→μ+μ- decays with the Run 1 and Run 2 datasets*. Aug. 2020. URL: https://cds.cern.ch/record/2725940.

[75] Mathieu Perrin-Terrin and Giampiero Mancinelli. *Optimisation of the binning of the discriminating variables used in the computation of $\mathcal{B}$(Bs0 \to \mu+ \mu-)$ upper limits with the modified frequentist approach*. Feb. 2012. URL: https://cds.cern.ch/record/1419784.

[76] R. Aaij et al. "Prompt charm production in pp collisions at $$\sqrt{s}=7TeV$$". In: *Nuclear Physics B* 871.1 (June 2013), pp. 1–20. ISSN: 0550-3213. DOI: 10.1016/J.NUCLPHYSB.2013.02.010.

[77] Benjamin M. Dassinger et al. "Model-independent analysis of lepton flavour violating τ decays". In: *Journal of High Energy Physics* 2007.10 (Oct. 2007), p. 039. ISSN: 1126-6708. DOI: 10.1088/1126-6708/2007/10/039. URL: https://iopscience.iop.org/article/10.1088/1126-6708/2007/10/039.

[78] LHCb collaboration. *LHCb Trigger Scheme*. URL: https://lhcb.web.cern.ch/speakersbureau/html/TriggerScheme.html.

[79] LHCb Collaboration. *LHCb VELO Upgrade Technical Design Report*. Nov. 2013. ISBN: 9789290833925. URL: https://cds.cern.ch/record/1624070.

[80] M. Calvi et al. "Characterization of the Hamamatsu H12700A-03 and R12699-03 multi-anode photomultiplier tubes". In: *Journal of Instrumentation* 10.09 (Sept. 2015), P09021. ISSN: 1748-0221. DOI: 10.1088/1748-0221/10/09/P09021. URL: https://iopscience.iop.org/article/10.1088/1748-0221/10/09/P09021.

[81] L. Cassina. "Photodetectors and front-end electronics for the LHCb RICH upgrade". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 876 (Dec. 2017), pp. 217–220. DOI: 10.1016/J.NIMA.2017.03.008.

[82] *Vol. 10 (2020): High-Luminosity Large Hadron Collider (HL-LHC): Technical design report | CERN Yellow Reports: Monographs*. URL: https://e-publishing.cern.ch/index.php/CYRM/issue/view/127.

[83] "ATLAS Phase-II Upgrade Scoping Document". In: (Sept. 2015). DOI: 10.17181/CERN.7CRX.AJHP. URL: http://cds.cern.ch/record/2055248.

[84] D. Contardo et al. *Technical Proposal for the Phase-II Upgrade of the CMS Detector*. Tech. rep. DOI: 10.17181/CERN.VU8I.D59J.

[85] CERN (Meyrin) LHCb Collaboration. *Framework TDR for the LHCb Upgrade II $\\ \text{- Opportunities in flavour physics, and beyond, in the HL-LHC era} \\$*. July 2021. URL: https://cds.cern.ch/record/2776420.

[86] LHCb collaboration et al. "Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era". In: (Aug. 2018). URL: https://arxiv.org/abs/1808.08865v4.

[87] M. Calvi et al. "Single photon detection with SiPMs irradiated up to 1014 cm-2 1-MeV-equivalent neutron fluence". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 922 (Apr. 2019), pp. 243–249. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2019.01.013.

[88] A. Lehmann et al. "Tremendously increased lifetime of MCP-PMTs". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 845 (Feb. 2017), pp. 570–574. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2016.05.017.

[89] K. Matsuoka et al. "Extension of the MCP-PMT lifetime". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 876 (Dec. 2017), pp. 93–95. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2017.02.010.

[90] Yu Melikyan et al. "Load capacity and recovery behaviour of ALD-coated MCP-PMTs". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 949 (Jan. 2020), p. 162854. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2019.162854.

[91] M. J. Charles and R. Forty. "TORCH: Time of flight identification with Cherenkov radiation". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 639.1 (May 2011), pp. 173–176. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2010.09.021.

[92] S. Bhasin et al. "Test-beam studies of a small-scale TORCH time-of-flight demonstrator". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 961 (May 2020), p. 163671. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2020.163671.

[93] Bernhard W. Adams et al. "A Brief Technical History of the Large-Area Picosecond Photodetector (LAPPD) Collaboration". In: (Mar. 2016). DOI: 10.48550/arxiv.1603.01843. URL: https://arxiv.org/abs/1603.01843v1.

[94] *LAPPD™ / Large Area Picosecond Photodetector*. URL: https://incomusa.com/lappd/.

[95] A. Yu Barnyakov et al. "Response of microchannel plates in ionization mode to single particles and electromagnetic showers". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 879 (Jan. 2018), pp. 6–12. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2017.10.002.

[96] A. Ronzhin et al. "Direct tests of micro channel plates as the active element of a new shower maximum detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 795 (Sept. 2015), pp. 52–57. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2015.05.029.

[97] P. Jenni et al. *The high resolution spaghetti hadron calorimeter*. 1987. URL: http://inis.iaea.org/Search/search.aspx?orig_q=RN:19028958.

[98] M Lucchini et al. "Test beam results with LuAG fibers for next-generation calorimeters". In: *Journal of Instrumentation* 8.10 (Oct. 2013), P10017. ISSN: 1748-0221. DOI: 10.1088/1748-0221/8/10/P10017. URL: https://iopscience.iop.org/article/10.1088/1748-0221/8/10/P10017.

[99] Marta Calvi et al. *LHCb RICH: Potential physics performance*. Oct. 2021. URL: https://cds.cern.ch/record/2783211.

[100] S. Gómez et al. "FastIC: a fast integrated circuit for the readout of high performance detectors". In: *Journal of Instrumentation* 17.05 (May 2022), p. C05027. ISSN: 1748-0221. DOI: 10.1088/1748-0221/17/05/C05027. URL: https://iopscience.iop.org/article/10.1088/1748-0221/17/05/C05027.

[101] Claudio Arnaboldi et al. *A digital ASIC for sub-ns timing with the LHCb RICH detectors in Run 4*. May 2022. URL: https://cds.cern.ch/record/2809218.

[102] M Calvi et al. "Single photon time resolution of photodetectors at high rate: Hamamatsu R13742 MaPMT and R10754 MCP-PMT". In: *Journal of Instrumentation* 15.10 (Oct. 2020), P10031–P10031. DOI: 10.1088/1748-0221/15/10/p10031. URL: https://doi.org/10.1088%2F1748-0221%2F15%2F10%2Fp10031.

[103] Hamamatsu Photonics K.K. *PHOTOMULTIPLIER TUBES Basics and Applications FOURTH EDITION*. fourth. 2017. URL: https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/etd/PMT_handbook_v4E.pdf.

[104] Thomas M. Conneely, James S. Milnes, and Jon Howorth. "Extended lifetime MCP-PMTs: Characterisation and lifetime measurements of ALD coated microchannel plates, in a sealed photomultiplier tube". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 732 (Dec. 2013), pp. 388–391. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2013.07.023.

[105] D. R. Beaulieu et al. "Plastic microchannel plates with nano-engineered films". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 633.SUPPL. 1 (May 2011), S59–S61. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2010.06.121.

[106] E.V. Antamanova et al. "Anode current saturation of ALD-coated Planacon® MCP-PMTs". In: *Journal of Instrumentation* 13.09 (Sept. 2018), T09001. ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/09/T09001. URL: https://iopscience.iop.org/article/10.1088/1748-0221/13/09/T09001.

[107] A Giachero et al. "Current feedback operational amplifiers as fast charge sensitive preamplifiers for photomultiplier read out". In: *Journal of Instrumentation* 6.05 (May 2011), P05004. ISSN: 1748-0221. DOI: 10.1088/1748-0221/6/05/P05004. URL: https://iopscience.iop.org/article/10.1088/1748-0221/6/05/P05004.

[108] A Britting et al. "Lifetime-issues of MCP-PMTs". In: *Journal of Instrumentation* 6.10 (Oct. 2011), p. C10001. ISSN: 1748-0221. DOI: 10.1088/1748-0221/6/10/C10001. URL: https://iopscience.iop.org/article/10.1088/1748-0221/6/10/C10001.

Acknowledgements

I would like to thank my supervisor Prof. Marta Calvi and my co-supervisor Dr. Claudio Gotti, for their precious advices and their patient guidance through each stage of this thesis. The results of this work wouldn't be possible without the contribution of the other members of the analysis team: Flavio Archilli, Davide Fazzini, Giulia Frau and Maurizio Martinelli.

I must also express my gratitude to my father Stefano Capelli, my mother Nicoletta Clementi and my sister Chiara Capelli, for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis.

Simple acknowledgements couldn't express all what I should say to my loving girl, friend, and artist Aurora Battaglia. Our jokes and our laughs supported me mentally through these years, and your delicious cuisine supported me physically. Thank you for everything you do for me, I really hope that our new life together at "La Tana" will be legen.. wait for it.. dary.

Many encouragements came from the other members of our group "TuaMammaÈ-DaQuattro": Samuele "Cicci" Ciciriello, Alessandro Cravedi, Jessica Gogni, Amanda Granelli, Veronica "V" Guagni, Andrea "G" Guglielmetti, Carlo Marani, Lorenzo "Pera" Peroni, Miriam Righini and in particular a heartfelt thanks to Yuri Perina. Thank you all for the time that we spent together.

Then there are so many other relatives and friends to acknowledge, Francesco Clementi and Norina Bini, Daniele Tosca, Luca "Cava" Cavanna, Vincenzo "Ciaccia" Ferrari, Luca "Pane" Panelli, Lorenzo "Skale" Scaletti, Arianna Chiapparini, Alessandro Fendillo, Massimo Morabito, and even more to list, but the margin is too narrow to contain them.

To everyone... Thank you...