The University of Manchester

Searches for new physics in violation of strong *CP* symmetry and lepton universality with the LHCb experiment

A thesis submitted to The University of Manchester for the degree of

Doctor of Philosophy

(PhD)

Lorenzo Capriotti

Faculty of Science and Engineering
School of Physics and Astronomy

2018

Contents

Word count: 25368

Abstract

In this thesis two analyses are presented, both searches for hints of new physics in data recorded by the LHCb detector at CERN. In the first one, a search for the *CP*-violating strong decays $\eta \to \pi^+\pi^-$ and $\eta'(958) \to \pi^+\pi^-$ is made, analysing the $\pi^+\pi^-$ mass spectra coming from the decays $D^+ \to \pi^+\pi^+\pi^-$ and $D_s^+ \to \pi^+\pi^+\pi^-$. The dataset corresponds to an integrated luminosity of $3.0\,\text{fb}^{-1}$ of pp collision data recorded during LHC Run 1 (2011 and 2012), and $0.3\,\text{fb}^{-1}$ recorded in Run 2 (2015). No evidence is seen, and upper limits at 90% confidence level are set on the branching fractions of both channels.
The second analysis is a sensitivity study for a test of lepton universality in semileptonic charm decays, aiming to measure the ratio of branching fractions $R^{\mu/e} = \frac{\mathcal{B}(D^0 \to K^-\mu^+\nu_\mu)}{\mathcal{B}(D^0 \to K^-e^+\nu_e)}$. The dataset corresponds to an integrated luminosity of $2.0\,\text{fb}^{-1}$ of pp collision data recorded during LHC Run 2 (2015 and 2016). The study is performed on a small fraction of the total dataset using toy simulation to describe the signal and background components, and leads to an expected reduction of the uncertainty on the current measurement by about one order of magnitude, neglecting systematic effects. A determination of electron identification efficiencies, using a new method, is also presented.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Lorenzo Capriotti

University of Manchester
Faculty of Science and Engineering
School of Physics and Astronomy
Particle Physics Group
Oxford Road
Manchester
M13 9PL
2018

Copyright

Acknowledgements

There will be a lot of names in this section. As somebody once told me, I am Italian; therefore, the thing I can do best is to make new friends. And I most certainly did.

Three years and a half is quite a long period of time. During these years I met so many great people and I had wonderful experiences with old friends and new friends, in several different countries. I also worked, a bit, so I will start from the people who allowed this thesis to exist.

Thanks to George. You have been a fantastic supervisor, you corrected my thesis so quickly and you have definitely taught me a lot, not only physics but also how to be critical, how to be practical and how to write using proper grammar (well, you definitely tried, I hope I learnt!). The attitude that will shape my (potential) future career will have deep roots in what I learnt while working with you.

Thanks to both the LHCb Group Leaders, Chris and Marco. You are so professional, but at the same time you manage to create the most pleasant athmosphere in the workplace. I believe that this, more than the (outstanding!) physics results, is the true flagship of the Manchester group.

Thanks to Pedro and Yvonne, for the pleasant viva and for all your comments and corrections.

Thanks to Adam, for all the last-minute questions.

Thanks to Summer, for having convinced me to apply for a PhD in Manchester, and for the great time we had together (but I have not forgiven you yet for starting the panda nightmare).

Thanks to all the flatmates: Phil and Quake, Dominik, and Kritika and Sonia. I truly enjoyed living with you all, especially during our glorious parties.

Thanks to all the office mates with whom I had the pleasure to work during these years, and to all the people in the Manchester HEP group. A special mention goes to Jon, Giulio, Alex, Kevin, Steffi, Suzanne, Chris, Gedas, Mark,

Owen, Martha, Vincent, Donal, Krishan (and the duoplasmatron), Diego, Pawel, Paki, Julia, Agni, Jacob, Chester, Rafał. A super special mention goes to Fabian (thanks for the cookies and all the cheese), Ste (thanks for all the existential questions and the deep conversations), Rhys (thanks for all the beers) and Johnny (thanks for all the 80's music).

Thanks to all the people in the LHCb collaboration who helped solve all the (many) problems I encountered. No names here, as there would be too many!

Thanks to the Italian crew in Geneva: Gasta, Genoveffa (STUPIDA), Sfettuccia, Deferara, Sirvia, Matteo, Ayatollah Khomeyni, Vingenzo, The Horrorist, Chanteclaire (OH FRECOTI), Giada, Ascoli, il buon Giuli, Tovarish, BETTOOOOO, Arianal, Daniele, Bea, Barbara, Adili, Sexy, Killer, Valeria, Elena, L'Aedo, Aghetta, Frankie, L'Ariete, Gigi Dag, Shak, Aneddotman, Lucrezia, Guido, Renado, Manfre, Phelipe, Scodella, Gioacchino (aka Gioacchino), Holymary, Benedetto (and all the Sardinians!), Silvia, String, Elisa. The bonds we created will surely last forever. Also, I am ready for Baccanali 2018.

Thanks to all the other CERN and non-CERN people that I met during my LTA, especially Andrea (aka Culo), the Dinner Club (Alex, Emma, Fabio, Andrea), Nicola, Fabrizio, Vieri, Emma, Karri and everybody else I am forgetting now. It would have been much harder to survive in Geneva without you.

Please bear with me, we are almost done.

To all my friends living around the world, as our job often requires, thanks for giving me shelter during all my trips and for having visited me: Ilaria, Giovanni, Carlotta, Flavia C., Flavia P., Valeria, Margherita, Vallo, Cava, Ambra, Annalisa, Elisa, Laura. A special mention goes to Belle Arti 42 (Giulia, Martina, Piero, Nico) and to your Brendona.

Thanks to all my long-time friends: Simone, Alberto, Gianluca (aka Netto), Alessandra, Giulia G., Valeria, Valerio, Beatrice, Giulia S., Giorgia, Gianluca. Siete il mio punto fisso in tutti i miei esodi.

Last but not least, thanks to my brother, my mother and my father, for all your support, emotional and practical (especially during illness!).

The author

Education

- **I.S.C. "Adolfo de Carolis" (1995 – 2000)**
  Primary School
- **Scuola Media Statale "Prof. Mignini Don Pacifico" (2000 – 2003)**
  Secondary School
- **Liceo Scientifico Statale "Benedetto Rosetti" (2003 – 2008)**
  Scientific High School
  Grade: 100/100
- **Sapienza Università di Roma (2008 – 2012)**
  University, Bachelor of Science, B.Sc
  Faculty of Mathematical, Physical and Natural Sciences
  Thesis title: *The "strange" world of neutral mesons with flavour*
  Original title: *Lo "strano" mondo dei mesoni neutri con sapore*
  Supervisor: Prof. C. Dionisi
  Grade: 110/110
- **Sapienza Università di Roma (2012 – 2014)**
  University, Master of Science, M.Sc
  Faculty of Mathematical, Physical and Natural Sciences
  Curriculum: Nuclear and subnuclear physics
  Thesis title: *Search for the decay $X(3872) \to J/\psi\omega$ with the LHCb experiment at CERN*
  Original title: *Ricerca del decadimento $X(3872) \to J/\psi\omega$ nell'esperimento LHCb al CERN*
  Supervisors: Dr. R. Santacesaria, Dr. A. A. Alves Jr
  Grade: 108/110
- **The University of Manchester (2014 – 2018)**,
  Doctor of Philosophy, PhD
  Faculty of Engineering and Physical Sciences
  Thesis title: *Searches for new physics in violation of strong CP symmetry and lepton universality with the LHCb experiment*
  Supervisors: Prof. G. Lafferty, Prof. C. Parkes

Introduction

With the discovery of the Higgs boson in 2012 by the ATLAS and CMS collaborations, after almost 50 years since it was theorised, the last big piece missing from the Standard Model was found. This confirmed the extraordinary success of the highly-predictive theory describing the fundamental particles and their interactions. However, there are still a number of phenomena which cannot be explained satisfactorily from the Standard Model: for instance, the hierarchy problem and the fine tuning of the Higgs mass, the problem of the neutrino masses and the anomalous magnetic dipole moment of the muon. For this reason, experiments are currently searching for what is commonly known as "physics beyond the Standard Model", or "new physics". This means to search for unambiguous new phenomena which are predicted by a set of extensions of the Standard Model, in order to modify the theoretical structure and to allow new, non-standard predictions to be tested in future experiments, and, possibly, to open a door into a previously unexplored region of particle physics.

In this thesis, two analyses, adressing two phenomena that are not well described by the Standard Model, are presented, from studies perfomed with data collected by the LHCb experiment at CERN.

The first one investigates the so-called "strong *CP* problem", and it is linked to the problem of the matter-antimatter asymmetry in the Universe. The analysis consists of a search for the decays $\eta \to \pi^+\pi^-$ and $\eta' \to \pi^+\pi^-$, which would both violate the *CP* symmetry, from a sample of $D^+ \to \pi^+\pi^+\pi^-$ and $D_s^+ \to \pi^+\pi^+\pi^-$ decays.

The second analysis is a test of lepton universality, which in recent years has been found by measurements to be violated at the level of 2-4 standard deviations, depending on the decay channel. The analysis, not yet completed, will measure the ratio of branching fractions $R^{e/\mu} = \frac{\mathcal{B}(D^0 \to K^- e^+ \nu_e)}{\mathcal{B}(D^0 \to K^- \mu^+ \nu_\mu)}$, where the D^0 comes from $D^{*+} \to D^0 \pi^+$ decays.

This thesis is structured as follows:

- In Chapter 1, a theoretical introduction to the Standard Model and to the problem of matter-antimatter asymmetry is given.
- In Chapter 2, the latest experimental results on both the search for strong *CP* violation and the tests of lepton universality, focussing on the hints of deviations from the predicted values, are described.
- In Chapter 3, the LHCb detector at CERN is described in detail, including the hardware components of all subdetectors and the physics performances.
- Chapter 4 presents the search for the strong *CP*-violating decays $\eta \to \pi^+\pi^-$ and $\eta' \to \pi^+\pi^-$.
- In Chapter 5, a test of lepton universality in semileptonic charm decays is described, along with sensitivity studies and determination of particle identification efficiencies.

1 The Standard Model

The Standard Model (SM) of particle physics is the theory that describes the phenomenology of elementary particles and their interactions. It is a relativistic quantum field theory based on the gauge symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, where $SU(3)_C$ is the symmetry group of the strong interactions and $SU(2)_L \otimes U(1)_Y$ is the symmetry group of the electroweak interactions. In this chapter, a brief overview of the elements of the Standard Model is given, with a description of all the elementary particles (Section 1.1) and their fundamental interactions (Section 1.2). The theory of quark mixing and lepton mixing is described in Section 1.3, and the C and P symmetries are described in Section 1.4. Finally, Sections 1.5 and 1.6 introduce the problem of matter-antimatter asymmetry in the universe. A schematic view is given in Figure 1.1.

1.1 Elementary particles

All particles in the SM are divided into two macrogroups, according to their spin:

- All the particles with half-integer spin are called fermions. They obey the Fermi-Dirac statistics and the Pauli exclusion principle and their dynamics are described by the Dirac equation:

$$(i\gamma^\mu \partial_\mu - m)\psi(x) = 0, \tag{1.1}$$

  where γ^μ are the Dirac gamma matrices, ∂_μ is the partial derivative with respect to the μ-coordinate, m is the mass, $\psi(x)$ is the fermion field and the μ index runs over the four spacial-temporal coordinates. Positive-energy solutions of the Dirac equations are identified as fields of matter particles, and negative-energy solutions with antimatter particles. All the elementary fermions in the SM have spin $\frac{1}{2}$.

- All the particles with integer spin are called bosons. They obey the Bose-Einstein statistics and a description of their dynamics is based on the Klein-Gordon equation:

$$(\partial_\mu \partial^\mu + m^2)\phi(x) = 0, \tag{1.2}$$

where $\phi(x)$ is the boson field. In the SM, elementary bosons fields are associated with fundamental interactions (see Section 1.2) and the spin-1 bosons act as force carriers; the only scalar boson is the Higgs boson, which gives mass to the elementary particles through the mechanism of spontaneous symmetry breaking.

Elementary fermions are further divided into two categories:

- Quarks: particles which are subject to all fundamental forces of nature (although gravity is neglected in the SM). There are six types (or flavours) of quarks, divided into three families. Quarks can combine into bound states to create composite particles (or hadrons), which according to their content are labelled mesons ($q\bar{q}$), baryons (qqq), tetraquarks ($qq\bar{q}\bar{q}$) [1] or pentaquarks ($qqqq\bar{q}$) [2]. Other combinations are in principle allowed [3] but have never been observed.

- Leptons: particles which do not interact via the strong force. The three electrically charged leptons are subject to the electromagnetic and weak forces, while the three neutrinos only interact via the weak force.

1.2 Fundamental interactions

The Dirac and Klein-Gordon equations describe the dynamics of non interacting particles. In Lagrangian formalism, the total Lagrangian can be split into the free particles term and an interaction term:

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_I. \tag{1.3}$$

From the principle of least action applied to the free action $S_0 = \int \mathrm{d}^4 x \mathcal{L}_0$ one can retrieve the equation of motion of free particles, Equations (1.1) and (1.2). The effect of the interaction term on scattering processes is generally evaluated perturbatively.

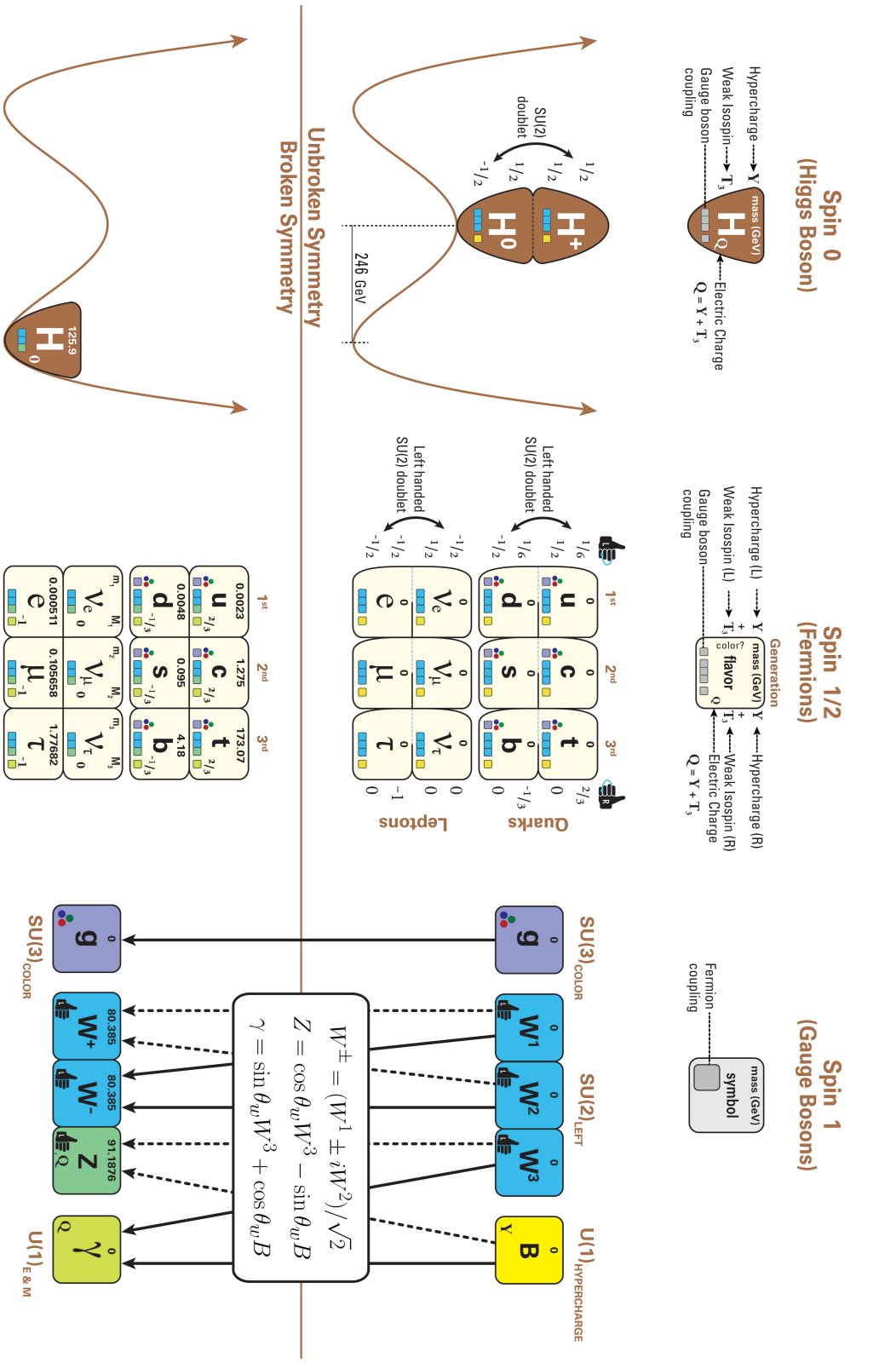


Figure 1.1: The particles of the Standard Model and their interactions, before and after the spontaneous symmetry breaking [4].

The requirement of gauge invariance plays a central role in the theory of interacting fields. A generic gauge transformation must not have any effect on any physical process or on any measurement. Therefore, the terms that can contribute to the Lagrangian of interaction depend on this requirement. Gauge transformations are also of primary importance in defining the structure of the non-Abelian Yang-Mills vacuum described in Section 1.6.

1.2.1 Quantum electrodynamics

Maxwell's equations can be summarised using the Lorentz invariant electromagnetic field tensor, with the Lagrangian

$$\mathcal{L}_0^{EM} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \tag{1.4}$$

where $F^{\mu\nu} = \partial^\nu A^\mu - \partial^\mu A^\nu$ and A^μ is the electromagnetic four-potential. Equation (1.4) describes the kinematics of the A^μ field; to introduce interactions between the electromagnetic field and the matter fields it is necessary to resort to gauge invariance requirements.

In order to preserve the invariance of the Dirac Lagrangian for free particles,

$$\mathcal{L}_0 = \overline{\psi}(x)(i\gamma_\mu\partial^\mu - m)\psi(x), \tag{1.5}$$

under a local gauge transformation of the fields,

$$\psi(x) \to \psi'(x) = e^{i\zeta(x)}\psi(x), \tag{1.6}$$

it is sufficient to define a covariant derivative, analogous to the minimal substitution in classical electrodynamics:

$$\partial^\mu \to D^\mu = \partial^\mu - ieA^\mu. \tag{1.7}$$

By combining Equations (1.5) and (1.7), and taking into account the gauge transformation of the electromagnetic potential $A^\mu \to A'^\mu = A^\mu + \frac{1}{e}\partial^\mu\Lambda(x)$ under which $F^{\mu\nu}$ is invariant, one obtains

$$\overline{\psi}(i\gamma_\mu D^\mu - m)\psi - \overline{\psi}\left(\gamma_\mu\partial^\mu\zeta\right)\psi = \overline{\psi}(i\gamma_\mu\partial^\mu - m)\psi + \overline{\psi}e\gamma_\mu A^\mu\psi = \mathcal{L}_D + \mathcal{L}_I^{EM}, \tag{1.8}$$

by setting $\Lambda(x) = \zeta(x)$. Therefore, by imposing gauge invariance under a local phase transformation, an interaction term has risen from the Dirac Lagrangian; the complete Lagrangian of QED is given by

$$\mathcal{L} = \mathcal{L}_0^D + \mathcal{L}_0^{EM} + \mathcal{L}_I^{EM} = \overline{\psi}(i\gamma_\mu\partial^\mu - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \overline{\psi}e\gamma_\mu A^\mu\psi. \tag{1.9}$$

After the second quantisation procedure, the excitations of the A^μ field are identified as photons.

1.2.2 Weak interactions and electroweak unification

A first description of weak interactions was given by Fermi [5], who attempted to define a qualitative theory of β decays of the neutron as a contact interaction of two vector currents,

$$G_F \left(\bar{n}\gamma^\mu p\right) \left(\bar{\nu}\gamma_\mu e\right). \tag{1.10}$$

To account for parity violation [6], the currents of the weak interactions must be described by a specific combination of vector and axial quantities, namely the $V - A$ structure, and this is achieved using the γ_5 Dirac matrix:

$$G_F \left(\bar{n}\gamma^\mu \left(1-\gamma_5\right) p\right) \left(\bar{\nu}\gamma_\mu \left(1-\gamma_5\right) e\right). \tag{1.11}$$

It is convenient to define the projection operators,

$$\hat{P}_L = \left(\frac{1-\gamma_5}{2}\right), \tag{1.12}$$

$$\hat{P}_R = \left(\frac{1+\gamma_5}{2}\right), \tag{1.13}$$

where the L and R subscripts refer to the left and right chirality eigenstates. From Equation (1.11) it can be seen how only the chiral left components of the fermion spinors enter into the weak interactions, which at this stage only include charged currents.

To obtain a unified model of the electromagnetic and weak interactions it it necessary to define the minimal symmetry group which can describe both the weak and electromagnetic currents: it is found to be $SU(2)_L \otimes U(1)_Y$, where the $SU(2)$ group acts on the left-chirality doublets of the weak isospin I_W,

$$\Psi_L = \begin{pmatrix} e \\ \nu \end{pmatrix}_L, \tag{1.14}$$

producing the currents

$$L^i_\mu = \overline{\Psi}_L \gamma_\mu \frac{\sigma^i}{2} \Psi_L \tag{1.15}$$

where σ^i, $i = 1,2,3$ are the Pauli matrices, i.e. the generators of the $SU(2)$ algebra. The weak current can then be defined as

$$L^1_\mu + iL^2_\mu = \bar{e}_L \gamma_\mu \nu_L = \frac{1}{2}\bar{e}\gamma_\mu \left(1-\gamma_5\right)\nu \tag{1.16}$$

and the electromagnetic current is combined with L^3_μ,

$$J^{EM}_\mu - L^3_\mu = -\bar{e}_R \gamma_\mu e_R - \frac{1}{2}\left(\bar{\nu}_L \gamma_\mu \nu_L + \bar{e}_L \gamma_\mu e_L\right) = \frac{1}{2} Y_\mu, \tag{1.17}$$

to define the weak hypercharge current Y_μ, which commutes with all the generators of $SU(2)_L$, hence being another symmetry of the system.

By applying the general Yang-Mills theory [7] and results from the previous section, the covariant derivatives can be defined,

$$D^\mu \Psi_L = \left[\partial^\mu + ig W^\mu_i \frac{\sigma_i}{2} + ig' \left(-\frac{1}{2}\right) B^\mu\right] \Psi_L \tag{1.18}$$

$$D^\mu e_R = \left[\partial^\mu + ig' \left(-1\right) B^\mu\right] e_R \tag{1.19}$$

$$D^\mu \nu_R = \left[\partial^\mu + ig' \left(0\right) B^\mu\right] \nu_R = \partial^\mu \nu_R \tag{1.20}$$

where the three W^i_μ are the gauge fields of $SU(2)_L$, B^μ is the gauge field of $U(1)_Y$, the numbers in parentheses are the hypercharge values of the different particle species and g and g' are the coupling constants of the interactions. It is worth noticing that the right-handed neutrinos, being neutral under both the weak isospin and the weak hypercharge, do not participate in the electroweak interactions. The Lagrangian at this point reads

$$\mathcal{L}^{EW} = \overline{\Psi}_L \gamma_\mu D^\mu \Psi_L + \bar{e}_R \gamma_\mu D^\mu e_R - \frac{1}{4}\left[W^i_{\mu\nu} W^{\mu\nu i} + B_{\mu\nu} B^{\mu\nu}\right], \tag{1.21}$$

$$W^i_{\mu\nu} = \partial_\nu W^i_\mu - \partial_\mu W^i_\nu + g \epsilon^{ijk} W^j_\mu W^k_\nu, \tag{1.22}$$

$$B_{\mu\nu} = \partial_\nu B_\mu - \partial_\mu B_\nu, \tag{1.23}$$

and all the fields are massless. An important consequence of the gauge invariance requirements in the electroweak Lagrangian is that the couplings must not depend on the nature of the fermion which is involved in the process. In other words, electroweak interactions described as a Yang-Mills gauge theory are universal.

Introucing mass terms in Equation (1.21) is not trivial [8], as they would couple the left and right chirality eigenstates, thus not being invariant under a $SU(2) \otimes U(1)$ transformation. The solution is found by allowing a new $SU(2)$ doublet scalar field to activate the mechanism of spontaneous symmetry breaking, leaving the symmetry of electromagnetic interactions untouched [9, 10],

$$SU(2)_L \otimes U(1)_Y \to U(1)_{em}. \tag{1.24}$$

This is called the Brout-Englert-Higgs mechanism [11, 12], which introduces mass terms for all particles and a new scalar field, the Higgs field. At this point, the only thing left is to identify the physical fields, i.e. the eigenstates of the mass matrix. The charged states are combined as

$$W_\mu = \frac{W_\mu^1 + iW_\mu^2}{\sqrt{2}}, \tag{1.25}$$

$$W_\mu^\dagger = \frac{W_\mu^1 - iW_\mu^2}{\sqrt{2}}, \tag{1.26}$$

while the neutral states are combined and rotated by an angle θ_W (Weinberg angle),

$$Z_\mu = \cos(\theta_W)W_\mu^3 - \sin(\theta_W)B_\mu, \tag{1.27}$$

$$A_\mu = \sin(\theta_W)W_\mu^3 + \cos(\theta_W)B_\mu, \tag{1.28}$$

where A_μ is identified as the electromagnetic field by imposing its mass eigenvalue to be zero.

1.2.3 Quantum chromodynamics

Quantum chromodynamics (QCD) is a Yang-Mills theory of strong interactions based on the symmetry group $SU(3)$. The generic Lagrangian is written as

$$\mathcal{L}_{QCD} = \sum_f \bar{q}_f \left(i\gamma_\mu D^\mu - m_f \right) q_f - \frac{1}{4} G_{\mu\nu}^a G^{\mu\nu a}, \tag{1.29}$$

where the sum runs over all the quark flavours, the label a denotes the eight massless gluon fields and the covariant derivative and the gluon field strength are defined as

$$D^\mu = \partial^\mu + ig_S \frac{\lambda_a}{2} G_a^\mu, \tag{1.30}$$

$$G_i^{\mu\nu} = \partial^\nu G_i^\mu - \partial^\mu G_i^\nu + g_S f^{abc} G_\mu^b G_\nu^c. \tag{1.31}$$

Analogously to the $SU(2)$ part of the weak Lagrangian, the gauge fields are described via the adjoint representation of the algebra, which in the case of $SU(3)$ has dimension 8. The covariant derivative is defined through the generators of the algebra, the eight 3x3 Gell-Mann matrices λ_a; therefore, G_μ^a describes the eight massless gluon fields which mediate the strong interactions.

As shown in Section 1.6, *CP* violation is theoretically possible within QCD, although not seen so far.

1.3 Quark mixing and lepton mixing

The masses of quarks and leptons are generated in the Standard Model by the Brout-Englert-Higgs mechanism of spontaneous symmetry breaking. This gives the following quark mass Lagrangian:

$$\mathcal{L}_{QM} = \overline{D}_L M^D D_R + \overline{U}_R M^U U_L + h.c., \tag{1.32}$$

where $U_{L,R}$ ($D_{L,R}$) indicate the left and right fields of type up (down). The two mass matrices include the couplings between the quark fields and the Higgs field and they are, in general, not diagonal. A diagonalisation procedure can be performed, which leaves one unitary matrix, called the Cabibbo-Kobayashi-Maskawa matrix (CKM) [13, 14], acting on the down fields:

$$\mathcal{L}_{QM} = \overline{D}_L V_{CKM} m^D D_R + \overline{U}_R m^U U_L + h.c., \tag{1.33}$$

where the $m^{D,U}$ mass matrices are diagonal.

The CKM matrix acts on the down flavour eigenstates and rotates them into the down physical states (i.e. the mass eigenstates) as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.34}$$

Therefore, in a flavour changing process involving an up quark q_i and a down quark q_j, the weak coupling is modified by a factor V_{ij}.

The elements of the CKM matrix are often parametrised with three rotation angles and a complex phase as

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{13} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{13} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{1.35}$$

where $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$ and δ is the complex phase.

The discovery of neutrino oscillations [15, 16] has proven that neutrinos are not massless; therefore, the same argument applies in the lepton sector, and an analogous matrix can be defined, i.e. the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [17]:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = V_{PMNS} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} = \begin{pmatrix} V_{e1} & V_{e2} & V_{e3} \\ V_{\mu 1} & V_{\mu 2} & V_{\mu 3} \\ V_{\tau 1} & V_{\tau 2} & V_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{1.36}$$

Likewise, when evaluating amplitudes of processes where a charged weak current is involved, the coupling is modified by a factor V_{ab}, where a is the charged lepton present in the process and b is the neutrino in the mass basis.

1.4 Parity, charge conjugation and *CP* symmetries

The discrete symmetries P (parity inversion) and C (charge conjugation) and, in particular, their combination CP play an important role in the Standard Model. In fact, the violation of the CP symmetry by the weak interactions is responsible for several phenomena, like the decay $K_L^0 \to \pi^+\pi^-$ and particle-antiparticle mixing in the neutral flavoured meson systems, and it is one of the necessary conditions for baryogenesis described in Section 1.5.1.

The parity transformation inverts the axes of the space coordinates of a system, thus reversing the momentum of a particle without flipping its spin,

$$\hat{P}\phi(\vec{x},t)\hat{P}^\dagger = \eta_P\phi(-\vec{x},t). \tag{1.37}$$

Since two consecutive applications of the parity operator result in the original state, the generic phase η_P, or intrinsic parity, can only assume the values ± 1. In the space of Dirac spinors, one gets

$$\hat{P}\psi(\vec{x},t)\hat{P}^\dagger = \eta_P\gamma_0\psi(-\vec{x},t). \tag{1.38}$$

As the spin is not affected by a parity transformation, a left-handed particle becomes a right-handed particle under parity. This can be seen from Equation (1.38) and Equations (1.12) and (1.13) by noticing that $\{\gamma_0, \gamma_5\} = 0$.

The charge conjugation transforms a particle into its antiparticle, and acts in the Dirac spinor space as

$$\hat{C}\psi(x)\hat{C}^\dagger = -i\gamma_2\gamma_0\bar{\psi}(x). \tag{1.39}$$

A left(right)-handed particle remains left(right)-handed after a charge conjugation transformation.

The combined action of the C and P transformations is known simply at CP and has the effect of transforming a particle into its own antiparticle while at the same time changing its chirality. So, for instance,

$$\begin{aligned} \hat{CP}|B^0\rangle &= |\bar{B}^0\rangle, \\ \hat{CP}|\nu_L\rangle &= |\bar{\nu}_R\rangle. \end{aligned} \tag{1.40}$$

The property of transforming left-handed fermions into right-handed antifermions make *CP* a good symmetry candidate, as the weak interactions, due to their chiral structure, violate both *P* and *C*; however, evidence that the *CP* symmetry is broken in the Standard Model was provided in 1964 [18] by the observation of $K_L^0 \to \pi^+\pi^-$ decays. The violation of *CP* is caused by the presence of an irreducible phase in the CKM matrix, which cannot be transformed into a real matrix by a redefinition of the fields. When evaluating decay amplitudes, this additional phase gives rise to differences in the rates of processes involving particles and antiparticles.

1.5 Matter-antimatter asymmetry in the Universe

All the observations from radioastronomy and cosmic ray telescopes indicate that all known celestial bodies in the observed universe, as well as most of the interstellar space, are made out of matter, and there is currently no evidence of any antimatter cluster, star or galaxy, although there are ongoing experiments to look for excess of antimatter in the cosmic ray flux in space (for instance, the PAMELA experiment on the Resurs-DK1 satellite [19] or the AMS-02 experiment on the International Space Station [20]). Therefore, it is safe to assume for now that the baryon asymmetry of the universe, defined as

$$\eta = \frac{N_B - N_{\bar{B}}}{N_B + N_{\bar{B}}} \sim \frac{N_B}{N_\gamma}, \tag{1.41}$$

where N_B ($N_{\bar{B}}$) is the number of baryons (antibaryons) in the universe, must be positive. In the last term, it is pointed out that, since the products of low energy annihilation processes are mainly photons and the number of antibaryons in the Universe is much smaller than the number of baryons, the baryon asymmetry η can be approximated by the baryon-to-photon density ratio. This quantity has been measured by the WMAP satellite [21] to be

$$\eta = (6.1^{+0.3}_{-0.2}) \times 10^{-10}. \tag{1.42}$$

From this measurement, it can be concluded that the baryon number *B*, defined as

$$B = \frac{1}{3}\left(N_q - N_{\bar{q}}\right), \tag{1.43}$$

where N_q ($N_{\bar{q}}$) is the number of quarks (antiquarks), is strictly positive for the observable universe. There are three possible scenarios to take this into account in the current theoretical description of the early universe:

- **B has always been positive.** This means that immediately after the Big Bang, a larger amount of matter has been generated with respect to antimatter. Although this would immediately solve the baryon-antibaryon asymmetry problem, it violates the naturalness principle.

- **B is actually zero, and has always been.** The observed dominance of matter over antimatter might simply be a local feature, and clusters of antimatter actually exist in regions outside the range excluded by current experiments.

- **B was zero at the beginning but $\frac{dB}{dt} > 0$ at some point.** The excess of matter might have been generated by some dynamical process happening after the Big Bang. This theory is called baryogenesis.

The third scenario is the only one relevant for this work. The next section describes some conditions which are necessary for any baryogenesis to occur, and how it is linked to the amount of *CP* violation in the Standard Model.

1.5.1 Sakharov conditions for baryogenesis

In order to include a dynamical baryogenesis mechanism, any theory for which $B = 0$ during the Big Bang and $B > 0$ in the present time must satisfy three necessary, but not sufficient, conditions first described by Andrei Sakharov in 1967 [22]:

- There must be at least one B-number violating process

- C and CP symmetries must be violated

- The B-number violating processes must occur outside of thermal equilibrium

To explain and prove the three conditions it is convenient to use a simple quantum statistical mechanics formalism [23, 24]. Let ρ be the density matrix of the universe, defined by the elements

$$\rho_{mn} = \sum_i p_i \langle u_m | \psi_i \rangle \langle \psi_i | u_n \rangle , \tag{1.44}$$

where $|u_n\rangle$ is an orthonormal basis, $|\psi_i\rangle$ is one of the states of the system and p_i is the probability of the system to be in the state i. The time evolution is

described by the von Neumann equation in the Schrödinger picture:

$$\frac{\partial \rho(t)}{\partial t} = -i \left[\hat{H}, \rho(t)\right]. \tag{1.45}$$

The ensemble average of a generic operator, which in this context will be the baryon number operator $\hat{B}$, is given by $\left\langle \hat{B} \right\rangle (t) = \mathrm{tr}(\hat{B}\rho(t)) = B(t)$. As pointed out by the last term, this corresponds to a measurement of the baryon number B on the ensemble with density matrix ρ. At this point, to prove the first Sakharov condition, it is sufficient to notice that if $\left[\hat{B}, \hat{H}\right] = 0$ and $B(t_0) = 0$, then $B(t) = 0$ for all times.

The second condition arises by noticing that, even if the first condition is satisfied, if C is conserved then, for every process that violates B, its width would be equal to the width of the C-conjugate process, resulting in a net B conservation over long times. Moreover, even allowing C violation but assuming that CP is conserved, a hypothetical B-violating process $X \to q_L q'_L$, where X is a state with zero baryonic number, could create left-handed baryons so that the C-conjugate decay would have a different rate, but

$$\Gamma(X \to q_L q'_L) + \Gamma(X \to q_R q'_R) = \Gamma(\overline{X} \to \bar{q}_L \bar{q}'_L) + \Gamma(\overline{X} \to \bar{q}_R \bar{q}'_R), \tag{1.46}$$

hence CP must be violated as well. To use the same formalism, given a generic discrete transformation T such that $\left[\hat{T}, \hat{H}\right] = 0$, if $\left[\hat{T}, \rho(t_0)\right] = 0$ at some initial time t_0, then $\rho(t)$ is symmetric under T for all times; therefore, in order to have a C-(CP-)asymmetric state, C (CP) must be violated.

The third condition is trivial since, by definition, in thermal equilibrium ρ is invariant under time translations and, therefore, $B = \mathrm{tr}(\hat{B}\rho)$ is constant.

The Standard Model of particle physics and the cosmological model satisfy all three conditions: baryon number violation can be achieved via the so-called "sphaleron processes" [25], i.e. transitions between different electroweak vacua which might have been possible with the temperature of the early stage of the universe (see Section 1.6 for a description of vacuum in Yang-Mills gauge theories), C and CP are violated by the weak interactions and the thermal nonequilibrium is given by the expansion of the universe. However, quantitatively the amount of CP violation seems not to be consistent with the measurement in 1.41. A rough estimate can be calculated using a convention-independent measure of CP violation, the Jarlskog determinant [26], which is the determinant of the commutator of the quark mass matrices and it is found to be, in terms of quark masses and mixing angles,

$$D = \sin(\theta_{12}) \sin(\theta_{23}) \sin(\theta_{13}) \delta_{CKM} \Delta^2_{tc} \Delta^2_{tu} \Delta^2_{cu} \Delta^2_{bs} \Delta^2_{bd} \Delta^2_{sd}, \tag{1.47}$$

where $\Delta_{ij}^2 = (m_i^2 - m_j^2)$. To construct a dimensionless quantity it is necessary to divide D by a temperature to the 12^{th} power; the relevant temperature would be the sphaleron temperature $T_{sph} \simeq 100 - 300$ GeV, below which any baryon to antibaryon asymmetry freezes out [23]. Calculating this ratio gives

$$\frac{D}{T_{sph}^{12}} \sim 10^{-20} \ll \eta, \tag{1.48}$$

about 10 orders of magnitude below the measured value. Other, more meticulous, calculations give quantities within a similar range [27].

Measurements are ongoing to search for other sources of *CP* violation in the SM. It is possible that the solution to this problem lies in the neutrino sector, and experiments are now being planned to search for *CP* violation in neutrino interactions and to measure the value of any associated complex phase in the PMNS matrix. Another source should come from the strong interactions, as the QCD Lagrangian contains a term which should give rise to *CP* violation, although no evidence for this has yet been seen. This is described in more detail in the next section.

1.6 The strong *CP* problem

The strong CP problem arises from the structure of the degeneracy of the vacuum when passing from a classical non-Abelian Yang-Mills gauge theory to a quantised one. In fact, for a classical Yang-Mills theory, the ground state is defined as the state for which the field strength configuration

$$F_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu + \left[A_\mu, A_\nu\right] \tag{1.49}$$

is zero. In terms of fields A_μ, this translates to

$$A_\mu = (\partial_\mu g(x)) g(x)^{-1}, \tag{1.50}$$

i.e. the vacuum state is realised by $A_\mu = 0$ and all the local gauge transformations $g(x)$ of zero. It is convenient to label all the possible gauge transformations into different classes based on topological categories. This can be done by means of the winding number of the transformation: in fact, it can be proven that gauge transformations with the same winding number belong to the same homotopy class, i.e. they can all be continuosly deformed from one to another. In other words, gauge transformations of the same homotopy class identify the

same vacuum state. The winding number, which is defined as

$$\nu = \frac{g^2}{16\pi^2}\epsilon^{\mu\nu\rho\sigma}\int \mathrm{d}^4x\ \mathrm{Tr}\left(F_{\rho\sigma}F_{\mu\nu}\right), \tag{1.51}$$

is an integer; this implies that a discrete infinity of homotopy classes exists, each one corresponding to a topologically distinct vacuum state of the theory.

In a quantum theory, tunnelling through the potential barriers can connect these states: this means that no one of the infinite $|\nu\rangle$ vacua is an acceptable physical ground state of the theory. Instead, a gauge invariant superposition of n-vacua can be defined, called θ-vacuum:

$$|\theta\rangle = \sum_\nu e^{i\nu\theta}\,|\nu\rangle\,. \tag{1.52}$$

When calculating vacuum to vacuum transitions, one obtains

$$\langle\theta'|e^{-iHt}\,|\theta\rangle = \delta(\theta'-\theta)\sum_n e^{in\theta}\,\langle n|\,e^{-iHt}\,|0\rangle\,, \tag{1.53}$$

$$\langle n|e^{-iHt}\,|0\rangle = \int [\mathrm{d}A_\mu]e^{iS}, \tag{1.54}$$

where the second equation is expressed in terms of a path integral. By combining 1.54 with 1.53 and substituting the definition of winding number in 1.51, it can be seen that the $e^{in\theta}$ factor can be absorbed in the path integral as

$$e^{iS}e^{in\theta} = \exp\left[i\int \mathrm{d}^4x\left(\mathcal{L} + \theta\frac{g^2}{32\pi^2}\epsilon^{\mu\nu\rho\sigma}F^a_{\rho\sigma}F^a_{\mu\nu}\right)\right] = \exp\left[i\int \mathrm{d}^4x\,(\mathcal{L}+\mathcal{L}_\theta)\right] \tag{1.55}$$

This new term in the Lagrangian is a product of the Levi-Civita symbol, which is antisymmetric, with FF, which is symmetric; therefore, the whole term changes sign under P and CP.

The θ phase could be easily removed by a chiral rotation of the fields, keeping the total Lagrangian invariant; however, this transformation would affect also the quark mass matrix phase. In fact, the invariant, and physically observable, phase is

$$\bar{\theta} = \theta - \mathrm{Arg(Det(M))}. \tag{1.56}$$

As is shown in Chapter 2, the latest measurements constrain $\bar{\theta}$ to be less than 10^{-10}. Since both angles are free, independent parameters, their cancellation is regarded as a fine tuning problem. Several solutions have been proposed, which involve axions [28], extra space-time dimensions [29], massless up quarks [30], string theory [31] or quantum gravity [32].

Finally, in this Section the θ term is obtained in a generic quantum non-Abelian Yan-Mills gauge theory, and hence in principle also the electroweak Lagrangian should contain an analogous term. However, in this case it can be demonstrated that the dependence on the θ parameter vanishes and no additional *CP* violation is induced [33].

2 Latest experimental results

In this Chapter, the most up-to-date results on both the strong *CP* violation sector, focussing on the decays $\eta^{(\prime)} \to \pi^+\pi^-$, and tests of lepton universality are discussed, excluding the work presented in this thesis (which has been published [34]).

2.1 Strong *CP* violation

2.1.1 Neutron electric dipole moment measurement

A non-zero QCD θ term would induce an electric dipole moment in the neutron (nEDM), as the *P*- and *CP*-violating interaction arising from the structure of the vacuum would modify the charge distribution inside the neutron. This effect can be modelled by virtual charged pion and kaon clouds surrounding the neutron, where one of the vertices connecting the neutron and the meson loop is the *CP*-violating interaction given by the QCD θ term. The total effect, obtained by summing the pion and kaon loops, is given by [35]

$$d_n = (1.4 \pm 0.1) \times 10^{-16} \times \overline{\theta}, \tag{2.1}$$

in units of $e \cdot$ cm. Measurements of the nEDM are possible by measuring the Larmor precession of the neutron spin under magnetic and electric fields [36]; the precession frequency is given by

$$h\nu_{\pm} = 2\mu_n B \pm 2d_n E, \tag{2.2}$$

where $\mu_n = -1.91304273 \pm 0.00000045$ [37] is the neutron magnetic dipole moment, B and E are the intensities of the magnetic and electric fields and the sign depends on whether the electric and magnetic fields are parallel or antiparallel. The nEDM is simply obtained by the difference between the two

frequencies,

$$d_n = \frac{h(\nu_+ - \nu_-)}{4E}. \tag{2.3}$$

The most precise measurement of the nEDM [38] sets the limit

$$|d_n| < 0.30 \times 10^{-25} e \cdot \text{cm} \tag{2.4}$$

at 90% Confidence Level (CL), which translates to a limit on the QCD $\overline{\theta}$ angle of

$$\overline{\theta} < 2 \times 10^{-10} \tag{2.5}$$

at 90% CL.

As anticipated in Section 1.6, the observed limit on $\overline{\theta}$, which could in principle assume any value in the range $[0, 2\pi]$, is regarded as a fine tuning problem; moreover, the expected branching fractions of the *CP*-violating decays $\eta \to \pi^+\pi^-$ and $\eta' \to \pi^+\pi^-$ mediated by the QCD θ term would be $BR \lesssim 3 \times 10^{-17}$ [39], which is about 12 orders of magnitude smaller than the value currently accessible by experiments. Any observation of BRs larger than this value would indicate a new source of *CP* violation in the strong interactions. More details are given in Chapter 4.

2.1.2 The decay $\eta \to \pi^+\pi^-$

The best experimental limit on the branching fraction of the decay $\eta \to \pi^+\pi^-$ comes from the KLOE experiment [40] at the Frascati ϕ factory DAΦNE, based on an integrated luminosity of 350 pb^{-1} collected during 2001 and 2002. KLOE looked for $\eta \to \pi^+\pi^-$ candidates in radiative $\phi(1020) \to \eta\gamma$ decays by fitting to the $\pi^+\pi^-$ mass spectrum from candidate $\phi(1020) \to \pi^+\pi^-\gamma$ events, and using a sample of 1.55×10^7 events of the type $\phi(1020) \to \eta\gamma$, $\eta \to \pi^0\pi^0\pi^0$ as normalisation channel. Figure 2.1 shows the fit to the dipion invariant mass used to extract the observed number of $\eta \to \pi^+\pi^-$, which is found to be $N_s < 33$ at 90% CL. This leads to a 90% CL limit on the branching fraction of

$$\mathcal{B}(\eta \to \pi^+\pi^-) = \frac{N_s}{N_\eta \epsilon} < 1.3 \times 10^{-5}, \tag{2.6}$$

where N_η is the total number of η candidates in the sample, evaluated from the normalisation channel, and ϵ is the total efficiency, evaluated from simulations.

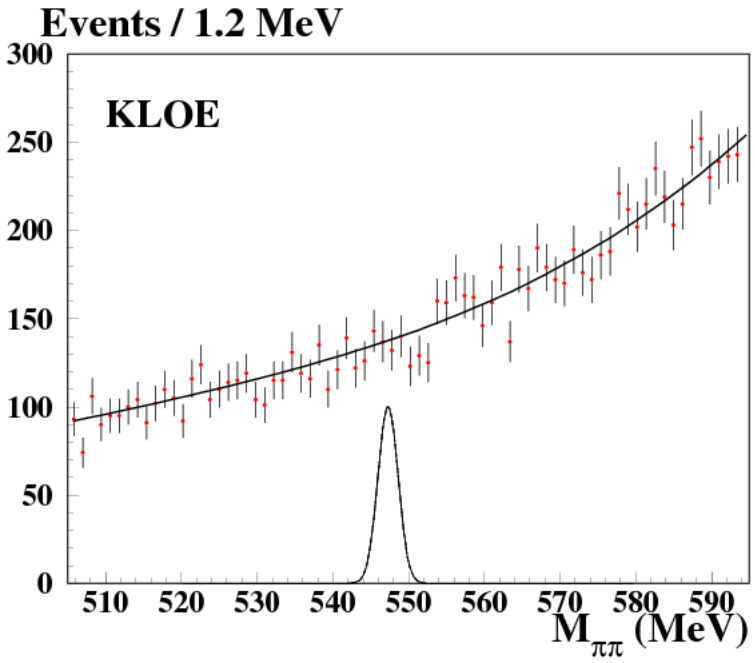


Figure 2.1: Search for $\eta \to \pi^+\pi^-$ at KLOE [40]. The signal peak, with an arbitrary normalisation, is drawn for illustration purposes.

2.1.3 The decay $\eta' \to \pi^+\pi^-$

The best experimental limit on the branching fraction of the decay $\eta' \to \pi^+\pi^-$ comes from the BESIII experiment [41] at the BEPCII accelerator in Beijing. BESIII searched for $\eta' \to \pi^+\pi^-$ from radiative $J/\psi \to \eta'\gamma$ decays, in a sample of $(225.2 \pm 2.8) \times 10^6$ J/ψ events. A fit to the dipion invariant mass spectrum, shown in Figure 2.2, using a signal shape from simulation, gives an upper limit on the number of observed signal candidates of $N_s < 32$ at 90% CL. The upper limit on the branching fraction at 90% CL is then given by

$$\mathcal{B}(\eta' \to \pi^+\pi^-) = \frac{N_s}{N_{J/\psi}\epsilon\mathcal{B}(J/\psi \to \eta'\gamma)} < 5.5 \times 10^{-5}, \tag{2.7}$$

where $N_{J/\psi}$ is the number of J/ψ candidates in the sample and ϵ is the total efficiency, evaluated from simulation.

2.2 Tests of lepton universality

Lepton universality has been tested meticulously since the discovery of the W and Z bosons in the 1980s, by measuring the partial widths or branching fractions of leptonic or semileptonic decays of various particles. They have

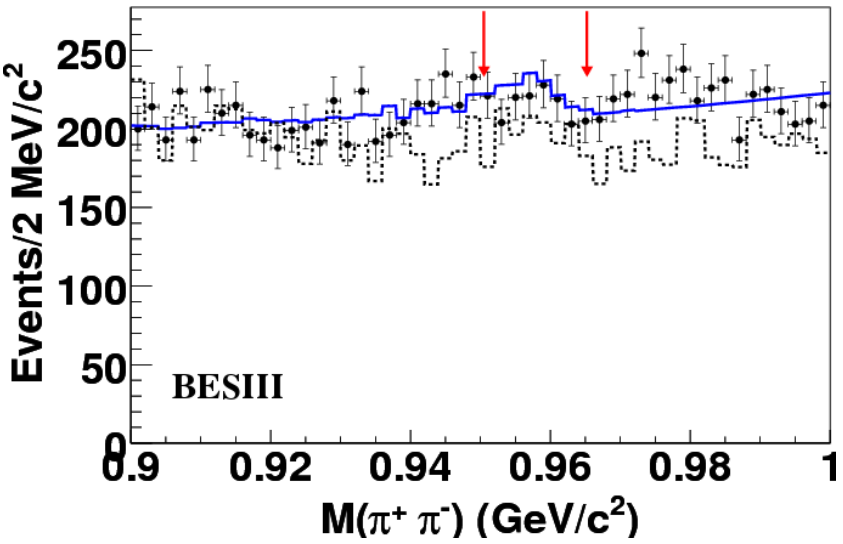


Figure 2.2: Search for $\eta' \to \pi^+\pi^-$ at BESIII [41]. The dotted line is the sum of all simulated backgrounds, which is not used in the fit (represented by the blue line). The red arrows define the shortest interval containing 95% of signal.

been tested with $Z \to ll$ and $W \to l\nu$ decays (see Section 2.2.1), with kaon decays [42, 43], with pion decays [44], with precise measurements of the τ leptonic decays [45] and in leptonic and semileptonic B decays [46]; these last studies have resulted in observations that seem to challenge the SM assumption of lepton universality. They are discussed in detail in Section 2.2.2 and Section 2.2.3.

2.2.1 Measurements of the widths of the weak gauge bosons

An indirect method to test the universality of weak interactions is to measure the total width of the W boson. In fact, the total width, ignoring neutrino mixing, is given by the sum of the individual widths which are

$$\Gamma_W = \Gamma(l^-\bar{\nu}_l) + \Gamma(d'u) + \Gamma(s'c), \tag{2.8}$$
$$\Gamma(l\nu) = \Gamma(W^- \to e^-\bar{\nu}_e) + \Gamma(W^- \to \mu^-\bar{\nu}_\mu) + \Gamma(W^- \to \tau^-\bar{\nu}_\tau),$$
$$\Gamma(d'u) = N_c\left[\Gamma(W^- \to d\bar{u})|V_{ud}|^2 + \Gamma(W^- \to s\bar{u})|V_{us}|^2 + \Gamma(W^- \to b\bar{u})|V_{ub}|^2\right],$$
$$\Gamma(s'c) = N_c\left[\Gamma(W^- \to d\bar{c})|V_{cd}|^2 + \Gamma(W^- \to s\bar{c})|V_{cs}|^2 + \Gamma(W^- \to b\bar{c})|V_{cb}|^2\right],$$

where the elements of the CKM matrix associated to the rotation of the states $|d(s)'\rangle = V_{ud(cd)}|d\rangle + V_{us(cs)}|s\rangle + V_{ub(cb)}|b\rangle$ have been made explicit. Assuming

that the W boson has the same coupling to all leptons and quarks, one gets

$$\begin{aligned}
\Gamma_l &= 3\Gamma(W^- \to e^-\bar{\nu}_e), \\
\Gamma(d'u) &= N_c\Gamma(W^- \to e^-\bar{\nu}_e)\left[|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2\right], \\
\Gamma(s'c) &= N_c\Gamma(W^- \to e^-\bar{\nu}_e)\left[|V_{cd}|^2 + |V_{cs}|^2 + |V_{cb}|^2\right].
\end{aligned} \tag{2.9}$$

Using the unitarity of the CKM matrix and setting the number of colours to $N_c = 3$, then

$$\Gamma_W = 9\,\Gamma(W^- \to e^-\bar{\nu}_e) = 9\frac{G_F M_W^3}{6\sqrt{2}\pi} = 2043 \pm 12\,\text{MeV}, \tag{2.10}$$

to be compared to the measured width from combined results of LEP and Tevatron experiments [47],

$$\Gamma_W^{exp} = 2085 \pm 42\,\text{MeV}, \tag{2.11}$$

which is in agreement with the assumptions used in Equation (2.10).

Decays of the Z boson into charged leptons provide another environment to test lepton universality. The partial width $\Gamma_{Z\to ll}$ for the decay $Z \to l^+l^-$, in the limit of massless leptons, is given by

$$\Gamma_{Z\to ll} = \frac{G_F M_Z^3}{12\sqrt{2}\pi}\left[\left(1 - 2\sin^2(\theta_W)\right)^2 + 4\sin^4(\theta_W)\right] \simeq 83.84\,\text{MeV}, \tag{2.12}$$

and does not depend on the lepton species. The experimental values of the partial widths, measured by LEP experiments [47], are

$$\begin{aligned}
\Gamma_{Z\to ee} &= (83.91 \pm 0.10)\,\text{MeV}, \\
\Gamma_{Z\to \mu\mu} &= (83.99 \pm 0.17)\,\text{MeV}, \\
\Gamma_{Z\to \tau\tau} &= (84.09 \pm 0.20)\,\text{MeV},
\end{aligned} \tag{2.13}$$

which are consistent with the hypothesis of lepton universality.

2.2.2 The $b \to sll$ anomalies

In the SM, $b \to sll$ decays proceed via flavour-changing neutral current (FCNC) transitions, which cannot occur at tree level; Figure 2.3 shows, as an example, the lowest-order, one-loop diagrams allowed in the Standard Model for the decay $B^0 \to K^{*0}l^+l^-$, where the antiquark line in the loops can be $\bar{t}, \bar{c}, \bar{u}$. Other $b \to sll$ decays proceed with similar diagrams, and they are highly sensitive

to the possible contribution of new virtual particles. The LHCb experiment has measured the differential branching fractions in bins of q^2 (the square of the 4-momentum transfer to the lepton pair, $q^2 = (p(l) + p(l'))^2$) for several decay channels, and the results, shown in Figure 2.4, are consistently lower than the SM predictions in the low q^2 regions [48–51], although there is not a single channel with a statistically significant deviation yet.

Moreover, the angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$ (see Figure 2.5) shows a discrepancy with respect to the SM prediction on the obsevable P_5', which is defined as

$$P_5' = \frac{S_5}{\sqrt{F_L(1-F_L)}}, \tag{2.14}$$

where F_L is the fraction of longitudinal polarisation of the K^{*0} meson, and S_5 is a bilinear combination of the K^{*0} decay amplitudes, defined in [52]. This discrepancy is quantified at 3.4 σ (standard deviatons) from a global maximum likelihood fit.

A last set of anomalies in the $b \to sll$ transitions, of particular interest for this work, is given by the measurement of the branching fractions ratios $R(K)$ and $R(K^*)$, shown in Figures 2.6 and 2.7 and defined as a double ratio of the branching fractions of $B \to hl^+l^-$ and $B \to h(J/\psi \to l^+l^-)$ decays:

$$R(h) = \frac{\mathcal{B}(B \to h\mu^+\mu^-)}{\mathcal{B}(B \to h(J/\psi \to \mu^+\mu^-))} \Bigg/ \frac{\mathcal{B}(B \to he^+e^-)}{\mathcal{B}(B \to h(J/\psi \to e^+e^-))}, \tag{2.15}$$

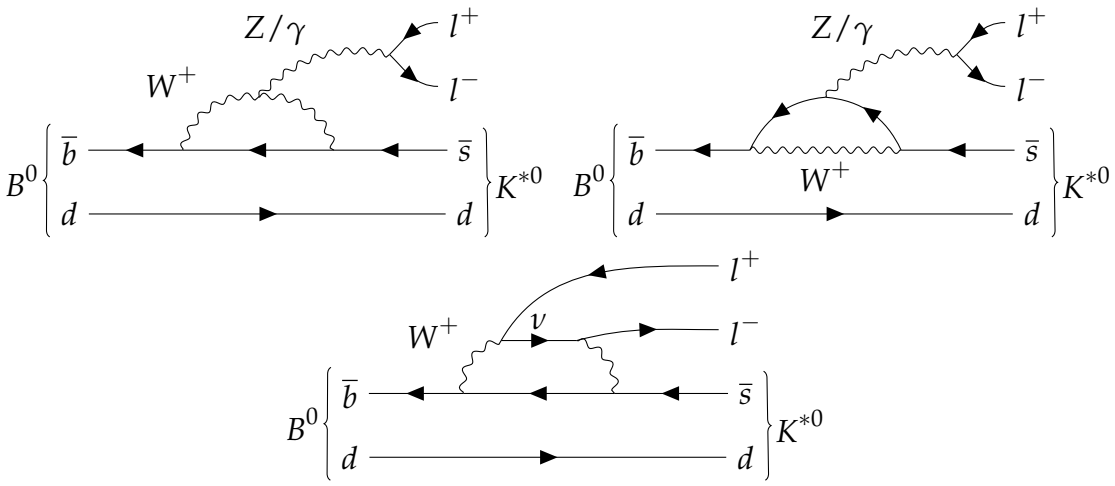
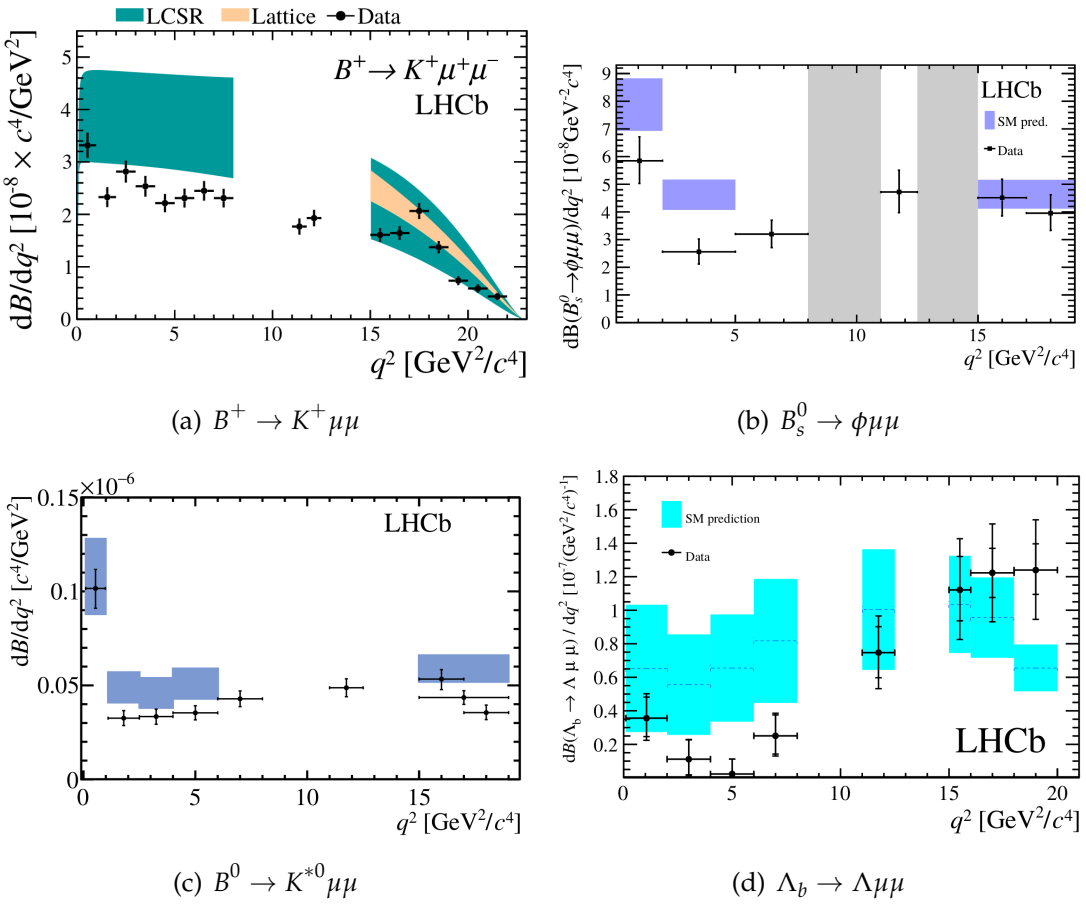


Figure 2.3: $b \to sll$ decays via FCNC at one-loop order in the SM.



(a) $B^+ \to K^+\mu\mu$ (b) $B_s^0 \to \phi\mu\mu$

(c) $B^0 \to K^{*0}\mu\mu$ (d) $\Lambda_b \to \Lambda\mu\mu$

Figure 2.4: Differential branching fractions for different channels measured by LHCb [48–51].

where $h = K^+, K^{*+}$. The results of the LHCb measurements are [55, 56]

$$R(K) = 0.745^{+0.090}_{-0.074} \pm 0.036, \tag{2.16}$$

$$R(K^*) = \begin{cases} 0.66\ ^{+\ 0.11}_{-\ 0.07} \pm 0.03 & \text{for } 0.045 < q^2 < 1.1\ \mathrm{GeV}^2c^4, \\ 0.69\ ^{+\ 0.11}_{-\ 0.07} \pm 0.05 & \text{for } 1.1 \quad < q^2 < 6.0\ \mathrm{GeV}^2c^4, \end{cases} \tag{2.17}$$

where the first uncertainties are statistical and the second are systematic. In both cases, a tension with the Standard Model prediction is observed, at 2.6 σ for $R(K)$, at $2.1 - 2.3\ \sigma$ for $R(K^*)$ in the low-q^2 region and $2.4 - 2.5\ \sigma$ in the high-q^2 region. There are several models developed to explain the $b \to sll$ anomalies observed by LHCb [57–59], including: a new gauge boson, called Z' [60–62], which induces FCNC at tree level and has different coupling with the different lepton flavours; a scalar leptoquark (LQ) [63–65], which induces FCNC at tree level as well and presents quark-lepton Yukawa interaction vertices. Their

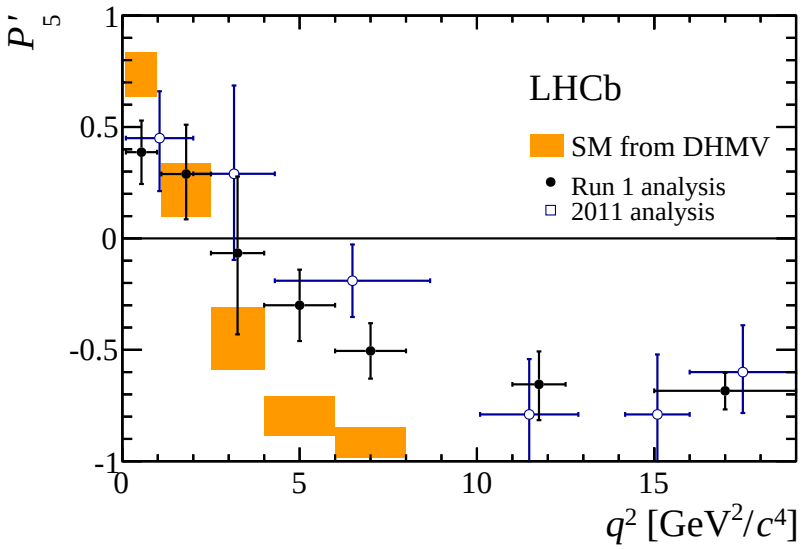


Figure 2.5: Distribution of the angular variable P'_5 as a function of q^2. The blue points are obtained from the 2011 data sample [53], while the black points are from the full Run 1 dataset [54].

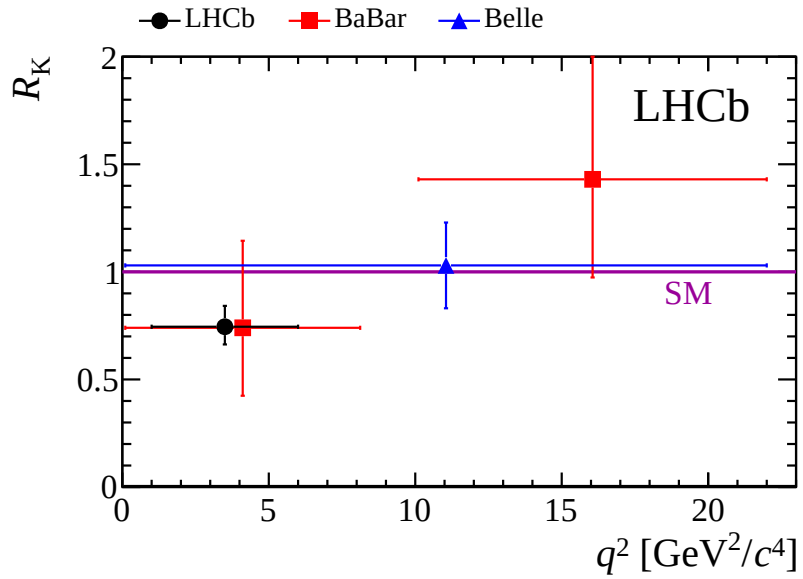


Figure 2.6: $R(K)$, binned in q^2, with results from Belle and BaBar and the SM prediction [55].

action is schematised by the Feynman diagrams in Figure 2.8. These same models can explain the other lepton universality deviations observed by LHCb, Belle and BaBar in other channels which are presented in Section 2.2.3 and Section 2.2.4.

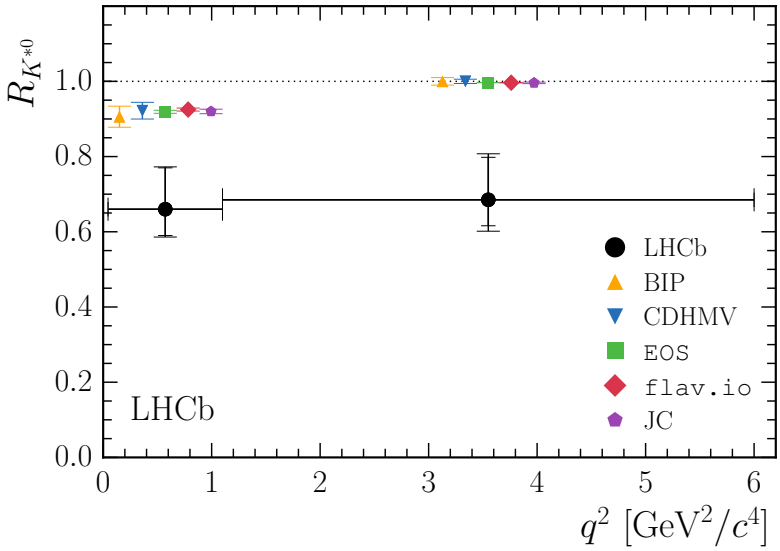


Figure 2.7: $R(K^*)$, binned in q^2, with different SM predictions [56].

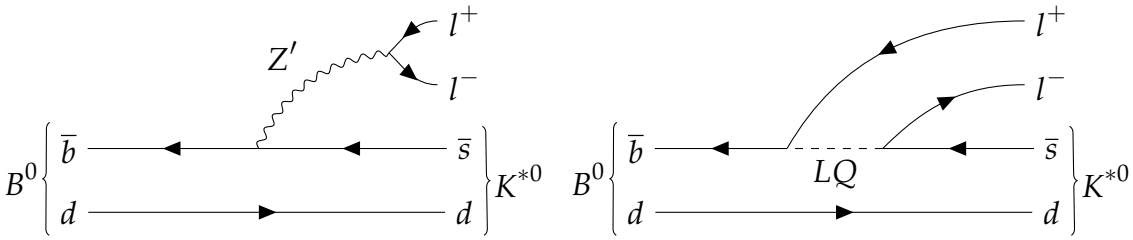


Figure 2.8: Z' and leptoquark possible contributions at tree level.

2.2.3 $R(D)$, $R(D^*)$ and $R(J/\psi)$

Apart from the hints of lepton universality deviation with $b \to sll$ channels ($R(K), R(K^*)$), other double ratios have been measured by LHCb and other experiments, showing a similar behaviour. Both Belle and BaBar have measured the excess of semileptonic $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays as

$$R(D) = \frac{\mathcal{B}(\bar{B} \to D\tau^-\bar{\nu}_\tau)}{\frac{1}{2}\left[\mathcal{B}(\bar{B} \to De^-\bar{\nu}_e) + \mathcal{B}(\bar{B} \to D\mu^-\bar{\nu}_\mu)\right]}, \tag{2.18}$$

$$R(D^*) = \frac{\mathcal{B}(\bar{B} \to D^*\tau^-\bar{\nu}_\tau)}{\frac{1}{2}\left[\mathcal{B}(\bar{B} \to D^*e^-\bar{\nu}_e) + \mathcal{B}(\bar{B} \to D^*\mu^-\bar{\nu}_\mu)\right]}, \tag{2.19}$$

while LHCb has measured only the latter, with a slightly different definition:

$$R(D^*) = \frac{\mathcal{B}(\bar{B} \to D^*\tau^-\bar{\nu}_\tau)}{\mathcal{B}(\bar{B} \to D^*\mu^-\bar{\nu}_\mu)}. \tag{2.20}$$

While Belle and BaBar take the average of the muonic and electronic channels, LHCb only considers the muonic channel as the detection efficiency and momentum resolution are both higher with respect to the channel with electrons. In total, BaBar has measured both $R(D)$ and $R(D^*)$ simultaneously [66] with purely leptonic τ decays; Belle has measured both $R(D)$ and $R(D^*)$ simultaneously [67] with purely leptonic τ decays and has published two more measurements of $R(D^*)$, one with semileptonic tag and purely leptonic τ decays [68], the other with hadronic tag and semileptonic $\tau \to \pi\nu_\tau$, $\tau \to \rho\nu_\tau$ decays [69]; LHCb has measured $R(D^*)$ twice, once with purely muonic τ decays [70] and another with 3-prong τ decays [71]. All these measurements are summarised by the Heavy Flavour Averaging Group (HFLAV) in Figure 2.9, where it is shown that the global average is 4.1 σ away from the SM prediction. Figure 2.10 shows all the measurements of the individual channels; it is worth noticing how, even though none of the cited measurements alone can claim an observation of lepton universality violation, they all deviate in the same direction with respect to the SM prediction.

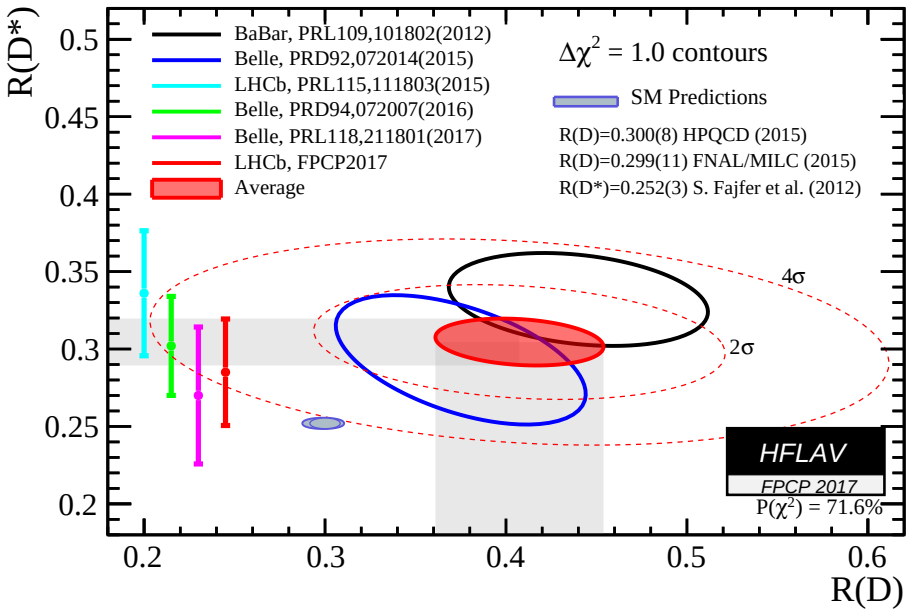
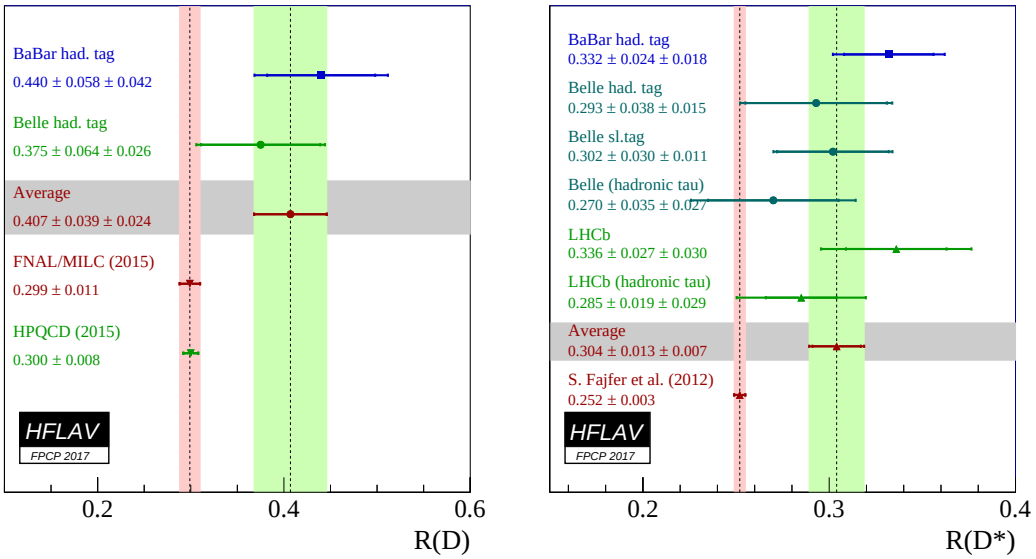


Figure 2.9: All measurements of $R(D)$ vs $R(D^*)$, with the HFLAV average and the SM prediction [72].

Another lepton universality test performed by LHCb is the measurement of



(a) $R(D)$ (b) $R(D^*)$

Figure 2.10: All mesurements of $R(D^*)$ and $R(D^*)$ separately, with the HFLAV average and the SM prediction [72].

the ratio [71]

$$R(J/\psi) = \frac{\mathcal{B}(B_c^+ \to J/\psi \tau^+ \nu_\tau)}{\mathcal{B}(B_c^+ \to J/\psi \mu^+ \nu_\mu)}, \tag{2.21}$$

where the τ lepton is reconstructed with the purely muonic channel $\tau \to \mu \nu_\mu \bar{\nu}_\tau$. This is the first test of lepton universality using a B_c^+ decay and the first study of the semitauonic decay channel. The ratio is measured to be

$$R(J/\psi) = 0.71 \pm 0.17 \pm 0.18, \tag{2.22}$$

where the first uncertainty is statistical and the second is systematic. The result, shown in Figure 2.11, is about 1.3σ away from the SM prediction, and again deviates in the same direction as the $R(D)$ and $R(D^*)$ measurements.

2.2.4 Tests of lepton universality in the charm sector

Given the strong hints of violation of lepton universality in semileptonic B decays, it is natural to wonder if a similar disagreement can be observed in other semileptonic decays accessible by LHCb. In the charm sector, the individual branching fractions of $D^0 \to h^- \mu^+ \nu_\mu$ and $D^0 \to h^- e^+ \nu_e$, where $h = K, K^*, \pi$, have been already measured, but their ratio has never been directly measured by a single experiment. By taking the most up-to-date world averages from [47],

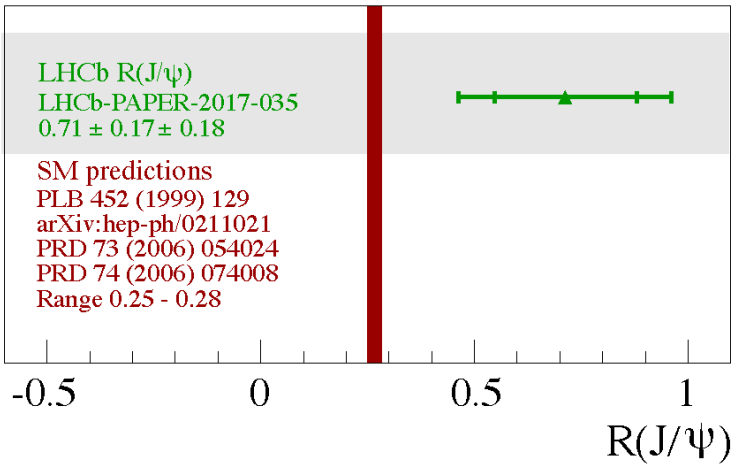


Figure 2.11: $R(J/\psi)$ from semileptonic B_c^+ decays, with SM predictions.

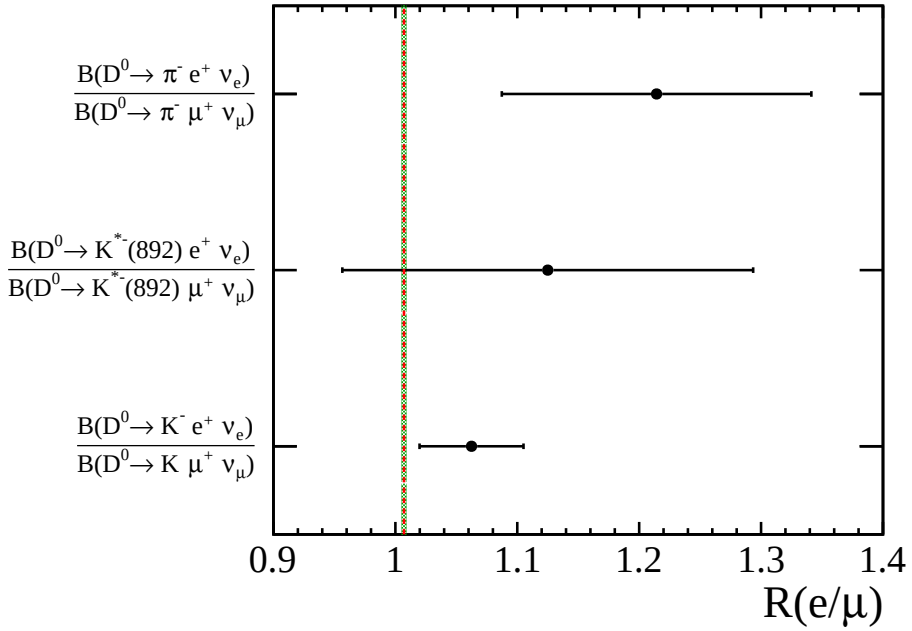


Figure 2.12: Branching fractions ratios for different semileptonic charm decays [73].

it is possible to calculate the branching fraction ratios for the modes involving the final states $K^{*-}l^+\nu_l$, $K^-l^+\nu_l$, $\pi^-l^+\nu_l$, which are summarised in Figure 2.12. Also in this case it can be noted that, even though all the measurements are consistent with the SM prediction, they all deviate in the same direction and also in the same direction as $R(K)$ and $R(K^*)$ (notice that the observables in Figure 2.12 are defined as the inverse with respect to the previously mentioned definitions of $R(K)$ and $R(K^*)$).

3 The LHCb experiment

The LHCb experiment is, along with ATLAS, CMS and ALICE, one of the main experiments at CERN (Conseil Européen pour la Recherche Nucléaire), outside the city of Geneva, and the only one specifically designed to study decays of beauty hadrons. The experiment aims to analyse data from proton-proton collisions created by the LHC circular accelerator (Large Hadron Collider), focusing on the decays of heavy mesons (containing charm and bottom quarks), to explore the asymmetry between matter and antimatter and to search for rare decays and rare particles, as well as possible new processes which are not predicted by the Standard Model of particle physics. The LHCb experiment is located along the LHC ring at the Intersection Point 8 (IP8), where, from 1989 to 2000, the DELPHI experiment was installed.

3.1 Large Hadron Collider

The particle accelerator LHC [74,75] is situated inside a 26.7 km circumference tunnel at an average depth of about 100 m underground, under Swiss and French soil, between the city of Geneva and the Jura mountain range. It is the most powerful particle accelerator ever built, operating at a centre of mass energy of 7 TeV from 2010 to 2011, 8 TeV during 2012 and 13 TeV from 2015, when the machine was restarted after an almost two years long scheduled interruption (Long Shutdown 1). The CERN accelerator complex is shown in Figure 3.1. The protons used for the collisions come from hydrogen, which is ionised with a duoplasmatron [76] and sent into the first accelerator of the CERN complex, the LINAC (LINear ACcelerator). The proton beam is accelerated using resonating electromagnetic fields, oscillating at a frequency of 400 MHz, generated by radiofrequency cavities; the waves are synchronised so that the protons are also divided into small bunches in this stage. After the

LINAC, they are sent to a series of increasingly larger circular accelerators: first the Proton Synchrotron Booster, then the Proton Synchrotron (PS), and then the Super Proton Synchrotron (SPS); finally, after the SPS, they are injected into the LHC. At each step, the protons are accelerated up to a certain energy, and when they leave the SPS they have an energy of 450 GeV and are travelling at 99.9998% of the speed of light.

The LHC is composed of two rings (beam pipes) where the proton bunches travel in opposite directions in ultra-high vacuum. The centripetal force and the beam focussing are given by a system of superconducting magnets placed alongside the ring, with an operating temperature of -271.25 $^{\circ}$C (1.9 K). The protons are then forced to collide at four points, corresponding to the positions of the four main experiments (ATLAS, ALICE, CMS, LHCb) with a bunch crossing rate of 40 MHz, i.e. 25 ns between two consecutive collisions. The LHC provides a nominal peak luminosity of $L = 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$; the LHCb experiment has been designed to record about 1/50 of such luminosity, to reduce the probability of misreconstruction of secondary vertices and particle misidentification due to pile-up. To obtain a lower luminosity independently from the other experiments, the beam crossing angle and offset at IP8 are adjusted; this also ensures that the luminosity level at LHCb is constant, which stabilises the data-taking conditions and reduces the impact on systematic uncertainties due to trigger performance (see Section 3.2.7) [77].

3.2 The LHCb detector

The LHCb detector is a single-arm forward spectrometer, roughly shaped like a pyramid with the vertex at the interaction point and the base perpendicular to the proton beam axis. Its angular acceptance ranges from $\approx$10 mrad to 300 mrad in the bending plane of its dipole magnet and to 250 mrad in the non-bending plane [79]. In Figure 3.2 the geometry of the detector and the position of each subdetector are shown. Its peculiar geometry is very different from the cylindrical design which characterises the general purpose detectors at CERN. For highly energetic pp collisions the b hadrons (and $\bar{b}$ hadrons) are produced mainly in the forward region. In fact, as shown in Figure 3.3, the b and $\bar{b}$ quarks are mainly produced at very small angles with respect to the beam axis and in the same hemisphere.

The LHCb detector is composed of several subdetectors: the VErtex LOcator

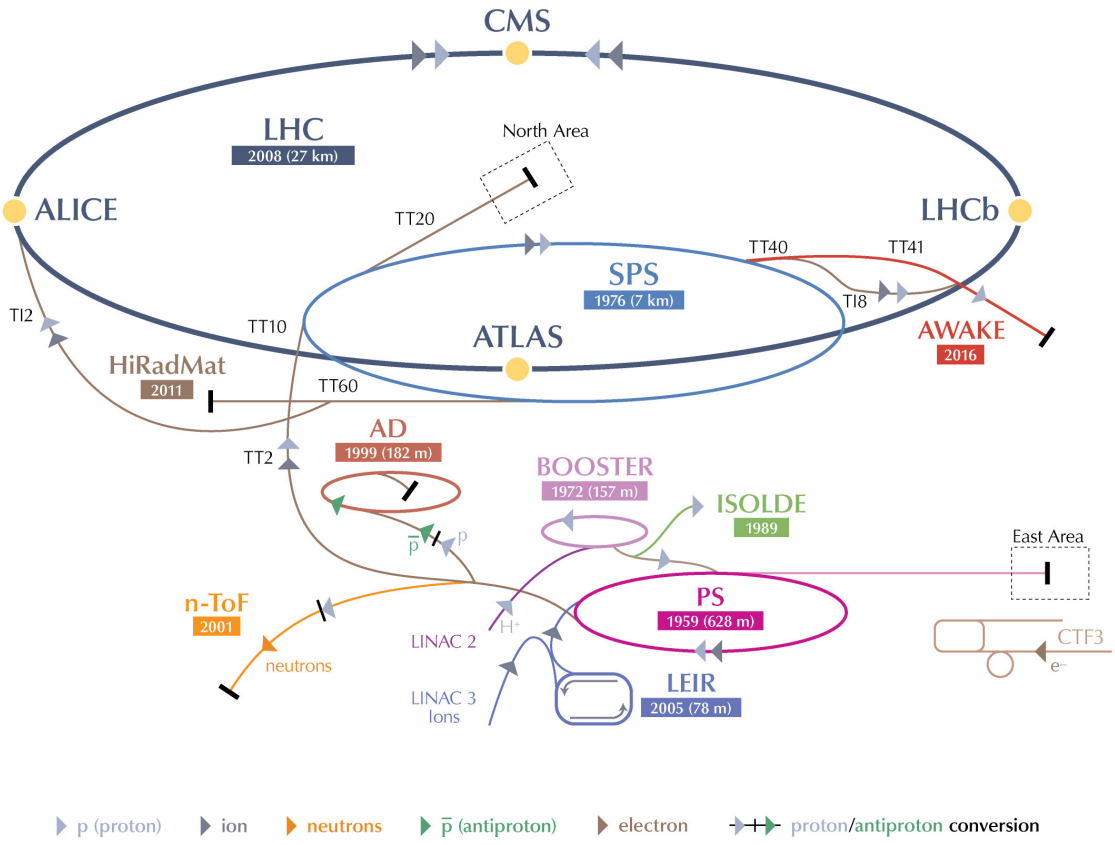


Figure 3.1: The CERN accelerator complex [78].

(VELO) in the close proximity of the interaction point, two Ring Imaging CHerenkov detectors (RICH), a warm dipole magnet, a tracking system, an electromagnetic and a hadronic calorimeter, and muon chambers at the end of the detector.

3.2.1 Vertex locator (VELO)

The closest subdetector to the interaction point is the VErtex LOcator (VELO) [82, 83]. Measuring the displacement of secondary vertices with respect to the point of interaction is of paramount importance for many of the physics analyses in LHCb, as a measurable flight distance (ranging from tens of microns to some centimeters) between the vertices of production and decay is a distinctive feature of heavy hadrons; therefore, given that vertices are reconstructed by combining tracks which originate from the same point, one of the fundamental requirements for the LHCb experiment is to have the capability to measure the position of track coordinates in the proximity of the interaction point as precisely as possible.

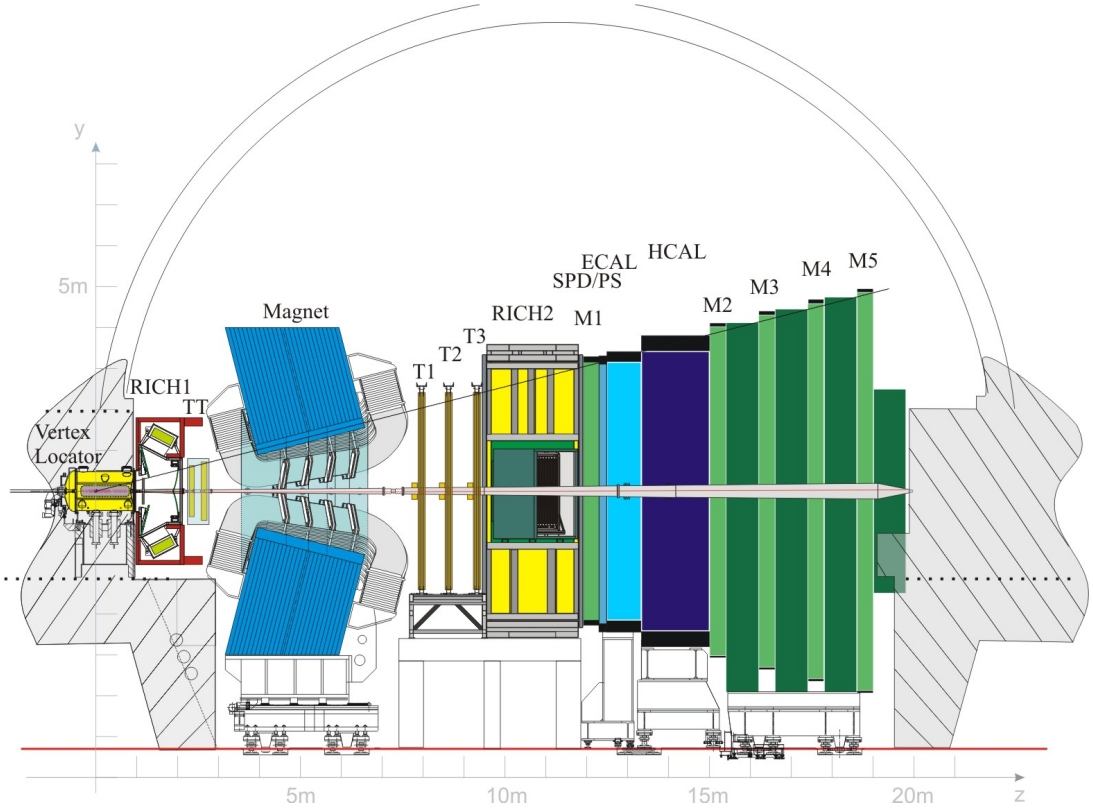


Figure 3.2: LHCb detector, lateral view (non-bending plane) [80].

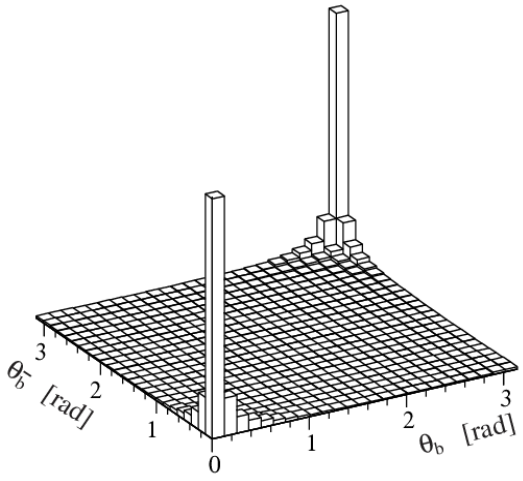


Figure 3.3: Polar angles of the b- and $\bar{b}$-hadrons calculated with PYTHIA [81].

The VELO system configuration and a schematic view of the sensors are shown in Figure 3.4. The VELO is a silicon tracker system, consisting of 25 stations placed around the beam pipe. Each station is composed of a pair of modules (left and right), and each module is composed of a radial sensor (R sensor), where the silicon strips are placed radially, and an angular sensor (Φ

sensor), where the silicon strips are parallel and concentric. When a charged particle travels through a module, it will deposit an electrical charge in both a radial and an angular strip, and from the combination of the two hits a 3D point is obtained; a 3D track can be extrapolated combining at least three points from different modules. The tracking reconstruction algorithm requires the track to traverse at least four modules in order to allow for the possibility of a missing hit. The azimuthal coverage for each sensor is about 182°, to give a small overlap between the right and left modules; this simplifies the relative alignment and guarantees a full azimuthal acceptance.

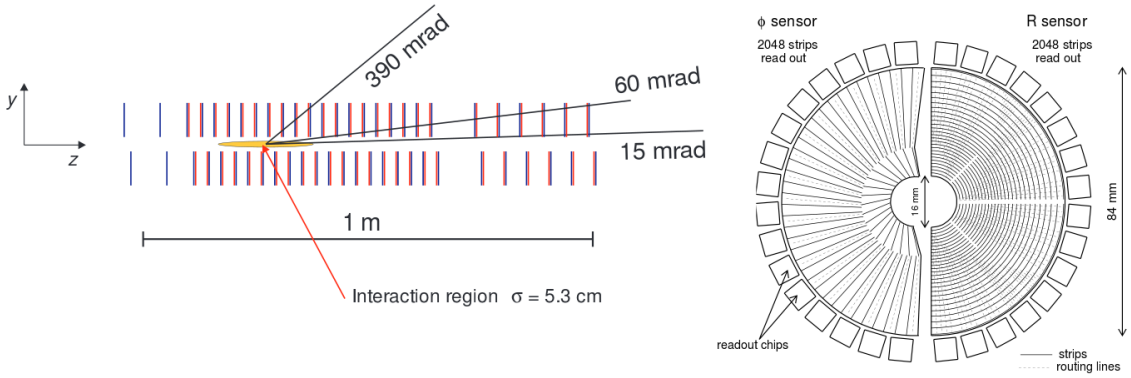


Figure 3.4: VELO system configuration and VELO sensors [81].

Apart from the position of primary and secondary vertices, another extremely important quantity used in data analysis is the impact parameter (IP) of a track. It is defined as the 3D Euclidean distance of closest approach with respect to a specific point, usually the primary or secondary vertex. The VELO sensors must be placed as close to the beam axis as possible: this corresponds to a distance of 8.2 mm, taking into account the LHC safety margin for the transversal spread of the proton beams (5 mm) plus the distance to the active area of the VELO sensors (3 mm). During the LHC injection phase, the position of the beam fluctuates and its transversal spread is larger; therefore, to avoid any risk of damage to the sensors, the VELO modules are installed onto a retractable support which puts the left and right sensors back together once the beam is declared to be stable. The sensors must operate in vacuum to achieve optimum performance. The whole VELO system is enclosed by a thin, corrugated aluminium sheet foil, which allows to create a separate vacuum with respect to the LHC one, and also shields the sensors against RF pickup and beam wakefields.

The VELO system measures tracks in the full LHCb angular acceptance; in

addition to that, the backwards hemisphere presents two upstream R sensors which are used to enhance the resolution on the primary vertex reconstruction and as a pile-up veto counter for the Level-zero (L0) trigger, which is described in Section 3.2.7. The uncertainties on the primary vertex position determination are linked mainly to the number of tracks produced in a pp-collision, as can be seen in Figure 3.5. For an average event, the resolution in the z-direction (the beam direction) is 42 μm and 15 μm perpendicular to the beam. The resolution on the decay length ranges from 220 μm to 370 μm, depending on the decay channel. The resolution on the IP is $(15 + 29/p_T)$ μm, where $p_T = \sqrt{p_x^2 + p_y^2}$ is the transverse momentum of the track in units of GeV/c.

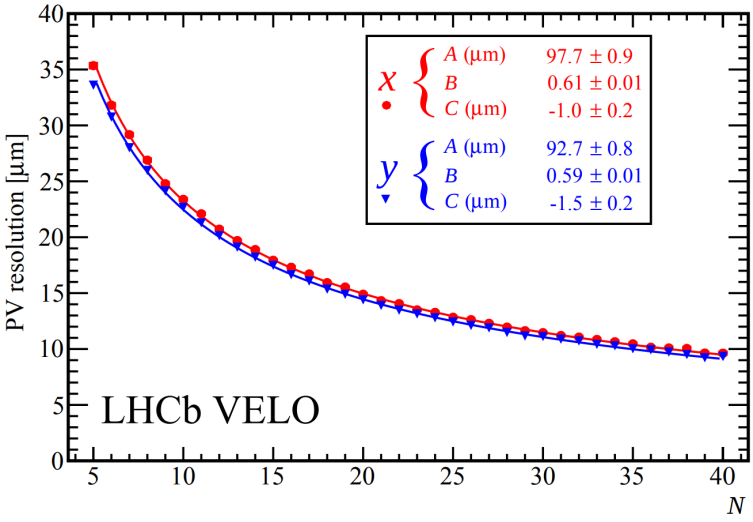
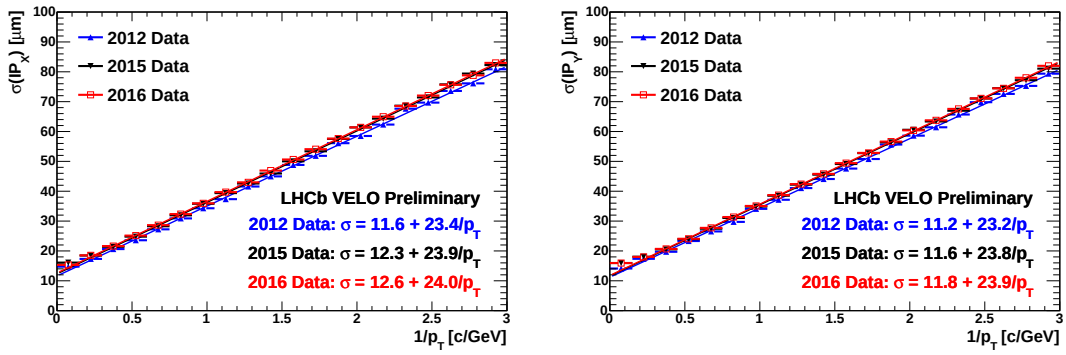


Figure 3.5: PV resolution in the transverse coordinates as a function of track multiplicity [83].

3.2.2 The Ring-Imaging Cherenkov detectors (RICH)

Charged hadron identification in LHCb is achieved with a high performance Ring-Imaging CHerenkov (RICH) system, composed of two detectors aiming at different momentum ranges [84, 85]. RICH1 is located upstream of the magnet and covers low momentum particles (from 1 GeV/c up to about 60 GeV/c) using silica aerogel and C_4F_{10} gas radiators with a polar angle acceptance from 25 to 300 mrad. RICH2, located downstream of the magnet and the tracking stations, has a more limited angular acceptance (from 15 to 120 mrad in the horizontal plane and from 15 to 100 mrad in the vertical plane); it covers the high momentum range, from about 15 GeV/c up to about 100 GeV/c, using a



(a) IP_x (b) IP_y

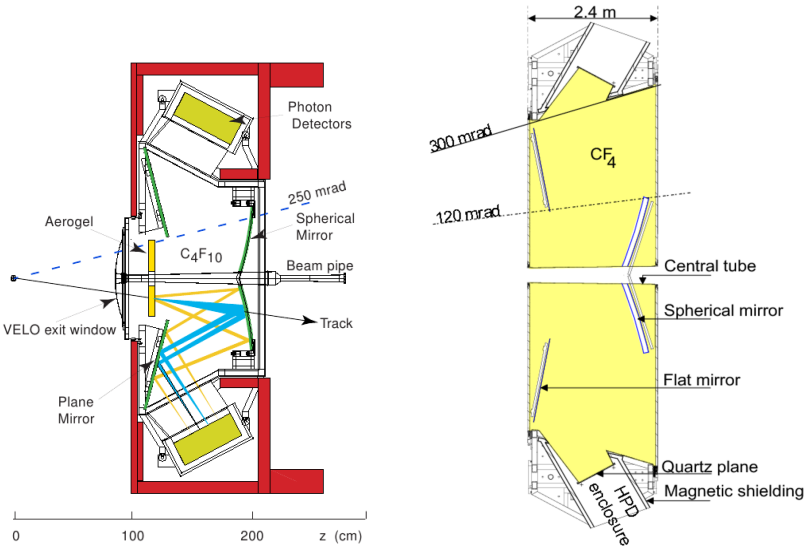
Figure 3.6: IP resolutions of the components as a function of the inverse of the transverse momentum. The resolution quoted in the text refers to the total $\mathrm{IP} = \sqrt{\mathrm{IP}_x^2 + \mathrm{IP}_y^2}$, where the IP_z component is neglected due to the LHCb forward geometry [83].

CF_4 radiator. Both layouts are shown in Figure 3.7.

The RICH system uses Cherenkov light, emitted by charged particles travelling through the radiators, to calculate the likelihood of a specific mass hypothesis to assign to a given track. The Cherenkov angle distribution for isolated tracks, as a function of track momentum, presents distinct bands according to their mass, as can be seen in Figure 3.8. More details on particle identification in LHCb are given in Section 3.3.2. Cherenkov light is focused onto the photon detector planes using tilted spherical mirrors and secondary plane mirrors, in order to reflect the image out of the spectrometer acceptance. The baseline photon detectors are multianode photomultiplier tubes (MaPMT). The anodes are arranged in an 8×8 array of pixels, each 2 mm $\times$ 2 mm, separated by 0.3 mm gaps.

3.2.3 Tracking system

The tracking system (see Figure 3.2) consists of four stations (TT, T1, T2, T3) placed between the VELO and the calorimeters: three of them (T1, T2 and T3) are identical and placed after the magnet, while the other one (TT, i.e. Tracker Turicensis) is located between the VELO and the magnet [86, 87]. T1, T2 and T3 are divided into two different sectors, where two different detectors are used: the Inner Tracker (IT) [88], shown in Figure 3.9, is a silicon microstrip detector that uses about 130000 microstrips with a strip pitch of 198 μm, capable of achieving a 50 μm single hit resolution; the Outer Tracker [89] (OT), shown in



(a) RICH1 (b) RICH2

Figure 3.7: RICH detectors layout [81].

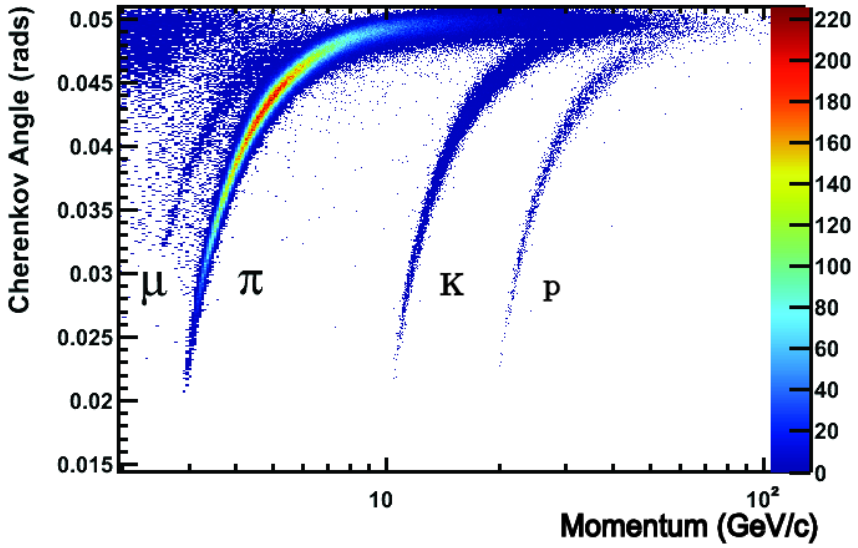


Figure 3.8: RICH1 Cherenkov angle (isolated tracks) as a function of momentum showing different particle species [85].

Figure 3.10, is a drift-time detector, consisting of straw tube modules with two staggered layers of drift tubes having an inner diameter of 4.9 mm, filled with Argon and CO_2, with less than 50 ns drift time and a drift-coordinate resolution of 200 μm. For each of the three downstream stations, IT modules cover the 120 cm by 40 cm central region, where the particle fluence is higher and a higher

spatial resolution is required, while OT modules cover the rest of the angular acceptance. To improve track reconstruction, all the stations are composed of four layers arranged in an x-u-v-x layout, such that the silicon microstrips in the IT and the straw tubes in the OT are vertically aligned in the first and in the last layer, whereas the other two (u,v) layers are rotated by stereo angles of $\pm 5^{\circ}$, providing the required sensitivity in the vertical direction.

The Tracker Turicensis (TT), shown in Figure 3.11, is quite similar to the IT for composition and performance; the only important differences are in the active area (about 3.6 times larger than IT) and the shape (square instead of elliptical). Given its reduced dimensions, the TT is composed only of slicon microstrips and, as it is located closer to the interaction point, it covers the full LHCb acceptance. As the other stations, TT also presents four layers in an x-u-v-x layout. In Figure 3.12 the full tracking system layout is shown.

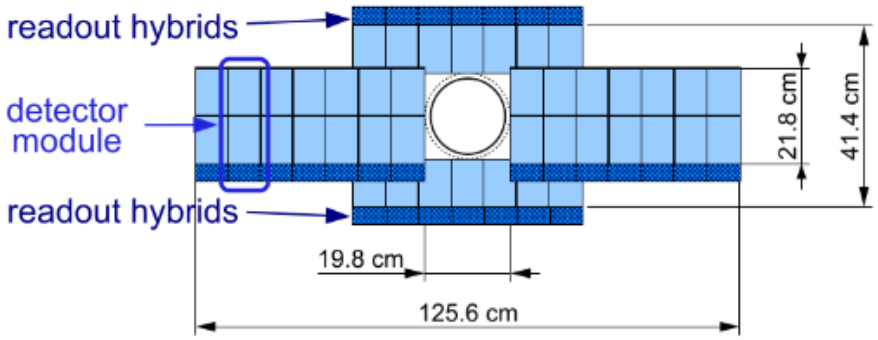


Figure 3.9: Inner Tracker layout [88].

3.2.4 Magnet

A warm dipole magnet is used to bend the tracks of charged particles in the horizontal plane, in order to allow the measurement of their momentum and their charge from the curvature of the trajectories [92]; a perspective view is given in Figure 3.13. The magnet consists of two trapezoidal coils, sloping along the z direction in order to match the detector acceptance, bent at 45° on the two transverse sides. The bending power, given by the integrated magnetic field, is 4 Tm. This is enough to measure momenta of charged particles with a resolution of $\Delta p/p = 0.4\%$ at low momentum (from approximately 250 MeV/c up to 10 GeV/c), up to $\Delta p/p = 1\%$ at 200 GeV/c. To reduce the impact of any potential asymmetric imperfection of the detector on measurements of

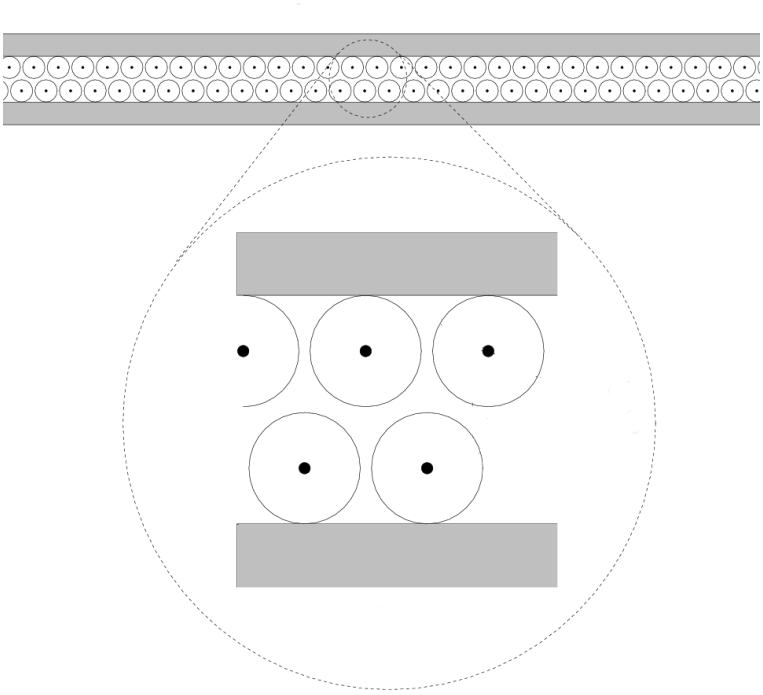


Figure 3.10: Outer Tracker straw tubes layout [89].

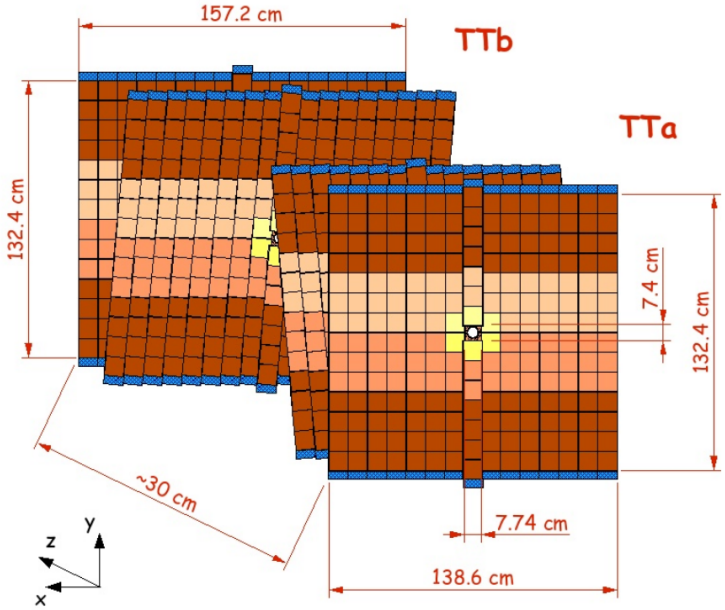


Figure 3.11: Trigger Tracker layout [90].

physics asymmetries, which needs to be taken into account in the estimation of systematic uncertainties, the direction of the magnetic field is periodically switched. The entire LHCb dataset contains roughly the same amount of data collected at both polarities.

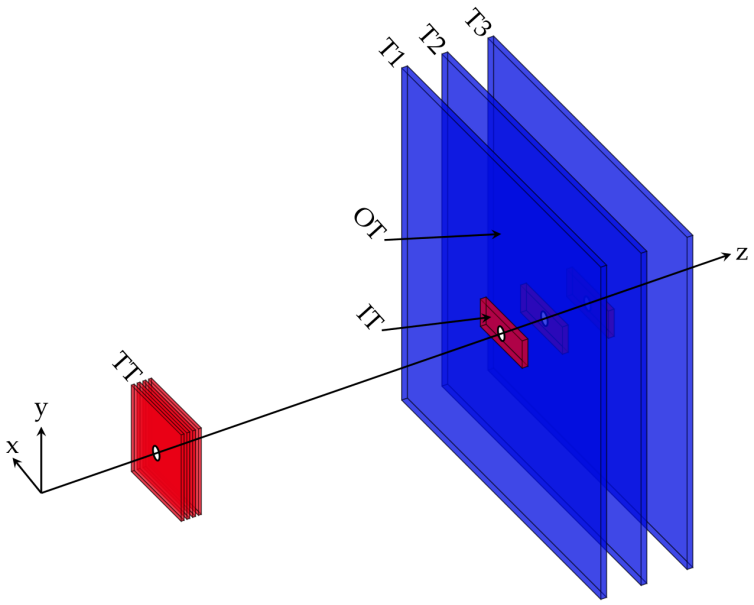


Figure 3.12: Tracking system layout [91].

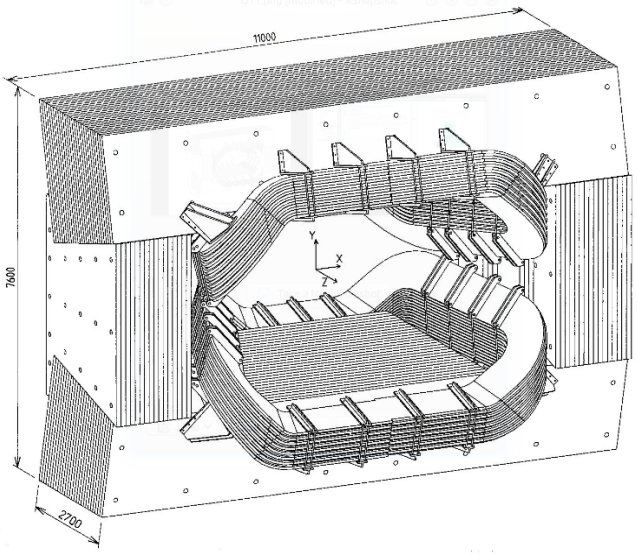


Figure 3.13: LHCb magnet, perspective view.

3.2.5 Calorimeters

The LHCb calorimeter system [93] is used for the identification of high transverse energy hadrons, electrons and photons. It measures their energy and selects candidates for the L0 trigger; at the offline analysis level, it provides also particle identification and position. Figure 3.14 shows the layout of the calorimeter system. It is composed of an Electromagnetic CALorimeter (ECAL) followed by a Hadronic CALorimeter (HCAL). To reject the high background

of charged pions, a PreShower detector (PS) is installed before the main ECAL, which ensures longitudinal segmentation of the electromagnetic shower detection. A Scintillator Pad Detector (SPD), which detects the passage of charged particles, is installed before the PS, in order to separate electrons from photons and neutral pions. A 15 mm lead converter, corresponding to 2.5 X_0 (radiation lenght) and 0.1 λ_I (interaction length), is placed between the SPD and the PS.

The ECAL is a sampling calorimeter. The absorber layers are made of 2 mm thick lead plates, while the detector layers are 4 mm thick polystyrene-based scintillator plates, and they are combined to obtain a total depth of 83.5 cm corresponding to 25 X_0 and 1.1 λ_I, with a Molière radius of 3.5 cm. The active area is segmented into three zones with different granularity, given the larger particle fluence in the central region: the dimensions of the inividual cells are 4x4 cm^2 for the inner part, and 6x6 cm^2 and 12x12 cm^2 moving outwards radially. The energy resolution of the ECAL is $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$, with E in units of GeV.

The HCAL consists of 16 mm thick iron plates interspaced with 4 mm thick scintillating tiles arranged parallel to the beam pipe. The length corresponds to 5.6 λ_I, which is enough as the trigger performances do not depend strongly on the hadronic energy resolution. It is transversely segmented into only two zones, where the cells have dimensions of 13x13 cm^2 for the inner zone and 26x26 cm^2 for the outer zone. The energy resolution for the HCAL is $\sigma_E/E = 69\%/\sqrt{E} \oplus 9\%$, with E in units of GeV.

3.2.6 Muon system

The muon system [94,95] consists of five tracking stations, the first (M1) between RICH2 and the calorimeters and the other four downstream of the calorimeters. The layout of the system is shown in Figure 3.15. Each muon station is divided into 276 chambers, which are composed of logical pads of various dimensions, depending on the distance from the beam axis and the station number: in the inner regions, they vary in size between 6.3x31.3 mm^2 in M2 and 31x39 mm^2 in M5; for all stations, the pads' dimensions in the subsquent regions are double the ones in the previous region. This ensures that the occupancy is roughly the same in each region; also, the use of large pads in the outer region is justified by multiple scattering of muons with low momentum, which dominates the spatial resolution in these regions. The muon stations are equipped with Multi Wire Proportional Chambers (MWPCs) with a mixture of Ar, CO_2 and CF_4, except

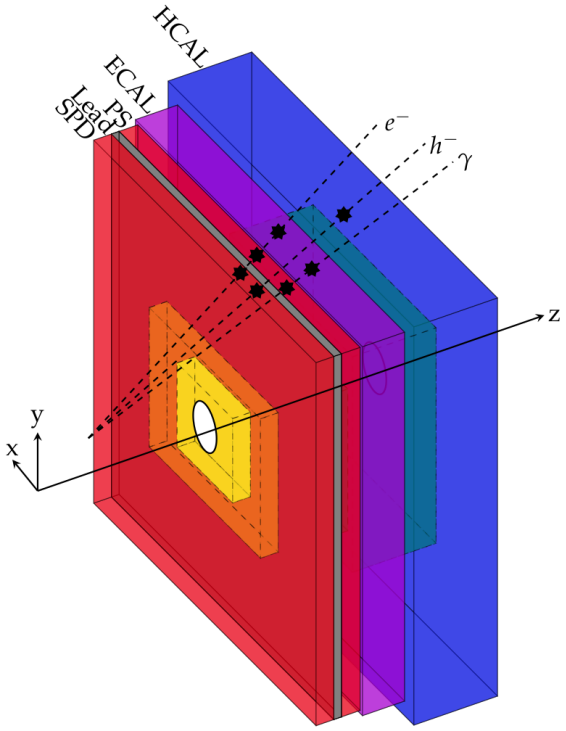


Figure 3.14: Layout of the calorimeter system, with different particle interactions. Relative dimensions along the z axis are not to scale [91].

for the inner region of M1, which is equipped with Gas Electron Multiplier (GEM) chambers; this is necessary to cope with the higher charged particle rate, as the MWPCs would not be sufficiently resistant to radiation damage for the full period of data taking. Stations M2 to M5 are interleaved with 80 cm thick iron absorbers, corresponding to a total of 15 λ_I. Each station is required to provide an efficiency high enough to achieve a 95% total trigger efficiency, which requires a quintuple hit coincidence. The angular acceptances of the muon system ranges between 20 (16) mrad and 306 (258) mrad in the bending (non-bending) plane, similar to that of the tracking system. This provides a geometrical acceptance of about 20% for muons from b decays relative to the full solid angle.

The muon stations are also used online for particle identification, using the position of hits to define a binary decision on whether a track is a muon or not. This information is included in the global PID variables described in Section 3.3.2.

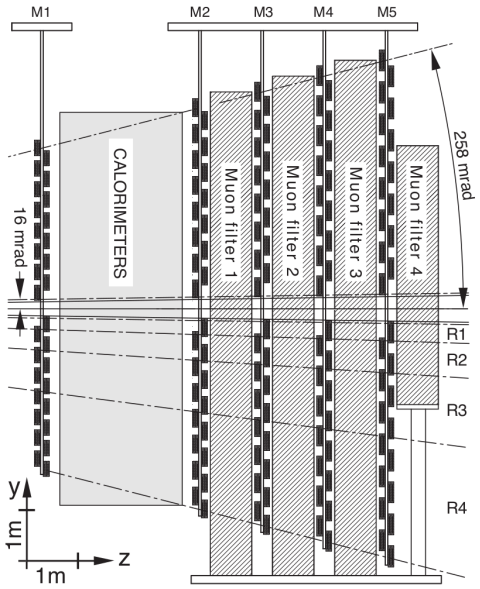
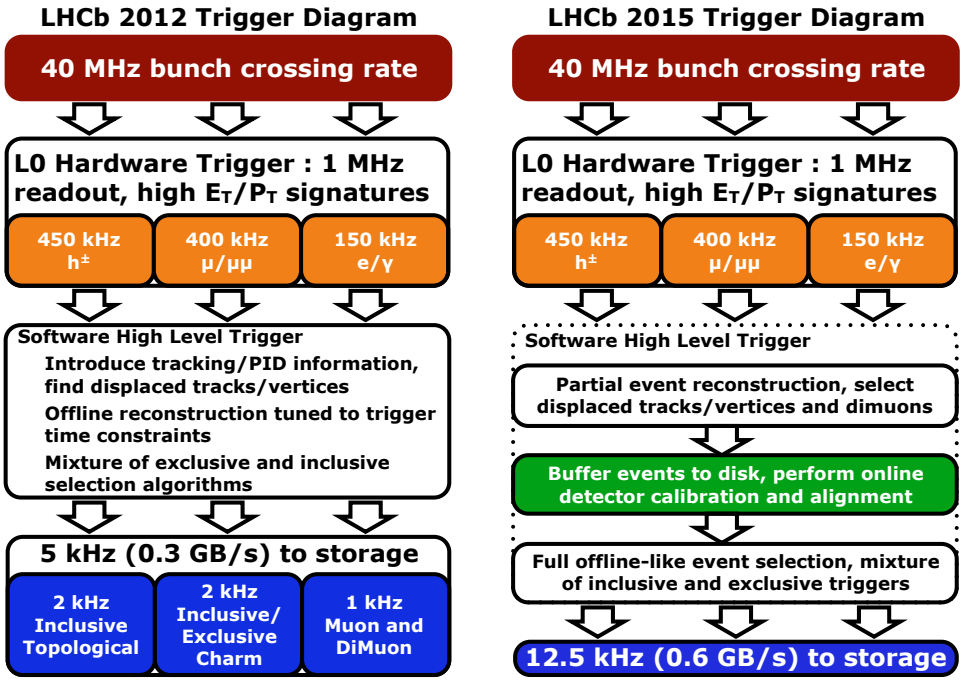


Figure 3.15: Muon system configuration [95].

3.2.7 Trigger, real-time alignment and calibration and the Turbo stream

Data collected at the LHC bunch crossing rate of 40 MHz needs to be reduced to 5 kHz in order to be stored to disk. To do so, a two-stage trigger strategy is required in LHCb [96]: the Level-0 trigger (L0) is a hardware trigger which reduces the rate to about 1 MHz; the High Level Trigger (HLT) is a software-based trigger that uses a combination of C++ algorithms, taking as input the L0 output and reducing its rate to 50 kHz in a first step (HLT1) and finally to 5 kHz (or 12.5 kHz in Run 2) in a second step (HLT2). Figure 3.16 (left) shows the LHCb trigger scheme for Run 1.

The L0 trigger uses custom-made electronics embedded in the calorimeters, in the muon system and in the upstream VELO R-sensors, which act as pile-up veto. It attempts to select the highest E_T hadron, electron and photon cluster, and the two highest-p_T muons: this is justified by the fact that final-state particles coming from a b- or c-hadron decay are characterised by high momenta, given the large mass of their mother particles. Events with too many tracks are rejected at this stage, as they would require an excessively large amount of computing resources to be processed: this is achieved by combining the information of the VELO pile-up system, that estimates the number of primary interaction points, with the number of hits in the SPD and the total energy in the calorimeters, which are proportional to the track



(a) Run 1 (b) Run 2

Figure 3.16: LHCb trigger scheme for Run 1 and Run 2 [97].

multiplicity. The final L0 trigger decision is given by a logical OR between all the 25 L0 channels [98]; as an example, here is a list of the most common ones, which cover most of the physics cases in LHCb, with an explanation of the requirements (typical values of the thresholds in the 2012 data taking period are given in parentheses):

- L0Muon: one high-p_T track reconstructed from a hit in each muon station ($p_T > 1.76$ GeV/c)
- L0DiMuon: a pair of tracks with a high product of p_T, each reconstructed from a hit in each muon station ($p_T^1 \times p_T^2 > 1.6\ (\text{GeV}/c)^2$)
- L0Hadron: a cluster in the HCAL with high E_T ($E_T > 3.5$ GeV)
- L0Photon: a cluster in the ECAL with high E_T, several corresponding PS hits and no corresponding SPD hits ($E_T > 2.72$ GeV)
- L0Electron: a cluster in the ECAL with high E_T, several corresponding PS hits and at least one corresponding SPD hit ($E_T > 2.72$ GeV)

Furthermore, for all quoted channels a cut on the SPD multiplicity of < 600 is applied, apart from the L0DiMuon, for which the cut is < 900.

The HLT1 takes L0 decisions as input and proceeds further by refining the candidates with harder cuts and introducing selections based on reconstructed tracks, using information from the VELO and the tracking stations. High-p_{T} and high-p tracks are reconstructed by matching hits in the tracking stations and, possibly, in the muon stations with hits in the VELO; at this stage, the position of the primary vertex is determined using VELO tracks and it is used to calculate the IP of the high-momentum tracks. The output rate of HLT1 is 30 kHz, which is sufficiently low to allow a full offline track reconstruction using information from the whole detector. This is achieved with HLT2, where the decisions are made by looking at track combinations into composite particles, event topology, quality of the track fit or the vertex fit, particle identification and, sometimes, the output of some multivariate algorithm, in addition to harder cuts on the variables already used in the previous stages of the trigger. The exact number and mix of variables used in each line heavily depends on the nature of the process being taken into consideration and whether the trigger line is inclusive or exclusive.

In 2012, the HLT CPU farm was not fully used in the periods between different LHC proton fills, and therefore it was decided to implement a deferred HLT by temporarily storing about 20% of events accepted by L0 [99], in order to have them processed during the dead times, effectively providing a 20% increase in available processing time. This led to an increase in trigger efficiency following a reduction of the p_{T} thresholds in the HLT. Given the success of this strategy, it was decided to expand the size of the local disks in the LHCb Event Filter Farm (EFF) and to double the CPU power, in order to introduce a full buffering layer in the Run 2 trigger scheme, as can be seen in Figure 3.16 (right). Furthermore, the buffered data are used to perform a fully automated alignment and calibration of the detector, which, in Run 1, was done offline by reprocessing data at the end of the data-taking year: this includes the alignment of the tracking stations and the VELO, the alignment of the RICH mirrors, the calibration of the global time of the straw drift tubes, the calibration of the photon detector refractive index (which changes with temperature, pressure and gas composition) and the HV settings of the calorimeter photomultipliers. The online alignment and calibration procedure, along with the increased computing time in the HLT level, provide a trigger output with reconstructed objects and physical quantities which are comparable in quality to events processed offline. This prompted the creation of the Turbo stream [99], a framework in which physics

analyses are performed directly on the trigger reconstruction output: for each event sent to the Turbo stream, only the candidate reconstructed and selected by the trigger is saved, and the rest of the event is discarded. While this cannot be applied to all LHCb analyses, as some will always need information from the full event, it provides a reduction of an order of magnitude in the size of the event saved on disk. In Run 2, about 20% of the trigger selections were sent to the Turbo stream, occupying less than 2% of the total available bandwidth; it is foreseen that this will be extended to more analyses starting from LHC Run 3 in 2021.

The search for strong *CP* violation in η and η' decays to two pions, described in Chapter 4, was the first to be published of the four analyses chosen to validate the Turbo procedure, using data from the Turbo stream collected during 2015.

3.3 Analysis tools and LHCb software

3.3.1 Reconstruction and stripping

Data selected by the trigger are stored on disk and analysed offline, combining hits in the tracking system, calorimeter clusters and information from the RICH detectors into objects that are then used in the analysis: tracks, particle identification (PID) hypotheses, primary and secondary vertices, decay vertices, IPs, decay times and so on. This procedure creates Data Summary Tape (DST) files, which include all the observables that are subsequently used by other algorithms to further select events. These files are then processed again, as they are still too large to be used by LHCb users, by running a collection of algorithms known as "stripping selection": the purpose of such algorithms is to further select events by applying different sets of loose cuts. This provides a generic categorisation of interesting candidates into different streams, according to their topology, kinematics and PID, to fit macrogroups of different analysis needs. The output of this procedure is a set of new DSTs of a more manageable size, which contain only events selected by stripping lines sharing some general properties. For example, the `DiMuon` stream contains events selected by a set of stripping lines which attempt to reconstruct events with two muons in the final state.

3.3.2 Particle identification

Particle identification (PID) at LHCb is performed by a combination of information from the RICH detectors, the calorimeters and the muon stations.

- RICH detectors: the Cherenkov effect connects the angle of emission of light coming from a superluminal charged particle in a dielectric medium with its speed and the refraction index of the medium, which is known for the RICH system. Combining the measurement of the emission angle (and hence the speed) with the track momentum, given by the tracking system, a measurement of the mass of the particle can in principle be extracted. Effectively, this is not the method used to assign a mass hypothesis, as it is extremely ineffective in regions of high track density, but rather a likelihood is calculated ($\mathcal{L}^{\mathrm{RICH}}$), which assumes that the number of observed photoelectrons in a single pixel is Poisson distributed and the Cherenkov angle is Gaussian distributed [100]. Since pions are the most abundant long-lived particles in p-p interactions, all tracks are assumed to be pions at first; then, for a single track, the Cherenkov angle of the hypothesis is changed to another particle type (e, μ, K, p), leaving the other tracks unchanged and calculating the likelihood at each change of hypohesis; this is repeated for all tracks in the event, and the configuration that gives the largest increase in likelihood is selected. Finally, the algorithm is iterated until no further improvement is found.

- Calorimeters: a description of the typical energy and SPD hit patterns from hadrons, photons and electrons has already been given in Section 3.2.7. Also in this case, a likelihood is calculated ($\mathcal{L}^{\mathrm{CALO}}$). The presence of energy deposits in either the ECAL or the HCAL, hits in the SPD and the amount of energy in the PS are the most discriminating variables for the different particle species.

- Muon stations: the muons at the momentum range of LHC processes lie close to the minimum of the Bethe-Bloch formula [101]. The total interaction length provided by the calorimeters and the iron filters between the muon stations is about 20 λ_I (the total radiation length for the same interval is roughly one order of magnitude larger). Therefore, by requiring a quintuple hit coincidence, the muon system alone ensures a misidentification rate of less than 2% for most of the kinematic range [102]. In order for

this information to be combined with the other two likelihoods, a third likelihood is calculated ($\mathcal{L}^{\text{MUON}}$).

At this point, a global PID likelihood is defined as the product of the three contributions:

$$\mathcal{L}^{\text{PID}}(K) = \mathcal{L}^{\text{RICH}}(K) \cdot \mathcal{L}^{\text{CALO}}(!e) \cdot \mathcal{L}^{\text{MUON}}(!\mu), \tag{3.1}$$

$$\mathcal{L}^{\text{PID}}(\pi) = \mathcal{L}^{\text{RICH}}(\pi) \cdot \mathcal{L}^{\text{CALO}}(!e) \cdot \mathcal{L}^{\text{MUON}}(!\mu), \tag{3.2}$$

$$\mathcal{L}^{\text{PID}}(e) = \mathcal{L}^{\text{RICH}}(e) \cdot \mathcal{L}^{\text{CALO}}(e) \cdot \mathcal{L}^{\text{MUON}}(!\mu), \tag{3.3}$$

$$\mathcal{L}^{\text{PID}}(\mu) = \mathcal{L}^{\text{RICH}}(\mu) \cdot \mathcal{L}^{\text{CALO}}(!e) \cdot \mathcal{L}^{\text{MUON}}(\mu), \tag{3.4}$$

where $\mathcal{L}(x)$ is the likelihood for a given track of being x, and $\mathcal{L}(!x)$ is the likelihood for a given track of not being x. As mentioned before, to each charged track a pion hypothesis is immediately assigned; the PID variables used in the physical analyses are then computed with respect to this hypothesis:

$$\text{PID}(x) = \Delta\text{LL}(x - \pi) = \ln\left(\frac{\mathcal{L}^{\text{PID}}(x)}{\mathcal{L}^{\text{PID}}(\pi)}\right). \tag{3.5}$$

Apart from the PID likelihoods, another class of PID variables is used, called ProbNN variables. These variables are evaluated from the output of a neural network, combining all information from the subdetectors but taking into account also correlations between different signatures; they perform better than their counterpart above but are generally unsuitable for online processing. The multivariate classifier is trained on simulated events. It considers all tracks in the event, including fake tracks (ghosts) reconstructed from matching random hits in the tracking system.

3.3.3 LHCb software

The LHCb software is mainly written in C++ and configurable through Python, and it is divided into several projects, each with a specific task. They are all executed within the GAUDI [103] framework. Some of the most important projects are:

- GAUSS [104]: the LHCb simulation software. This is used to produce simulated events using Monte Carlo (MC) methods, as well as to reproduce the detector response. The p-p collision and the hadronisation phase are generated with PYTHIA [105], and subsquently the EVTGEN package [106] is

used to generate the decays of the particles of interest; final-state radiation can also be simulated via the PHOTOS package [107]. The simulation of the interaction of the generated particles with the detector is implemented using the GEANT4 toolkit [108, 109].

- BOOLE [110]: it is the last step of the LHCb detector simulation. This applies the detector response to hits previously generated with GAUSS, and simulates the readout electronics and the L0 trigger hardware.

- BRUNEL [111]: the LHCb reconstruction software. This is used to create tracks, the fundamental object used in physics analysis, from hits in the detector. It also provides particle identification on the reconstructed tracks.

- MOORE [112]: the LHCb HLT application software. This runs in the online trigger farm or offline, starting from real data or from the BOOLE digitisation output.

- DAVINCI [113]: the LHCb physics analysis software. This is used to combine the final state particles, with a set of selection cuts, from the output of the trigger selection or the stripping categorisation, in order to create the decay chain of interest for the analysis, generating ntuples from DST files, readable by ROOT [114], which store all the interesting high-level observables. The list of variables, or functions of one or more variables, stored in the ntuples is completely customisable using standard or user-defined tools. DAVINCI is also used to produce the DST files during the stripping campaigns.

4 Search for strong *CP* violation

As already shown in Section 1.5.1, and in particular in Equation (1.48), the strength of *CP* violation in weak interactions in the quark sector is well below what would be required to provide an explanation for the imbalance that arose in the early universe between the amounts of matter and antimatter. It is possible that the solution lies in the neutrino sector, and experiments are now being planned to search for *CP* violation in neutrino interactions and to attempt to measure the value of any associated complex phase in the PMNS matrix.

According to the description of the structure of the Yang-Mills vacuum given in Section 1.6, the QCD lagrangian should contain a natural term, the so-called θ term, that would give rise to *CP* violation in the strong interactions, but no evidence for this has yet been seen. The measured upper limit on the neutron electric dipole moment (nEDM) implies a limit $\bar{\theta} < 2 \times 10^{-10}$ [35,38]. There is no reason, a priori, why the value of $\bar{\theta}$ should be so close to zero. This is seen as a fine-tuning problem in QCD, the so-called "strong *CP* problem".

The strong decay modes $\eta \to \pi^+\pi^-$ and $\eta'(958) \to \pi^+\pi^-$ would both violate *CP* symmetry, and searches for these decays represent the most fruitful way to look for strong *CP* violation in particle physics experiments. In the Standard Model these decays could take place through mediation by a virtual K^0_S, with expected branching fractions (BF) $\mathcal{B}(\eta \to \pi^+\pi^-) < 2 \times 10^{-27}$ and $\mathcal{B}(\eta' \to \pi^+\pi^-) < 4 \times 10^{-29}$ [39]. Strong decays mediated by the QCD θ term would have $\mathcal{B} \lesssim 3 \times 10^{-17}$, based on the limit from the nEDM [39]. Any observation of BFs larger than this value would indicate a new source of *CP* violation in the strong interactions, and could help solve the problem of the origin of the matter-antimatter asymmetry in the early universe.

4.1 Outline of the analysis method

In this analysis, decays of the type $D^+ \to \pi^+\pi^+\pi^-$ and $D_s^+ \to \pi^+\pi^+\pi^-$, from the large samples of charm mesons recorded by the LHCb experiment in the full 3 fb^{-1} dataset of LHC Run 1 (2011-2012) and in the 0.3 fb^{-1} dataset of Run 2 (2015, referred in the following simply as Run 2), are used to look for evidence of the presence of the η and η' resonances in the $\pi^+\pi^-$ mass spectra. They could in principle come from the known decays $D^+_{(s)} \to \pi^+\eta^{(\prime)}$, $\eta^{(\prime)} \to \pi^+\pi^-$ (inclusion of charge-conjugate modes is implied throughout). For $N(\eta^{(\prime)})$ observed $\eta^{(\prime)}$ signal decays in the $\pi^+\pi^-$ mass spectrum from a total of $N(D^+_{(s)})$ mesons reconstructed in the $\pi^+\pi^+\pi^-$ final state, the measured branching fraction for $\eta \to \pi^+\pi^-$ and $\eta' \to \pi^+\pi^-$ decays would be

$$\mathcal{B}(\eta^{(\prime)} \to \pi^+\pi^-) = \frac{N(\eta^{(\prime)})}{N(D^+_{(s)})} \times \frac{\mathcal{B}(D^+_{(s)} \to \pi^+\pi^+\pi^-)}{\mathcal{B}(D^+_{(s)} \to \pi^+\eta^{(\prime)})} \times \frac{1}{\epsilon(\eta^{(\prime)})}, \tag{4.1}$$

where $\epsilon(\eta^{(\prime)})$ accounts for any variaton in efficiency across the $\pi^+\pi^-$ mass spectrum from $D^+_{(s)} \to \pi^+\pi^+\pi^-$ decays, as discussed in Section 4.7. The branching fractions $\mathcal{B}(D^+_{(s)} \to \pi^+\pi^+\pi^-)$ and $\mathcal{B}(D^+_{(s)} \to \pi^+\eta^{(\prime)})$ are taken from the Particle Data Group [47]. Since the analysis starts from a given number of selected $D^+_{(s)} \to \pi^+\pi^+\pi^-$ candidates, there are no normalisation channels and only one efficiency correction as a function of the $\pi^+\pi^-$ invariant mass. All selections are finalised and expected sensitivities are evaluated before the η and η' signal regions in the $\pi^+\pi^-$ mass spectra are examined. In practice, no signals are expected to be seen and upper limits on the branching fractions are derived using the $\mathrm{CL_s}$ method [115].

The $D^+_{(s)} \to \pi^+\pi^+\pi^-$ events are selected from the output of dedicated stripping lines, described in detail in Section 4.2.3. For Run 1, a single line selects charged combinations of $\pi^+\pi^+\pi^-$, compatible with originating from the same vertex, in an invariant mass interval large enough to include both the D^+ and the D_s^+ peaks. For Run 2, there are two separate Turbo lines which are optimised to select candidates coming from D^+ or D_s^+. Further selections to reduce the backgrounds under the D^+ and D_s^+ peaks and the choice of selection mass windows around the peaks are jointly optimised to obtain the smallest expected uncertainties on $\mathcal{B}(\eta^{(\prime)} \to \pi^+\pi^-)$ in Equation (4.1). A given $\pi^+\pi^+\pi^-$ mass window will contain $N_{sig} + N_{bkg}$ events, where $N_{sig} = N(D^+_{(s)} \to \pi^+\pi^+\pi^-)$ is the number of signal $D^+_{(s)}$ events and N_{bkg} is the

number of background events. The values of $N(\eta^{(\prime)})$ are obtained by fitting to the inclusive $\pi^+\pi^-$ mass spectra from the selected events (which have two entries per event). In the absence of an $\eta^{(\prime)} \to \pi^+\pi^-$ signal, the statistical error on $N(\eta^{(\prime)})$ will therefore be proportional to $\sqrt{N_{sig} + N_{bkg}}$. The optimisation therefore is designed to maximise the value of $N_{sig}/\sqrt{N_{sig} + N_{bkg}}$.

4.2 Datasets and event selection

4.2.1 Data and Monte Carlo simulation samples

The results described in this chapter are obtained using the data collected at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ in 2011, corresponding to 1.0 fb^{-1} of integrated luminosity, at $\sqrt{s} = 8\,\text{TeV}$ in 2012, corresponding to 2.0 fb^{-1} of integrated luminosity, and at $\sqrt{s} = 13\,\text{TeV}$ in 2015, corresponding to 0.3 fb^{-1} of integrated luminosity.

The Monte Carlo (MC) samples used in this analysis are listed in Table 4.1. The full simulation samples are created using the latest description of the detector conditions throughout 2011 and 2012 data taking. Although this analysis involves the use of both 2011 and 2012 data, for $D^+_{(s)} \to \pi^+\pi^+\pi^-$ samples only 2012 conditions are used. The analysis is relatively insensitive to the differences in beam energy, trigger, and other effects. However, for $\eta^{(\prime)} \to \pi^+\pi^-$ signal MC, luminosity-weighted averages of 2011 and 2012 samples are used. All samples are generated using PYTHIA 8 with a 'phase-space' model, based on the phase-space decay of the relevant mother particles. To save CPU time, cuts are applied to all the "Full" samples at the generation level of the simulation, as described in Section 4.2.2. The "Gen. level" samples do not include the simulation of the detector response.

No 2015 simulated samples for 2015 are used, as they were not available at the time this analysis was done.

4.2.2 Generator-level cuts and Monte Carlo filtering

In order to reduce the number of simulated events that are produced but not accepted into the final selected samples, generator-level cuts are introduced based on cuts that would subsequently be applied in the reconstruction. Conservative cut values are chosen to safely allow for momentum smearing after

Table 4.1: MC simulation samples used in the analysis. The MC production type and the number of events generated are also given. All "Full" samples are produced with an approximately equal amount of both magnet polarities.

Channel	Production type	# events
$D^+ \to \pi^+\pi^+\pi^-$	Full 2012	1,220,378
$D_s^+ \to \pi^+\pi^+\pi^-$	Full 2012	1,206,823
$D^+ \to \pi^+\pi^+\pi^-$	Gen. level 2012	980,000
$D_s^+ \to \pi^+\pi^+\pi^-$	Gen. level 2012	980,000
$D^+ \to \eta(\pi^+\pi^-)\pi^+$	Full 2012	68,646
$D^+ \to \eta'(\pi^+\pi^-)\pi^+$	Full 2012	65,182
$D^+ \to \eta(\pi^+\pi^-)\pi^+$	Full 2011	63,063
$D^+ \to \eta'(\pi^+\pi^-)\pi^+$	Full 2011	59,420
$D_s^+ \to K^+\pi^+\pi^-$	Toy	10^7
$D_s^+ \to \pi^+\pi^+\pi^-\pi^0$	Toy	10^7
$D^+ \to \eta(\pi^+\pi^-\gamma)\pi^+$	Full 2012	42,957
$D^+ \to \eta'(\pi^+\pi^-\gamma)\pi^+$	Full 2012	30,574
$D^+ \to \eta(\pi^+\pi^-\gamma)\pi^+$	Toy	10^9
$D^+ \to \eta'(\pi^+\pi^-\gamma)\pi^+$	Toy	10^9
$D_s^+ \to \eta(\pi^+\pi^-\gamma)\pi^+$	Toy	10^9
$D_s^+ \to \eta'(\pi^+\pi^-\gamma)\pi^+$	Toy	10^9

the detector simulation is performed. These cuts, which are summarised in the second column of Table 4.2, are applied to all final-state particles in the signal samples. For the $D_{(s)}^+ \to \pi^+\pi^+\pi^-$ samples the cuts are tightened, as listed in the third column of Table 4.2. All final-state particles are also required to be within the LHCb detector acceptance in all cases. No generator-level cuts are imposed on the $D_{(s)}^+ \to \pi^+\pi^+\pi^-$ generator-level-only samples.

In order to save disk space, all the samples listed in Table 4.1, with the exception of the generator-level and toy samples, are filtered so that only events that pass a specific stripping requirement are written to disk. Toy events are simulated with a simple n-body decay model, without any additional intermediate resonance, and without simulating the detector response. The filtering on the $D^+ \to \eta^{(\prime)}\pi^+$ signal samples corresponds to the stripping

selection described in Section 4.2.3 and Table 4.3, with the exception of the mass cuts on π pairs. For the $D^+ \to \pi^+\pi^+\pi^-$, $D_s^+ \to \pi^+\pi^+\pi^-$, $D^+ \to \eta(\pi^+\pi^-\gamma)\pi^+$ and $D^+ \to \eta'(\pi^+\pi^-\gamma)\pi^+$ samples the filtering also includes the requirement that trigger lines selecting high-momentum tracks, detached from the primary vertex and coming from a charm hadron, have fired.

Table 4.2: MC generator-level cuts. All final-state particles are also required to be in the LHCb acceptance.

	$D^+_{(s)} \to (\eta^{(\prime)} \to \pi^+\pi^-)\pi^+$	$D^+_{(s)} \to \pi^+\pi^+\pi^-$
$\pi^\pm\, p_{\mathrm{T}}$	$> 225\,\mathrm{MeV}/c$	$> 250\,\mathrm{MeV}/c$
$\pi^\pm\, p$	$> 1800\,\mathrm{MeV}/c$	$> 2000\,\mathrm{MeV}/c$
$D^\pm_{(s)}\, p_{\mathrm{T}}$	$> 900\,\mathrm{MeV}/c$	$> 2100\,\mathrm{MeV}/c$
$D^\pm_{(s)}\, p$	–	$> 14\,000\,\mathrm{MeV}/c$

4.2.3 Stripping and offline selections for Run 1 data

The combined set of stripping and offline cuts for $D^+ \to \pi^+\pi^+\pi^-$ and $D_s^+ \to \pi^+\pi^+\pi^-$ candidates in 2011 and 2012 data are described in Table 4.3. All the pion candidates come from a standard selection, which contains all the tracks which pass some fiducial loose criteria. The LHCb-specific variables listed in the table have the following definitions:

- $\chi^2_{\mathrm{IP}}/\mathrm{ndf}$ is the reduced χ^2 of the impact parameter of the pion candidate track with respect to the best primary vertex
- PIDK is the likelihood of the pion candidate being a K with respect to the likelihood of it being a π
- χ^2_{DV} is the χ^2 of the fit to the decay vertex of the $D^+_{(s)}$ candidate
- $\cos\alpha$ is the cosine of the angle between the momentum vector of the $D^+_{(s)}$ candidate and a line connecting the primary vertex to the decay vertex. α is also called DIRection Angle (DIRA).
- χ^2_{FD} is the χ^2 of the flight distance between the decay vertex and the origin vertex of the candidate, weighted by the sum of the covariance matrices

of the two points[1]

All dipion mass combinations are required to be larger than $300\,\mathrm{MeV}/c^2$ to remove clone tracks and misidentified $\gamma \to e^+e^-$ conversions associated with a random pion. An additional cut, $m(\pi^+\pi^-) < 1650\,\mathrm{MeV}/c^2$ on opposite sign dipion masses, is applied to remove events of the type $D^0 \to K^-\pi^+$ where the kaon is misidentified as a pion and an additional random pion is associated with the decay vertex.

Figure 4.1 shows the $\pi^+\pi^+\pi^-$ mass spectrum for the full $3.0\,\mathrm{fb}^{-1}$ Run 1 dataset directly after the stripping has been applied.

Table 4.3: Stripping selections for the Run 1 data. Additional offline cuts are indicated by †.

$\pi^\pm$	
p_T	$> 250\,\mathrm{MeV}/c$
p	$> 2000\,\mathrm{MeV}/c$
$\chi^2_\mathrm{IP}/\mathrm{ndf}$	> 4
PIDK	< 3
π pairs	
$m(\pi^+\pi^\pm)$	$> 300\,\mathrm{MeV}/c^{2\,\dagger}$
$m(\pi^+\pi^-)$	$< 1650\,\mathrm{MeV}/c^{2\,\dagger}$
$D^\pm_{(s)}$	
p_T	$> 1000\,\mathrm{MeV}/c$
$\sum$ daughter p_T	$> 2800\,\mathrm{MeV}/c$
# daughters $\chi^2_\mathrm{IP}/\mathrm{ndf} > 10$	≥ 2
χ^2_DV	< 10
$\cos\alpha$	> 0.98
χ^2_IP	< 12
χ^2_FD	> 125
m	$\in [1820, 2020]\,\mathrm{MeV}/c^{2\,\dagger}$

[1] To be specific: $\chi^2_\mathrm{FD} = [(\vec{p}_1 - \vec{p}_2)]^T[(\mathrm{cov}(\vec{p}_1) + \mathrm{cov}(\vec{p}_2))]^{-1}[(\vec{p}_1 - \vec{p}_2)]$, where $\vec{p}_1$ and $\vec{p}_2$ are the vectors of coordinates of the two vertices and $\mathrm{cov}(\vec{p}_1)$, $\mathrm{cov}(\vec{p}_2)$ are the two covariance matrices.

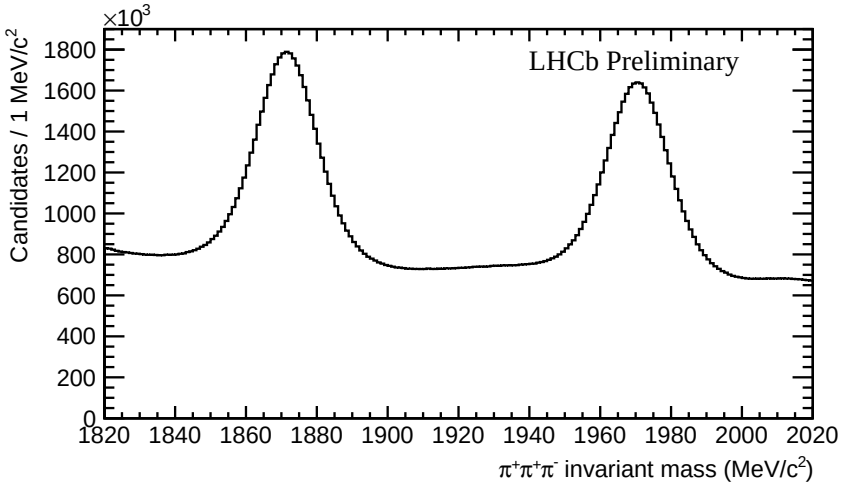


Figure 4.1: Invariant mass spectrum of $\pi^+\pi^+\pi^-$ candidates from the output of the Run 1 stripping.

4.2.4 Turbo stream selection for 2015 data

The combination of Turbo stream and offline cuts for $D^+ \to \pi^+\pi^+\pi^-$ and $D_s^+ \to \pi^+\pi^+\pi^-$ candidates in 2015 data is described in Table 4.4. The Turbo stream cuts present slightly different selection criteria for D^+ and D_s^+ decays, as can be seen from the table. The same offline cuts are applied to the Turbo stream as for the stripping. Also in this case, the pion candidates come from a standard selection.

Figure 4.2 shows the $\pi^+\pi^+\pi^-$ mass spectrum for the full 0.3 fb^{-1} dataset directly after the Turbo stream selection has been applied.

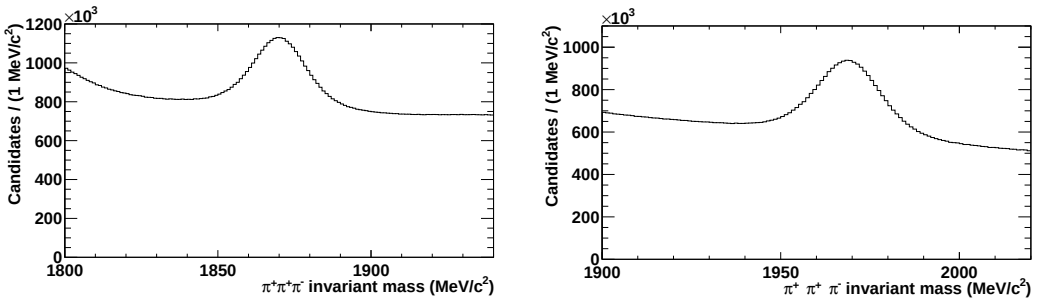


Figure 4.2: Invariant mass spectrum of $\pi^+\pi^+\pi^-$ candidates from the Run 2 Turbo stream: D^+ region on the left, D_s^+ region on the right.

Table 4.4: Turbo stream selection (Run 2 data). Where two cut values are given the first refers to the D^+ selection and the second refers to the D_s^+ selection. Additional offline cuts are indicated by †.

$\pi^\pm$	
p_T	$> 250\,\mathrm{MeV}/c$
$\chi^2_\mathrm{IP}/\mathrm{ndf}$	> 4
π pairs	
$m(\pi^+\pi^\pm)$	$> 300\,\mathrm{MeV}/c^{2\dagger}$
$m(\pi^+\pi^-)$	$< 1650\,\mathrm{MeV}/c^{2\dagger}$
$D^\pm_{(s)}$	
$\sum$ daughter p_T	$> 3200\,\mathrm{MeV}/c$
# daughters $\chi^2_\mathrm{IP}/\mathrm{ndf} > 10$	≥ 2
# daughters $\chi^2_\mathrm{IP}/\mathrm{ndf} > 50$	≥ 1
# daughters $p_\mathrm{T} > 400\,\mathrm{MeV}/c$	≥ 2
# daughters $p_\mathrm{T} > 1000\,\mathrm{MeV}/c$	≥ 1
χ^2_DV	< 6
$\cos\alpha$	> 0.99995
χ^2_FD	$> 150/100$
Lifetime	$> 0.4\,\mathrm{ps}$ / $> 0.2\,\mathrm{ps}$
m	$\in [1800, 1949]\,\mathrm{MeV}/c^{2\dagger}$ / $\in [1889, 2020]\,\mathrm{MeV}/c^{2\dagger}$

4.3 Comparison of data and simulation

In Sections 4.3.1 and 4.4 the use of multivariate classification to discriminate between signal and background is discussed. Before this classification is performed, potential differences between simulation and data are examined, as the training phase of the multivariate analysis utilises samples of both types (for signal and background respectively). Whilst performing the training without taking into account these differences will not yield an incorrect result, it may lead to a suboptimal performance of the classifier; therefore, in Sections 4.3.2 and 4.3.3 a method to reduce discrepancies is described.

4.3.1 Multivariate classifier input variables

In order to reduce the backgrounds under the $D^+_{(s)}$ peaks, the events are processed through a Boosted Decision Tree (BDT), using the Toolkit for Multivariate Data Analysis (TMVA) [116], provided by the ROOT framework. The choice of the classifier comes from an optimisation study which is described in Appendix A. The variables chosen as input for the classifier are:

- $\log|\text{ProbNNghost}(\min_{p_T}\pi)|$
- $\log|\text{ProbNNghost}(\text{mid}_{p_T}\pi)|$
- $\log|\text{ProbNNghost}(\max_{p_T}\pi)|$
- $\log|\text{ProbNNk}(\min_{p_T}\pi)|$
- $\log|\text{ProbNNk}(\text{mid}_{p_T}\pi)|$
- $\log|\text{ProbNNk}(\max_{p_T}\pi)|$
- $\log|1\text{-DIRA}(D^+)|$
- $\log|\chi^2_{\text{IP}}(D^+)|$
- $\log|\chi^2_{\text{DV}}(D^+)|$
- $\log|1.02-\text{ProbNNpi}(\min_{p_T}\pi)|$
- $\log|1.02-\text{ProbNNpi}(\text{mid}_{p_T}\pi)|$
- $\log|1.02-\text{ProbNNpi}(\max_{p_T}\pi)|$

- $\log|\chi^2_{\text{IP}}(\min_{p_T}\pi)|$
- $\log|\chi^2_{\text{IP}}(\text{mid}_{p_T}\pi)|$
- $\log|\chi^2_{\text{IP}}(\max_{p_T}\pi)|$
- $\log|\min_{p_T}\text{DOCA}+0.005|$
- $\log|\text{mid}_{p_T}\text{DOCA}+0.005|$
- $\log|\max_{p_T}\text{DOCA}+0.005|$
- $\chi^2_{track}(\min_{p_T}\pi)/\text{ndf}$
- $\chi^2_{track}(\text{mid}_{p_T}\pi)/\text{ndf}$
- $\chi^2_{track}(\max_{p_T}\pi)/\text{ndf}$.

For the variables where there is a value for each of the final-state tracks the values are ordered by their transverse momentum. Two new variables are introduced in this list: the reduced χ^2_{track} of the track fit and the Distance Of Closest Approach (DOCA) between two tracks, where $\min_{p_T}\text{DOCA} = \text{DOCA}(t_1, t_2)$, $\text{mid}_{p_T}\text{DOCA} = \text{DOCA}(t_2, t_3)$ and $\max_{p_T}\text{DOCA} = \text{DOCA}(t_1, t_3)$ and t_1, t_2, t_3 are the final-state tracks ordered by their p_T so that $p_T(t_i) < p_T(t_{i+1})$. A description of the ProbNN variables is given in Section 3.3.2. Logarithms are taken to have smoother distributions; even though this does not have any impact on the classifier performances, it is useful to have a better visualisation of the signal and background distributions, since several of them would otherwise present rather narrow peaks.

Figures 4.3 and 4.4 show the distributions of the input variables to the multivariate classifier described in Section 4.4. The MC sample used for this study consists of the $D^+ \to \pi^+\pi^+\pi^-$ signal sample listed in Table 4.1. The data comprises 5% of the combined 2011 and 2012 data sample, background-subtracted around the D^+ peak using the mass ranges $m(\pi^+\pi^+\pi^-) \in [1856.4, 1886.4]$ MeV for the peak region and $m(\pi^+\pi^+\pi^-) \in [1910, 1940]$ MeV for the background.

Differences between data and simulation are found, particularly for the ProbNN variables shown in Figure 4.4; therefore, the ProbNN variables are corrected by the prescription described in the next section, whilst the other variables are kept with their original values.

4.3.2 Scaling of classifier input variables in the simulation

In order to reduce the differences between simulation and data two methods were considered. In the first method, the MC ProbNN values were randomly resampled from a calibration dataset, using $D^{*+} \to D^0\pi^+$ decays, in eight bins of pseudorapidity and transverse momentum, as listed in Table 4.5. In fact, the discrimination power in particle identification strongly depends on η and p_T of the track, and, in general, also on the total number of tracks in the event, which was not taken into account in this study.

Table 4.5: Binning scheme for ProbNN corrections.

η	p_T (MeV)
< 3.2	< 5000
< 3.2	> 5000
3.2 – 3.7	< 2000
3.2 – 3.7	2000 – 6000
3.2 – 3.7	> 6000
> 3.7	< 1400
> 3.7	1400 – 5000
> 3.7	> 5000

Whilst this procedure generally gives good agreement between corrected simulation and data, it removes correlations between the values of the variables for the three final-state tracks, which significantly degrades the performance of the classifier.

To avoid this problem, the values of the ProbNN variables are instead scaled with a function $f : [0,1] \to [0,1]$, which at the lowest order is taken to be

$$\text{ProbNN}^{\eta,p_T}_{\text{MC,scaled}} = a(\text{ProbNN}^{\eta,p_T}_{\text{MC}})^2 + (1-a)\text{ProbNN}^{\eta,p_T}_{\text{MC}}, \tag{4.2}$$

where the value of the parameter a is calculated such that the average scaled ProbNN values in each (η, p_T) bin match those of the corresponding distributions in data (this method is referred to in the following as 'MC scaling'); therefore, calling $A_{\text{MC}} = \left\langle \text{ProbNN}^{\eta,p_T}_{\text{MC}} \right\rangle$, $A^{\text{sq}}_{\text{MC}} = \left\langle (\text{ProbNN}^{\eta,p_T}_{\text{MC}})^2 \right\rangle$ and $A_{\text{data}} = \left\langle \text{ProbNN}^{\eta,p_T}_{\text{data}} \right\rangle$, where the η and p_T indices are implicit,

$$\begin{aligned} &\left\langle \text{ProbNN}^{\eta,p_T}_{\text{MC,scaled}} \right\rangle = aA^{\text{sq}}_{\text{MC}} + (1-a)A_{\text{MC}} = A_{\text{data}} \\ &\implies a = \frac{A_{\text{data}} - A_{\text{MC}}}{A^{\text{sq}}_{\text{MC}} - A_{\text{MC}}}. \end{aligned} \tag{4.3}$$

Figure 4.4 shows the result of the scaling for the variables of interest. With this procedure, better agreement between simulation and data is achieved, thus improving the discrimination of the classifier.

4.3.3 Scaling of classifier input variables for Run 2

Since no simulated sample was available for Run 2 at the time this analysis was done, the multivariate training is performed with the same MC dataset used for the Run 1 training. In order to improve the agreement between Run 1 simulation and Run 2 data, the MC scaling method described in Section 4.3.2 is used on all the input variables. Figures 4.5 and 4.6 show the results of the application of this method for the input variables for Run 2.

4.4 Signal and background classification

As already mentioned in Section 4.3.1, a BDT is used as multivariate classifier. The training signal sample for Run 1 consists of 8.17×10^5 $D^+ \to \pi^+\pi^+\pi^-$ phase-space, truth-matched Monte Carlo events, while the training background sample consists of 10.21 $\times10^6$ events from data in the $\pi^+\pi^+\pi^-$ invariant mass sidebands, in 20 MeV/c^2 mass windows on either side of the D^+ peak. Both samples are rescaled to the same number of events before the beginning of the training. Figures 4.7 and 4.8 show the signal and background distributions of all the input variables, while in Figures 4.9 and 4.10 the Receiver Operating Characteristic (ROC) curve [117] and the classifier response are shown.

For Run 2 the same signal sample is used, after having applied the MC scaling procedure described in Section 4.3.2, while the background sample consists of 3.56×10^6 events from the $\pi^+\pi^+\pi^-$ invariant mass sidebands in 20 MeV/c^2 mass windows on either side of the D^+ peak. The BDT response for Run 2 shows a better separation with respect to the Run 1 BDT response, which is largely due to a cut in the distribution of $\chi^2_{IP}(D^+)$ that is present in the Run 1 stripping lines but absent in the Run 2 turbo selection. Figure 4.15 shows how the background level is dramatically reduced by cutting on this variable, for both the D^+ and D_s^+ peak regions in Run 2. The optimised BDT cut, obtained with an optimisation procedure described in Section 4.5, further reduces the background level, and the effect on data is also shown in the plot.

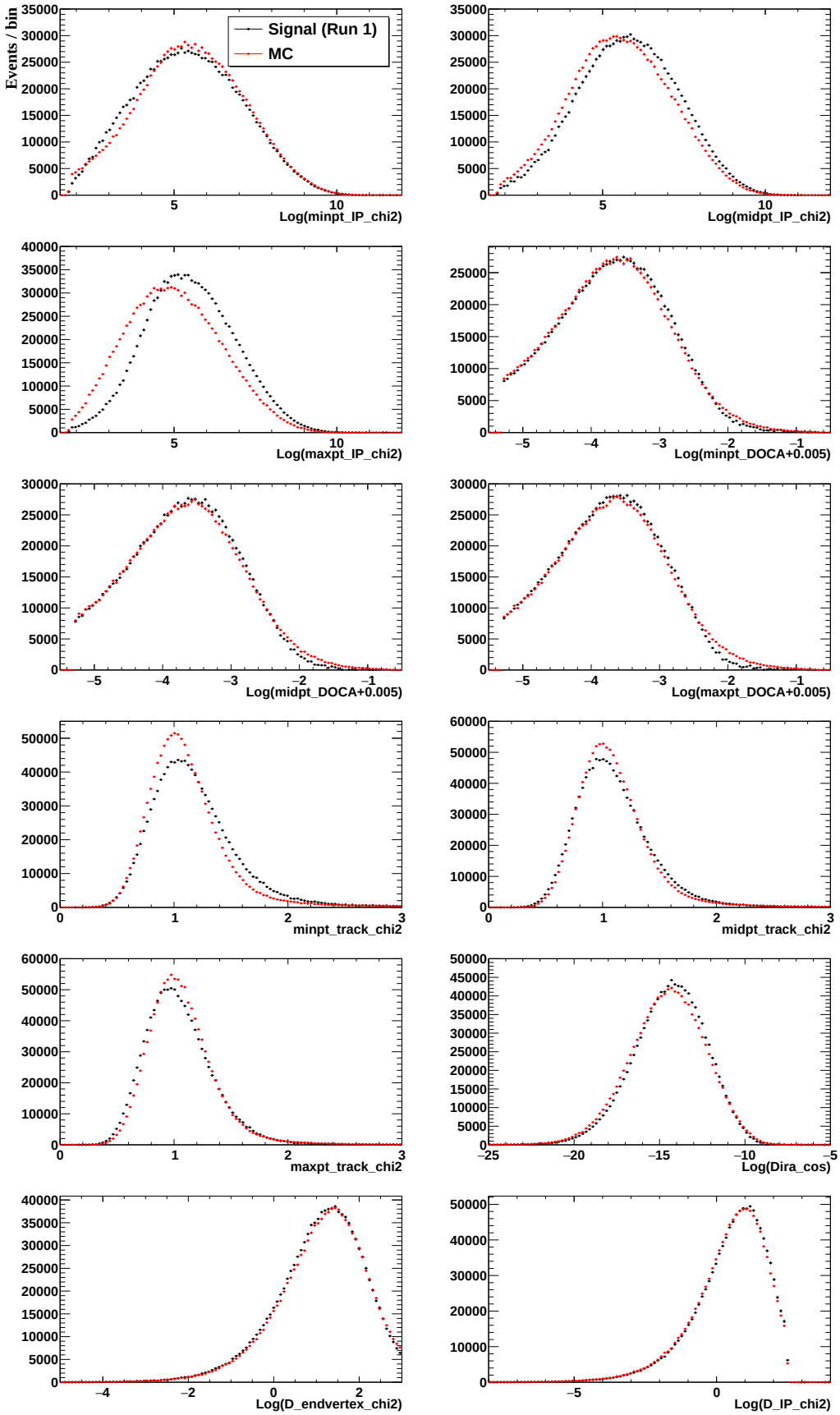


Figure 4.3: Comparison between simulation and background-subtracted data for the input variables to the multivariate classifier for Run 1. These variables are not rescaled.

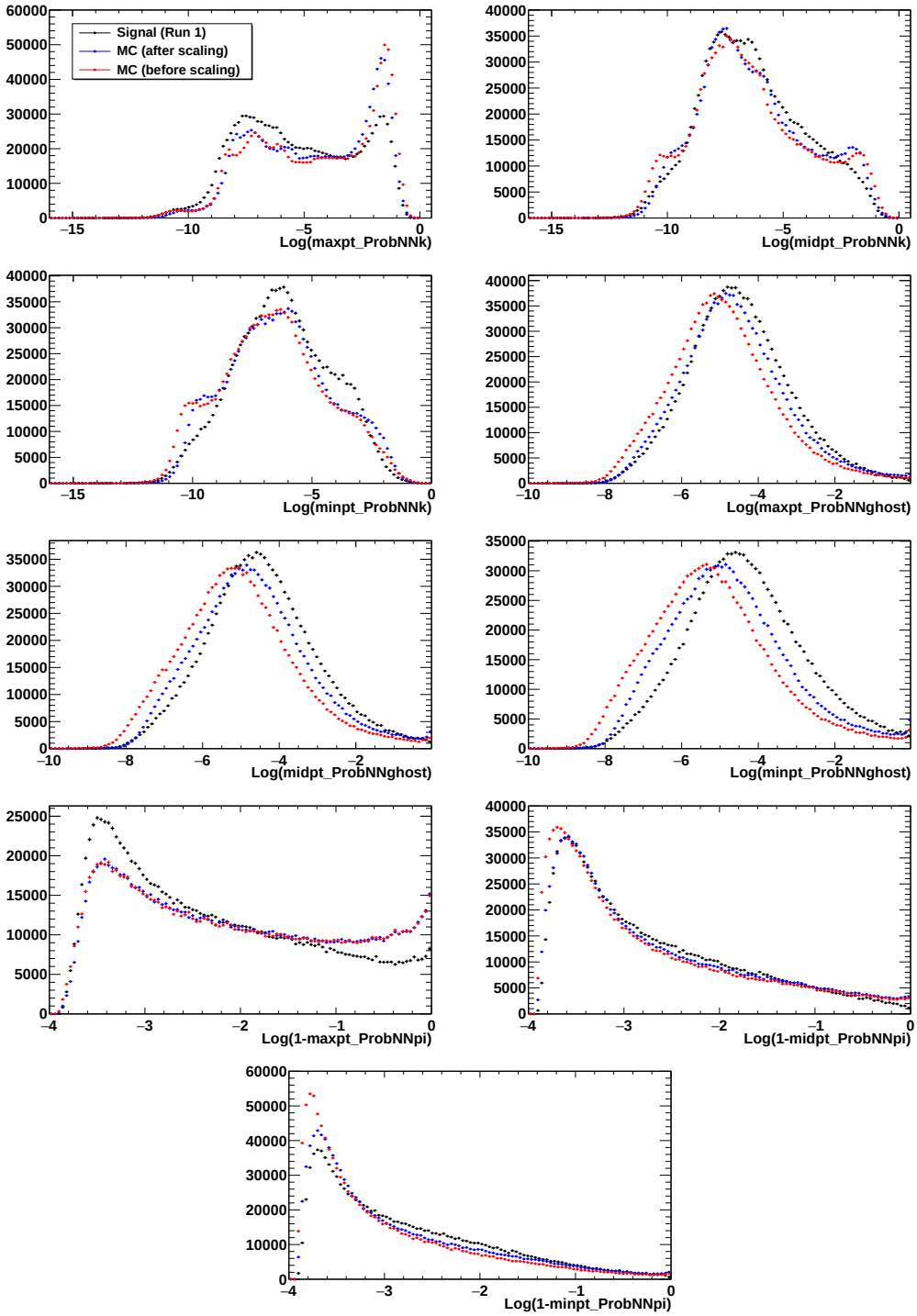


Figure 4.4: Comparison between simulation, scaled simulation and background-subtracted data for the input variables to the multivariate classifier with the largest discrepancy between data and simulation, for Run 1.

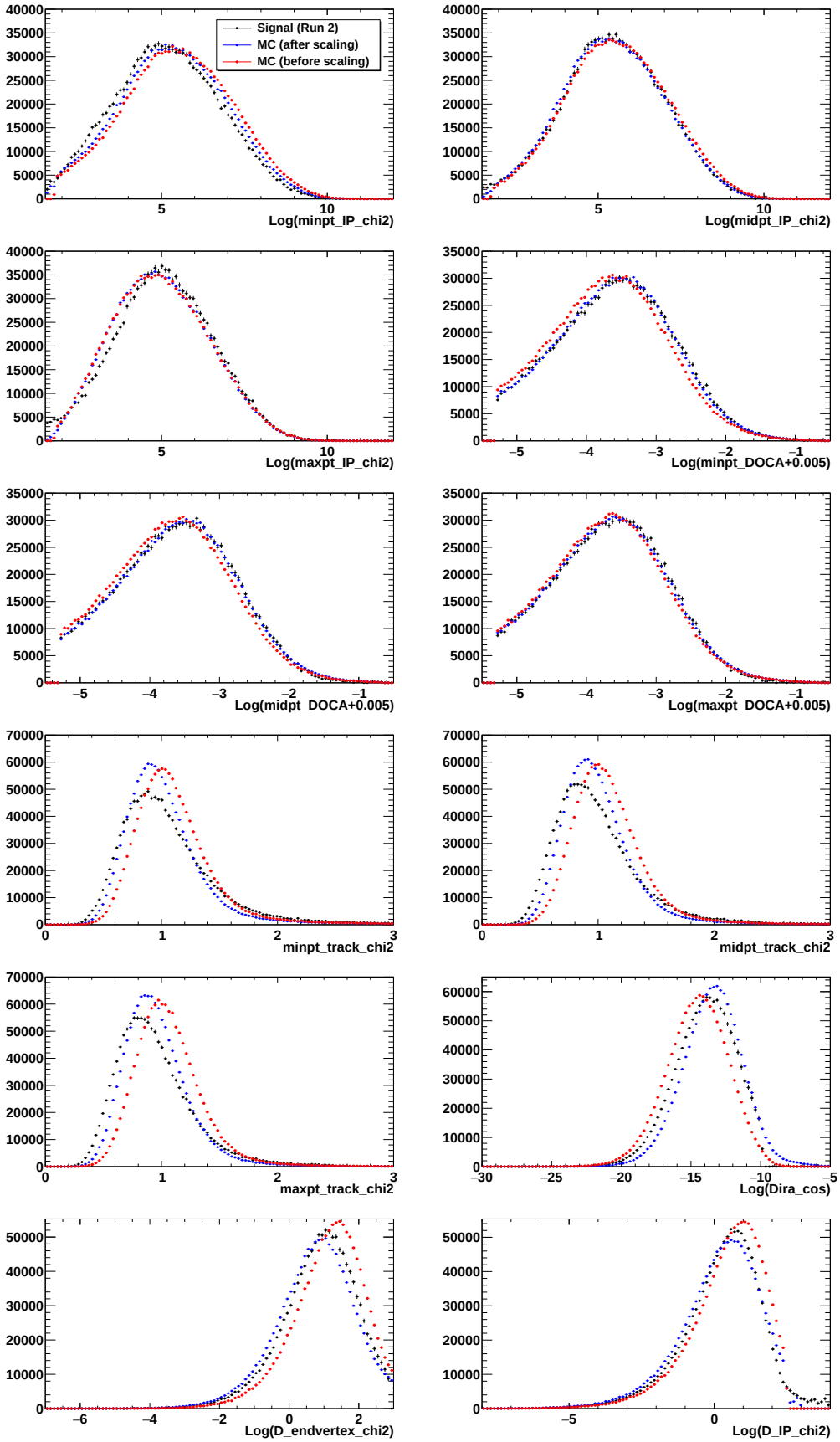


Figure 4.5: Comparison between simulation and background-subtracted data for the input variables to the multivariate classifier, for Run 2.

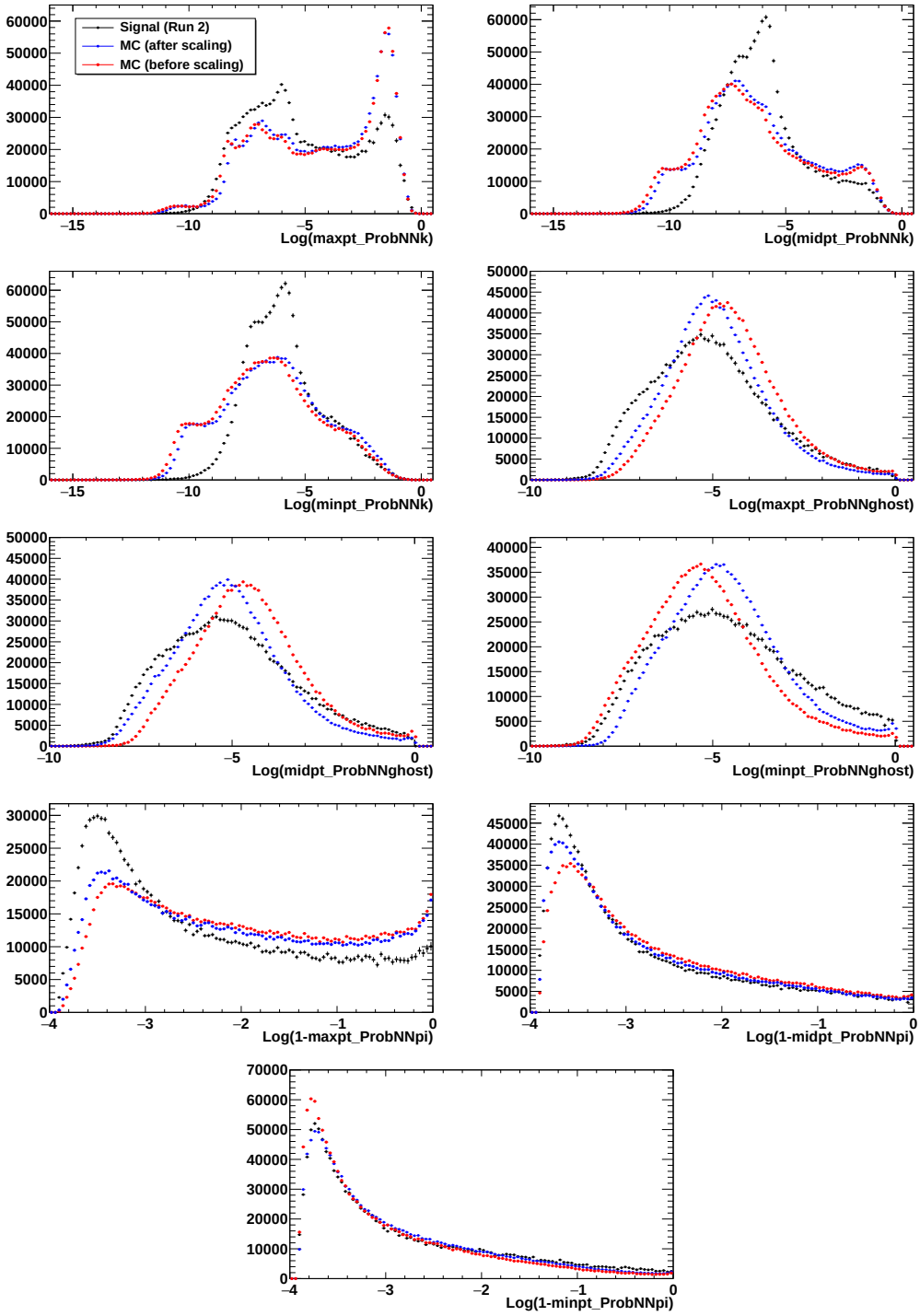


Figure 4.6: Comparison between simulation and background-subtracted data for the input variables to the multivariate classifier, for Run 2.

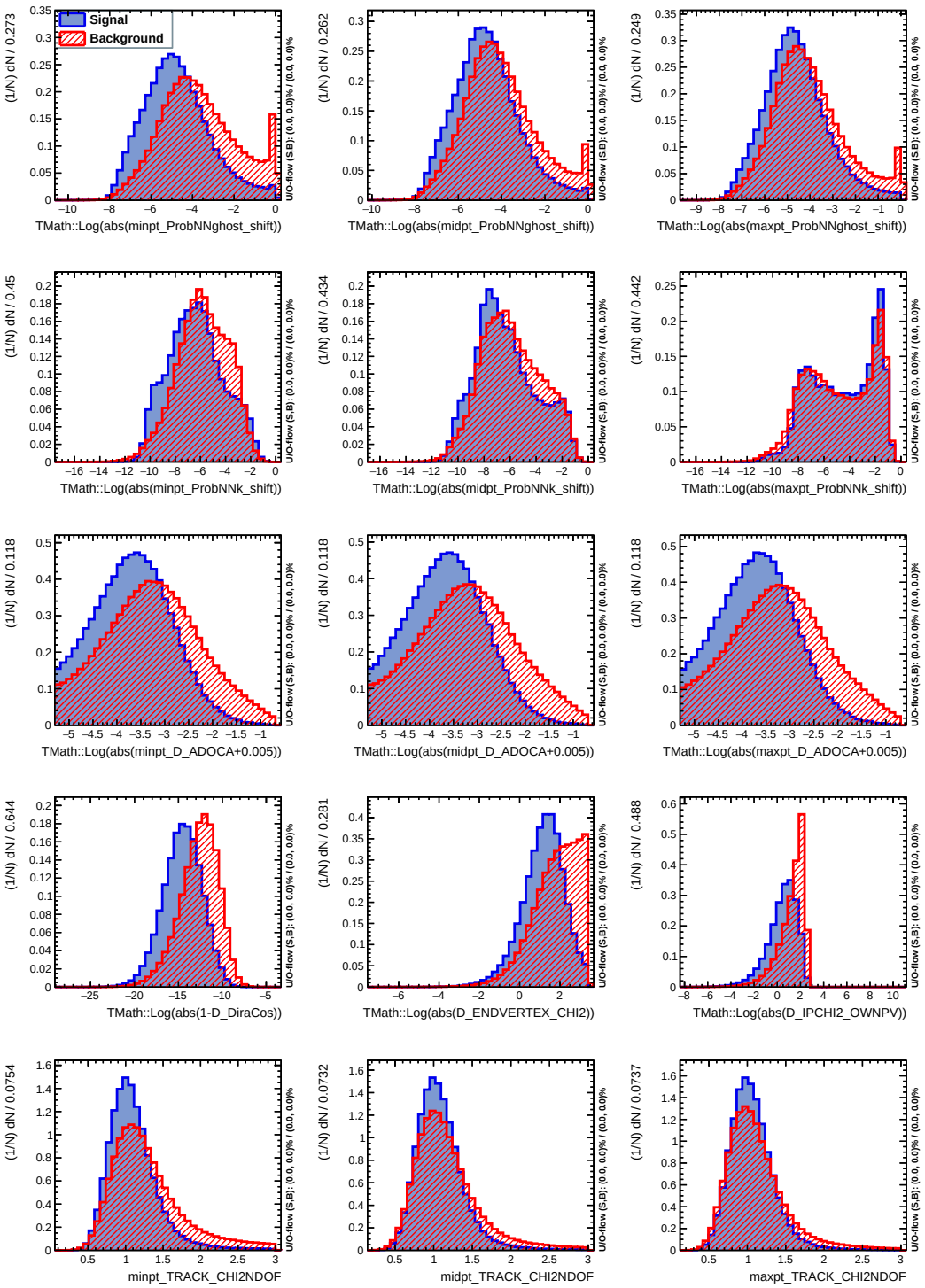


Figure 4.7: Input variables for the BDT training for Run 1. The label “shift" indicates a variable scaled according to the procedure described in Section 4.3.2.

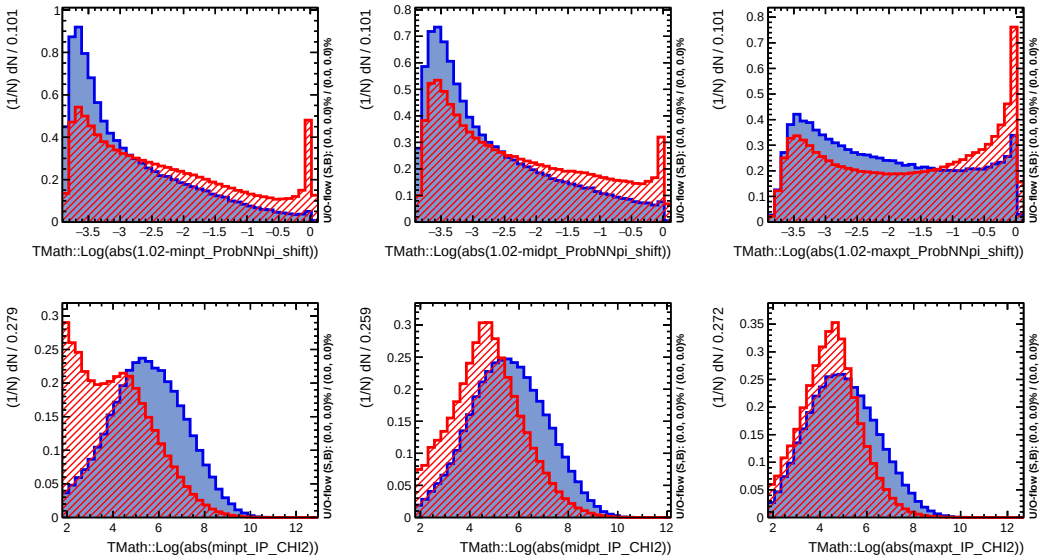


Figure 4.8: Input variables for the BDT training for Run 1. The label “shift" indicates a variable scaled according to the procedure described in Section 4.3.2.

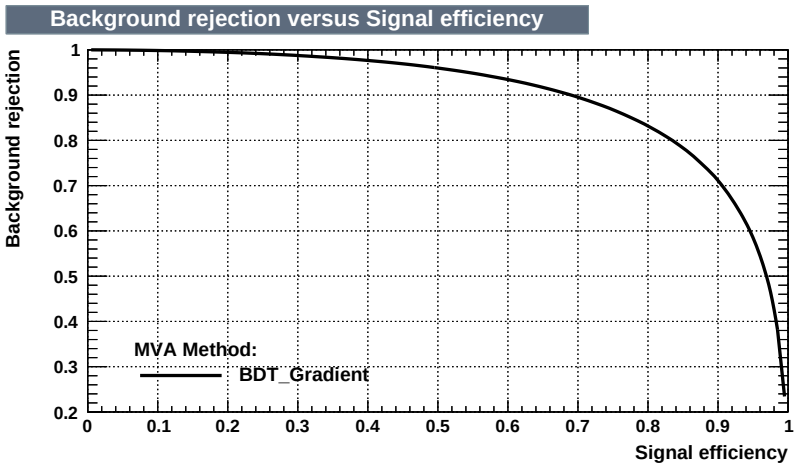


Figure 4.9: ROC curve for the BDT training for Run 1.

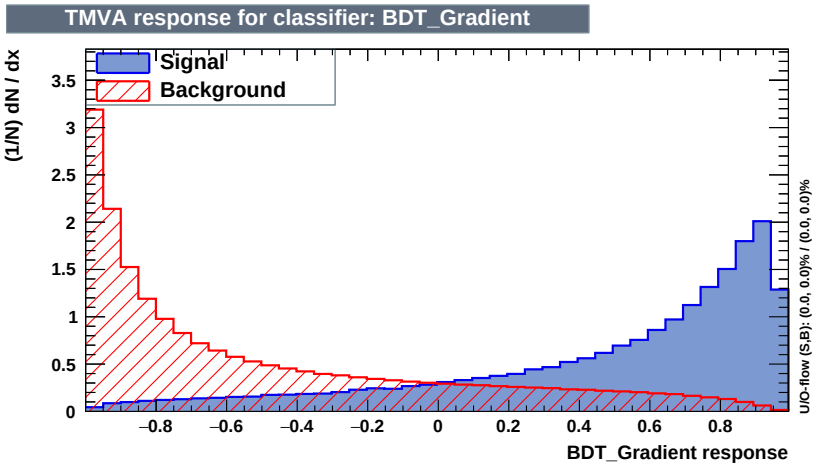


Figure 4.10: Classifier response, signal and background distribution for Run 1.

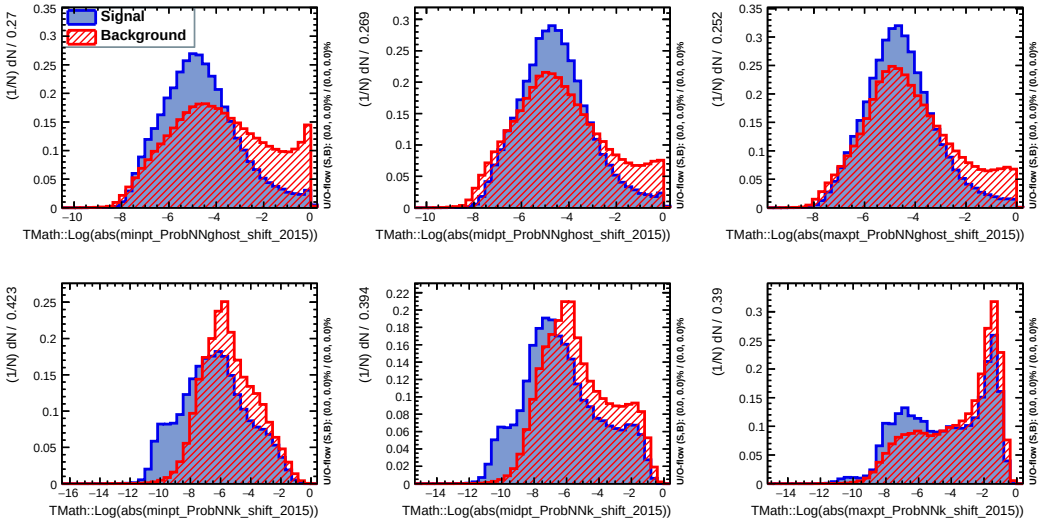


Figure 4.11: Input variables for the BDT training for Run 2. The label “shift" indicates a variable scaled according to the procedure described in Section 4.3.2.

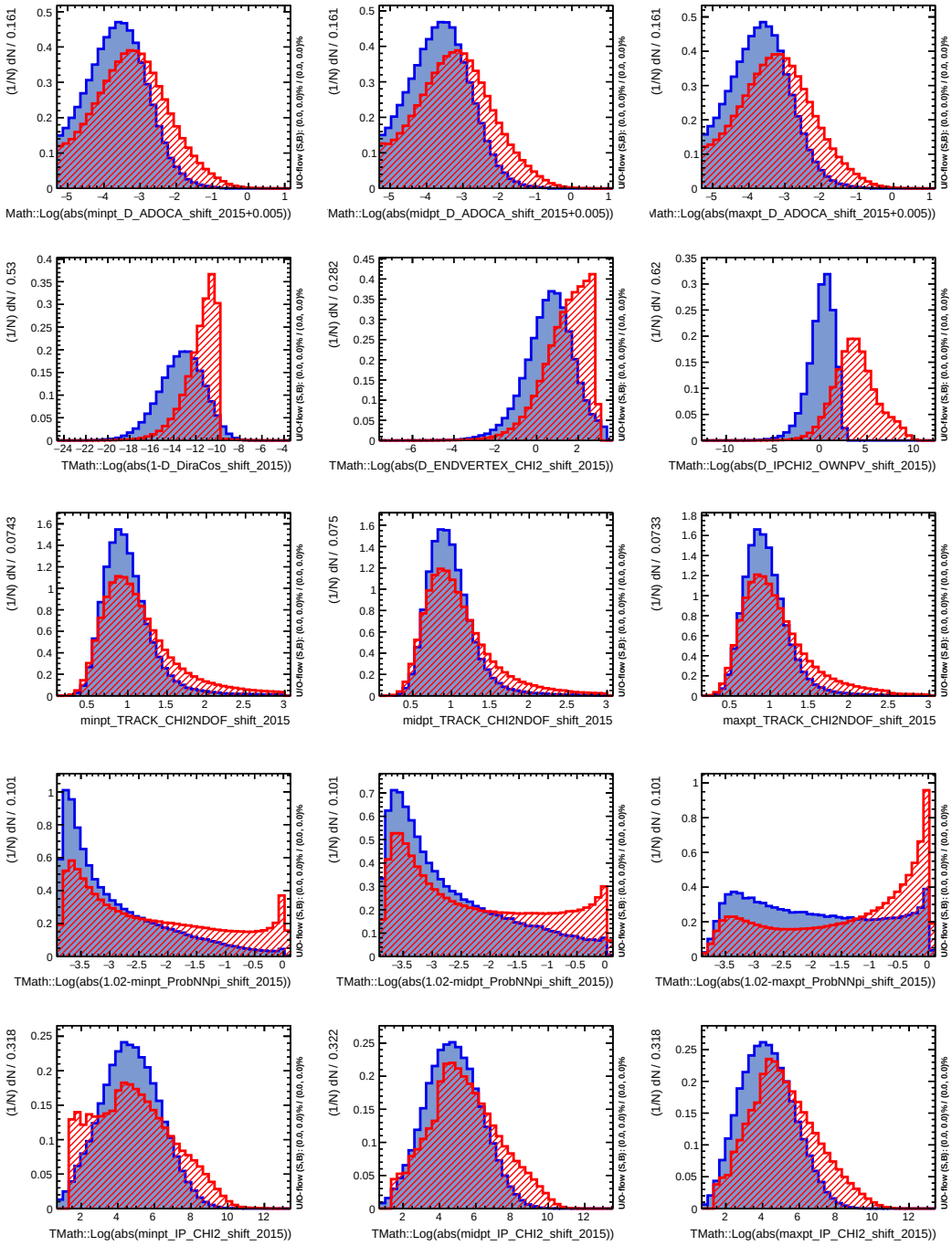


Figure 4.12: Input variables for the BDT training for Run 2. The label “shift" indicates a variable scaled according to the procedure described in Section 4.3.2.

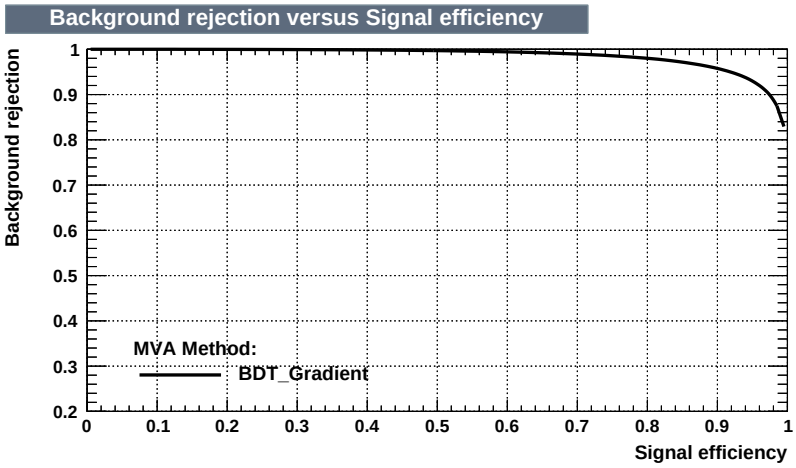


Figure 4.13: ROC curve for the BDT training for Run 2.

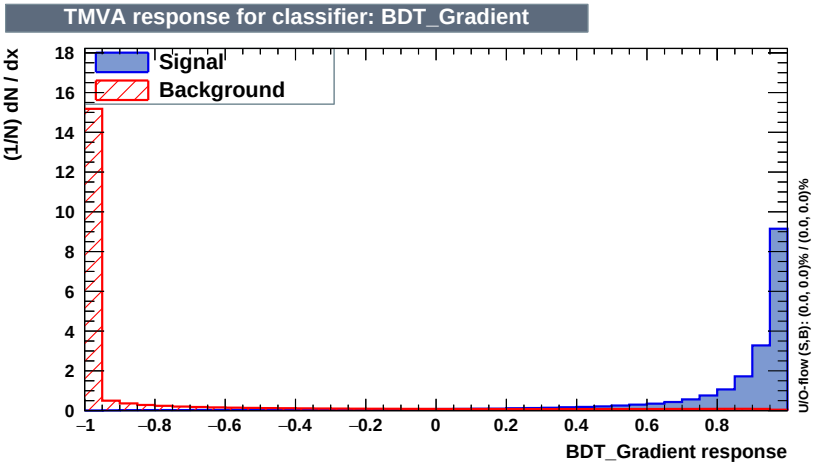


Figure 4.14: Classifier response, signal and background distribution for Run 2.

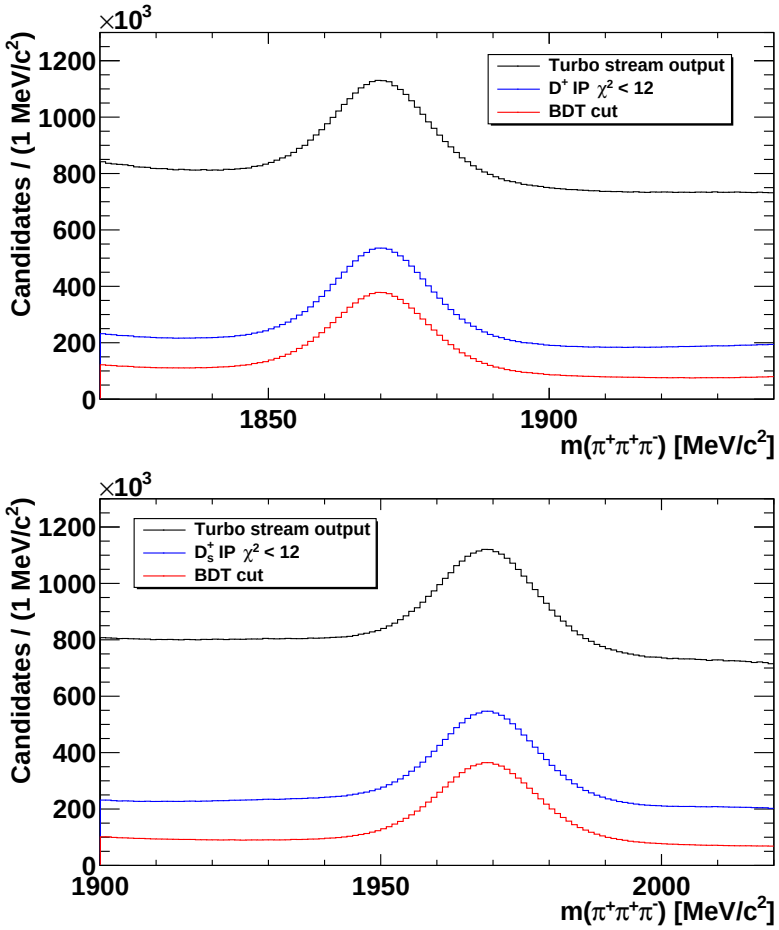


Figure 4.15: Comparison of the 3π mass spectra for D^+ (top) and D_s^+ (bottom) for Run 2, before and after the cut on $\chi^2_{IP}(D^+)$, and after the application of the BDT (but without the cut on $\chi^2_{IP}(D^+)$).

4.5 Optimisation of selections

The best cut values for the discriminating variables are found by studying the figure of merit

$$\text{FoM} = \frac{N_{sig}}{\sqrt{N_{sig} + N_{bkg}}} \tag{4.4}$$

as a function of both the cut on the BDT response and the width of the mass window, symmetrical around the D^+ or D_s^+ peak values, for each $D_{(s)}^+$ peak. In Equation (4.4) N_{sig} and N_{bkg} are the numbers of fitted signal and background events in the $\pi^+\pi^+\pi^-$ mass spectrum. As is outlined in Section 4.1, from quantitative considerations on Equation (4.1) it is shown that maximizing the value of the FoM corresponds to minimizing the statistical error and, consequently, the limit on $\mathcal{B}(\eta^{(\prime)} \to \pi^+\pi^-)$; therefore, the best cut values for the BDT response and the optimal mass window are found at the maximum of the FoM distribution.

The numbers of signal and background events used to calculate the FoM for each cut value of the BDT response are extracted from a binned extended likelihood fit to the mass spectrum, using a third-order Chebyshev polynomial for the background and a sum of a Gaussian and a Crystal Ball function [118] for the D^+ and D_s^+ peaks, and integrating the signal and background functions across the appropriate mass window.

In Figs. 4.16 and 4.17 the distributions of the FoMs for D^+ and D_s^+ from Run 1 and Run 2 are shown, binned in BDT response and width of the mass window.

Figure 4.18 and Figure 4.19 show the $\pi^+\pi^+\pi^-$ invariant mass distributions before and after the application of the cut on the BDT response, found separately for the two peaks by maximising the FoM. For Run 1, it is found to be the same for both D^+ and D_s^+. The mass windows which maximise the FoM are also shown, with vertical blue lines, for each peak. The values are:

- Run 1: D^+ BDT > -0.5, D^+ mass window ± 20 MeV/c^2
- Run 1: D_s^+ BDT > -0.5, D_s^+ mass window ± 20 MeV/c^2
- Run 2: D^+ BDT > -0.5, D^+ mass window ± 21 MeV/c^2
- Run 2: D_s^+ BDT > -0.6, D_s^+ mass window ± 21 MeV/c^2

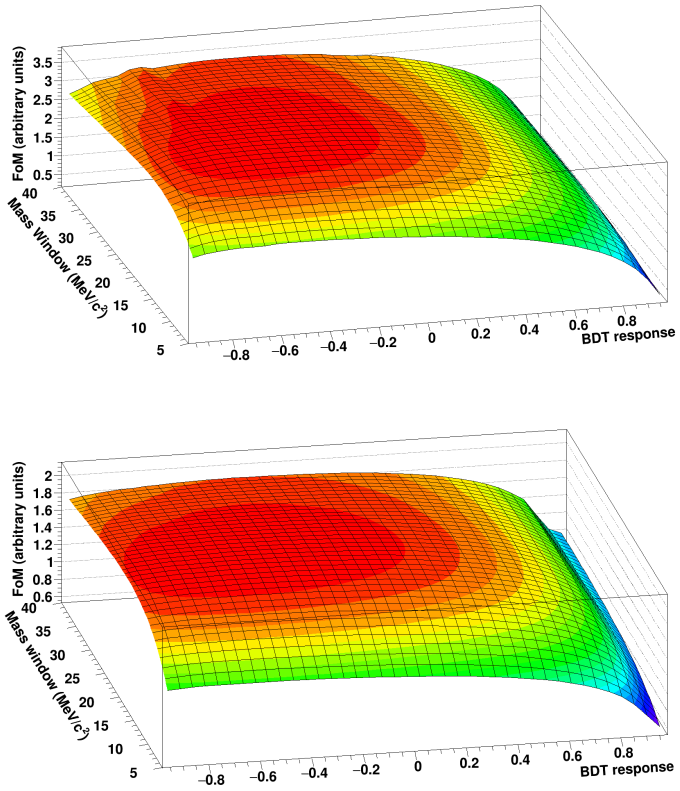


Figure 4.16: FoM distribution for D^+ from Run 1 (top) and Run 2 (bottom).

4.5.1 Fits to the optimised mass spectra

The optimal cut values for the BDT response, given in the previous section, are applied to obtain the $\pi^+\pi^+\pi^-$ invariant mass distributions, which are fitted with a binned extended likelihood fit to extract the numbers of signal events in the D^+ and D_s^+ mass windows. The total PDF for both the D^+ and D_s^+ regions is the sum of a double-sided Crystal Ball function and a Gaussian function for the peaks. The background is modeled with a fourth-order Chebyshev polynomial, plus six histogram templates containing toy Monte Carlo samples, to account for the contribution of the decays $D_s^+ \to K^+\pi^+\pi^-$, where the kaon is misidentified as pion, $D_s^+ \to \pi^+\pi^+\pi^-\pi^0$ and the four $D_{(s)}^+ \to (\eta^{(\prime)} \to \pi^+\pi^-\gamma)\pi^+$ channels. Figures 4.20 and 4.21 show the shapes of the first two contributions, obtained with 10^7 toy MC events each, generated with the `TGenPhaseSpace` class in ROOT; the experimental width of the D_s^+ has been neglected. For the four $D_{(s)}^+ \to (\eta^{(\prime)} \to \pi^+\pi^-\gamma)\pi^+$ contributions, shown in Figure 4.22, 10^9 events have been generated. The contribution of the decay $D^+ \to K^+\pi^+\pi^-$, with the kaon misidentified as a pion, lies outside the $\pi^+\pi^+\pi^-$ mass range used for

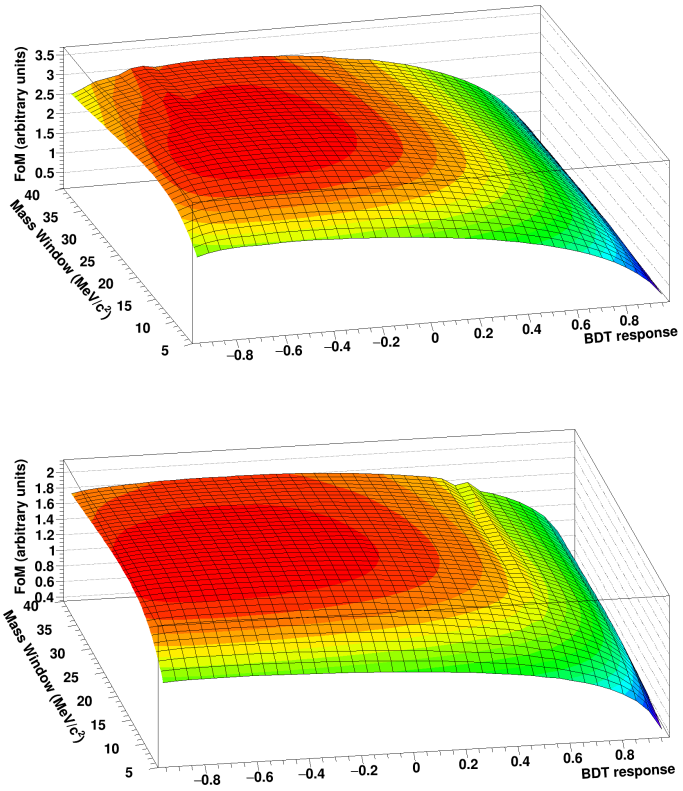


Figure 4.17: FoM distribution for D_s^+ from Run 1 (top) and Run 2 (bottom).

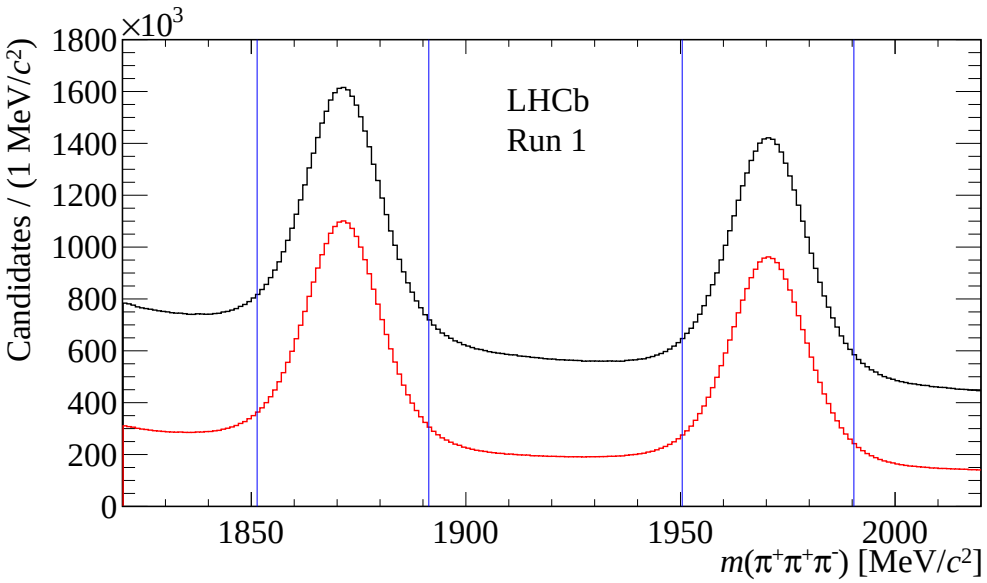


Figure 4.18: Invariant mass spectrum of $\pi^+\pi^+\pi^-$ from Run 1, before (black) and after (red) the cut on the BDT. The vertical blue lines define the optimum mass windows.

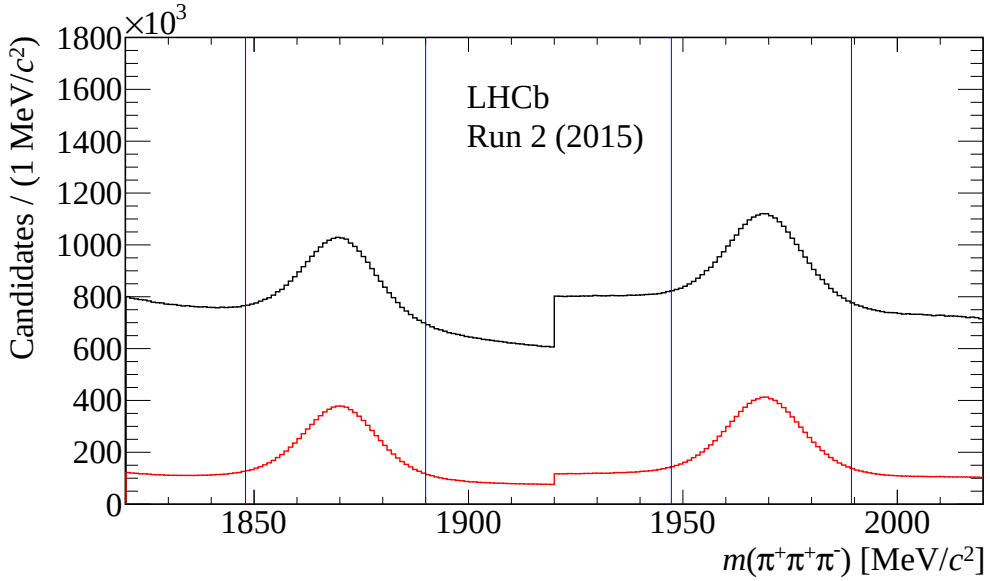


Figure 4.19: Invariant mass spectrum of $\pi^+\pi^+\pi^-$ for D^+ (top) and D_s^+ (bottom) from Run 2, before (black) and after (red) the cut on the BDT. The vertical blue lines define the optimum mass windows.

this analysis.

For each peak, the Gaussian and Crystal Ball functions are constrained to have the same peak value, which is allowed to vary by ± 3 MeV/c^2 around the nominal masses. The fraction of the Crystal Ball yield with respect to the Gaussian is given by a parameter, $f_{D_{(s)}^+}$, with a range between 0 and 1. The left and right tails of the Crystal Ball functions are allowed to vary independently and are constrained to be the same for both peaks. The yields of the $D_{(s)}^+ \to \eta(\pi^+\pi^-\gamma)\pi^+$ components are constrained to be a fraction of the respective yields of the $D_{(s)}^+ \to \eta'(\pi^+\pi^-\gamma)\pi^+$ components, which are calculated as the ratios of the branching fractions, taken from the Particle Data Group.

Figure 4.23 shows the fit to the mass spectrum for the BDT cut value which maximises the FoM for both the D^+ and D_s^+ for Run 1 data. The ratios of the fitted values and the data are also shown, with the two horizontal red lines indicating a $\pm 2\%$ variation between the data and the fits. To account for the mismodelling of the peak shapes, a conservative systematic uncertainty of $\pm 2\%$ on the numbers of fitted $D_{(s)}^+$ mesons in the sample is assumed. A summary of the fit results is given in Table 4.6.

Figure 4.24 shows the fits to the mass spectra for the Run 2 data. For these distributions the fluctuations of the ratio between the fits and the data lie within a $\pm 1\%$ variation, shown in the figures with two horizontal red lines; therefore,

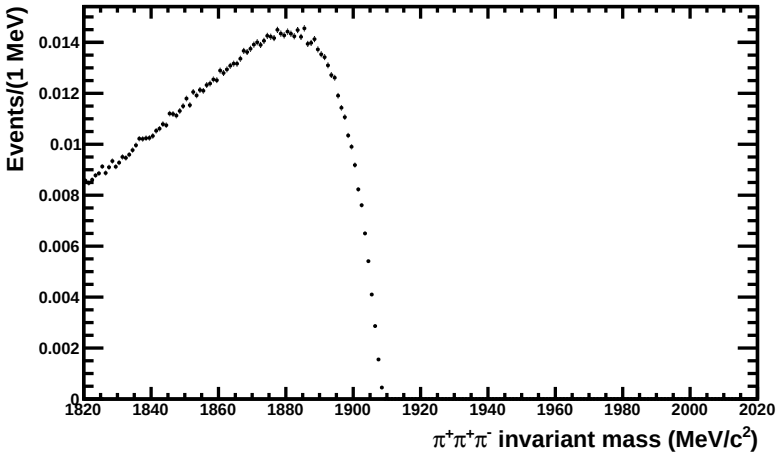


Figure 4.20: PDF for the mis-ID $D_s^+ \to K^+\pi^+\pi^-$ channel.

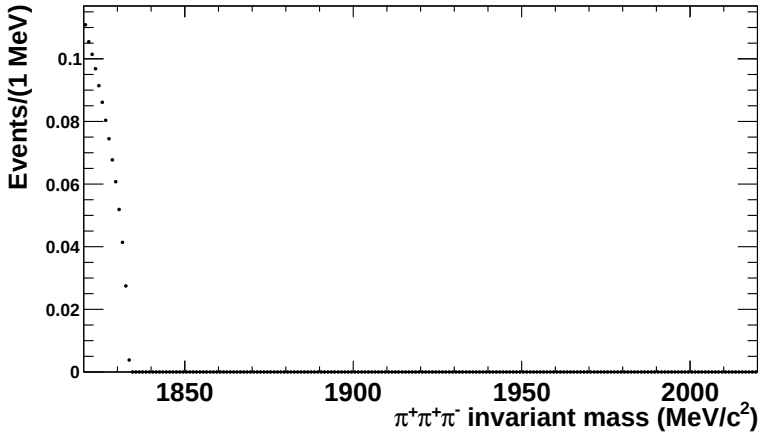


Figure 4.21: PDF for the $D_s^+ \to \pi^+\pi^+\pi^-\pi^0$ channel.

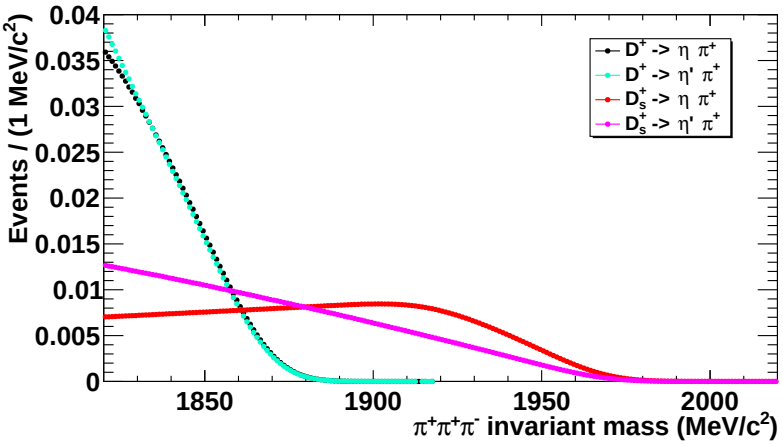


Figure 4.22: PDFs for the $\eta^{(\prime)} \to \pi^+\pi^-\gamma$ channels.

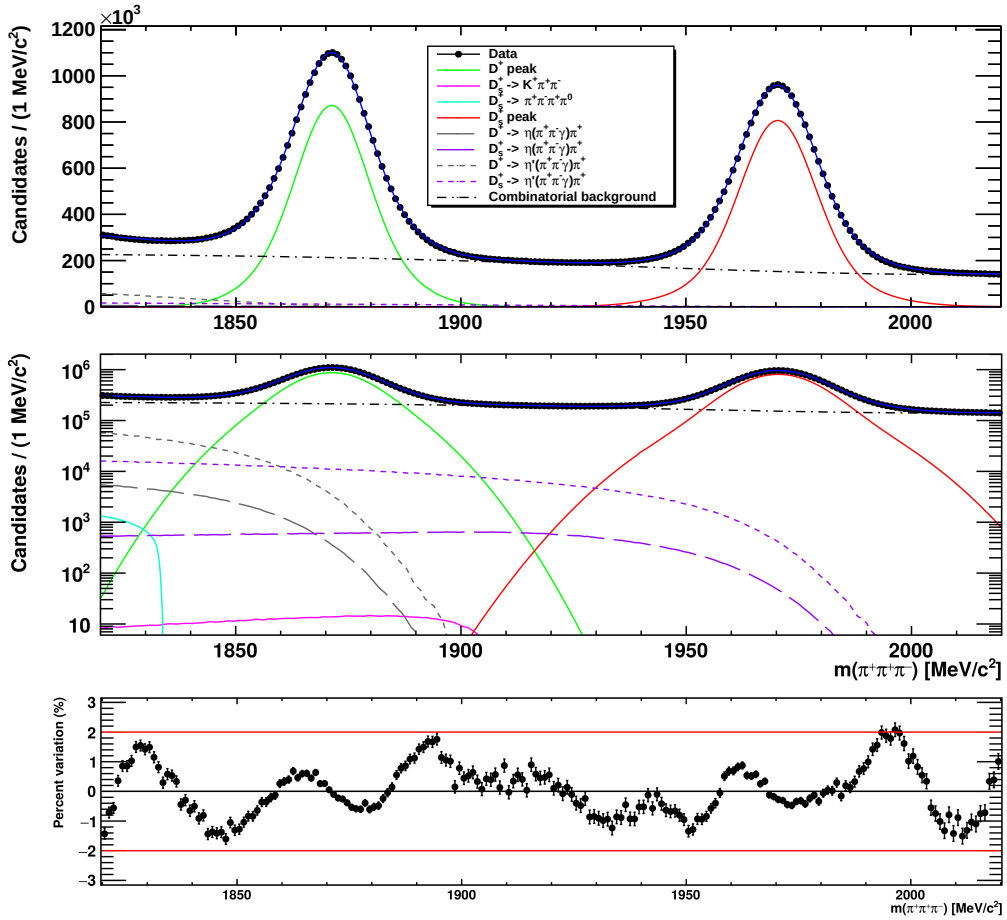


Figure 4.23: Fit to the mass spectrum from Run 1 for BDT > -0.5.

for Run 2 data a systematic uncertainty of $\pm 1\%$ is assumed on the number of fitted $D^+_{(s)}$ mesons. Since the two Run 2 datasets come from different turbo lines and have a different BDT cut, the Crystal Ball parameters are not constrained to be the same for both peaks. A summary of the fit results is given in Table 4.7. The uncertainty on several parameters seems to be quite small: this is actually expected, given the large sample size, and it was verified with fits to toy MC samples with variable statistics.

For Run 1, the fit results give 1.86×10^7 signal and 9.40×10^6 background events in the D^+ mass window, and 1.73×10^7 signal and 6.67×10^6 background events in the D_s^+ mass window; for Run 2, there are 6.1×10^6 signal and 4.0×10^6 background events in the D^+ mass window, and 6.3×10^6 signal and 3.8×10^6 background events in the D_s^+ mass window.

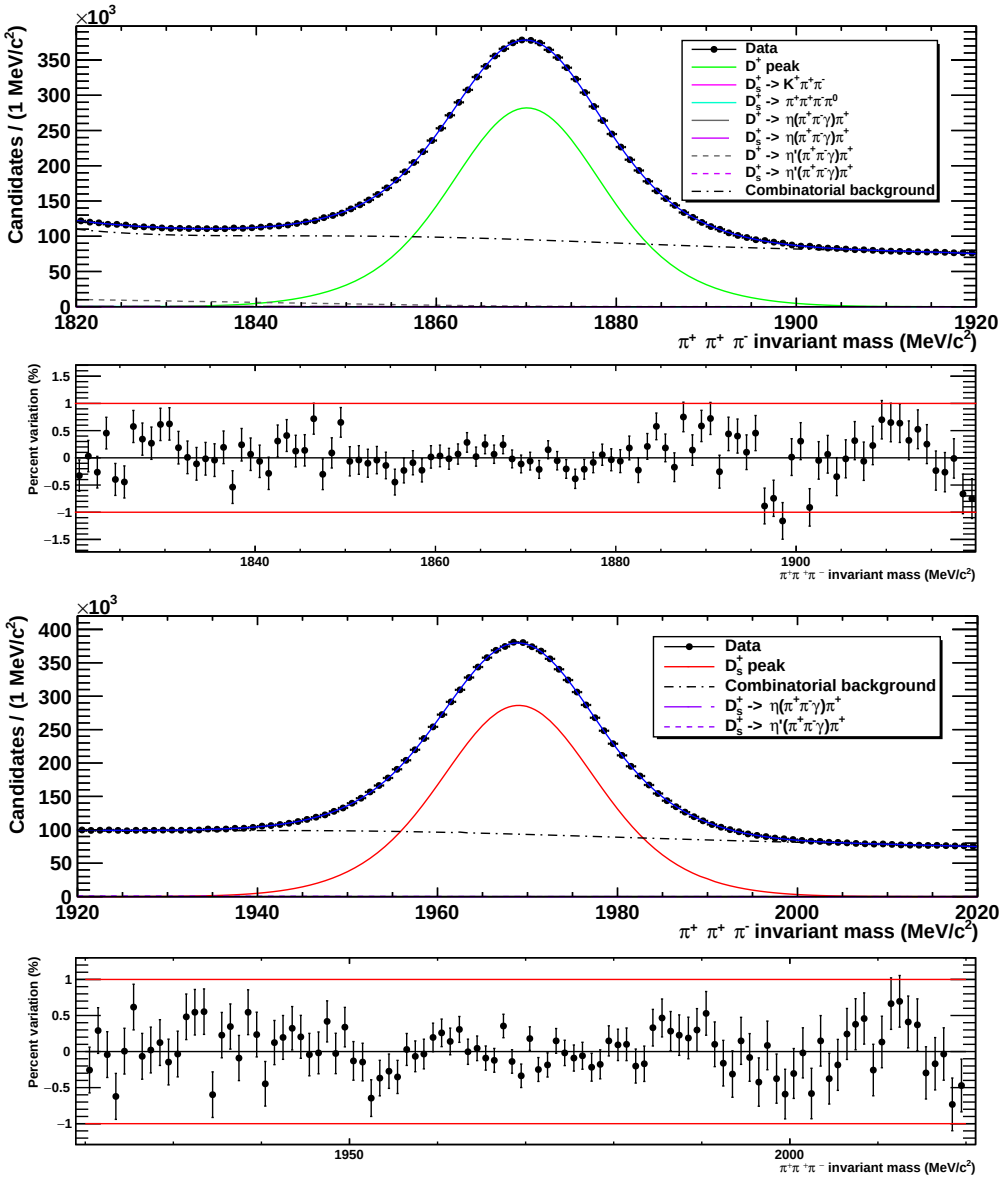


Figure 4.24: Fit to the mass spectra from Run 2 data for BDT > -0.5 (D^+, top plot) and BDT > -0.6 (D_s^+, bottom plot).

Table 4.6: Fitted parameters corresponding to Fig. 4.23. The peak values μ of the Gaussian functions and the standard deviations σ of the Gaussian and Crystal Ball functions are given in units of MeV/c^2.

Function	**Parameter**	**Value (BDT>-0.5)**
D^+ Gaussian	μ	1871.313 ± 0.003
	σ	11.850 ± 0.008
D^+ DS Crystal Ball	σ	7.188 ± 0.005
	α_L	7.0 ± 0.3
	n_L	5.3 ± 0.4
	α_R	6.5 ± 0.3
	n_R	4.5 ± 0.2
D_s^+ Gaussian	μ	1970.425 ± 0.003
	σ	14.309 ± 0.007
D_s^+ DS Crystal Ball	σ	7.949 ± 0.004
Yields	f_{D^+}	(45.1 ± 0.1) %
	$f_{D_s^+}$	(64.6 ± 0.1) %
	N_{D^+}	$(2.0033 \pm 0.0007) \times 10^7$
	$N_{D_s^+}$	$(1.9367 \pm 0.0006) \times 10^7$
	N_{comb}	$(3.674 \pm 0.001) \times 10^7$
	$N_{K\pi\pi}$	$(1.1 \pm 0.5) \times 10^3$
	$N_{4\pi}$	$(1.2 \pm 0.2) \times 10^4$
	$N_{D^+ \to \eta'(\pi\pi\gamma)\pi}$	$(1.50 \pm 0.03) \times 10^6$
	$N_{D_s^+ \to \eta'(\pi\pi\gamma)\pi}$	$(1.26 \pm 0.01) \times 10^6$
Signal window	N_{D^+}	$(1.8775 \pm 0.0006) \times 10^7$
	$N_{bkg}(D^+)$	$(9.767 \pm 0.003) \times 10^6$
	$N_{D_s^+}$	$(1.7461 \pm 0.0005) \times 10^7$
	$N_{bkg}(D_s^+)$	$(6.918 \pm 0.002) \times 10^6$

Table 4.7: Fitted parameters corresponding to Fig. 4.24. The peak values μ of the Gaussian functions and the standard deviations σ of the Gaussian and Crystal Ball functions are given in units of MeV/c^2.

Function	**Parameter**	**Value**
D^+ Gaussian	μ	1869.962 ± 0.006
	σ	12.06 ± 0.01
D^+ DS Crystal Ball	σ	7.244 ± 0.007
	α_L	7 ± 1
	n_L	5.2 ± 0.8
	α_R	7 ± 1
	n_R	1.8 ± 0.6
D^+ Yields	f_{D^+}	(52.8 ± 0.1) %
	N_{D^+}	$(6.418 \pm 0.003) \times 10^6$
	N_{comb}	$(9.2557 \pm 0.004) \times 10^6$
	$N_{K\pi\pi}$	$(6.9 \pm 0.4) \times 10^2$
	$N_{4\pi}$	$(5.0 \pm 0.2) \times 10^3$
	$N_{D^+ \to \eta'(\pi\pi\gamma)\pi}$	$(2.6 \pm 0.2) \times 10^5$
	$N_{D_s^+ \to \eta\prime(\pi\pi\gamma)\pi}$	$(2.4 \pm 0.1) \times 10^4$
D_s^+ Gaussian	μ	1969.146 ± 0.006
	σ	11.24 ± 0.02
D_s^+ DS Crystal Ball	σ	7.221 ± 0.008
	α_L	9 ± 1
	n_L	2 ± 1
	α_R	4.6 ± 0.7
	n_R	2 ± 1
D_s^+ Yields	$f_{D_s^+}$	(59.7 ± 0.3) %
	$N_{D_s^+}$	$(6.547 \pm 0.003) \times 10^6$
	N_{comb}	$(9.053 \pm 0.004) \times 10^6$
	$N_{D_s^+ \to \eta\prime(\pi\pi\gamma)\pi}$	$(1.8 \pm 0.1) \times 10^5$
Signal window	N_{D^+}	$(6.090 \pm 0.003) \times 10^6$
	$N_{bkg}(D^+)$	$(4.035 \pm 0.002) \times 10^6$
	$N_{D_s^+}$	$(6.260 \pm 0.003) \times 10^6$
	$N_{bkg}(D_s^+)$	$(3.900 \pm 0.002) \times 10^6$

4.6 Decay tree fit algorithm

The $\pi^+\pi^-$ mass spectra are obtained by selecting events in the optimised mass windows of the $D^+_{(s)}$ peak and then performing a kinematic fit [119] on the selected $D^+_{(s)}$ candidates (called the Decay Tree Fit (DTF) algorithm): the 4-momenta of the final state particles are modified so that they are constrained to originate from the same vertex, the $D^+_{(s)}$ candidate is constrained to originate from the best primary vertex in the event, and the experimental width on its mass is eliminated by constraining it to the $D^+_{(s)}$ mass from the Particle Data Group. Events for which the $D^+_{(s)}$ candidate does not originate from the primary vertex or its mass is not close to the nominal $D^+_{(s)}$ mass before the DTF algorithm, quantitatively evaluated with a χ^2 calculation, are discarded. This procedure leads to an increase in resolution of almost a factor of 2 in the $\pi^+\pi^-$ spectra, as can be seen in Figure 4.25 for the $\eta \to \pi^+\pi^-$ peak from MC.

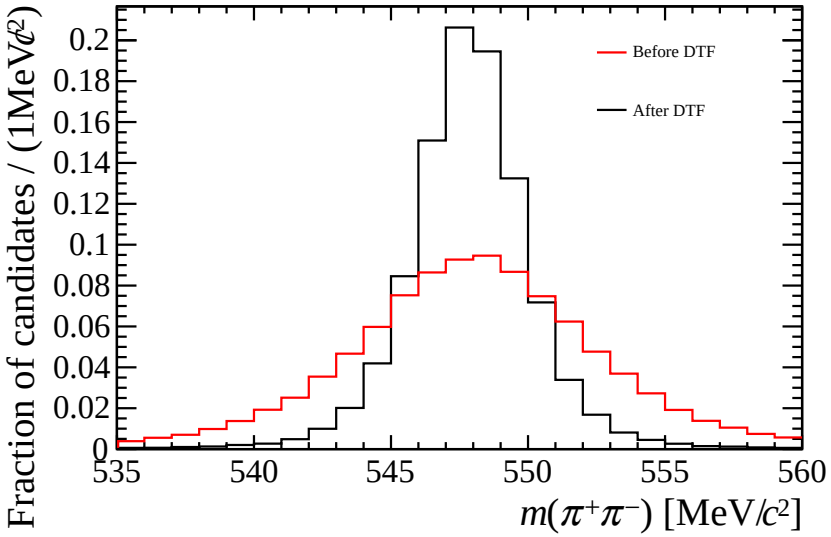


Figure 4.25: Effect of the DTF algorithm on the η peak from simulated $D^+ \to \eta\pi^+$, $\eta \to \pi^+\pi^-$.

4.7 Relative efficiency variation with $\pi^+\pi^-$ mass

4.7.1 Run 1

The term $\epsilon_{D^+_{(s)}}(\eta^{(\prime)})$ in Equation (4.1) corrects for any non-uniformity in efficiency over the $\pi^+\pi^-$ mass spectrum from $D^+_{(s)} \to \pi^+\pi^+\pi^-$ decays. A study

of the variation of these efficiencies with $\pi^+\pi^-$ mass has been made using simulated events of the type $D^+ \to \pi^+\pi^+\pi^-$ and $D_s^+ \to \pi^+\pi^+\pi^-$, with flat phase-space decays. For each channel two samples are used for the study: one at generator level and one after a full simulation of the detector response. The samples are listed in Table 4.1. The distributions of the simulated $\pi^+\pi^-$ invariant mass for $D^+ \to \pi^+\pi^+\pi^-$ and $D_s^+ \to \pi^+\pi^+\pi^-$ decays are shown, respectively, in Figures 4.26 and 4.27, in which the effect of the different analysis steps, each one of which can deform the mass spectrum, is shown.

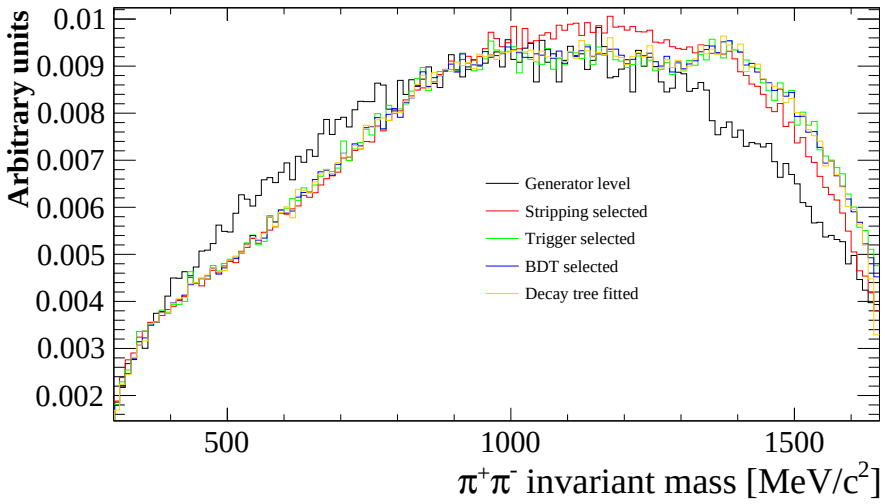


Figure 4.26: M($\pi^+\pi^-$) distribution for the $D^+ \to \pi^+\pi^+\pi^-$ generator-level sample (black), the reconstructed and selected sample (red), the triggered sample (green), the BDT-selected sample (blue) and the decay tree-fitted sample (orange). Each consecutive sample also contains the criteria of the previous samples.

The ratios of the generator level and the decay tree fitted spectra give the relative efficiency distributions for $D^+ \to \pi^+\pi^+\pi^-$ and $D_s^+ \to \pi^+\pi^+\pi^-$, as functions of the $\pi^+\pi^-$ mass, shown in Figures 4.28 and 4.29. The red bands show the results of fits using sixth-order polynomials with 68% CL error bands obtained from the uncertainties of the fitted curve parameters. The blue horizontal lines are included to help show the variation from uniform efficiency and the η and η' masses are highlighted with vertical lines. The percentage variations from uniform efficiency for $D^+ \to \pi^+\pi^+\pi^-$ are found to be $(-15 \pm 1)\,\%$ for the η and $(+1 \pm 1)\,\%$ for the η'. For $D_s^+ \to \pi^+\pi^+\pi^-$ they are found to be $(-20 \pm 1)\,\%$ for the η and $(+3 \pm 1)\,\%$ for the η'. The values $\epsilon_{D^+}(\eta) = 0.85 \pm 0.01$, $\epsilon_{D_s^+}(\eta) = 0.80 \pm 0.01$, $\epsilon_{D^+}(\eta') = 1.01 \pm 0.01$ and $\epsilon_{D_s^+}(\eta') = 1.03 \pm 0.01$ are therefore used as corrections to the measured

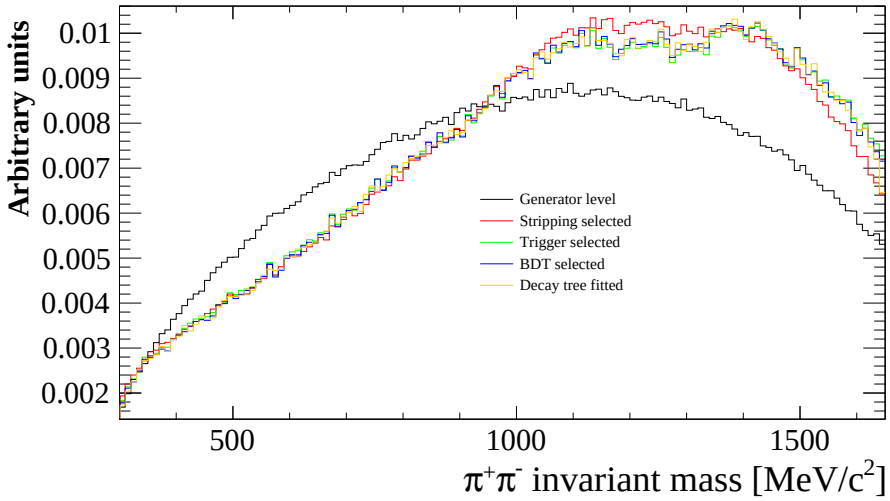


Figure 4.27: M($\pi^+\pi^-$) distribution for the $D_s^+ \to \pi^+\pi^+\pi^-$ generator-level sample (black), the reconstructed and selected sample (red), the triggered sample (green), the BDT-selected sample (blue) and the decay tree-fitted sample (orange). Each consecutive sample also contains the criteria of the previous samples.

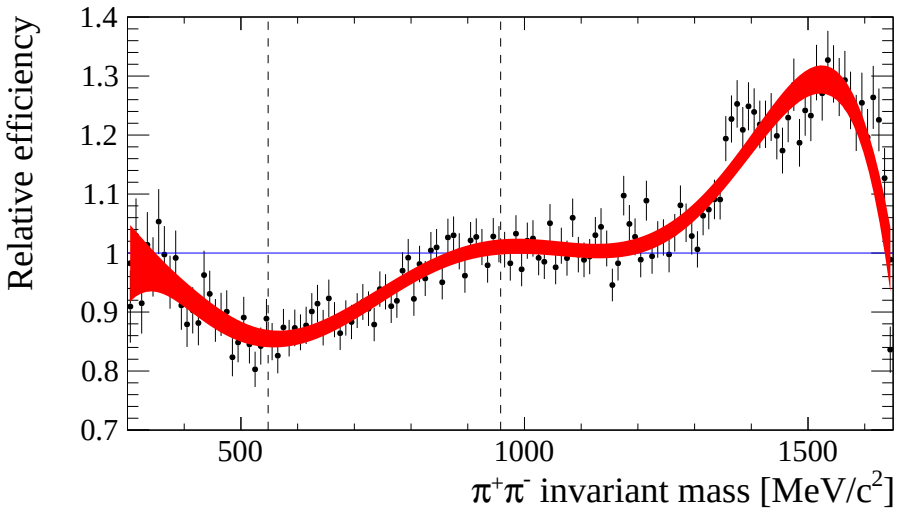


Figure 4.28: Relative efficiency distribution for $D^+ \to \pi^+\pi^+\pi^-$ as a function of M($\pi^+\pi^-$).

numbers of $\eta^{(\prime)} \to \pi^+\pi^-$ events when computing the limits on the branching fractions for the Run 1 data. These results are summarised in Table 4.10.

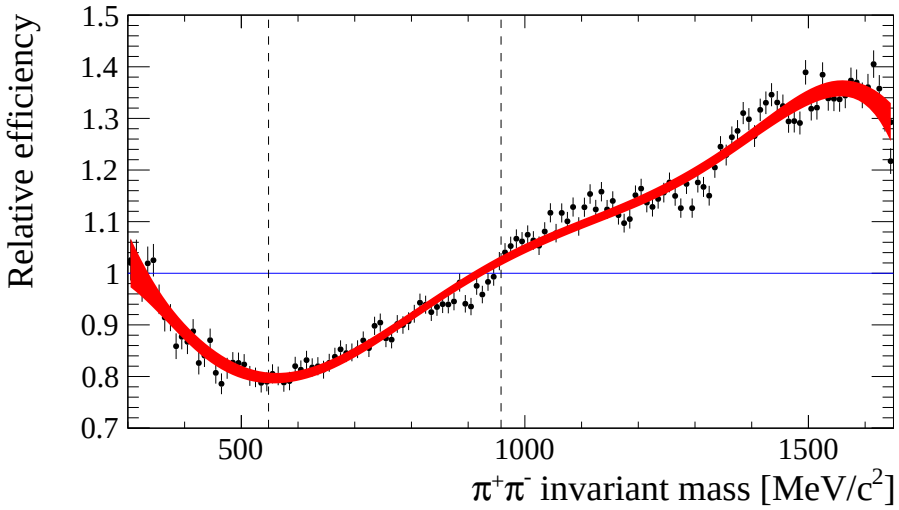


Figure 4.29: Relative efficiency distribution for $D_s^+ \to \pi^+\pi^+\pi^-$ as a function of M($\pi^+\pi^-$).

4.7.2 Run 2

Since no Monte Carlo sample for Run 2 was available at the time the analysis was performed, the relative efficiency is calculated by comparing the $\pi^+\pi^-$ mass spectra with those for Run 1 data, after all selections. Figures 4.30 and 4.31 show the $\pi^+\pi^-$ invariant mass spectra for Run 1 (blue) and Run 2 (red), from the D^+ and D_s^+ peaks separately, within the optimised mass windows. The differences are assumed to be mainly due to differences between the set of cuts in the Run 2 turbo stream selections and the one in the Run 1 stripping and trigger selections. In Figs. 4.32 and 4.33 the ratios of the Run 2 distributions to the Run 1 distributions are shown. The red bands show the results of fits, in intervals containing the η and η' regions, using third-order polynomials with 68% CL error bands obtained from the uncertainties of the fitted curve parameters. For the D^+ the relative efficiency for Run 2 is found to be (2.2 ± 0.5)% higher for the η and (1 ± 1)% higher for the η' than the relative efficiency for Run 1. For the D_s^+ the relative efficiency for Run 2 is found to be (5.0 ± 0.5)% higher for the η and (3.0 ± 0.5)% lower for the η' than the relative efficiency for Run 1. The relative efficiency factors for Run 2 are obtained by multiplying the corresponding Run 1 factors by the correction terms obtained from the ratios of Run 2 and Run 1 distributions. The resulting values are listed in Table 4.10.

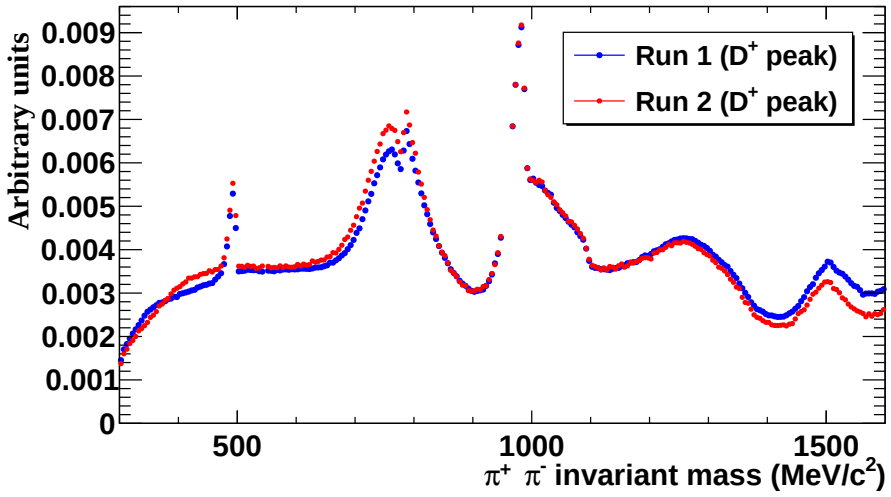


Figure 4.30: Distributions of $\pi^+\pi^-$ mass for Run 1 and Run 2, for the D^+ peak region.

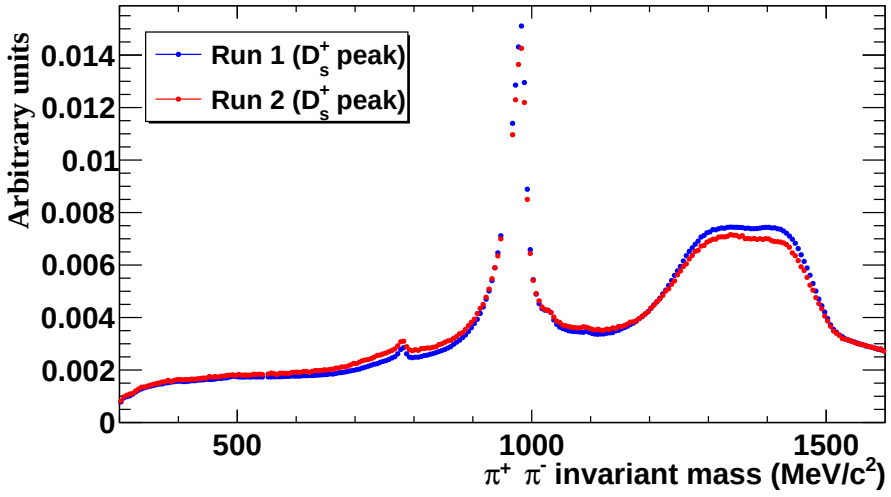


Figure 4.31: Distributions of $\pi^+\pi^-$ mass for Run 1 and Run 2, for the D_s^+ peak region.

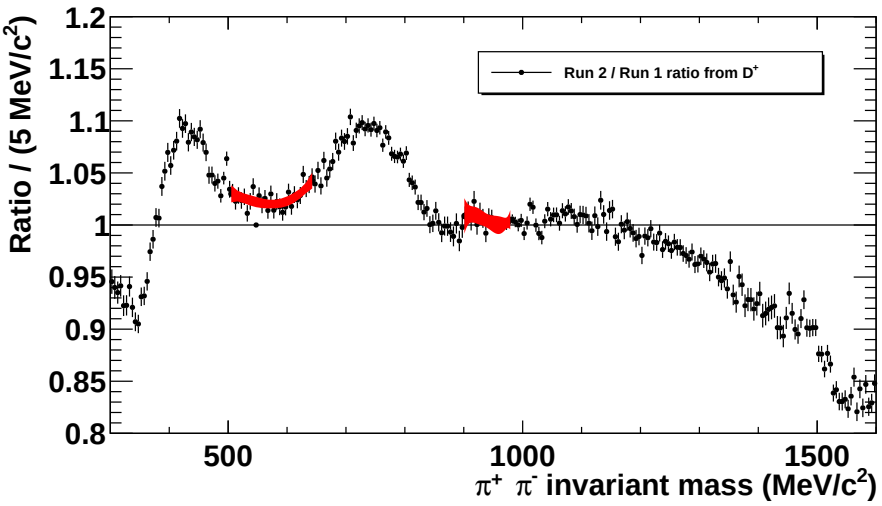


Figure 4.32: Ratio of $\pi^+\pi^-$ mass distributions for Run 2 with respect to Run 1, for the D^+ peak region.

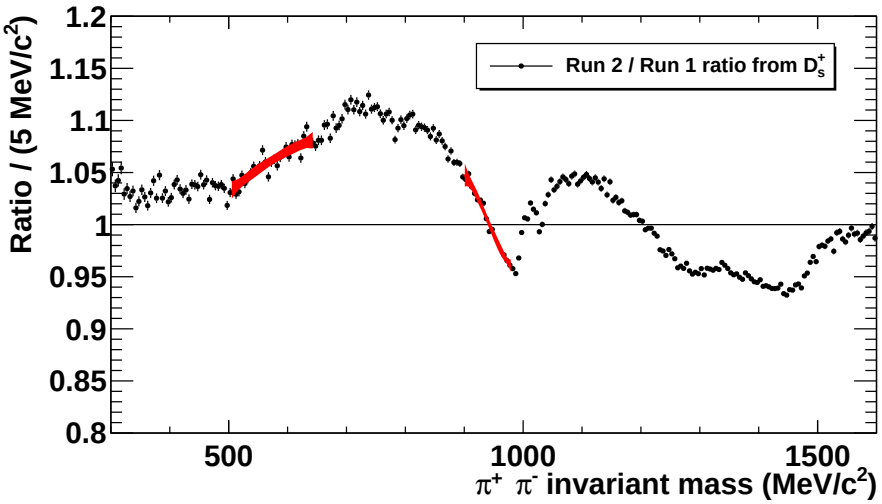


Figure 4.33: Ratio of $\pi^+\pi^-$ mass distributions for Run 2 with respect to and Run 1, for the D_s^+ peak region.

4.8 Expected limits

The expected limits on $\mathcal{B}(\eta^{(\prime)} \to \pi^+\pi^-)$ are calculated with the CL_s method, using fits to the $\pi^+\pi^-$ mass spectra for several values of the signal PDF yield. In order to do so, it is necessary to obtain the signal and background PDFs as functions of the $\pi^+\pi^-$ invariant mass.

4.8.1 Signal PDFs and mass resolution

For the signal PDFs a dedicated $D^+ \to \eta^{(\prime)}\pi^+$ Monte Carlo dataset is used, and the $\eta^{(\prime)}$ peaks in the $\pi^+\pi^-$ mass spectra are each fitted with a sum of two Gaussian functions. Figures 4.34 and 4.35 show the fitted η and η' peaks from these samples and Table 4.8 summarises the fit results. The effective resolution, obtained with a weighted sum of the two Gaussian widths, is 2.3 MeV/c^2 for the η mass region and 2.7 MeV/c^2 for the η'.

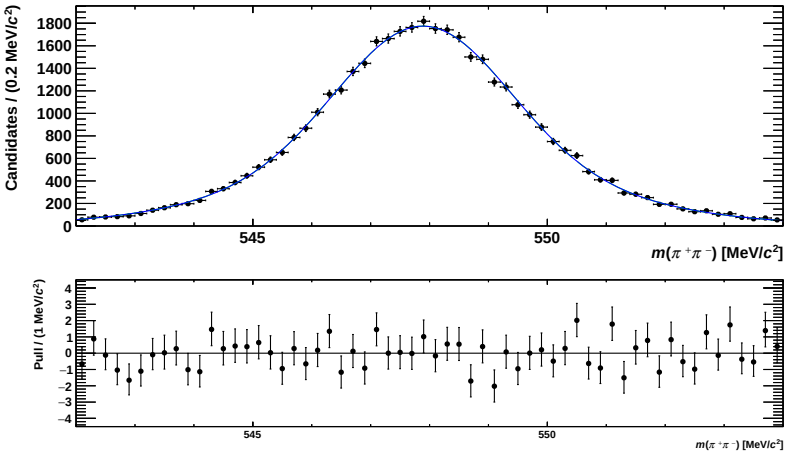


Figure 4.34: The η peak in the $\pi^+\pi^-$ invariant mass spectrum from $D^+ \to \eta\pi^+$ MC.

Whilst a difference between data and MC in the fitted central value for the $\eta^{(\prime)}$ mass would have a negligible effect on the limit, a systematic uncertainty to take into account small differences in mass resolution is estimated by comparing the fit to the η mass peak in the Monte Carlo, without the DTF requirement, with a fit to the $K^0_S \to \pi^+\pi^-$ peak in data using a sum of two Gaussian PDFs, shown in Figure 4.36 and Figure 4.37 for Run 1 and Run 2, respectively. The widths of the two Gaussians for the η are $\sigma_1(\eta) = 6.6\,\text{MeV}/c^2$ and $\sigma_2(\eta) = 3.4\,\text{MeV}/c^2$, with a yield ratio of 40%, to be compared with $\sigma_1(K^0) = 6.9\,\text{MeV}/c^2$ and $\sigma_2(K^0) = 3.8\,\text{MeV}/c^2$ and a yield ratio of 44% for Run 1 and $\sigma_1(K^0) =$

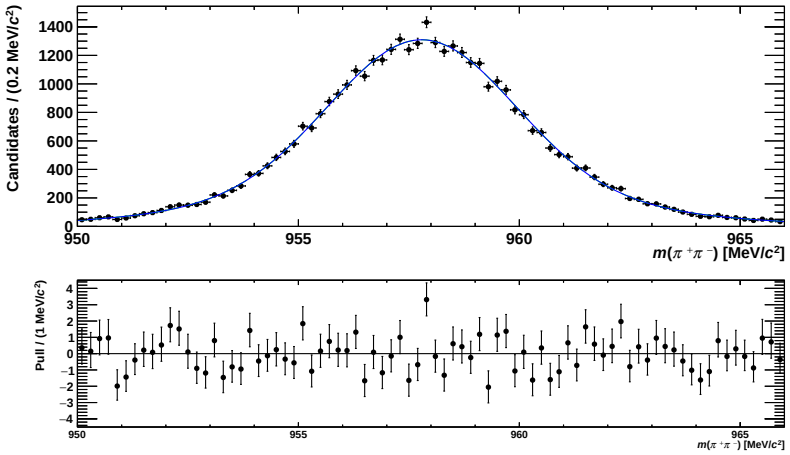


Figure 4.35: The η' peak in the $\pi^+\pi^-$ invariant mass spectrum from $D^+ \to \eta'\pi^+$ MC.

Table 4.8: Fitted parameters corresponding to Figs. 4.34 and 4.35. The peak values μ and standard deviations σ of the Gaussian functions and the effective resolution σ_{eff} are given in units of MeV/c^2. The parameter f_1 gives the relative contribution of the wider Gaussian.

Parameter	**Value ($\eta \to \pi^+\pi^-$)**	**Value ($\eta' \to \pi^+\pi^-$)**
μ	547.89 ± 0.01	957.79 ± 0.01
σ_1	2.8 ± 0.1	4.2 ± 0.2
σ_2	1.44 ± 0.04	2.08 ± 0.05
f_1	$(57 \pm 1)\%$	$(71 \pm 1)\%$
σ_{eff}	2.2 ± 0.1	3.6 ± 0.2

8.2 MeV/c^2 and $\sigma_2(K^0) = 4.1$ MeV/c^2 and a yield ratio of 25% for Run 2. The total widths are defined as half the range of the signal PDF, symmetric about the peak, which contains 68.27% of its integral. They are found to be $\sigma(\eta) = 5.25 \pm 0.05$ MeV/c^2, $\sigma(K^0) = 5.5 \pm 0.6$ MeV/c^2 for Run 1 and $\sigma(K^0) = 5.9 \pm 0.6$ MeV/c^2 for Run 2; therefore, a 5% (10%) systematic uncertainty on the mass resolution for Run 1 (Run 2) is included in the limit calculation. Figure 4.38 shows the distribution of the η mass resolution, after the application of the DTF, as a function of $\chi^2_{\text{IP}}(D^+)$. The red horizontal dotted line represents the value of the mass resolution obtained from the overall fit, as given in Table 4.8, while the solid lines define the $\pm 5\%$ interval, taken as systematic uncertainty. No significant deviation is observed and it is assumed that the η and η' mass resolutions do not depend on the fractions of prompt and non-prompt $D^+_{(s)}$

mesons in the sample.

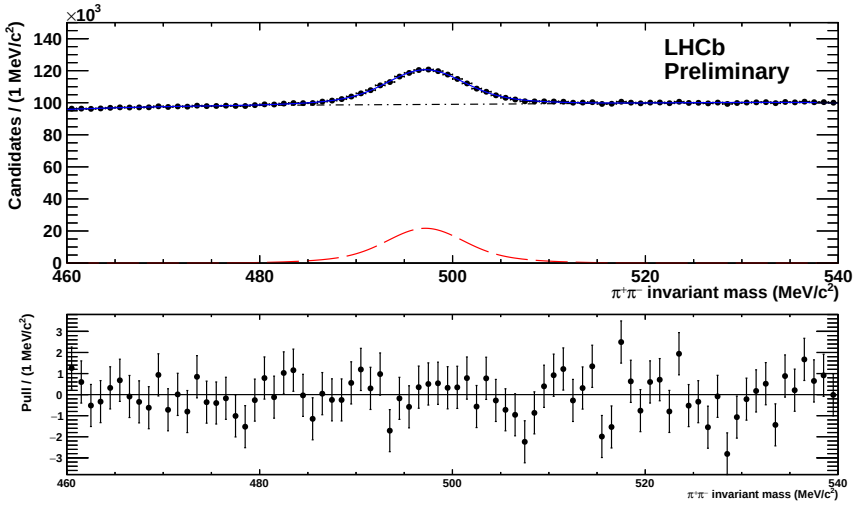


Figure 4.36: Fit to the K^0 peak in the $\pi^+\pi^-$ spectrum in data from Run 1.

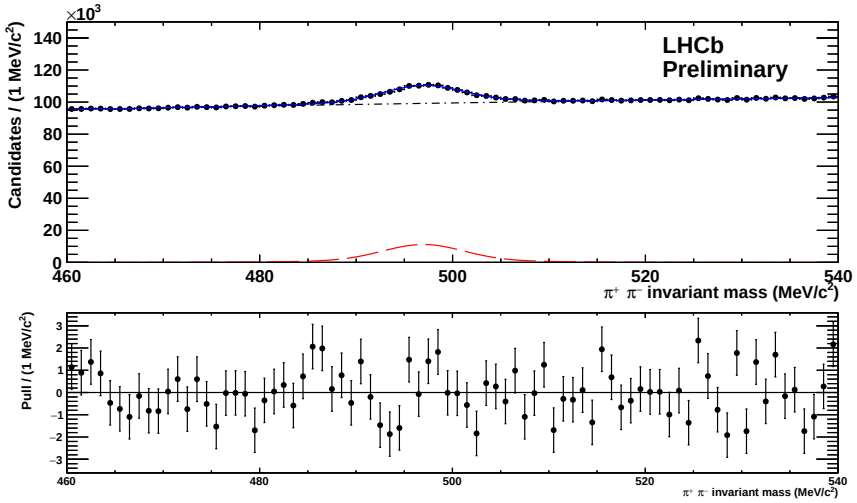


Figure 4.37: Fit to the K^0 peak in the $\pi^+\pi^-$ spectrum in data from Run 2.

4.8.2 Background PDFs

The background PDFs, i.e. the $\pi^+\pi^-$ invariant mass spectra under the hypothesis of no $\eta^{(\prime)} \to \pi^+\pi^-$ contribution, are taken from fits to data, over the ranges $m \in [515, 630]$ MeV/c^2 and $m \in [920, 980]$ MeV/c^2, with the the η and η' peak regions blinded. The fitting ranges have been chosen to avoid the peaks from the K^0_s, $\rho(770)$ and $f_0(980)$ present in the $\pi^+\pi^-$ spectrum. The intervals for the blinded regions are $m \in [544, 552]$ MeV/c^2 for the η and $m \in [952, 964]$ MeV/c^2

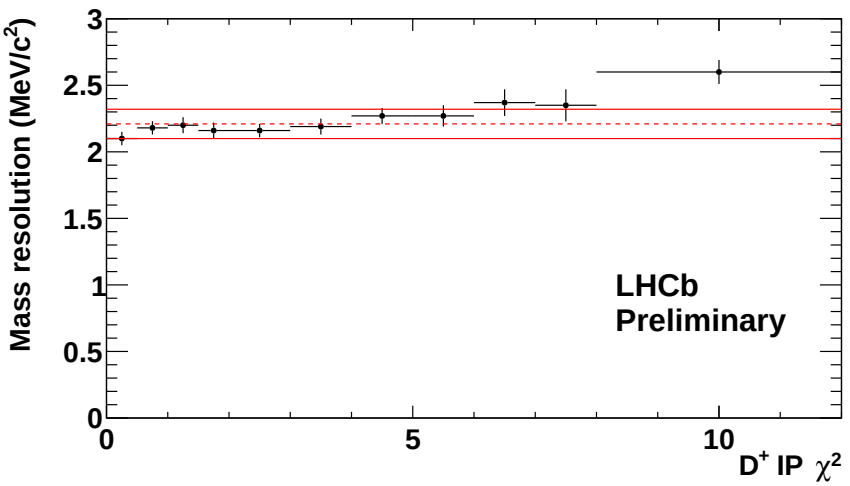


Figure 4.38: η mass resolution as a function of $\chi^2_{\rm IP}(D^+)$.

for the η', corresponding to approximately ± 2 times the mass resolution. Separate background PDFs are created for the D^+ and D_s^+ signal regions.

The distributions are fitted with the sum of a third-order Chebyshev polynomial. The impact of the decays $D^+_{(s)} \to \eta^{(\prime)}\pi^+$, $\eta^{(\prime)} \to \pi^+\pi^-\gamma$ is found to be negligible and completely absorbed by the polynomial background; in fact, simulation studies of $\eta^{(\prime)} \to \pi^+\pi^-\gamma$, using the full matrix element given in Ref. [120], showed that there is no peaking in these contributions within the fit ranges. Indeed, there are no contributions close to the signal regions, as is further verified for the η region by Refs. [121] and [122]. The $\eta' \to \pi^+\pi^-\gamma$ simulated distribution is shown in Figure 4.39.

Figures 4.40, 4.41, 4.42 and 4.43 show the distributions of the $\pi^+\pi^-$ mass spectra in Run 1 and Run 2 data, fitted with the aforementioned third-order Chebyshev polynomial. The resulting background PDFs are also used to generate toy Monte Carlo events over the entire mass ranges, including the η and η' regions, to give simulated expected mass spectra in the no $\eta^{(\prime)} \to \pi^+\pi^-$ hypothesis. These spectra are then used to obtain expected limits on the $\eta \to \pi^+\pi^-$ and $\eta' \to \pi^+\pi^-$ branching fractions, as described in Section 4.8.5.

4.8.3 Systematic uncertainties

The following sources of systematic uncertainty are included in the limit setting:

- Uncertainties on the $D^+_{(s)} \to \pi^+\pi^+\pi^-$ and $D^+_{(s)} \to \pi^+\eta^{(\prime)}$ branching fractions (from the Particle Data Group);

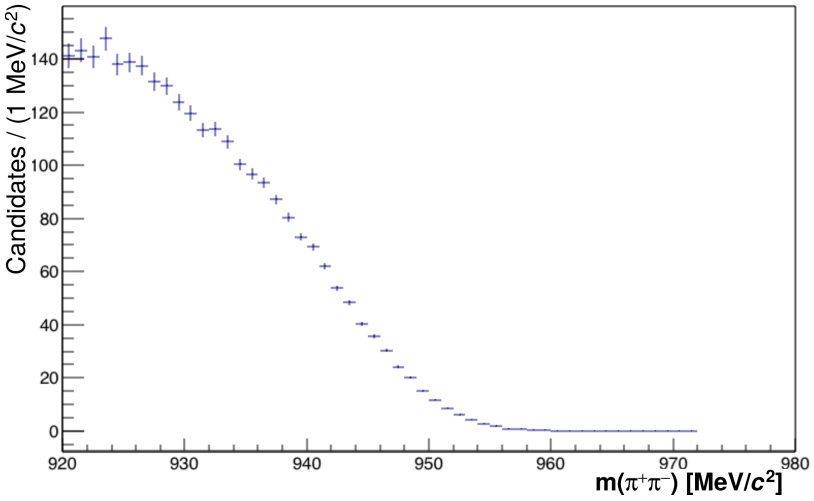


Figure 4.39: Dipion spectrum from $\eta' \to \pi^+\pi^-\gamma$, simulated with a full decay matrix element.

- Uncertainties on the fitted numbers of D^+ and D_s^+ mesons (from the fluctuations of the fit residuals, see Figures 4.23 and 4.24);
- Uncertainty on the relative efficiency variation with $\pi^+\pi^-$ mass (see Section 4.7);
- Uncertainties in the background PDF parameters (from the fits to the $\pi^+\pi^-$ mass spectra, see Section 4.8.2);
- Uncertainty in the η and η' mass resolution (see Section 4.8.1).

Table 4.9 gives the percentage values of each of the systematic uncertainties.

The first five sources are implemented as overall systematic errors in the scaling factors when calculating the CL_s (see Sec. 4.8). The contribution for the background uncertainties is implemented as two histograms, with the upper and lower distribution of the 68% CL error band given by the errors on the fitted parameters of the background PDF. The contribution for the mass resolution is implemented in an analogous way: a $\pm 5\%$ variation for Run 1 and a $\pm 10\%$ variation for Run 2 are applied to the signal PDF width to obtain the upper and lower distribution histograms.

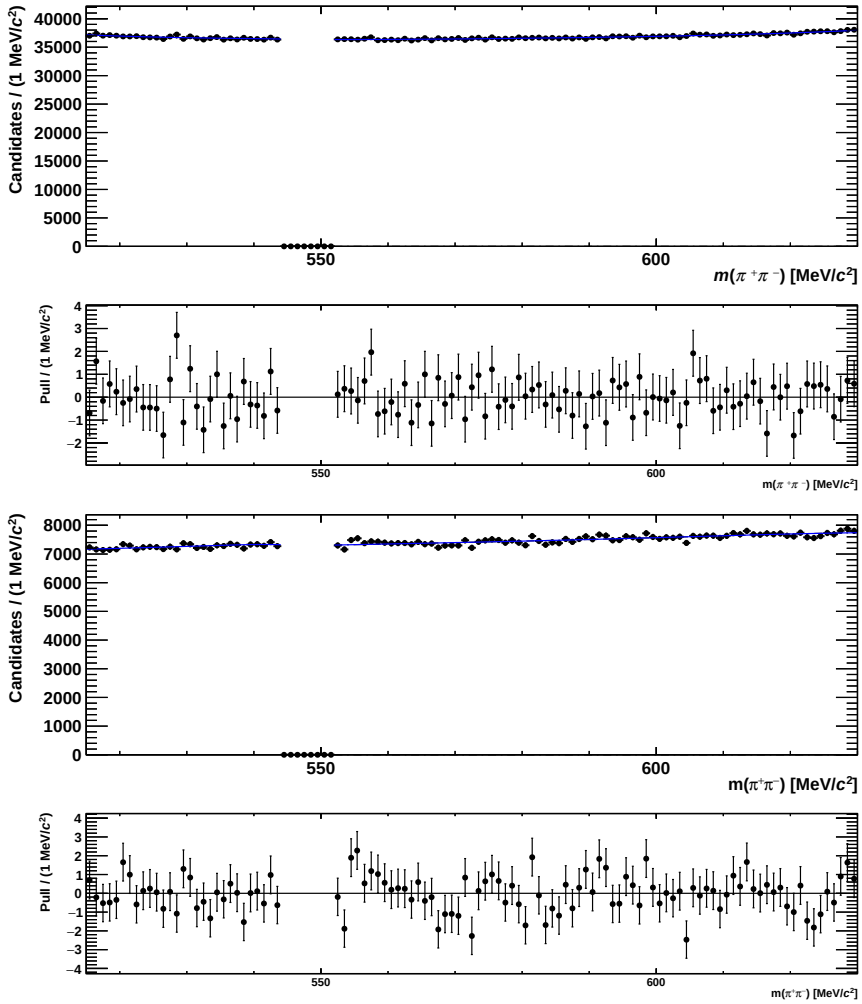


Figure 4.40: $\pi^+\pi^-$ invariant mass spectra from Run 1 data, with blinded η regions for the D^+ (top) and D_s^+ (bottom) peak regions.

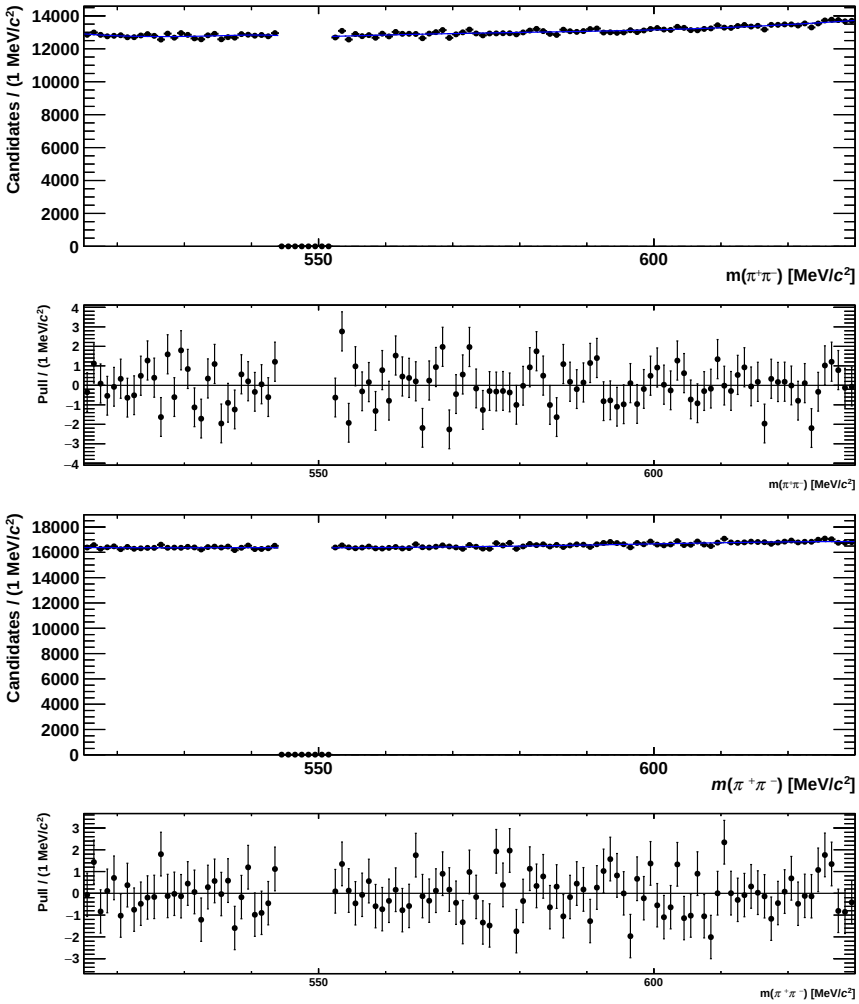


Figure 4.41: $\pi^+\pi^-$ invariant mass spectra from Run 2 data, with blinded η regions for the D^+ (top) and D_s^+ (bottom) peak regions.

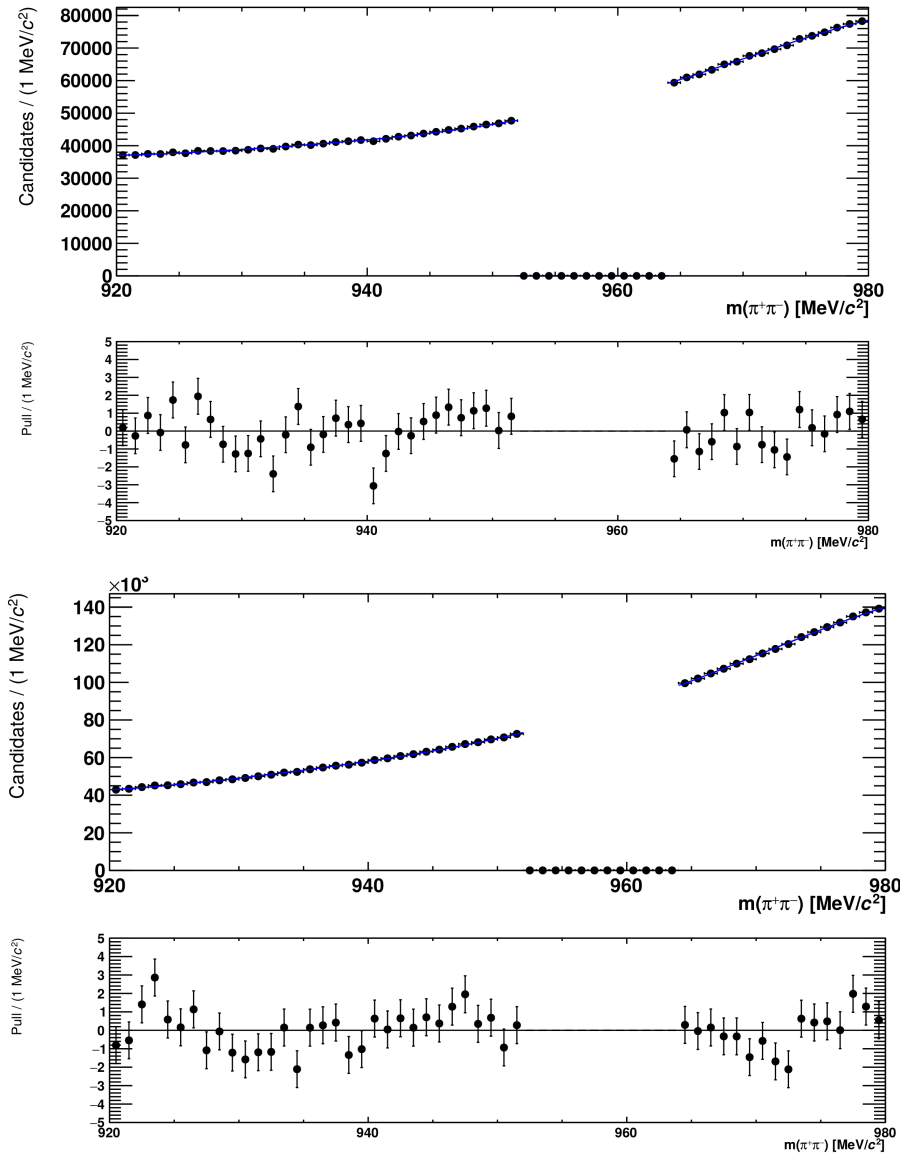


Figure 4.42: $\pi^+\pi^-$ invariant mass spectra from Run 1 data, with blinded η' regions for the D^+ (top) and D_s^+ (bottom) peak regions.

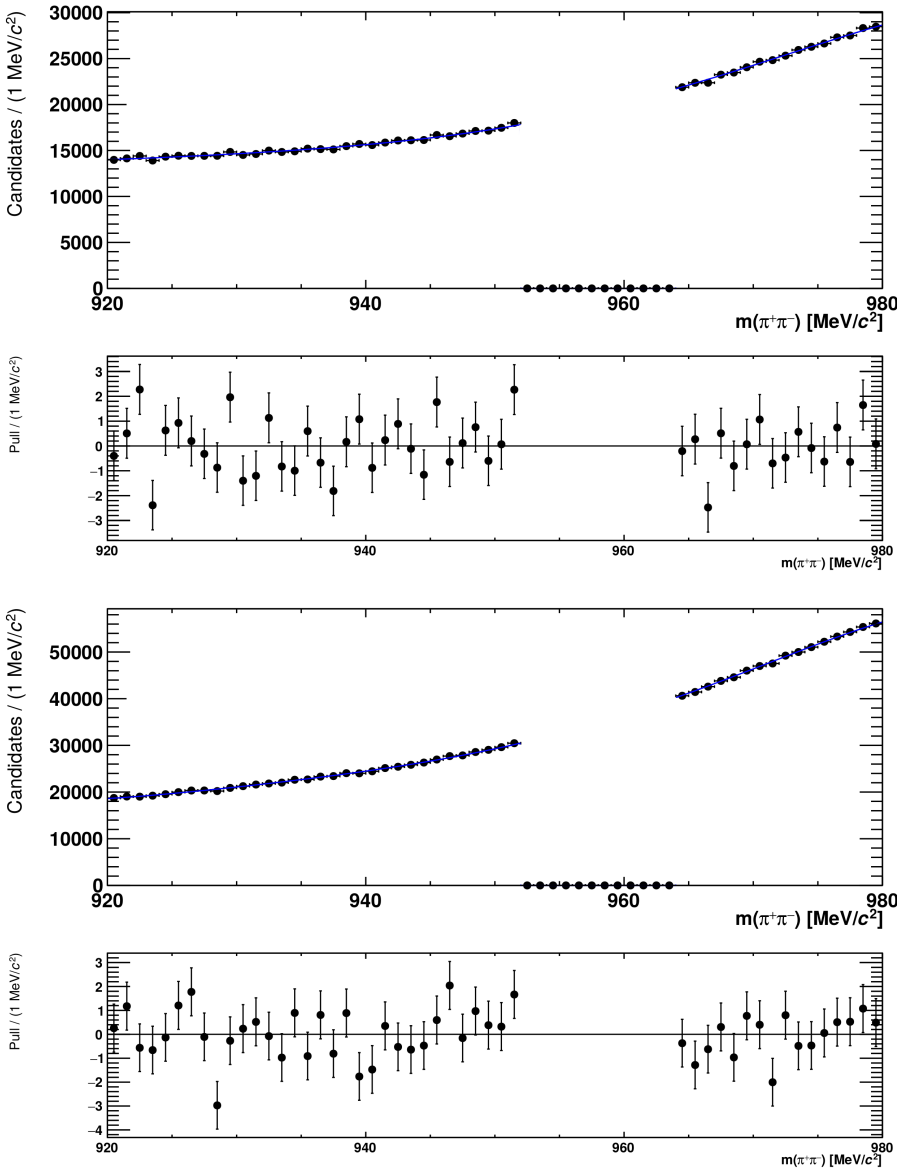


Figure 4.43: $\pi^+\pi^-$ invariant mass spectra from Run 2 data, with blinded η' regions for the D^+ (top) and D_s^+ (bottom) peak regions.

Table 4.9: Systematic uncertainties included in the limit calculation.

Source	**Systematic uncertainty**			
	$D^+ \to \eta\pi^+$	$D_s^+ \to \eta\pi^+$	$D^+ \to \eta'\pi^+$	$D_s^+ \to \eta'\pi^+$
BRs	8.5%	7.0%	8.8%	7.8%
Number of $D_{(s)}^+$ (Run 1)	2%	2%	2%	2%
Number of $D_{(s)}^+$ (Run 2)	1%	1%	1%	1%
Efficiency variation (Run 1)	1%	1%	1%	1%
Efficiency variation (Run 2)	1%	1%	1%	1%
Background fit (Run 1)	0.2%	0.2%	0.4%	0.3%
Background fit (Run 2)	0.3%	0.5%	0.4%	0.4%
Mass resolution (Run 1)	5%	5%	5%	5%
Mass resolution (Run 2)	10%	10%	10%	10%

4.8.4 Summary of terms in the branching fraction calculation

Table 4.10 gives the values of all of the terms entering into the calculation of $\mathcal{B}(\eta^{(\prime)} \to \pi^+\pi^-)$ together with their total uncertainties. The statistical uncertainties on $N(D^+ \to \pi^+\pi^+\pi^-)$ and $N(D_s^+ \to \pi^+\pi^+\pi^-)$ are found to be negligible in comparison to the systematic uncertainties.

Table 4.10: Values of terms used in the limit calculation and their uncertainties.

$\epsilon_{D^+}(\eta)$ (Run 1)	0.85 ± 0.01
$\epsilon_{D^+}(\eta)$ (Run 2)	0.87 ± 0.01
$\epsilon_{D_s^+}(\eta)$ (Run 1)	0.80 ± 0.01
$\epsilon_{D_s^+}(\eta)$ (Run 2)	0.84 ± 0.01
$\epsilon_{D^+}(\eta')$ (Run 1)	1.01 ± 0.01
$\epsilon_{D^+}(\eta')$ (Run 2)	1.02 ± 0.01
$\epsilon_{D_s^+}(\eta')$ (Run 1)	1.03 ± 0.01
$\epsilon_{D_s^+}(\eta')$ (Run 2)	1.00 ± 0.01
$N(D^+ \to \pi^+\pi^+\pi^-)$ (Run 1)	$(1.88 \pm 0.04) \times 10^7$
$N(D_s^+ \to \pi^+\pi^+\pi^-)$ (Run 1)	$(1.75 \pm 0.03) \times 10^7$
$N(D^+ \to \pi^+\pi^+\pi^-)$ (Run 2)	$(6.09 \pm 0.06) \times 10^6$
$N(D_s^+ \to \pi^+\pi^+\pi^-)$ (Run 2)	$(6.26 \pm 0.06) \times 10^6$
$\mathcal{B}(D^+ \to \pi^+\pi^+\pi^-)$	$(3.29 \pm 0.20) \times 10^{-3}$
$\mathcal{B}(D_s^+ \to \pi^+\pi^+\pi^-)$	$(1.09 \pm 0.05) \times 10^{-2}$
$\mathcal{B}(D^+ \to \eta\pi^+)$	$(3.66 \pm 0.22) \times 10^{-3}$
$\mathcal{B}(D_s^+ \to \eta\pi^+)$	$(1.70 \pm 0.09) \times 10^{-2}$
$\mathcal{B}(D^+ \to \eta'\pi^+)$	$(4.84 \pm 0.31) \times 10^{-3}$
$\mathcal{B}(D_s^+ \to \eta'\pi^+)$	$(3.94 \pm 0.25) \times 10^{-2}$

4.8.5 Expected limits

Expected limits are obtained using the CL_s method, using the RooStats framework for advanced statistical analysis, built on the ROOFIT toolkit. In the method, CL_s values are calculated using likelihoods evaluated for both background-only and signal+background hypotheses, scanning a range of branching fraction values. The η and η' blinded regions are populated with simulated data, obtained from the extrapolation of the polynomial fit to the $\pi^+\pi^-$ mass spectra. The total likelihoods are evaluated using the combined information from the four distributions $(D, D_s)\times$ (Run 1, Run 2). In Figs. 4.44 and 4.45 the CL_s distributions for η and η' are shown, along with the one and two sigma error bands.

The red vertical lines in both figures represent the current world best limit for the branching fractions at 90% CL, which are $\mathcal{B}(\eta \to \pi^+\pi^-) < 1.3 \times 10^{-5}$ and $\mathcal{B}(\eta' \to \pi^+\pi^-) < 5.5 \times 10^{-5}$. Expected limits at 90% CL are then obtained by calculating the limits on the signal branching fractions that correspond to $1 - CL_s = 0.9$. The expected limits are

$$\mathcal{B}(\eta \to \pi^+\pi^-) < 2.0 \times 10^{-5},$$
$$\mathcal{B}(\eta' \to \pi^+\pi^-) < 1.7 \times 10^{-5}.$$

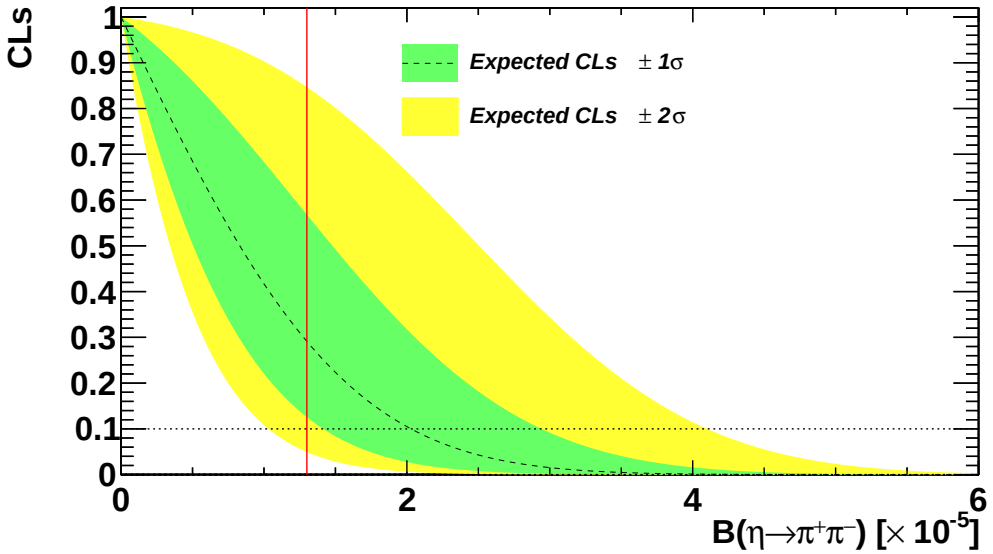


Figure 4.44: CL_s as a function of $\mathcal{B}(\eta \to \pi^+\pi^-)$. The horizontal dotted line defines the 90% CL.

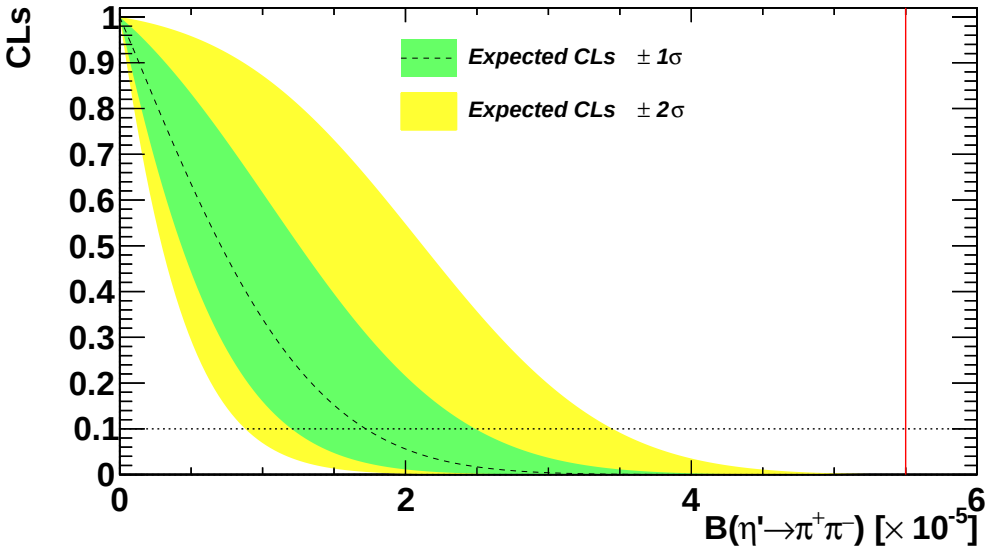


Figure 4.45: CL_s as a function of $\mathcal{B}(\eta' \to \pi^+\pi^-)$. The horizontal dotted line defines the 90% CL.

4.9 Results

After having estimated expected limits as described in the previous section, the toy MC events used to fill the blinded η and η' regions are replaced with real data. Before calculating the observed limits with the CL_s method, fits to the $\pi^+\pi^-$ mass spectra are performed. The fourth-order polynomial is used as background PDF, but allowing its parameters to vary. The signal PDF is composed of the shapes extracted from MC, described in Section 4.8.1. This procedure gives good fits to the data in all cases, and they are shown in Figures 4.46, 4.47, 4.49 and 4.50. For the η' fits, the plots shown represent the difference between data and the fitted background, in order to better show the signal contribution. Values of χ^2/ndf are all close to 1, and yields are all consistent with zero, with the largest of the eight deviations from zero being just over 2σ. They are summarised in Table 4.11. Figures 4.48 and 4.51 show the summed $\pi^+\pi^-$ mass spectra in the η and the η' regions after unblinding, with the sums of the four fits superimposed. The weighted average of the fitted branching fractions is $-0.9 \pm 1.8 \times 10^{-5}$ for $\eta \to \pi^+\pi^-$ and $0.8 \pm 1.6 \times 10^{-5}$ for $\eta' \to \pi^+\pi^-$.

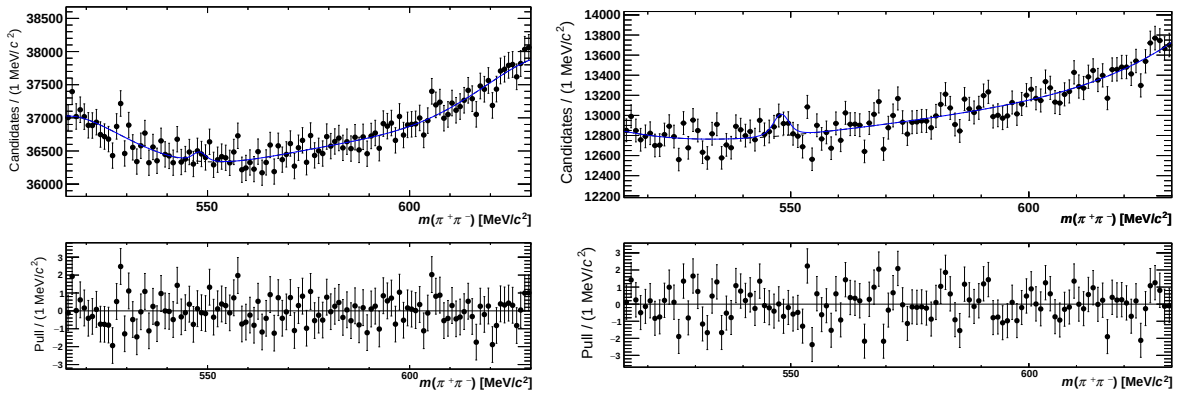


Figure 4.46: Fit to the unblinded $\pi^+\pi^-$ mass spectrum in the η range from $D^+ \to \pi^+\pi^+\pi^-$ for Run 1 (left) and Run 2 (right).

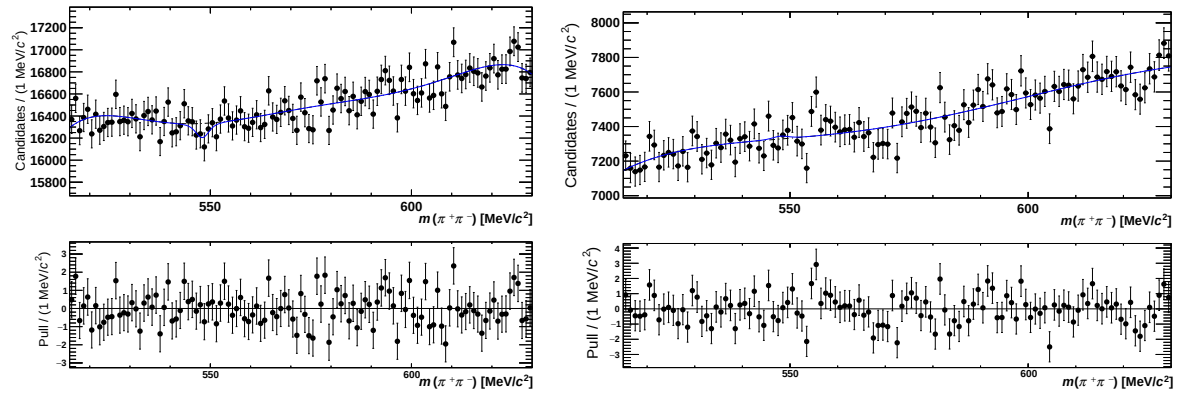


Figure 4.47: Fit to the unblinded $\pi^+\pi^-$ mass spectrum in the η range from $D_s^+ \to \pi^+\pi^+\pi^-$ for Run 1 (left) and Run 2 (right).

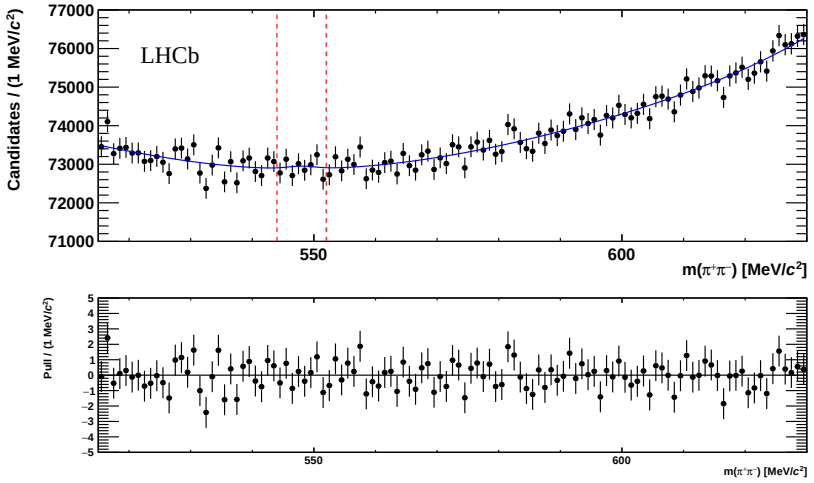


Figure 4.48: The unblinded $\pi^+\pi^-$ invariant mass distribution in the η mass region from the sum of the four contributions, showing also the sum of the fitted curves. The red vertical lines define the region that was initially blinded.

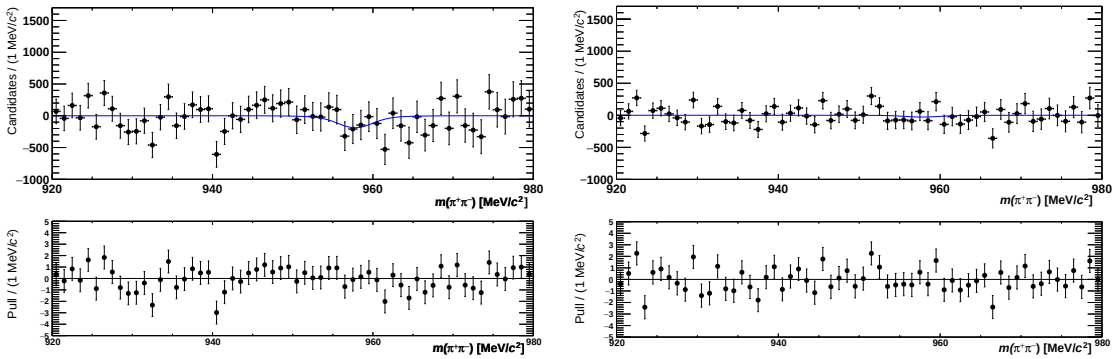


Figure 4.49: Fit to the unblinded $\pi^+\pi^-$ mass spectrum in the η' range from $D^+ \to \pi^+\pi^+\pi^-$ for Run 1 (left) and Run 2 (right). For clarity, the difference between the data and the fitted background is shown.

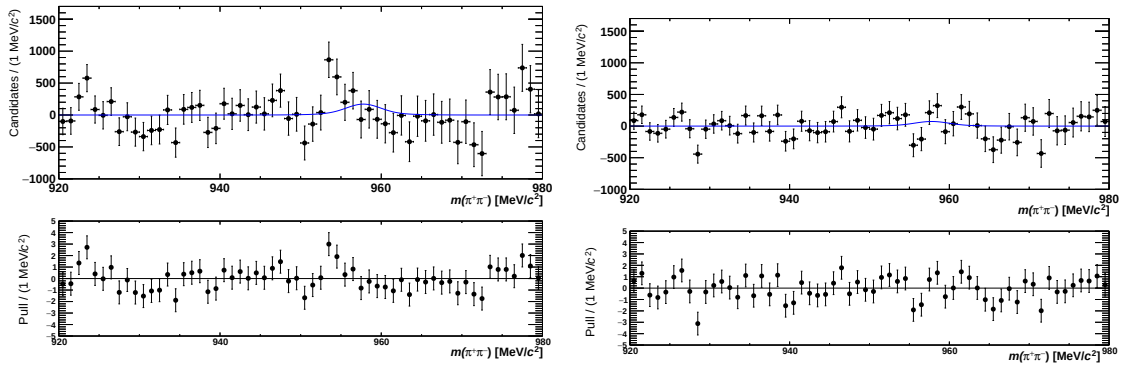


Figure 4.50: Fit to the unblinded $\pi^+\pi^-$ mass spectrum in the η' range from $D_s^+ \to \pi^+\pi^+\pi^-$ for Run 1 (left) and Run 2 (right). For clarity, the difference between the data and the fitted background is shown.

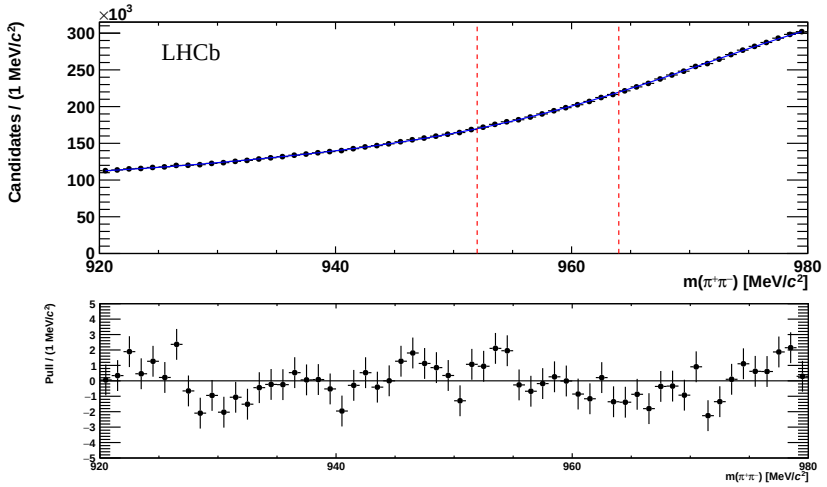


Figure 4.51: The unblinded $\pi^+\pi^-$ invariant mass distribution in the η' mass region from the sum of the four contributions, showing also the sum of the fitted curves. The red vertical lines define the region that was initially blinded.

Table 4.11: χ^2 of the fits to the unblinded datasets and fitted number of signal $\eta \to \pi^+\pi^-$ candidates (top) and $\eta' \to \pi^+\pi^-$ candidates (bottom).

Channel	**Signal yield**	χ^2 **/ ndof**
$D^+ \to \pi^+\pi^+\pi^-$ (Run 1)	410 ± 770	82/106
$D^+ \to \pi^+\pi^+\pi^-$ (Run 2)	970 ± 460	94/106
$D_s^+ \to \pi^+\pi^+\pi^-$ (Run 1)	-1020 ± 510	92/106
$D_s^+ \to \pi^+\pi^+\pi^-$ (Run 2)	150 ± 340	89/106
$D^+ \to \pi^+\pi^+\pi^-$ (Run 1)	-1070 ± 1040	57/51
$D^+ \to \pi^+\pi^+\pi^-$ (Run 2)	-130 ± 660	58/51
$D_s^+ \to \pi^+\pi^+\pi^-$ (Run 1)	1200 ± 1290	62/51
$D_s^+ \to \pi^+\pi^+\pi^-$ (Run 2)	580 ± 900	57/51

For the background-only hypothesis for the CL_s method, the unblinded distributions are fitted with a fourth-order Chebyshev polynomial, without any signal component. The expected limits at 90% CL are

$$\mathcal{B}(\eta \to \pi^+\pi^-) < 2.0 \times 10^{-5},$$
$$\mathcal{B}(\eta' \to \pi^+\pi^-) < 1.8 \times 10^{-5}.$$

The observed limits at 90% CL are:

$$\mathcal{B}(\eta \to \pi^+\pi^-) < 1.6 \times 10^{-5},$$
$$\mathcal{B}(\eta' \to \pi^+\pi^-) < 1.8 \times 10^{-5}.$$

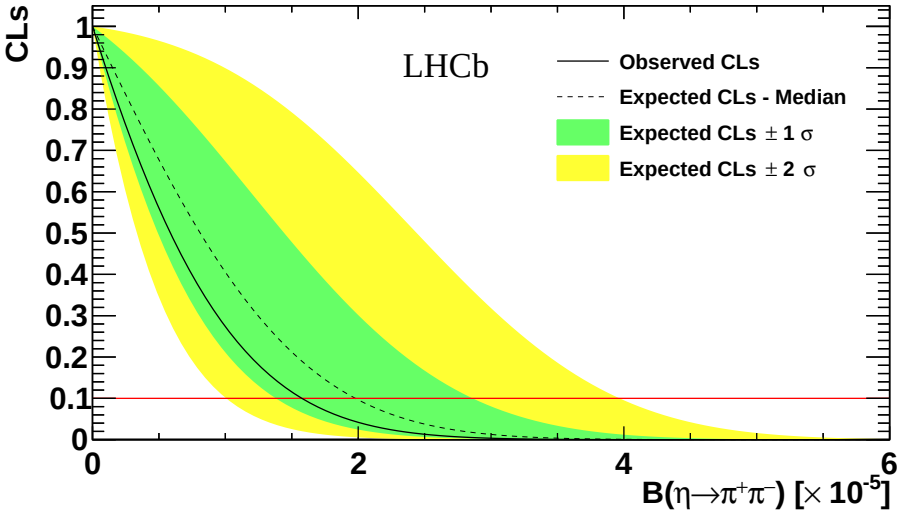


Figure 4.52: CL_s distribution for $\mathcal{B}(\eta \to \pi^+\pi^-)$

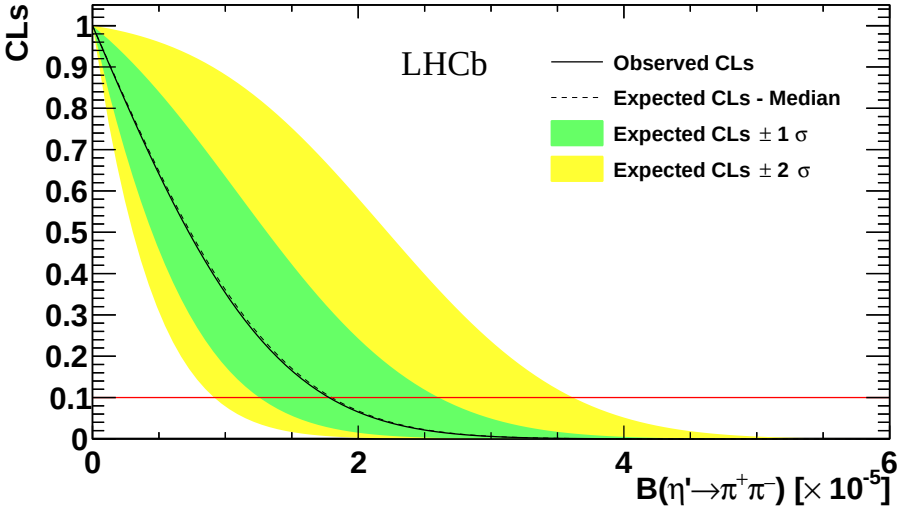


Figure 4.53: CL_s distribution for $\mathcal{B}(\eta' \to \pi^+\pi^-)$

4.10 Conclusions

In this analysis, a search for strong *CP* violation in the decays $\eta \to \pi^+\pi^-$ and $\eta' \to \pi^+\pi^-$, where the $\eta^{(\prime)}$ originates from a D^+ or D_s^+ meson, has been performed, and a new method is introduced, which relies on the large production rate of charm mesons at LHCb. As the limits are completely dominated by the statistical uncertainty, the sensitivity of this method will improve as more data are collected.

From the analysis of Run 1 data and data from the first year of Run 2 (2015), no signals are seen and upper limits on branching fractions are set, using the $\mathrm{CL_s}$ method. The limits at 90% CL on the BRs of the two decays are

$$\mathcal{B}(\eta \to \pi^+\pi^-) < 1.6 \times 10^{-5},$$
$$\mathcal{B}(\eta' \to \pi^+\pi^-) < 1.8 \times 10^{-5}.$$

The observed limits are compatible with the expected limits. The observed limit on $\mathcal{B}(\eta \to \pi^+\pi^-)$ is close to the world best limit, while that on $\mathcal{B}(\eta' \to \pi^+\pi^-)$ is more than three times more stringent than the previous world best.

5 Lepton universality violation in semileptonic charm decays

As already mentioned in Sections 1.2.2 and 2.2, in the Standard Model the only difference between the interactions of leptons with other forms of matter or antimatter is given by their different masses, as the coupling constants must be the same to preserve the gauge invariance of the electroweak Lagrangian. Recently, results from several experiments have pointed towards the possibility of non-universal leptonic interactions, in particular a difference between τ leptons and the lighter lepton flavours from semileptonic B decays. If confirmed, this may mean that the description of the Standard Model as a Yang-Mills gauge theory is not correct, or that hidden phenomena, such as leptoquarks or new gauge bosons, contribute to the amplitudes of such decays.

Measurements of BFs of other semileptonic decays might in principle show a similar disagreement with the predicted values. In particular, no dedicated search for lepton non-universality has ever been performed in semileptonic D decays. This analysis focuses on $D^0 \to K^- l \nu_l$ decays, where $l = e, \mu$; the individual channels have been already measured by many experiments [47] but the ratio has never been measured directly. In the large LHCb Run 2 dataset, the statistics alone are enough to reduce the error on the measurement of the ratio of $D^0 \to K^- e^+ \nu_e$ to $D^0 \to K^- \mu^+ \nu_\mu$ BFs by an order of magnitude, assuming some million signal events for both channels, which is confirmed by preliminary studies.

This chapter presents sensitivity studies for the measurement of the ratio $R^{\mu/e} = \frac{\mathcal{B}(D^0 \to K^- \mu^+ \nu_\mu)}{\mathcal{B}(D^0 \to K^- e^+ \nu_e)}$ using the prompt decay $D^{*+} \to D^0 \pi^+$. Unless explicitly stated, charge conjugate decays are implied.

5.1 Effective Lagrangian for $c \to s l \nu_l$ transitions

Assuming that the mass of any new particle contributing to the calculation of the BFs of the $D^0 \to K^- l^+ \nu_l$ decays is significantly larger than the typical hadronic energy scale, so that they can be integrated out, a four-fermion point interaction to describe $c \to s l \nu_l$ transitions can be defined with an operator product expansion as

$$\mathcal{L}_{eff} = 4 G_F V_{cs} \sum_{l=e,\mu,\tau} \sum_k C_k^l \hat{O}_k^l + h.c., \tag{5.1}$$

where the C_k^l are the complex Wilson coefficients [123] of the operators $\hat{O}_k^l$. In a low energy approximation of the electroweak theory, the leading four-fermion operator is

$$\hat{O}_{SM}^l = \left(\bar{s}\gamma_\mu \hat{P}_L c\right)\left(\bar{\nu}_l \gamma_\mu \hat{P}_L l\right), \tag{5.2}$$

and one obtains the Fermi effective Lagrangian already introduced in Equation (1.11) by setting $C_{SM}^l = 1$.

Possible new physics contributions in semileptonic decays $D \to K l \nu_l$ can be described by non-standard scalar quark and lepton densities with left-handed neutrinos [124, 125], with the effective operators

$$\hat{O}_{L,R}^l = \left(\bar{s}\hat{P}_{L,R} c\right)\left(\bar{\nu}_l \hat{P}_R l\right), \tag{5.3}$$

and the decay rates are affected by the scalar Wilson coefficient defined as

$$C_S^l = C_L^l + C_R^l. \tag{5.4}$$

The constraints on the scalar Wilson coefficient for the muonic channel is shown in Figure 5.1. For the electronic channel, $\left|C_S^e\right| < 0.2$ at 95% CL; therefore, it is assumed to be $C_S^e = 0$ in Ref. [125], i.e. new physics contributions are allowed to affect only the second generation of leptons.

The sensitivity to possible new physics contributions does not only come from the ratio of the integrated branching fractions, but also from the ratio of the differential branching fractions as a function of q^2, the square of the four-momentum transfer to the leptons in the rest frame of the D meson. This is illustrated in Figure 5.2, where the red band is the SM prediction and the grey band represents the allowed $R^{\mu/e}$ values from the constraint on the scalar Wilson coefficient shown in Figure 5.1.

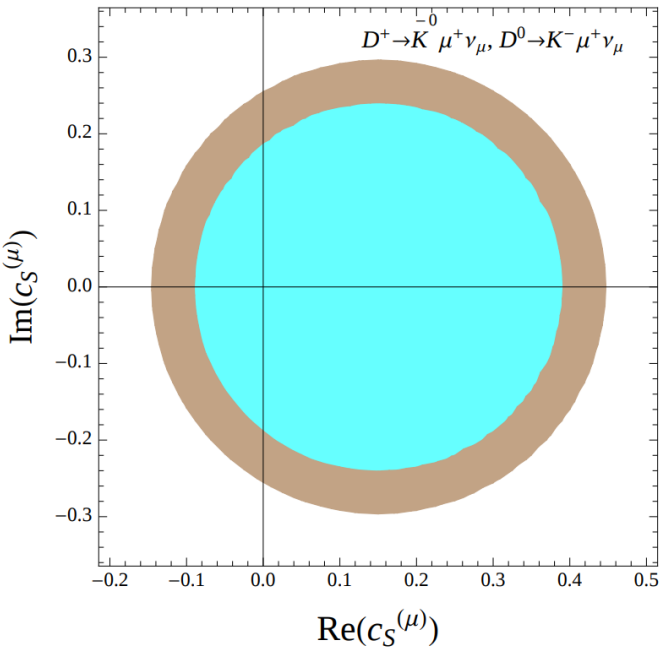


Figure 5.1: Allowed regions of the effective complex coupling C_S^μ from $D \to K\mu^+\nu_\mu$ decays. The 68% (95%) CL region is represented in brown (light blue) [125].

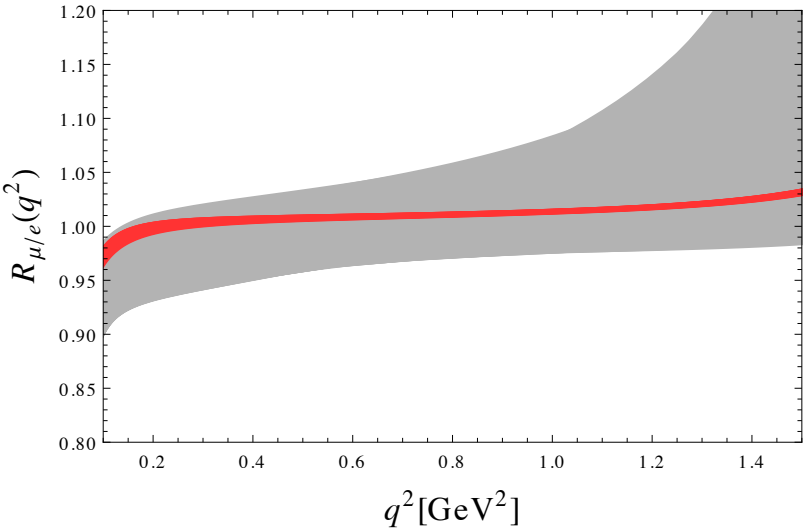


Figure 5.2: SM prediction (red) and currently allowed bands (gray) of the ratio of muon to electron $D^0 \to K^-\ell^+\nu_\ell$ decays, from the constraints to the non-standard scalar Wilson coefficient, as a function of q^2 [125].

5.2 Outline of the analysis method

Semileptonic charm decays of the type $D^0 \to K^- l^+ \nu_l$ are selected and analysed, from 2 fb^{-1} of Run 2 data collected by the LHCb experiment in 2015 and 2016. The objective of the analysis is to measure the branching fraction ratio of the muon channel over the electron channel, using the observed number of $D^0 \to K^-\mu^+\nu_\mu$ and $D^0 \to K^- e^+ \nu_e$ candidates coming from prompt $D^{*+} \to D^0\pi^+$ decays,

$$\begin{aligned} R^{\mu/e} &= \frac{\mathcal{B}(D^0 \to K^-\mu^+\nu_\mu)}{\mathcal{B}(D^0 \to K^- e^+ \nu_e)} \\ &= \frac{N(D^0 \to K^-\mu^+\nu_\mu)}{N(D^0 \to K^- e^+ \nu_e)} \times \frac{\epsilon_{tot}(e)}{\epsilon_{tot}(\mu)}, \end{aligned} \tag{5.5}$$

where $N(D^0 \to K^- l^+ \nu_l)$ represents the number of observed $D^0 \to K^- l^+ \nu_l$ events and ϵ_{tot} the total efficiencies.

The measurement is done in bins of q^2 to try to exploit the sensitivity to new physics shown in Figure 5.2. This is achieved by reconstructing the q^2 with the cone-closure method, described in Section 5.4. The efficiency terms include all the possible sources of signal loss coming from the event selection procedure,

$$\epsilon_{tot} = \epsilon_{acc} \cdot \epsilon_{rec|acc} \cdot \epsilon_{sel|rec} \cdot \epsilon_{PID|sel} \tag{5.6}$$

where ϵ_{acc} is the fraction of $D^{*+} \to D^0\pi^+$, $D^0 \to K^- l^+ \nu_l$ decays for which the final-state tracks are within the LHCb acceptance, with respect to the full solid angle; $\epsilon_{rec|acc}$ is the reconstruction efficiency, including single track reconstruction efficiency and the efficiency of forming the candidates from the decay, given the acceptance efficiency; $\epsilon_{sel|rec}$ is the selection efficiency, given all trigger, stripping and offline selection criteria; $\epsilon_{PID|sel}$ is the particle identification efficiency, given the selection efficiency. The only efficiencies taken into account in the calculation of (5.5) are those for the electrons and muons, as for the other final-state particles involved (the pion from the D^{*+} decay and the kaon) the efficiencies cancel in the ratio. Each of the individual selection efficiencies rely on the step before, and are calculated from MC, except for the PID efficiency

If there are no differences in the physics of the production of the electron versus muon modes, the acceptance efficiency should cancel explicitly. This is explicitly checked.

The numbers of $D^0 \to K^- l^+ \nu_l$ candidates are extracted from a bidimensional fit to ΔM_{vis}, the difference between the mass of the system of all the

visible particles and the mass of the K^-l^+ system,

$$\Delta M_{vis} = m(K^-l^+\pi_s) - m(K^-l^+), \tag{5.7}$$

where π_s is the slow pion coming from the decay of the D^{*+}, and M_{corr}, the corrected mass, defined as

$$M_{corr} = \sqrt{m(K^-l^+)^2 + (p'_T)^2} + p'_T, \tag{5.8}$$

where p'_T is the transverse momentum of the neutrino with respect to the direction of flight of the D^0, and illustrated in Figure 5.3. Effectively, M_{corr} corresponds to the D^0 mass, calculated from the decay products, where the longitudinal component of the momentum of the neutrino is ignored, as it cannot be measured in LHCb. By construction, $p'_T(\nu) = p'_T(K^-l^+)$.

Finally, it is noted that the typical momentum of the slow pion is approximately 3 GeV/c. Therefore, the momentum resolution of the typical slow pion is about 0.4%: this means that the momentum of such pions is very well measured.

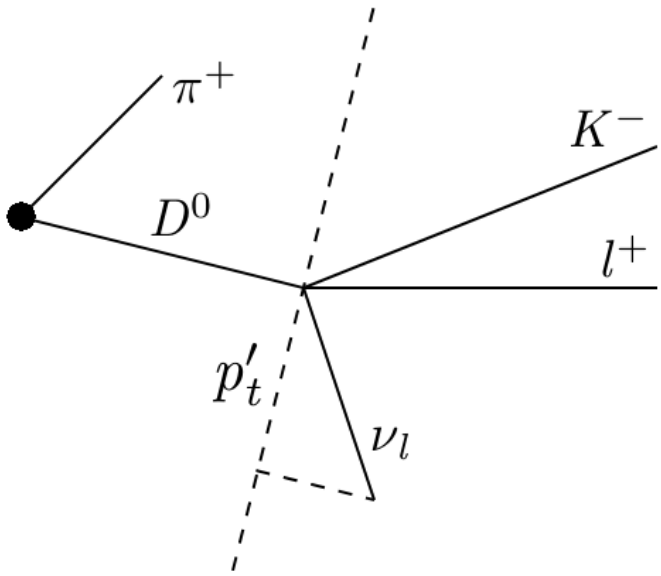


Figure 5.3: Definition of p'_T.

5.3 Datasets and event selection

5.3.1 Data and Monte Carlo simulation samples

The complete analysis will use data collected at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ in 2015, corresponding to 0.3 fb^{-1} of integrated luminosity, and at

$\sqrt{s} = 13$ TeV in 2016, corresponding to 1.7 fb^{-1} of integrated luminosity. For the studies presented here, a 0.3% subsample of 2016 data is used to determine the sensitivities of the analysis.

The Monte Carlo (MC) samples used in this analysis are divided into three groups and listed in Table 5.1. The first group contains unfiltered events (see Section 4.2.2), used for the determination of the reconstruction and selection efficiencies and for preliminary studies of sensitivity. The second group contains filtered events, to match the trigger and stripping requirements, and is used to build templates for the fit described in Section 5.5. All samples are generated using PYTHIA 8 with several decay models, depending on the specific final state and the presence or not of intermediate resonances. The third group contains toy MC samples, which are used for the sensitivity study described in Section 5.6. Apart from samples from the third group, all the others are produced with a full detector response simulation and with an approximately equal amount of both magnet polarities. In all samples, the D^0 comes from a $D^{*+} \to D^0\pi^+$ decay.

5.3.2 Stripping and offline selections

The set of stripping cuts for $D^0 \to K^-l^+\nu_l$ candidates is described in Table 5.2. All the pion, kaon, muon and electron candidates come from standard selections, according to some loose fiducial criteria. The large, asymmetric range of the cut on the reconstructed D^0 mass takes into account the low-mass tail of the distribution due to the missing neutrino and, in the electron case, due to non-recovered bremsstrahlung. As the processing rate of the stripping line is found to be too high, a prescaling factor of 0.1 is set, effectively processing only 10% of the recorded events.

5.4 q^2 reconstruction: the cone-closure method

The cone-closure method [126] is used to reconstruct the missing momentum of the neutrino. It relies on the fact that the D^0 meson comes from a $D^{*\pm}$ decay. From conservation of energy and momentum one gets

$$\begin{aligned} \vec{p}(D^0) &= \vec{p}(K) + \vec{p}(l) + \vec{p}(\nu_l), \\ E(D^0) &= E(K) + E(l) + E(\nu_l). \end{aligned} \tag{5.9}$$

Table 5.1: MC simulation samples used in the analysis. The MC production type and the approximate number of events generated are also given. When two numbers are given, the first refers to the unfiltered sample and the second to the filtered sample.

Channel	Production type	# events (2015)	# events (2016)
$D^0 \to K^- \mu^+ \nu_\mu$	Full	$0.5 \times 10^6 / 7 \times 10^6$	$3 \times 10^6 / 23 \times 10^6$
$D^0 \to K^- e^+ \nu_e$	Full	$0.5 \times 10^6 / 7 \times 10^6$	$3 \times 10^6 / 22 \times 10^6$
$D^0 \to K^{*-} \mu^+ \nu_\mu$	Full	$0.5 \times 10^6 / 3 \times 10^6$	$3 \times 10^6 / 8 \times 10^6$
$D^0 \to K^{*-} e^+ \nu_e$	Full	$0.5 \times 10^6 / 3 \times 10^6$	$3 \times 10^6 / 8 \times 10^6$
$D^0 \to K^- \mu^+ \nu_\mu \pi^0$	Full	$0.5 \times 10^6 / 3 \times 10^6$	$3 \times 10^6 / 8 \times 10^6$
$D^0 \to K^- e^+ \nu_e \pi^0$	Full	$0.5 \times 10^6 / 3 \times 10^6$	$3 \times 10^6 / 8 \times 10^6$
$D^0 \to K^- K^+$	Full	$0.5 \times 10^6 / 2 \times 10^6$	$3 \times 10^6 / 5 \times 10^6$
$D^0 \to K^- \pi^+ \pi^0$	Full	$0.5 \times 10^6 / 2 \times 10^6$	$3 \times 10^6 / 5 \times 10^6$
$D^0 \to K^- \pi^+ \pi^- \pi^+$	Full	$0.5 \times 10^6 / 3 \times 10^6$	$3 \times 10^6 / 5 \times 10^6$
$D^0 \to \pi^+ \pi^- \pi^0$	Full	$0.5 \times 10^6 / 3 \times 10^6$	$3 \times 10^6 / 7 \times 10^6$
$D^0 \to \pi^- \mu^+ \nu_\mu$	Full	$0.5 \times 10^6 / 3 \times 10^6$	$3 \times 10^6 / 7 \times 10^6$
$D^0 \to \pi^- e^+ \nu_e$	Full	$0.5 \times 10^6 / 3 \times 10^6$	$3 \times 10^6 / 5 \times 10^6$
$D^0 \to K^- \mu^+ \nu_\mu$	Toy	-	10^7
$D^0 \to K^- \mu^+ \nu_\mu \pi^0$	Toy	-	10^7
$D^0 \to K^{*-} \mu^+ \nu_\mu$	Toy	-	10^7
$D^0 \to K^- \pi^+$	Toy	-	10^7

Boosting the event to the Kl rest frame, the momentum of the neutrino and the momentum of the D^0 are equal,

$$\vec{p}(D^0) = \vec{p}(\nu_l). \tag{5.10}$$

Substituting Equation (5.10) into Equation (5.9), one gets

$$\begin{aligned} E(D^0) &= E(K) + E(l) + E(\nu) \\ &= \sqrt{|\vec{p}(K) + \vec{p}(l)|^2 + m^2(Kl)} + \sqrt{|\vec{p}(\nu)|^2 + m^2(\nu)} \\ &= m(Kl) + |\vec{p}(D^0)|, \end{aligned} \tag{5.11}$$

where $|\vec{p}(K) + \vec{p}(l)|^2 = 0$ in the Kl rest frame, $m(Kl) = m(K) + m(l)$ and the mass of the neutrino is neglected. Next, the D^0 mass constraint is used to

Table 5.2: Stripping selection.

K^-	
p_{T}	$> 800\,\mathrm{MeV}/c$
p	$> 3000\,\mathrm{MeV}/c$
$\chi^2_{\mathrm{IP}}/\mathrm{ndf}$	> 9
PIDK	> 5
PIDμ	< 5
PIDp	< 5
e^+	
p_{T}	$> 500\,\mathrm{MeV}/c$
PIDe	> 0
μ^+	
p_{T}	$> 500\,\mathrm{MeV}/c$
PIDK	< 0
PIDμ	> 3
PIDp	< 0
π_s^+	
p_{T}	$> 300\,\mathrm{MeV}/c$
p	$> 1000\,\mathrm{MeV}/c$
PIDe	< 5
D^0	
χ^2_{FD}	> 100
χ^2_{DV}	< 20
$\cos\alpha$	> 0.999
$\chi^2_{\mathrm{IP}}/\mathrm{ndf}$	< 100
m	$\in [500, 2000]\,\mathrm{MeV}/c^2$
ΔM_{vis}	$< 400\,\mathrm{MeV}/c^2$
All tracks	
ProbNNghost	< 0.35

eliminate the true momentum of the D^0,

$$\begin{aligned} m^2(D^0) &= E^2(D^0) - |\vec{p}(D^0)|^2 \\ &= m^2(Kl) + 2m(Kl)|\vec{p}(D^0)| + |\vec{p}(D^0)|^2 - |\vec{p}(D^0)|^2 \\ \Rightarrow |\vec{p}(D^0)| &= \frac{m^2(D^0) - m^2(Kl)}{2m(Kl)}. \end{aligned} \tag{5.12}$$

Finally, the D^* mass constraint allows to eliminate the remaining variables in favour of angles in the Kl mass frame:

$$\begin{aligned} m^2(D^*) &= (E(D^0) + E(\pi_s))^2 - (\vec{p}(D^0) + \vec{p}(\pi_s))^2 \\ &= E^2(D^0) + E^2(\pi_s) + 2E(D^0)E(\pi_s) - |\vec{p}(D^0)|^2 - |\vec{p}(\pi_s)|^2 - 2|\vec{p}(D^0)||\vec{p}(\pi_s)|\cos\theta \\ &= m^2(D^0) + m^2(\pi_s) + 2E(D^0)E(\pi_s) - 2|\vec{p}(D^0)||\vec{p}(\pi_s)|\cos\theta. \end{aligned} \tag{5.13}$$

Substituting Equations (5.12) and (5.11) into (5.13) one gets the angle θ between the slow pion and the neutrino momentum, in the Kl rest frame:

$$\cos\theta = -\frac{m^2(D^*) - m^2(D^0) - m^2(\pi_s) - 2E(\pi_s)\left(m(Kl) + \frac{m^2(D^0)-m^2(Kl)}{2m(Kl)}\right)}{\frac{m^2(D^0)-m^2(Kl)}{m(Kl)}|\vec{p}(\pi_s)|}. \tag{5.14}$$

While the polar angle θ can be calculated from all known quantities, there is still one degree of freedom which cannot be calculated, represented by the azimuthal angle around the direction of the slow pion momentum; however, the flight direction of the D^0 in the lab frame can be used to constrain this quantity numerically. This can be done by taking 1000 points along the cone around $\vec{p}(\pi_s)$ with polar angle θ (corresponding to $2\pi/1000 \approx 6$ mrad precision) and using the angles in the Kl rest frame to solve for the corresponding angles in the lab frame. The solution that best aligns the sum of the final-state momenta to the D^0 direction of flight gives the azimuthal angle. A pictorial representation of the momentum of the system for the $l = e$ sample is given in Figure 5.4.

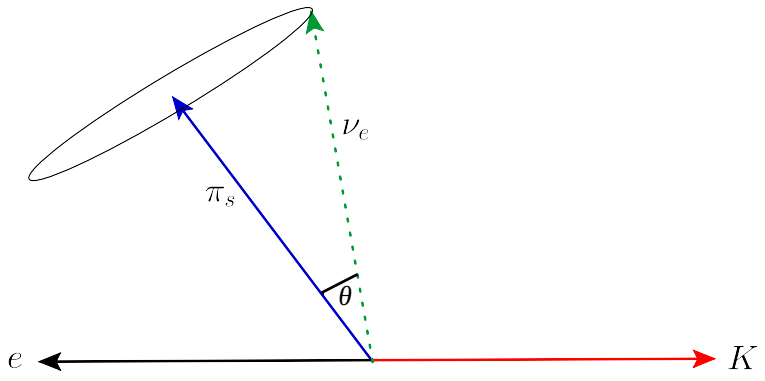


Figure 5.4: Momentum components of the π_s, ν_e, K and e in the Kl rest frame. The cone on which the neutrino momentum is restricted to lie is shown by the circle.

5.5 Fit strategy

The number of signal $D^0 \to K^- l^+ \nu_l$ candidates is extracted by a fit to the bidimensional ΔM_{vis} and M_{corr} distribution, using templates from all the MC samples described in Table 5.1. The background modes represent partially reconstructed backgrounds (for example, $D^0 \to K^- \mu^+ \nu_\mu \pi^0$ for the muon channel, where the π^0 is not reconstructed), misidentified backgrounds (for example, $D^0 \to \pi^- \mu^+ \nu_\mu$ for the muon channel, where the pion is misidentified as kaon) and a combination of both (for example, $D^0 \to K^- e^+ \nu_e \pi^0$ for the muon channel, where the π^0 is not reconstructed and the electron is misidentified as muon). Two additional templates are generated: a same-sign sample from data, i.e. $D^0 \to K^- l^- \nu_l$, used as a proxy for the combinatorial background, and a random slow pion sample from the signal sample, i.e. $D^0 \to K^- l^+ \nu_l$ associated to a random pion track coming from the primary vertex.

In order to take into account the finite size of the MC samples, the fitter algorithm is constructed to maximise the Barlow-Beeston log-likelihood [127],

$$\ln \mathcal{L} = \sum_{i=1}^{n} (d_i \ln(f_i) - f_i) + \sum_{j=1}^{n} \sum_{k=1}^{m} (a_{jk} \ln(A_{jk}) - A_{jk}), \tag{5.15}$$

where n is the number of bins in the bidimensional space spanned by ΔM_{vis} and M_{corr}, m is the combined number of signal and background sources, d_i is the number of events in data falling into bin i, a_{jk} is the number of MC events coming from the source k and falling in the bin j, A_{jk} is the predicted number of events from the source k falling into bin j and $f_i = \sum_{h=1}^{m} p_h A_{hi}$ is the predicted number of events in the i-th bin from all the m sources, each with yield p_h. In other words, the total likelihood is the combined probability of the observed set of d_i and the observed set of a_{jk}, which are both taken as Poisson-distributed, while the yields p_j and the true predicted number of events A_{jk} are unknown and can be found by maximising Equation (5.15). This method takes into account the statistical fluctuations of the MC predictions in bins with a small number of entries, which is to be expected in a multidimensional fit with limited-size MC samples.

The fits to the electron and muon channels are performed simultaneously, with the muon signal yield as shared parameter so that the electron signal yield is $N(D^0 \to K^- e^+ \nu_e) = N(D^0 \to K^- \mu^+ \nu_\mu)/R^{\mu/e}$. This ensures that the uncertainty on $R^{\mu/e}$ comes out directly from the fit using the procedure described above; moreover, the central value of $R^{\mu/e}$ can be easily blinded.

5.6 Sensitivity studies

Sensitivity studies with toy MC have been performed, with a limited number of background channels, to estimate an expected number of signal candidates and to check that there are no unexpected structures in data. This has been performed only for $D^0 \to K^- \mu^+ \nu_\mu$ decays, as the toy model used for this study does not provide any implementation of bremsstrahlung effects. For each channel, 10^7 events are generated and filtered according to the stripping selection described in Section 5.3.2, with a few differences: no PID cuts are present since there is no detector response simulation; for the same reason, no vertex quality requirement is applied; the cut on the flight distance significance, $\chi^2_{FD} > 100$, is replaced by a cut on the actual flight distance of FD$>$ 2 mm. In order to roughly simulate the momentum and vertex smearing due to the experimental resolution of the detector, all final-state momenta are smeared with a Gaussian resolution with width $\sigma_p = 0.006 \cdot p$, as this is the average momentum smearing factor for typical tracks in LHCb, while the three coordinates x_i of all decay vertices are smeared with a Gaussian resolution with width $\sigma_{x_i} = 0.03 \cdot x_i$; the 3% smearing factor is empirically chosen in the range between 0% and 5% as it is the one that gives the best fit to data.

The channels used for this study are the following:

- $D^0 \to K^- l^+ \nu_l$ (signal, Figure 5.5)
- $D^0 \to K^- l^+ \nu_l \pi^0$ (non-resonant, Figure 5.6)
- $D^0 \to (K^{*-} \to K^- \pi^0) l^+ \nu_l$ (resonant, Figure 5.7)
- $D^0 \to K^- \pi^+$ (pion is misidentified as muon, Figure 5.8)
- $D^0 \to K^- l^- \nu_l$ (same sign)
- $D^0 \to K^- l^+ \nu_l$, combined with a random π_s

The same-sign sample comes from data, while the random π_s sample is constructed by assigning the 4-momentum of a generated π_s to the next event generated. Bidimensional histograms in ΔM and M_{corr} are filled with the filtered and smeared events, normalised to unit area and then used as templates to fit the data, which consists of about 3700 candidates, corresponding to 4.6

fb^{-1}, or 0.3%, of the 2016 dataset. The bidimensional histograms and the unidimensional projections of the first four channels are shown in Figures 5.5, 5.6, 5.7 and 5.8.

Figure 5.9 shows the bidimensional distribution of the data used for this study. The total of the fitted templates is shown in Figure 5.10, and Figure 5.11 shows the bidimensional pull distribution of the fit, where the signal region has been highlighted. The fit projections on ΔM_{vis} and M_{corr} are shown, respectively, in Figures 5.12 and 5.13. Although it is evident from the pull distribution and from the individual fit projections that the fit quality can be improved as more background channels are added, by using the simple toy model described above it is possible to get an estimate of the analysis sensitivity.

The fitted number of signal candidates is $N(D^0 \to K^-\mu^+\nu_\mu) = 1090 \pm 85$; extrapolating to the full statistics and dividing by the 0.1 prescale factor mentioned in Section 5.3.2, the expected number of signal candidates is found to be $N(D^0 \to K^-\mu^+\nu_\mu) \approx 4 \times 10^6$. Assuming that the extrapolated number of signal events in the $D^0 \to K^-e^+\nu_e$ channel is of the same order of magnitude, ignoring all systematic uncertainties as well as statistical uncertainties on the efficiencies, the relative uncertainty on the measurement of $R^{\mu/e}$ can be estimated as

$$\frac{\Delta R^{\mu/e}}{R^{\mu/e}} = \frac{1}{R^{\mu/e}}\sqrt{\frac{N_\mu}{N_e^2}\left[1 + \frac{N_\mu}{N_e}\right]} \approx 0.1\%, \tag{5.16}$$

where $N_\mu = N(D^0 \to K^-\mu^+\nu_\mu)$ and $N_e = N(D^0 \to K^-e^+\nu_e)$. For the ratio shown in Figure 2.12 the relative error is 4%; therefore, the measurement of $R^{\mu/e}$ at LHCb could improve the relative error by more than an order of magnitude.

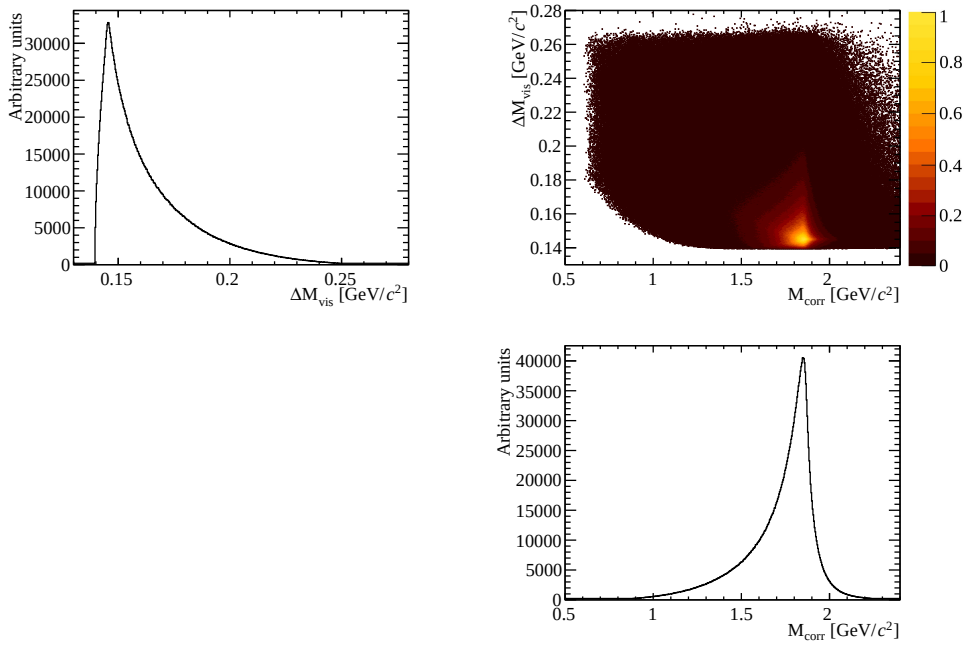


Figure 5.5: Distributions from toy MC of the channel $D^0 \to K^- \mu^+ \nu_\mu$ (signal)

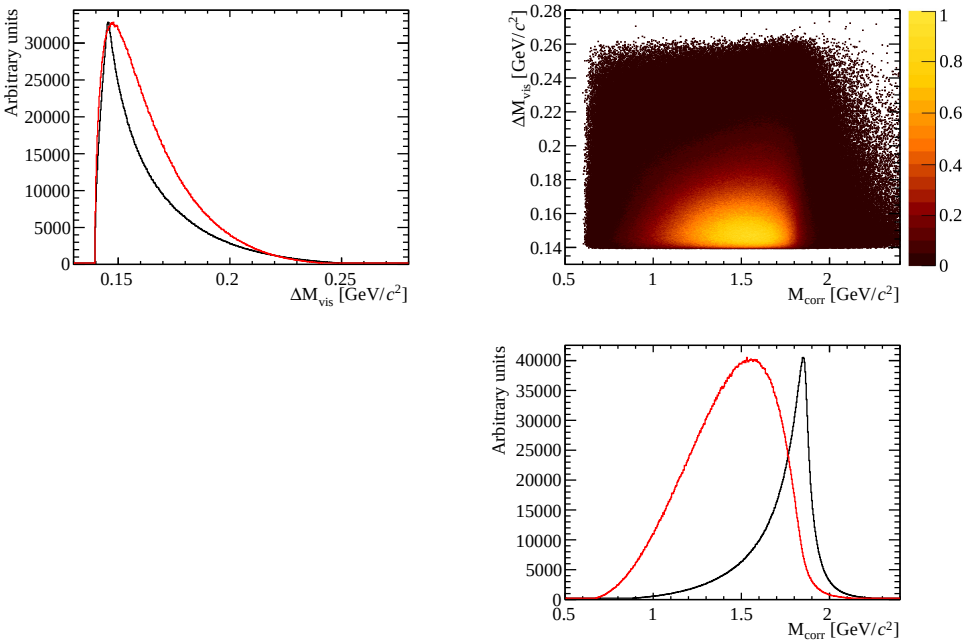


Figure 5.6: Distributions (red) from toy MC of the channel $D^0 \to K^- \mu^+ \nu_\mu \pi^0$. The black distributions, corresponding to the signal, are given as comparison.

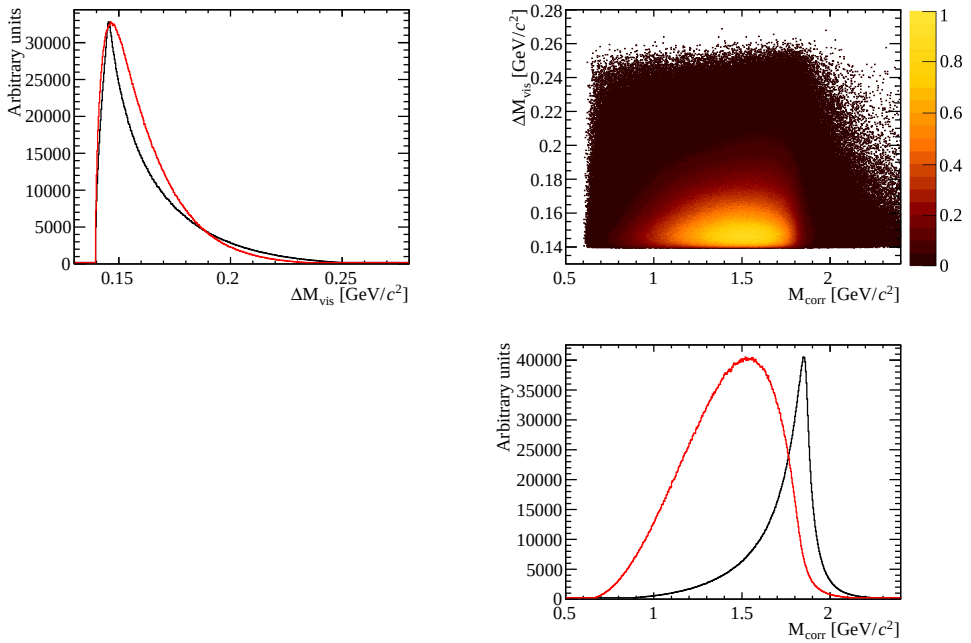


Figure 5.7: Distributions (red) from toy MC of the channel $D^0 \to (K^{*-} \to K^- \pi^0)\mu^+\nu_\mu$. The black distributions, corresponding to the signal, are given as comparison.

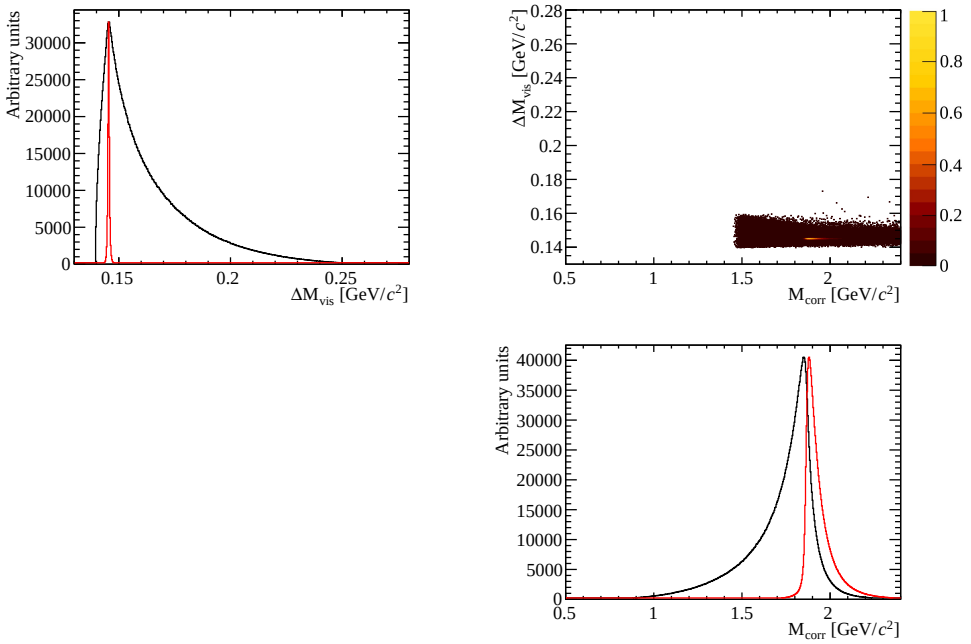


Figure 5.8: Distributions (red) from toy MC of the channel $D^0 \to K^- \pi^+$. The black distributions, corresponding to the signal, are given as comparison.

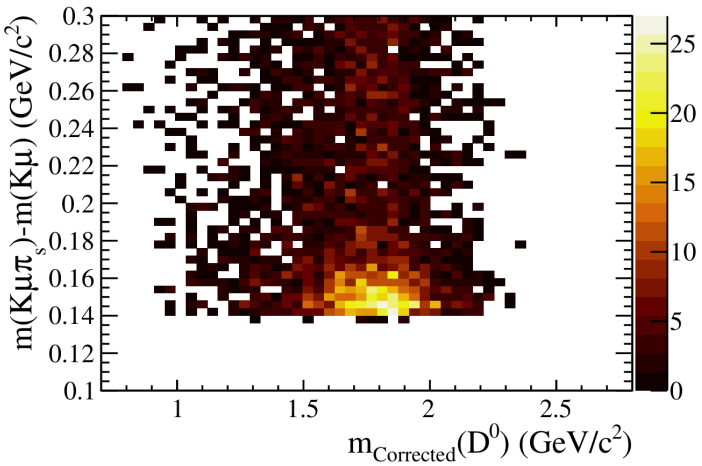


Figure 5.9: M_{corr} vs ΔM_{vis} data distribution.

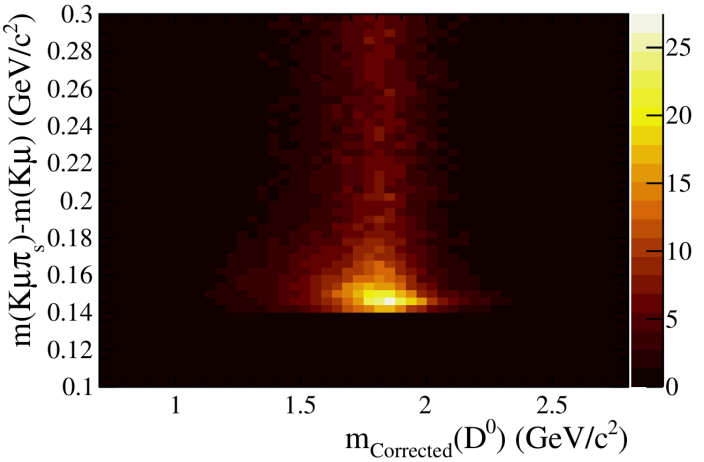


Figure 5.10: M_{corr} vs ΔM_{vis} fitted total template distribution.

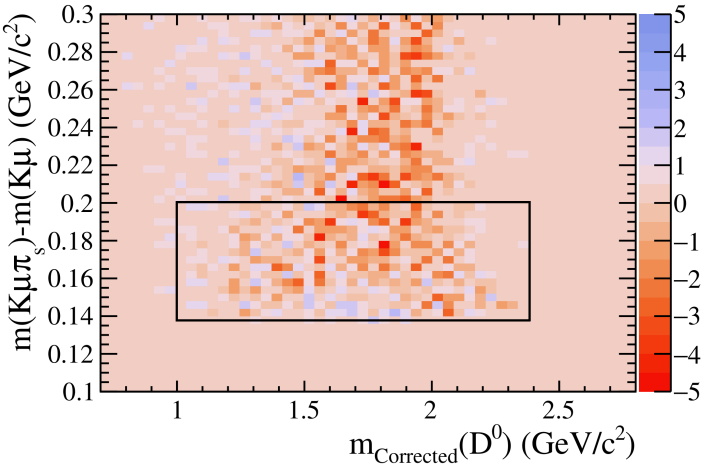


Figure 5.11: Bidimensional pulls distribution of the fit. The area in the black rectangle indicates the signal peak region.

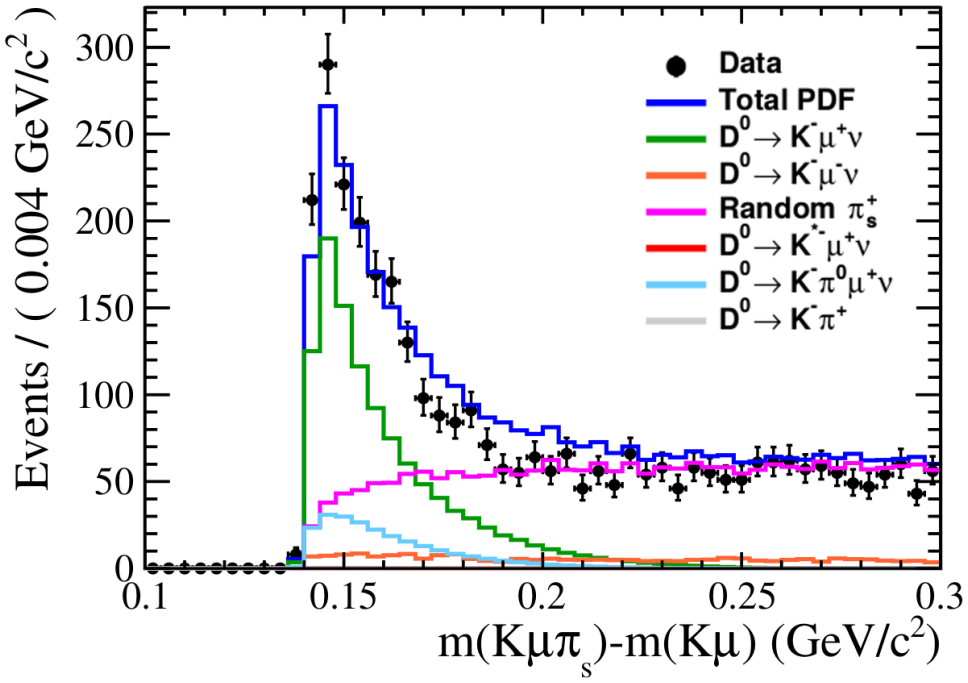


Figure 5.12: Fit projection on the ΔM_{vis} axis.

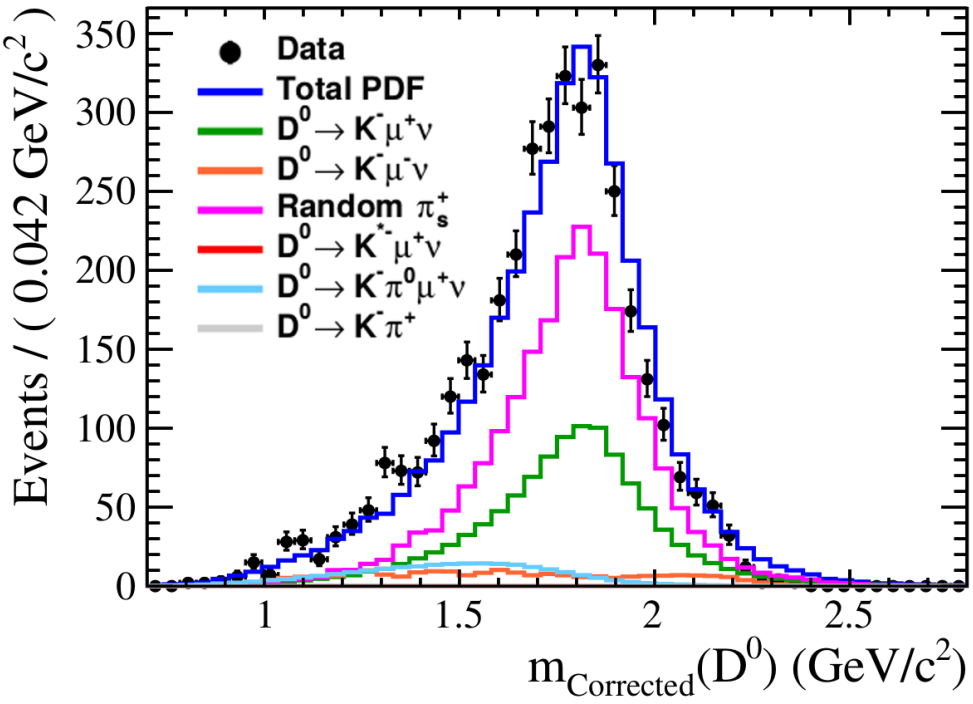


Figure 5.13: Fit projection on the M_{corr} axis.

5.7 PID efficiencies

The simulation of the response of the subdetectors which participate in the evaluation of the PID variables is sensitive to the experimental conditions of the system at the moment data are taken, such as gas pressure, temperature and detector alignment. In fact, in LHCb it is often found that the simulated PID variables do not correctly reproduce the data, as can be seen in Section 4.3. The LHCb collaboration has developed a tool [128], called PIDCalib, which can be used to calculate data-driven particle identification efficiencies from some calibration samples with no PID requirements in the event selection, which present high statistics and low background levels. To isolate the signal PID distributions, the *sPlot* technique [129] is used by performing a fit to the mass spectrum of the mother particle of the decay and extracting the weights (called sWeights) which are then applied to the PID variables. The efficiencies are evaluated by cutting on the PID variables and dividing the weighted number of candidates before and after the cut, in bins of track multplicity of the event and pseudorapidity (η) and transverse momentum of the track.

The calibration sample for muons comes from $B^+ \to K^+(J/\psi \to \mu^+\mu^-)$ decays, and consists of approximately 2 million signal candidates. Muon PID efficiencies have not yet been evaluated, and the PIDCalib method will be used.

The calibration sample for electrons comes from $B^+ \to K^+(J/\psi \to e^+e^-)$ decays. Although the BF of this channel is comparable with the previous one, the statistics are much lower due to the difference in the reconstruction and selection efficiency of muons over electrons in LHCb (about 430000 $B^+ \to K^+(J/\psi \to e^+e^-)$ candidates in total after the cuts described in the next section). As electrons have a higher probability of emitting bremsstrahlung photons, the shape of the e^+e^- mass spectra will be significantly more distorted. This poses a problem in the evaluation of the signal PID variables, since the quality of the fit to the B^+ mass spectrum is affected. Furthermore, the *sPlot* technique assumes that the fit variable (i.e. the B^+ mass) and the parametrising variables (i.e. the PID variables) are uncorrelated, which is not usually the case: this can cause biases or systematic effects that need to be corrected. Lastly, when binning in η and p_T (the track multiplicity dependence is ignored in this study), in some bins the statistics can be low, and the calculated efficiency might be unphysical, i.e. lying outside the range $\epsilon \in [0, 1]$: this is a consequence of the definition of sWeights, which can be negative or greater than 1, together with a

potentially non perfect fit used to extract them.

To summarise, the evaluation of electron PID efficiencies is affected by a series of problems, which do not have an easy solution; these are mostly negligible in the muon case where the fit to the B^+ mass spectrum is easier to parametrise and the statistics are high enough to cover the full multidimensional binning. To avoid a study of the several systematic effects or possible biases that would arise from the standard PIDCalib method, for this analysis another strategy is used, as described in the next section. As it is an independent technique to calculate PID efficiencies, the results obtained are compared with efficiencies calculated with sWeights, and will be implemented as a parallel method in the PIDCalib package.

5.7.1 The tag-and-probe method for electrons

In this section, a different method, which makes no use of sWeighted data, is described to calculate electron PID efficiencies. Instead of extracting signal distributions from a fit to the B^+ mass distribution, some offline cuts are applied to the calibration sample in order to minimise the non-combinatorial background under the J/ψ peak so that the shape of the tails depends mostly on bremsstrahlung effects. Fits to the J/ψ mass distribution are then performed by applying a hard PID cut on one of the two electrons and scanning the other PID variable with a large set of increasingly harder cuts.

Events are selected from $B^+ \to K^+(J/\psi \to e^+e^-)$ decays from 2016 data, using the calibration stripping line, which contains no PID cuts; a summary of the stripping and additional offline cuts is given in Table 5.3. The last cuts in the J/ψ and B^+ rows, when combined, represent a diagonal cut in the $m(J/\psi) - m(B^+)$ plane, shown in Figure 5.14, which is applied to isolate the B^+ peak and its radiative tail from the $\psi(2S)$ contribution and the partially reconstructed background contamination.

The PID efficiencies are calculated by splitting the data sample according to the magnet polarity (MagUp and MagDown) and a flag (HasBremAdded) which determines if the electron has been reconstructed with or without the addition of non-collinear bremsstrahlung photons: in fact, given the geometry of the LHCb detector, if an electron emits a photon after the magnet it will be mostly collinear and will not be resolvable from the electron shower in the electromagnetic calorimeter, although it will contribute to the measured energy of the electron; if an electron emits a photon before the magnet, or while passing

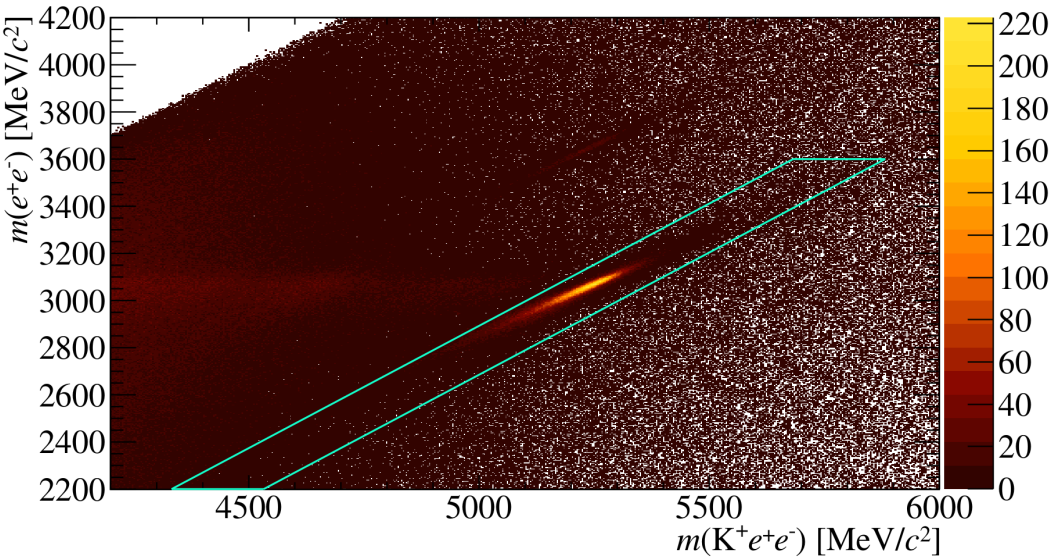


Figure 5.14: The $m(J/\psi) - m(B^+)$ plane for the electron calibration sample. The diagonal cut around the J/ψ peak is shown.

through the magnetic field, one can attempt to extrapolate the direction of flight of the electron before entering the magnetic field and look for energy deposits where the projected track hits the electromagnetic calorimeter, and add the photon energy back to the electron. The mass resolution of the J/ψ and the particle identification accuracy will be different in the two cases and so it is necessary to treat them separately.

The method to estimate the efficiencies, called "tag-and-probe" algorithm, is described in detail in the following:

1. One of the two candidate electron tracks is chosen as the electron "tag" by applying a tight cut on its PIDe, i.e. PID$e > 5$.

2. The other track is the "probe".

3. For the probe, a PIDe cut value is defined as $x_j = \min(\text{PID}e) + 1.1j$, where $j \in \mathbb{N}$ is initially set to be 1.

4. The numbers of candidates passing and being rejected by the cut are defined respectively as $P_j = N(\text{PID}e > x_j)$ and $R_j = N(\text{PID}e < x_j)$.

5. Let n_{low} be the minimum value of j so that $P_j > 1200$ and $R_j > 1200$; this is to ensure that there are enough events in each subsample to perform a meaningful fit.

Table 5.3: Stripping cuts for electron PID efficiency studies. Additional offline cuts, potentially superseding the stripping ones, are indicated with †.

$e^{\pm}$ candidate tracks	$p_T > 500\ \mathrm{MeV}/c$ $\chi^2_{track}/N_{dof} < 5$ $\chi^2_{IP}(PV) > 25^{\dagger}$ $m(e^+e^-) \in [2100, 4300]\ \mathrm{MeV}/c^2$
e_{tag}	$p_T > 1500\ \mathrm{MeV}/c$ $p > 6000\ \mathrm{MeV}/c$ PID$e > 5.0$
e_{probe}	$p_T > 500\ \mathrm{MeV}/c$ $p > 3000\ \mathrm{MeV}/c$
J/ψ	$\chi^2_{vertex} < 9$ $\chi^2_{FD}(PV) > 5$ $m(J/\psi) \in [2250, 3600]\ \mathrm{MeV}/c^{2\dagger}$
K^+	$p_T > 1000\ \mathrm{MeV}/c$ PID$K > 0$ $\chi^2_{track}/N_{dof} < 4$ $\chi^2_{IP}(PV) > 9$
B^+	$\chi^2_{vertex}/N_{dof} < 9$ $\chi^2_{IP}(PV) < 9^{\dagger}$ $m(B^+) \in [4200, 6000]\ \mathrm{MeV}/c^2$ $\left\lvert m(B^+) - m(J/\psi) - 2182.3 \right\rvert < 100\ \mathrm{MeV}/c^{2\dagger}$

6. For the same reason, let n_{high} be the maximum value of j so that the same condition is satisfied.

7. A set S of scan values of PIDe for the probe is then defined by all the x_j in the range $j \in [n_{low}, n_{high}]$.

8. For each $x_j \in S$, two $m(e^+e^-)$ plots are produced, one for the events that pass the cut and one for the events that are rejected by it.

9. All the $m(e^+e^-)$ distributions are fitted with a total PDF composed of the sum of an exponential PDF (for the combinatorial background) and a double-sided Crystal Ball PDF (for the J/ψ signal, including the bremsstrahlung tail).

10. From the sample of events that pass the first cut, i.e. PID$e > x_{n_{low}}$, a mass window is defined, corresponding to the shortest interval containing 90% of the integral of the signal PDF, and it is fixed for all the subsequent fits.

11. The signal yields, for each PID cut value, are extracted from the fits and multiplied by the fraction of the signal PDF in the chosen mass window; these quantities are labelled N_{pass} (from the sample of events that pass the cut) and N_{rej} (from the sample of events that get rejected by the cut).

12. For each $x_j \in S$, the total number of $J/\psi \to e^+e^-$ decays is calculated as $N_{tot} = N_{pass} + N_{rej}$.

13. The signal efficiency is calculated as $\epsilon_{PID} = \frac{N_{pass}}{N_{tot}}$.

14. The tag and the probe are swapped and the algorithm is repeated from step 3.

The uncertainties on the efficiencies are calculated using Bayes' theorem,

$$P\left(\epsilon_{PID}\middle|N_{pass}\right) = \frac{P\left(N_{pass}\middle|\epsilon_{PID}\right)P\left(\epsilon_{PID}\right)}{Z}, \tag{5.17}$$

where Z is a normalisation constant to be determined, $P(N_{pass}|\epsilon_{PID})$ is the binomial distribution,

$$P\left(N_{pass}\middle|\epsilon_{PID}\right) = \frac{N_{tot}!}{N_{pass}!(N_{tot}-N_{pass})!}\epsilon_{PID}^{N_{pass}}\left(1-\epsilon_{PID}\right)^{N_{tot}-N_{pass}}, \tag{5.18}$$

and $P(\epsilon_{PID})$ is the prior knowledge on the efficiency, which is reasonable to assume flat in the inclusive range [0,1]. The normalisation constant Z is

determined from

$$\begin{aligned}
\int_{-\infty}^{\infty} P\left(\epsilon_{PID}\middle|N_{pass}\right) d\epsilon_{PID} &= \int_{-\infty}^{\infty} \frac{P\left(N_{pass}\middle|\epsilon_{PID}\right) P\left(\epsilon_{PID}\right)}{Z} d\epsilon_{PID} \\
&= \frac{1}{Z}\binom{N_{tot}}{N_{pass}} \int_0^1 \epsilon_{PID}^{N_{pass}} (1-\epsilon_{PID})^{N_{tot}-N_{pass}} d\epsilon_{PID} \\
&= \frac{1}{Z}\binom{N_{tot}}{N_{pass}} \beta(N_{pass}+1, N_{tot}-N_{pass}+1) \\
&= \frac{1}{Z}\binom{N_{tot}}{N_{pass}} \frac{\Gamma(N_{pass}+1)\Gamma(N_{tot}-N_{pass}+1)}{\Gamma(N_{tot}+2)} \\
&= \frac{1}{Z}\frac{\Gamma(N_{tot}+1)}{\Gamma(N_{tot}+2)} \\
&= \frac{1}{Z}\frac{1}{N_{tot}+1} \\
&= 1, \qquad (5.19)
\end{aligned}$$

where $\beta(x,y)$ is the Euler beta function. When substituting Equations (5.18), (**??**) and (5.19) into (5.17), one obtains the full description of the probability distribution of ϵ_{PID} given N_{pass} events that pass the cut over N_{tot} total events,

$$P\left(\epsilon_{PID}\middle|N_{pass}\right) = \frac{\Gamma(N_{tot}+2)}{\Gamma(N_{pass}+1)\Gamma(N_{tot}-N_{pass}+1)} \epsilon_{PID}^{N_{pass}} (1-\epsilon_{PID})^{N_{tot}-N_{pass}}. \qquad (5.20)$$

Finally, the central value of ϵ_{PID} is taken as the mode of the distribution, while the uncertainty is obtained as the smallest 68% confidence interval around the mode.

The algorithm described above can be visualised in Figures 5.15 and 5.16, which show the probe PIDe distributions for the MagUp sample, with the e^- as the probe, separated by the HasBremAdded flag. The superimposed red lines represent the scan values used to obtain the subsamples of events passing and being rejected by the cut in the tag-and-probe method.

In order to compare this result with the standard PIDCalib method, electron PID efficiencies are calculated from the sWeights obtained from the fit to the B^+ mass. To do so, calling w_{tot} the sum of all the signal sWeights for the probe, for each of the $x_j \in S$ calculated as above, the sum of sWeights that pass the cut is defined as

$$w_{pass} = \sum_{i \in P_j} w_i, \qquad (5.21)$$

where w_i is the sWeight of the i-th candidate and P_j is the set of all events that pass the cut PID$e > x_j$. The efficiency for each x_j is then calculated as the ratio

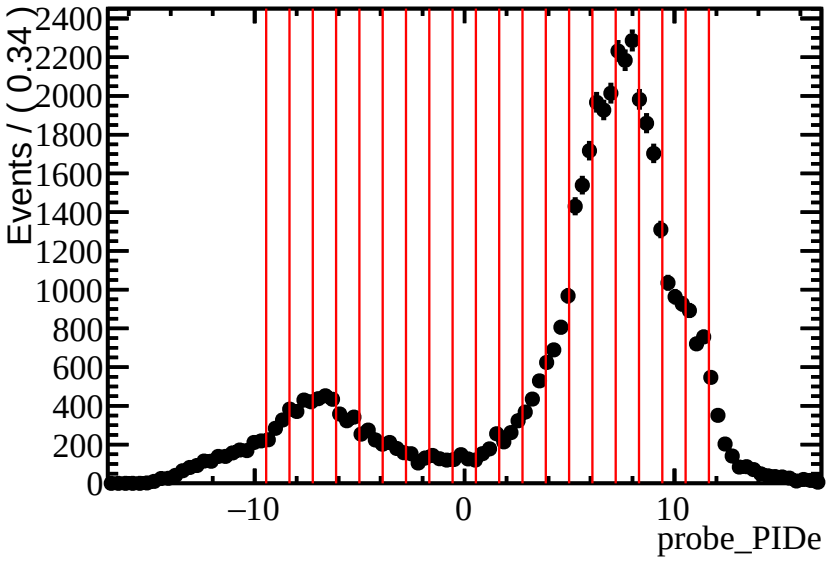


Figure 5.15: Probe PIDe distribution with no bremsstrahlung photons added. The red vertical lines indicate the cut values obtained according to the algorithm.

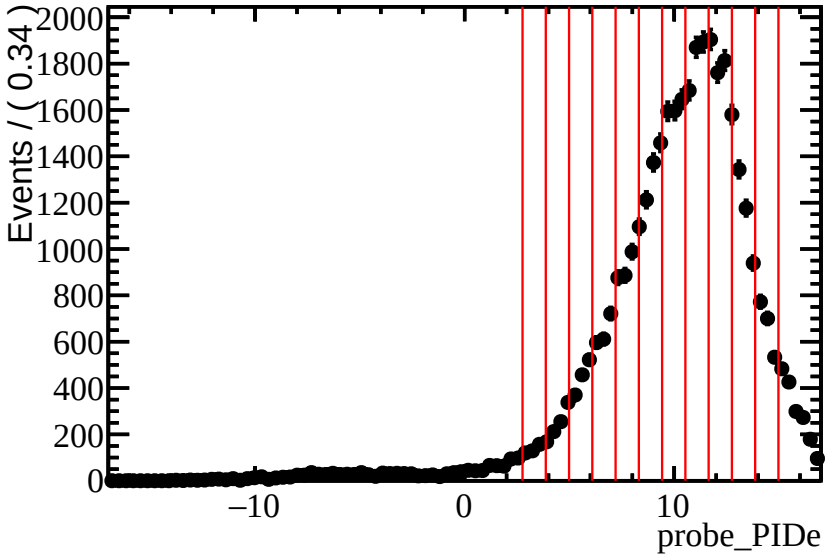


Figure 5.16: Probe PIDe distribution with bremsstrahlung photons added. The red vertical lines indicate the cut values obtained according to the algorithm.

between w_{pass} and w_{tot}. For this comparison, uncertainties are ignored and will be evaluated in a future study. Figures 5.17 and 5.18 show the distributions of the electron PID efficiency as a function of the cut value x_j for the two subsamples split according to the HasBremAdded flag. Each plot represents a combination of magnet polarity (MagUp or MagDown) and charge of the probe ("em" for e^- and "ep" for e^+). The largest relative differences are found to be

for high values of x_j, which correspond to the case of low statistics in the fits to the events that pass the cut: this leads to more unstable fits. However, this case is of no interest for this analysis as the current stripping cut for electrons is PID$e > 0$, which falls well within the stable region of the distribution.

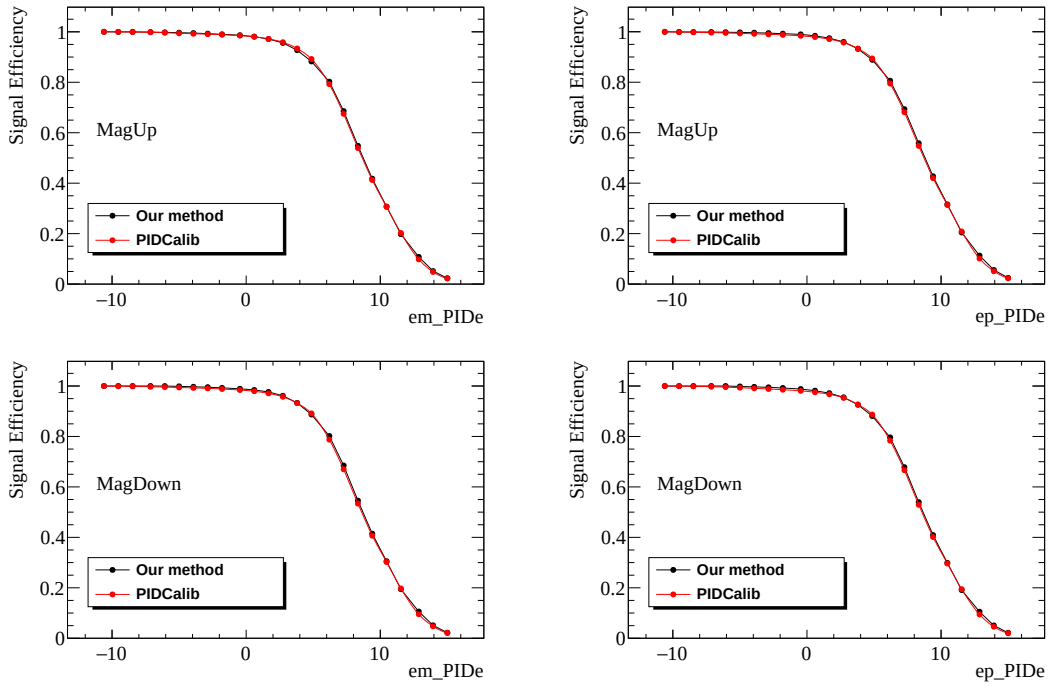


Figure 5.17: Comparison between the tag-and-probe ("Our method") and the standard ("PIDCalib") methods to calculate electron PID efficiencies, for both magnet polarities and both values of charge of the probe. No bremsstrahlung photons are added.

As a summary, in this section a new method to calculate electron PID efficiencies, which does not rely on the *sPlot* formalism, is described. Although some further studies are required, it has been shown that it is stable and can be used to validate the standard LHCb strategy or, in the future, as another independent standard method suitable for precision measurements involving electrons in the final state.

5.7.2 PID efficiency tables

As mentioned at the beginning of this section, particle identification performances in LHCb depend on the transverse momentum and pseudorapidity of the tracks. In order to be sensitive to this effect, the electron PID efficiencies in this analysis are calculated in bins of p_T and η of the probe. This allows also to compare the PID efficiencies with the track reconstruction efficiencies in

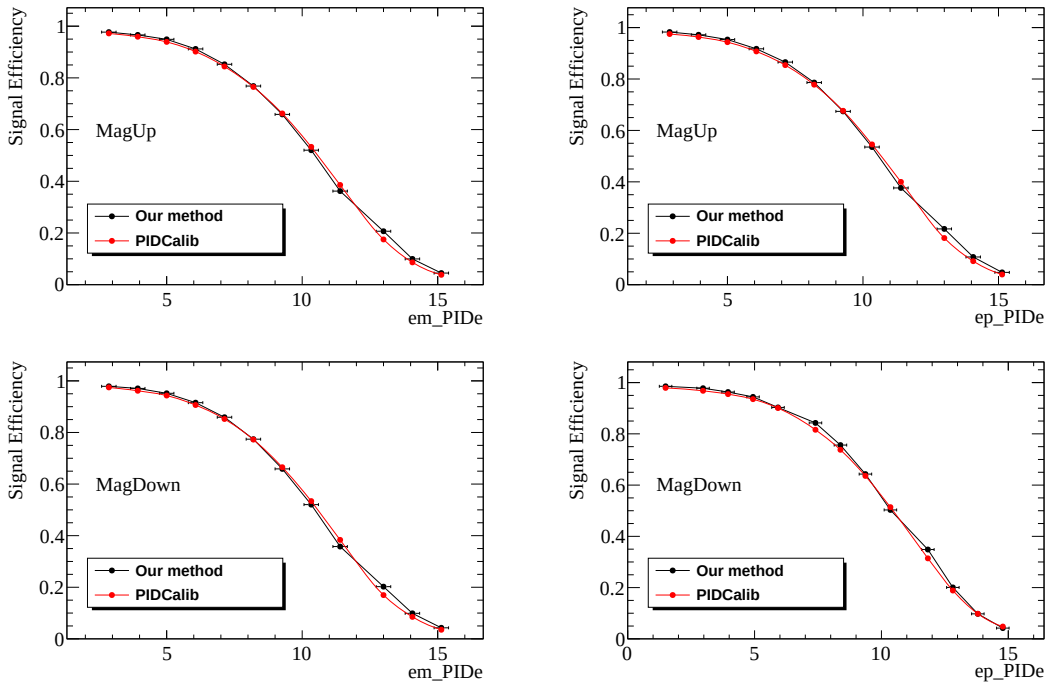


Figure 5.18: Comparison between the tag-and-probe method ("Our method") and the standard method ("PIDCalib") to calculate electron PID efficiencies, for both magnet polarities and both values of charge of the probe. Bremsstrahlung photons are added.

order to check their correlation. Figure 5.19 shows in blue the binning chosen for this analysis, based on the criterion of having roughly the same number of entries in each bin, for the MagUp polarity and the e^- as the probe. For each bin, the cut PID$e > 0$ is applied and the same procedure as the one described in the previous section is performed. The fit parameters of the double-sided Crystal Ball functions are constrained to vary under a Gaussian constraint. For each parameter, the mean is set as the average value of its distribution as a function of the PIDe cut, removing the points in the low statistics regions, and the width is set as the spread of the same distribution. Both are extracted from the previous study and ensure that, for the bins with low statistics, the tails of the Crystal Ball function do not absorb flat backgrounds. The mass windows containing 90% of the signal PDF are evaluated for each bin of the $p_T - \eta$ plane, from the sample of events that pass the cut. All the individual fits are shown in Appendix B.

From this procedure, two efficiency tables, for both values of the HasBremAdded flag, are produced for the cut PID$e > 0$, corresponding to the stripping cut. They are summarised in Table 5.4.

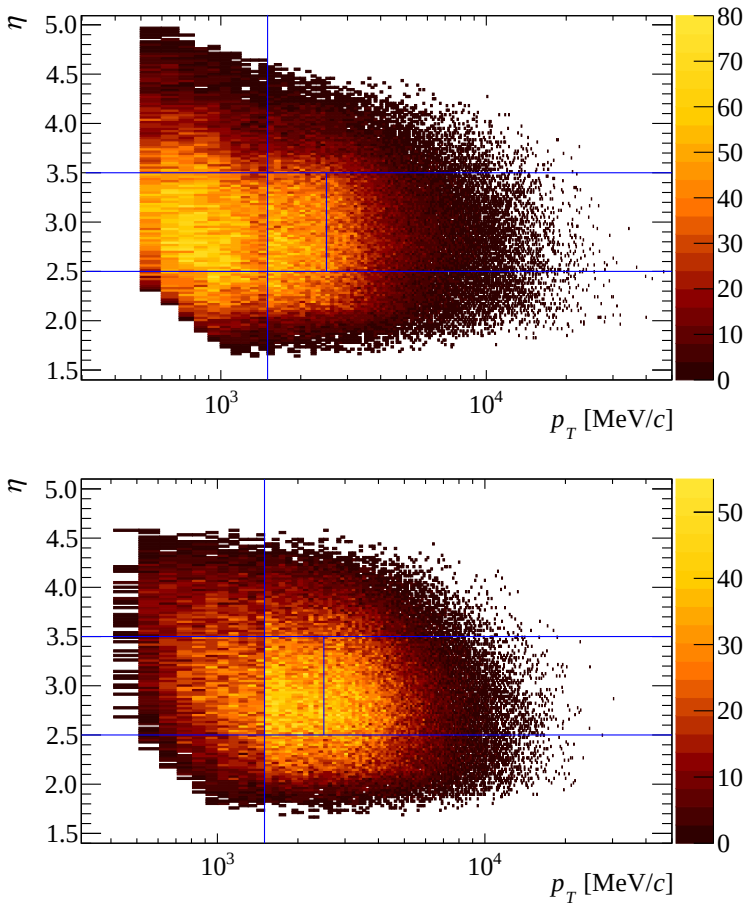


Figure 5.19: Distribution of probe tracks in the $p_T - \eta$ plane, without (top) and with (bottom) added bremsshtrahlung. The blue lines define the binning chosen to calculate binned PID efficiencies.

η	p_T [GeV/c]	ϵ^0_{PID}	ϵ^1_{PID}
$1.4 - 2.5$	$0 - 2$	$0.928^{+0.005}_{-0.006}$	$0.990^{+0.003}_{-0.004}$
	$2 - 50$	$0.9938^{+0.0009}_{-0.0011}$	$0.9993^{+0.0002}_{-0.0004}$
$2.5 - 3.5$	$0 - 2$	$0.951^{+0.002}_{-0.003}$	$0.996^{+0.001}_{-0.001}$
	$2 - 3.5$	$0.994^{+0.001}_{-0.001}$	$0.9976^{+0.0006}_{-0.0008}$
	$3.5 - 50$	$0.976^{+0.001}_{-0.002}$	$0.9965^{+0.0005}_{-0.0006}$
$3.5 - 5.1$	$0 - 2$	$0.870^{+0.006}_{-0.007}$	$0.962^{+0.004}_{-0.005}$
	$2 - 50$	$0.920^{+0.005}_{-0.005}$	$0.965^{+0.003}_{-0.003}$

Table 5.4: Efficiency table for the cut PID$e > 0$. The PID efficiency ϵ^1_{PID} (ϵ^0_{PID}) is relative to the sample where the bremsstrahlung photons are (are not) added

5.8 Conclusions and prospects

This study has shown how LHCb can perform a direct test of lepton universality in semileptonic charm decays and improve the precision on the current measurement of the ratio

$$R^{\mu/e} = \frac{\mathcal{B}(D^0 \to K^- \mu^+ \nu_\mu)}{\mathcal{B}(D^0 \to K^- e^+ \nu_e)} \tag{5.22}$$

by potentially a factor of more than 10. This would be the first direct measurement of the ratio $R^{\mu/e}$.

The precision on this measurement is expected to be limited by systematic uncertainties. As mentioned before, the statistical uncertainties are expected to be of the order of $\mathcal{O}(0.1\%)$. The largest source of systematic uncertainty will probably come from electron efficiencies, both PID, track reconstruction and their correlation. PID efficiencies are measured to be of the order of $\mathcal{O}(0.01\% - 1\%)$ depending on the bin in the $p_T - \eta$ plane.

This analysis is already in an advanced state. The PID efficiency tables for electrons have been produced; the fit strategy is defined; the framework has been written and tested and has passed a series of checks with toy data; a study on tracking reconstruction efficiencies is ongoing, for both electrons and muons; the q^2 reconstruction procedure is clear and is being tested against possible biases against simulated data; a first draft of the list of possible systematic effects that could affect the measurement has been written; the request for the large unfiltered MC sample was made and the sample is currently being generated. Once the MC samples are ready, one can proceed to the blinded template fits on data, to the measurement of the q^2 resolution and to the determination of the remaining efficiencies, checking explicitly that the acceptance and all the non-leptonic efficiencies cancel in the ratio $R^{\mu/e}$.

Appendices

A Choice of the multivariate classifier

In order to obtain the best separation between signal and background, the events are initially processed through several multivariate classifiers. The best performing algorithm is then chosen and optimised to be used in the analysis. The following algorithms were tested:

- Boosted Decision Tree (BDT)
- MLP Boosted Artificial Neural Network
- Likelihood with Kernel Density Estimator (KDE)
- Likelihood with Kernel Density Estimator and spline functions (MIX)
- PDE FoamBoost
- Fisher Discriminants
- Boosted Fisher

The training signal sample consists of 8.17×10^5 $D^+ \to \pi^+\pi^+\pi^-$ phase-space, truth-matched Monte Carlo events while the training background sample consists of 9.22 $\times10^6$ events obtained from the $\pi^+\pi^+\pi^-$ invariant mass sidebands, in 20 MeV/c^2 mass windows on either side of the D^+ peak. Both samples are rescaled to the same number of events before the beginning of the training. For the optimisation step the number of events in the background sample is increased to 10.21 $\times10^6$. The variables chosen as input for the comparison of the classifiers are:

- log |ProbNNghost(π_1)|
- log |ProbNNghost(π_2)|
- log |ProbNNghost(π_3)|

- $\log|\text{ProbNNk}(\pi_1)|$
- $\log|\text{ProbNNk}(\pi_2)|$
- $\log|\text{ProbNNk}(\pi_3)|$
- $\log|\text{1-DIRA}(D^+)|$
- $\log|\chi^2_{\text{IP}}(D^+)|$
- $\log|\chi^2_{\text{DV}}(D^+)|$

where π_2 and π_3 are the same sign pions.

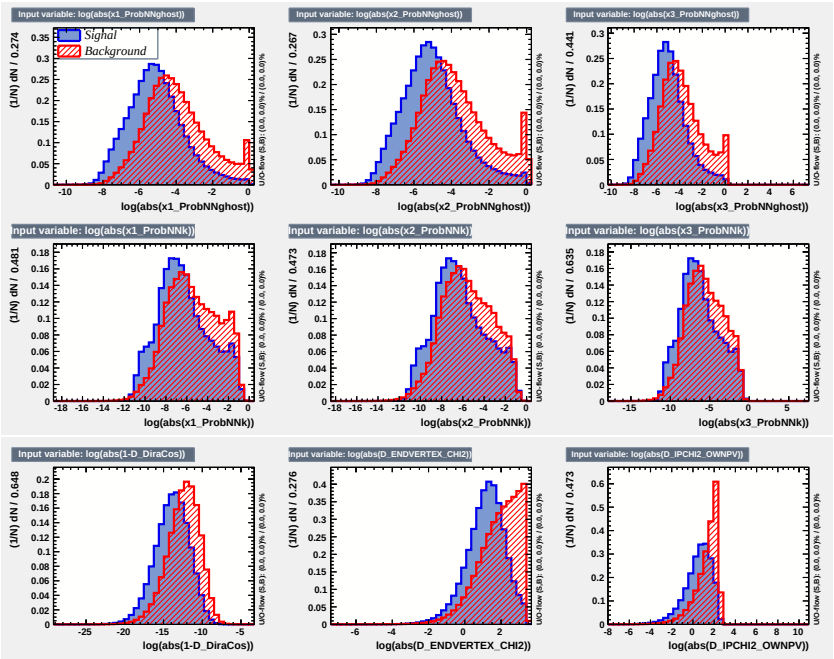


Figure A.1: Input variables for the MVA selection.

In Figure A.1 the signal and background distributions for the input variables are shown. The ROC curves for the different methods are shown in Figure A.2, where it can be seen that the BDT method gives the the best performance. As a result the BDT classifier is chosen for a dedicated tuning of the classifier parameters. Before the tuning is performed, the number of input variables is increased from 9 to 21, where the additional variables are:

- $\log|1.02-\text{ProbNNpi}(\pi_1)|$
- $\log|1.02-\text{ProbNNpi}(\pi_2)|$
- $\log|1.02-\text{ProbNNpi}(\pi_3)|$

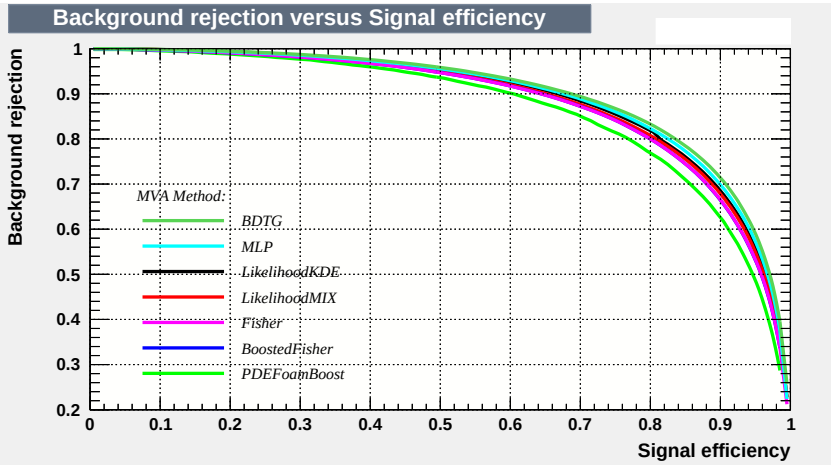


Figure A.2: ROC curves for the MVA selection.

- $\log|\chi^2_{\mathrm{IP}}(\pi_1)|$
- $\log|\chi^2_{\mathrm{IP}}(\pi_2)|$
- $\log|\chi^2_{\mathrm{IP}}(\pi_3)|$
- $\log|\mathrm{DOCA}(\pi_1,\pi_2)+0.005|$
- $\log|\mathrm{DOCA}(\pi_1,\pi_3)+0.005|$
- $\log|\mathrm{DOCA}(\pi_2,\pi_3)+0.005|$
- $\chi^2_{track}(\pi_1)/\mathrm{ndf}$
- $\chi^2_{track}(\pi_2)/\mathrm{ndf}$
- $\chi^2_{track}(\pi_3)/\mathrm{ndf}$

The signal and background distributions for the additional variables used in the BDT optimisation are shown in Figure A.3.

The resulting ROC curve shows a better performance than with the previous configuration, giving about 10% increase in signal-background separation, as can be seen in Figure A.4. Figure A.5 shows the BDT output distributions for signal and background in the MVA testing phase.

For the final classifier training two improvements are introduced. Firstly, for the variables where there is a value for each of the final-state tracks, the values are ordered by the p_T of the track. The improvement in the classifier performance as a result of the p_T ordering is shown in Figure A.6.

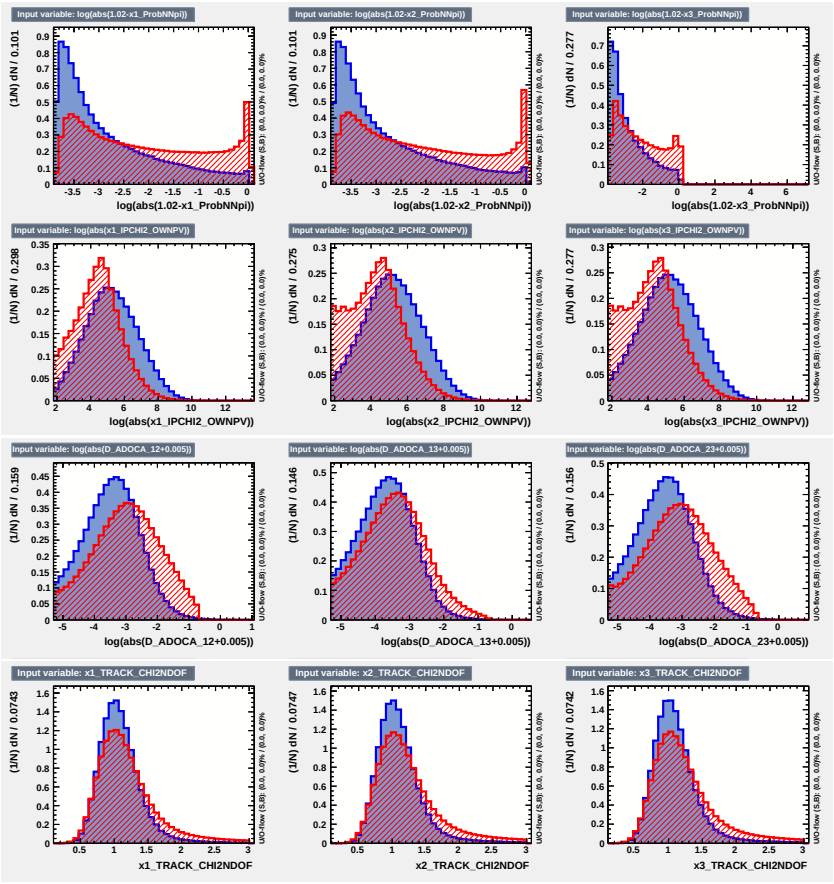


Figure A.3: Additional input variables for the BDT training.

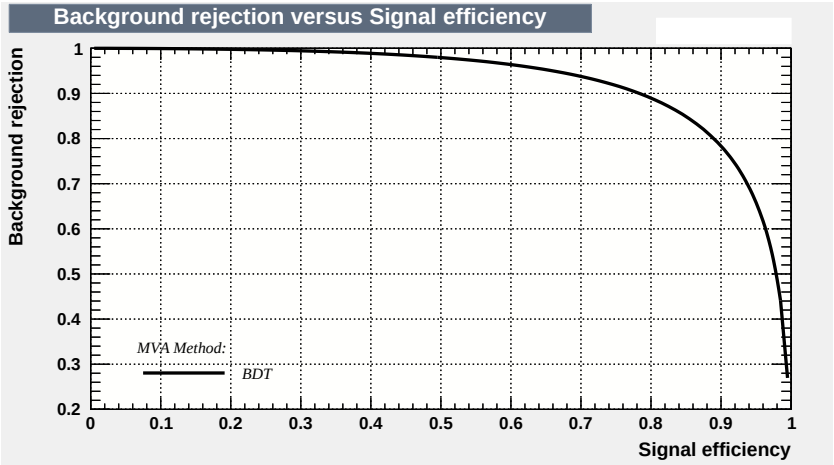


Figure A.4: ROC curve for the BDT training.

Secondly, the means of the ProbNN variables in MC are scaled according to the corresponding means in data, as described in Section 4.3. The difference in classifier performance as a result of the MC scaling is shown in Figure 4.4. Whilst the performance according to the ROC curve is worse after the MC scal-

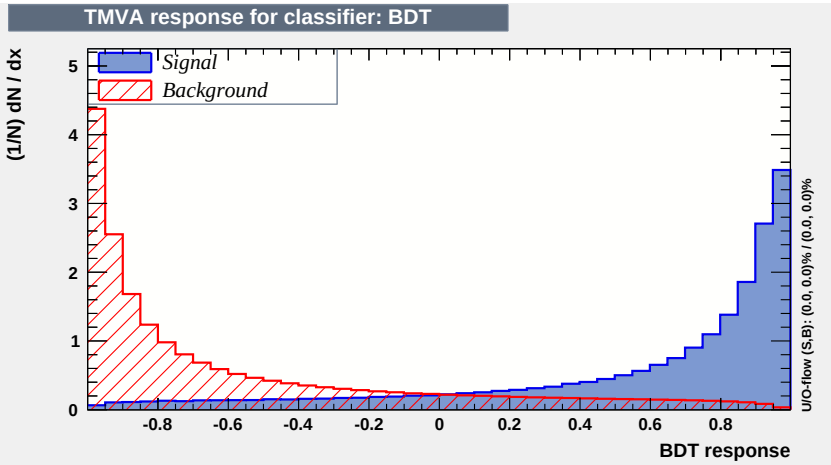


Figure A.5: BDT response, signal and background distribution (testing phase).

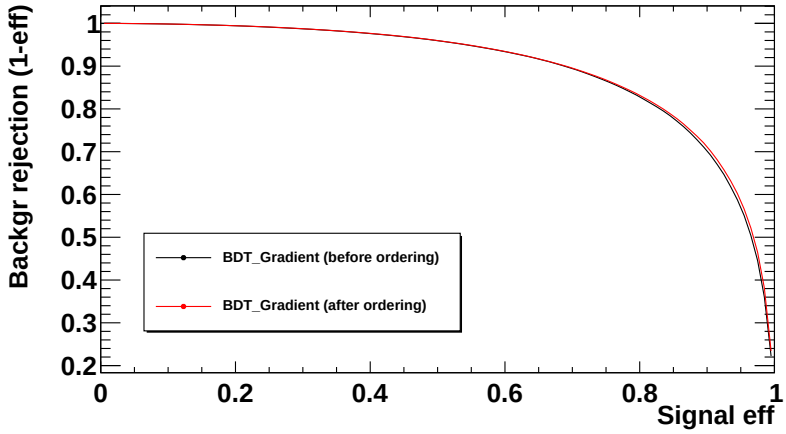


Figure A.6: Comparison of the BDT performance before and after p_T ordering.

ing, the FoM (see Section 4.5) improves due to the improved description of the training variables in the MC. Therefore the resulting limit is also improved with respect to the non-scaled case. The ROC curve with p_T ordering in Figure A.6 represents the classifier used throughout the remainder of the analysis. The signal MC distributions of all input variables to this classifier are shown in Figures 4.3 (non-scaled variables) and 4.4 (scaled variables).

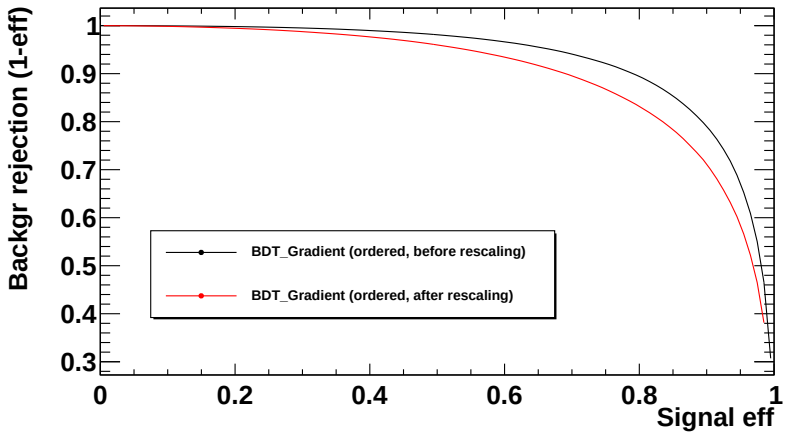


Figure A.7: Comparison of the BDT performance before and after MC scaling.

B Additional material on electron PID studies

B.1 Percent variations between the two methods

The plots shown in Figures 5.17 and 5.18 show the comparison between the tag-and-probe method to calculate electron PID efficiencies, described in Section 5.7, and the standard PIDCalib method using sWeighted data. Figures B.1 and B.2 show the relative difference, in percentage points, of the PIDCalib efficiencies with respect to the ones calculated with the tag-and-probe method. For most of the cut values, the difference is not larger than 2%.

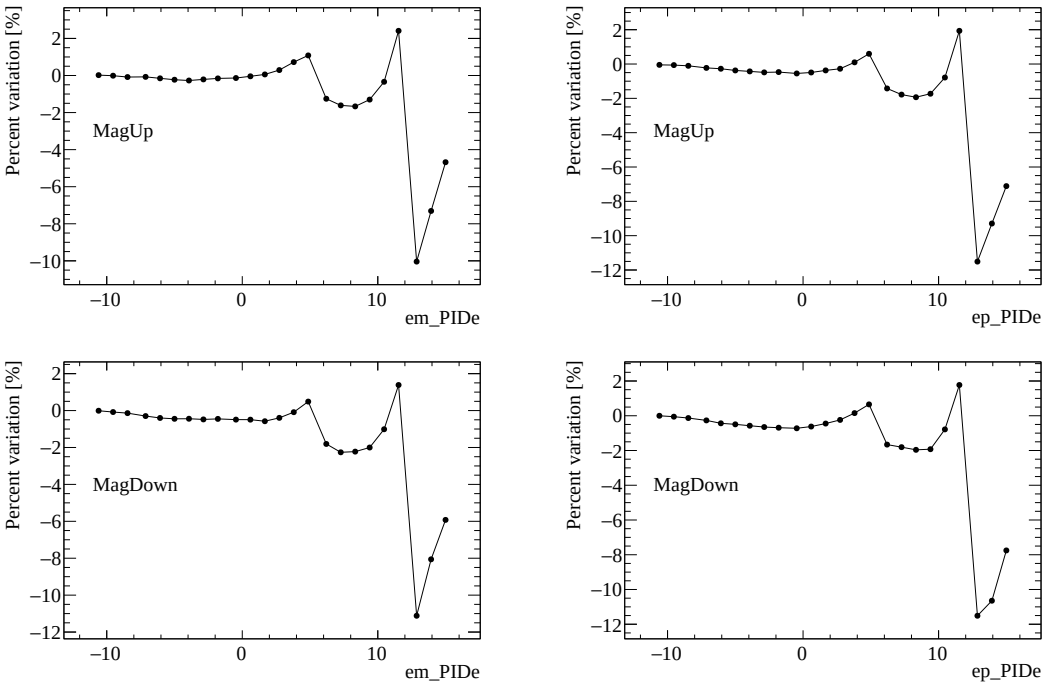


Figure B.1: Percent variation of the PIDCalib efficiencies with respect to the ones calculated with the tag-and-probe method, for both magnet polarities and both values of charge of the probe. No bremsstrahlung photons are added.

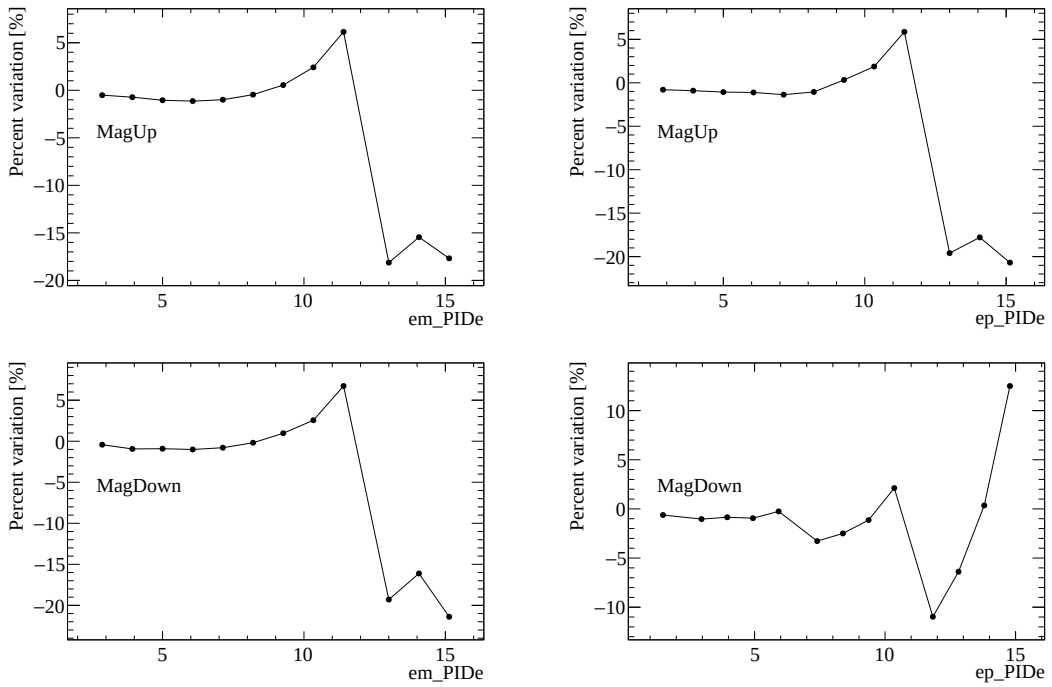


Figure B.2: Percent variation of the PIDCalib efficiencies with respect to the ones calculated with the tag-and-probe method, for both magnet polarities and both values of charge of the probe. Bremsstrahlung photons are added.

B.2 Fits to the J/ψ mass for the PID efficiency tables

This section contains all the fits, for events that pass or are rejected by the PID cut, separated by the HasBremAdded flag, for each bin of Figure 5.19. For practical purposes, the binning scheme is numbered according to the convention shown in Figure B.3 and the captions of the figures of the individual fits indicate the bin to which the fit refers. For all the fits, the red vertical lines indicate the smallest interval containing 90% of the signal PDF, which is determined, for each bin, from the sample of events passing the PID cut; the blue line represents the total PDF, while the background component is shown by a black dotted line and the signal component by a red line.

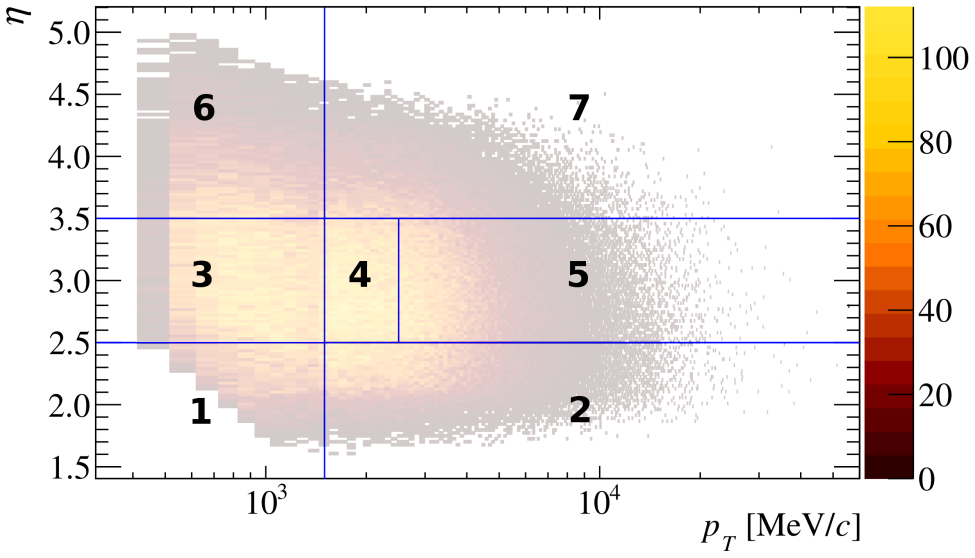


Figure B.3: Numbered binning scheme.

In the following pages, Figures from B.4 to B.17 refer to the sample with no bremsstrahlung added, while Figures from B.18 to B.31 refer to the sample with bremsstrahlung added. On each page, the figure on top shows the fit to the events that pass the PID cut for a given bin, while the figure at the bottom shows the fit to the events that are rejected by the PID cut for the same bin.

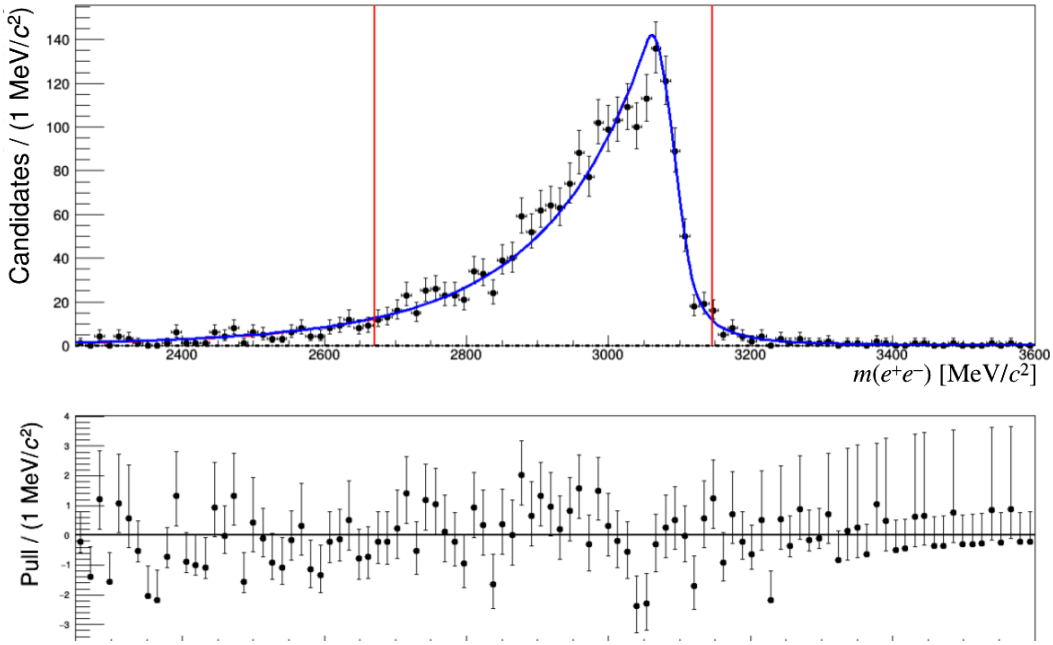


Figure B.4: Fit to the e^+e^- mass for bin 1, no bremsstrahlung added, PID$e > 0$.

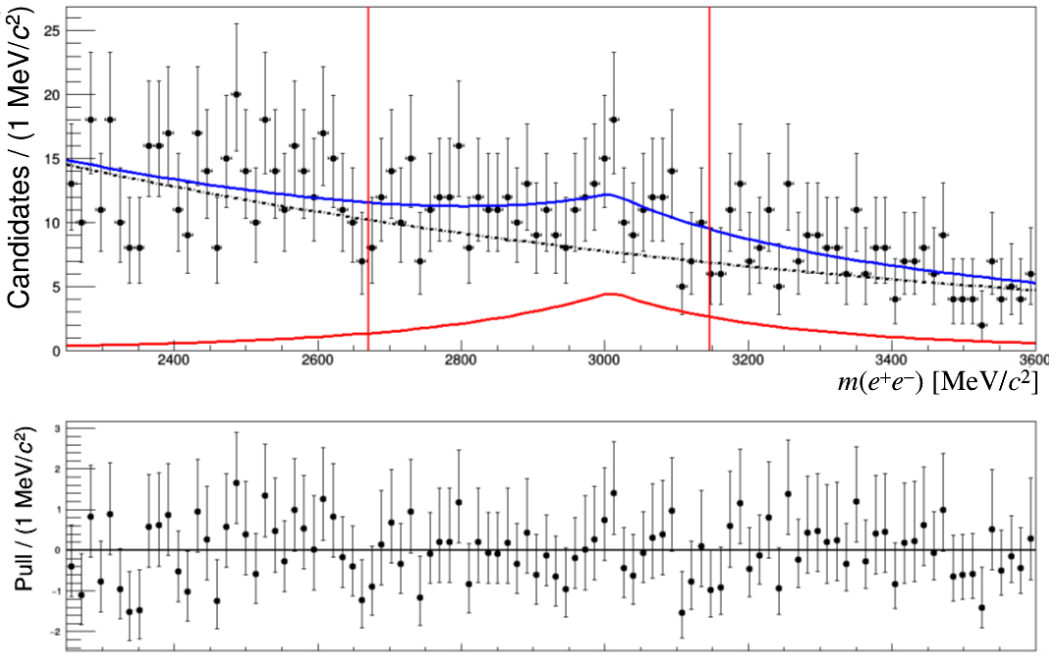


Figure B.5: Fit to the e^+e^- mass for bin 1, no bremsstrahlung added, PID$e < 0$.

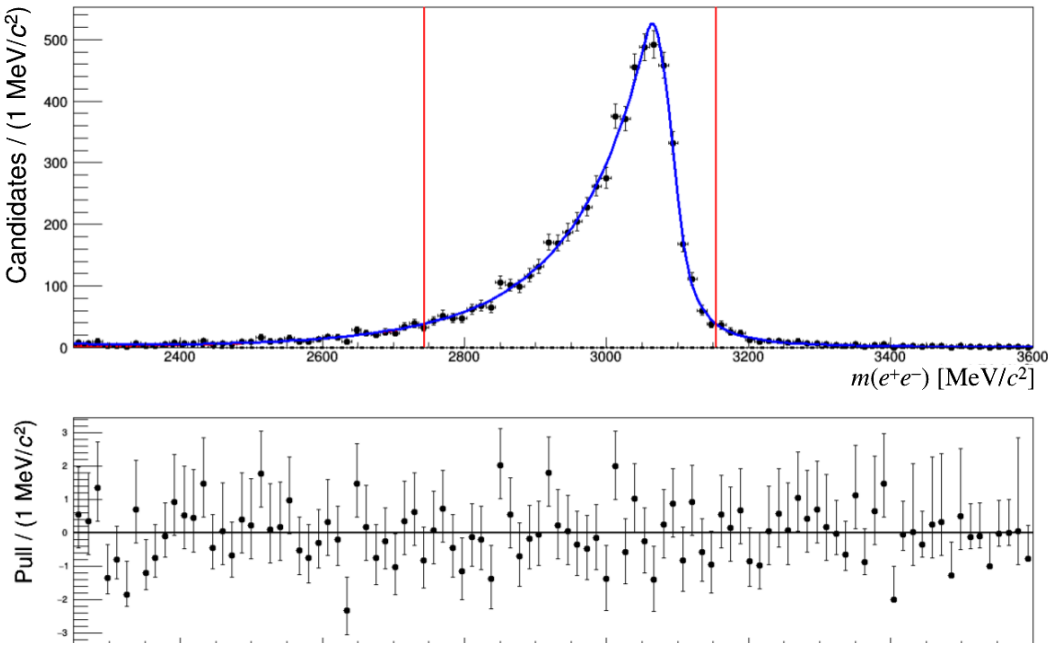


Figure B.6: Fit to the e^+e^- mass for bin 2, no bremsstrahlung added, PID$e > 0$.

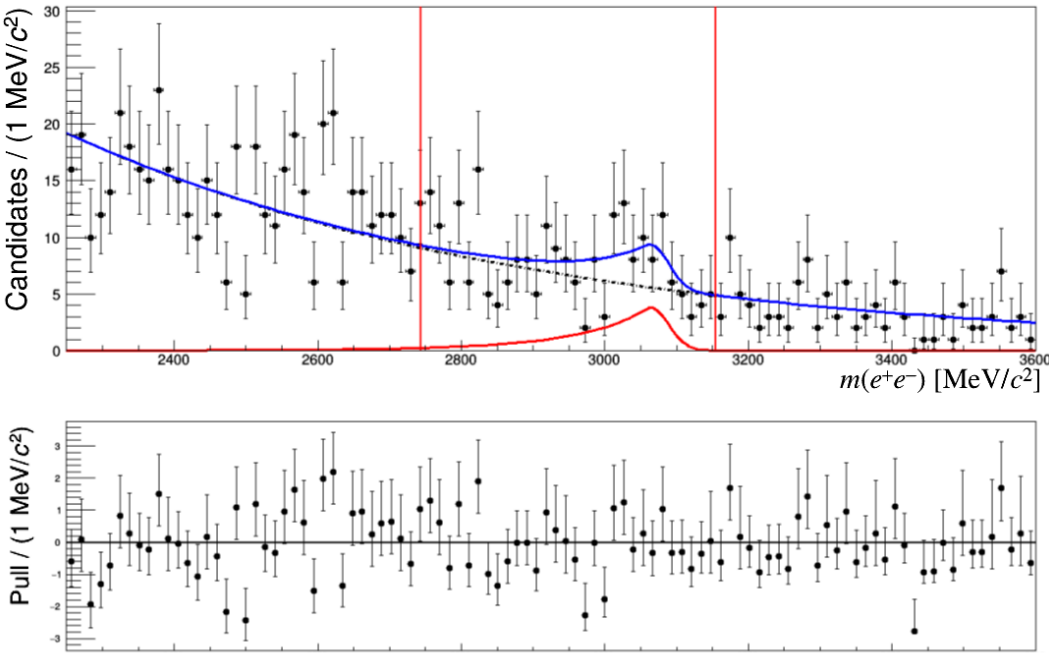


Figure B.7: Fit to the e^+e^- mass for bin 2, no bremsstrahlung added, PID$e < 0$.

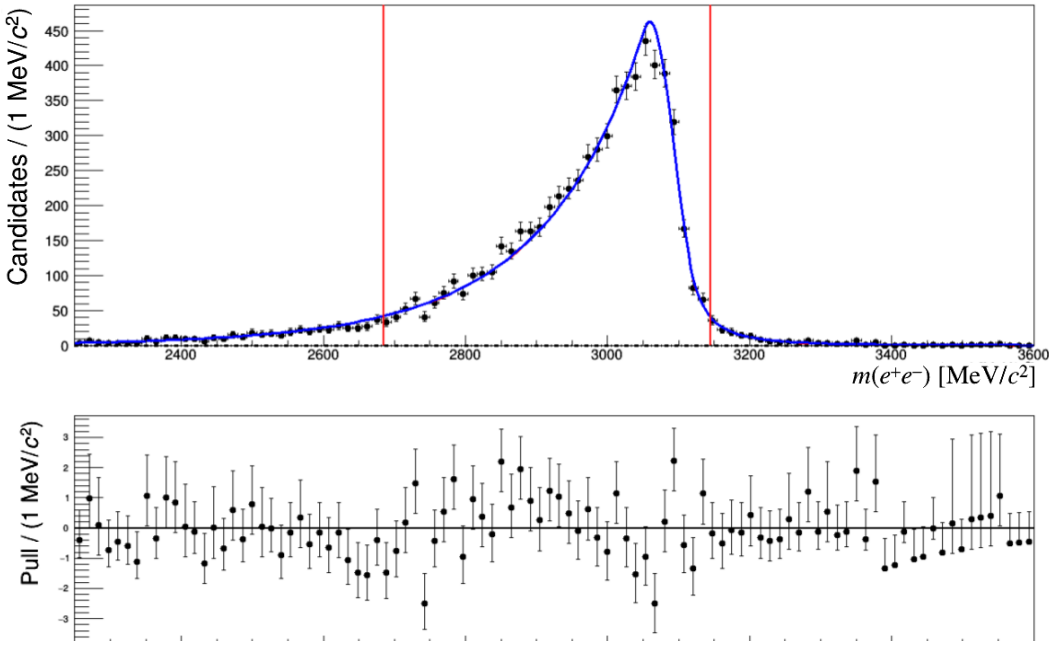


Figure B.8: Fit to the e^+e^- mass for bin 3, no bremsstrahlung added, PID$e > 0$.

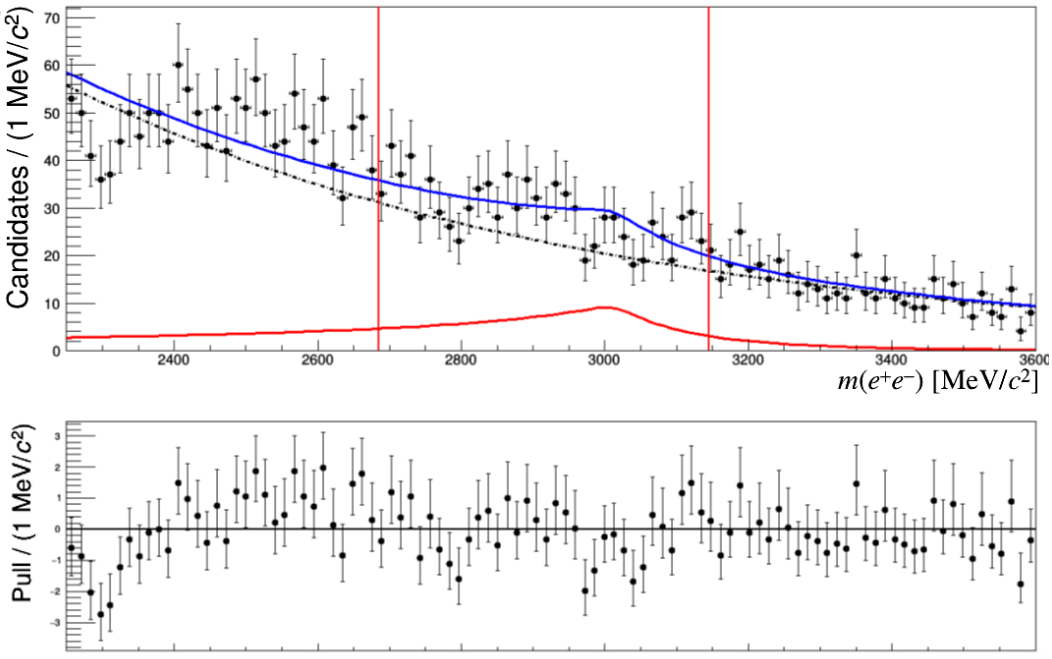


Figure B.9: Fit to the e^+e^- mass for bin 3, no bremsstrahlung added, PID$e < 0$.

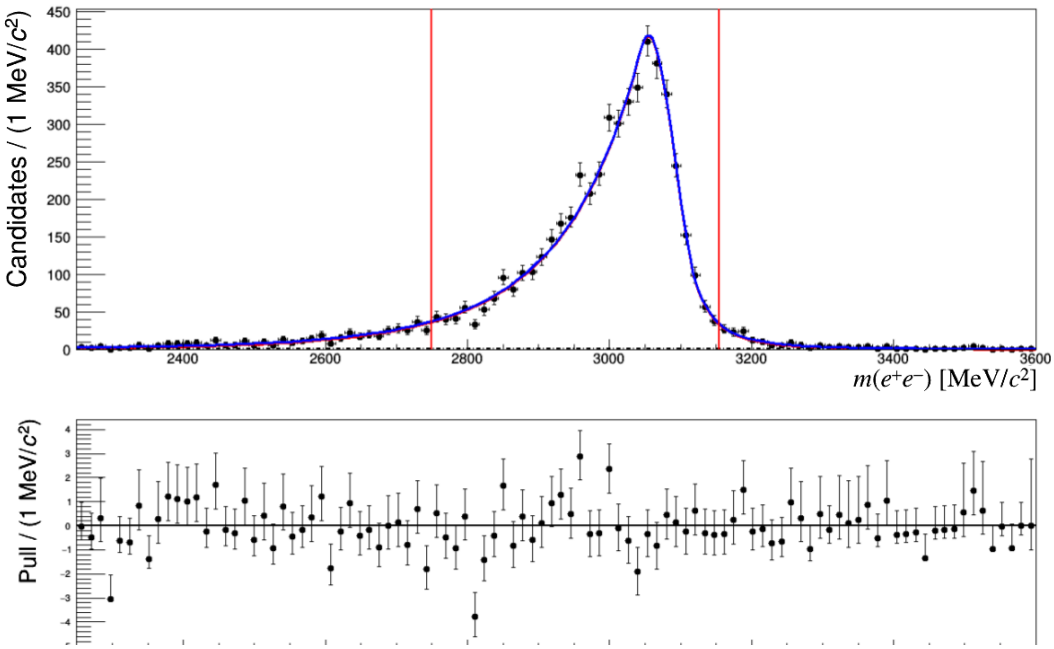


Figure B.10: Fit to the e^+e^- mass for bin 4, no bremsstrahlung added, PID$e > 0$.

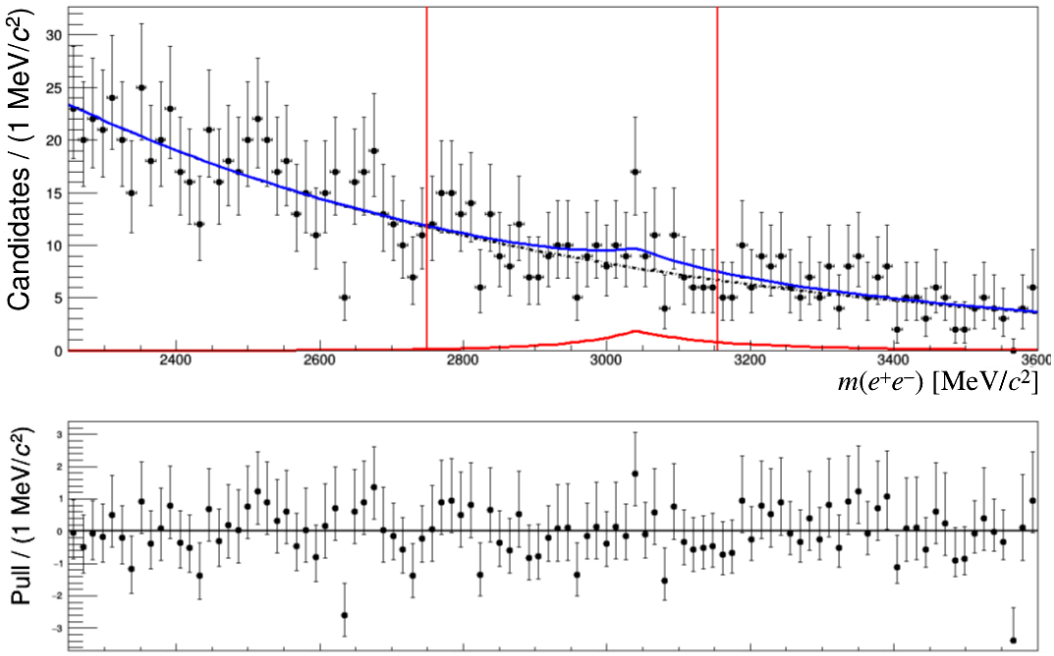


Figure B.11: Fit to the e^+e^- mass for bin 4, no bremsstrahlung added, PID$e < 0$.

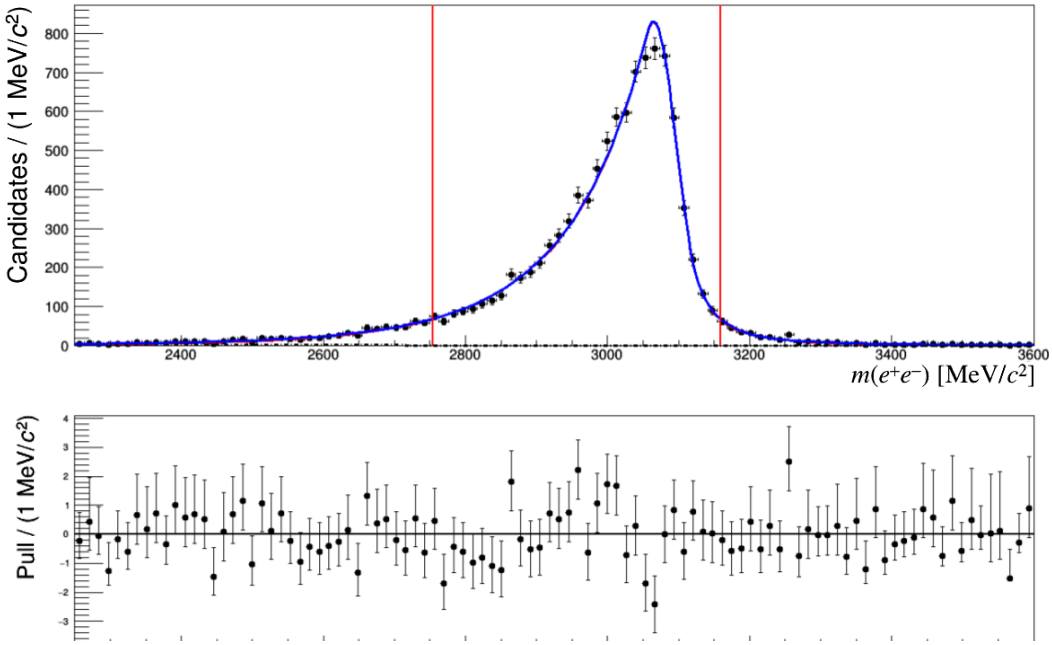


Figure B.12: Fit to the e^+e^- mass for bin 5, no bremsstrahlung added, PID$e > 0$.

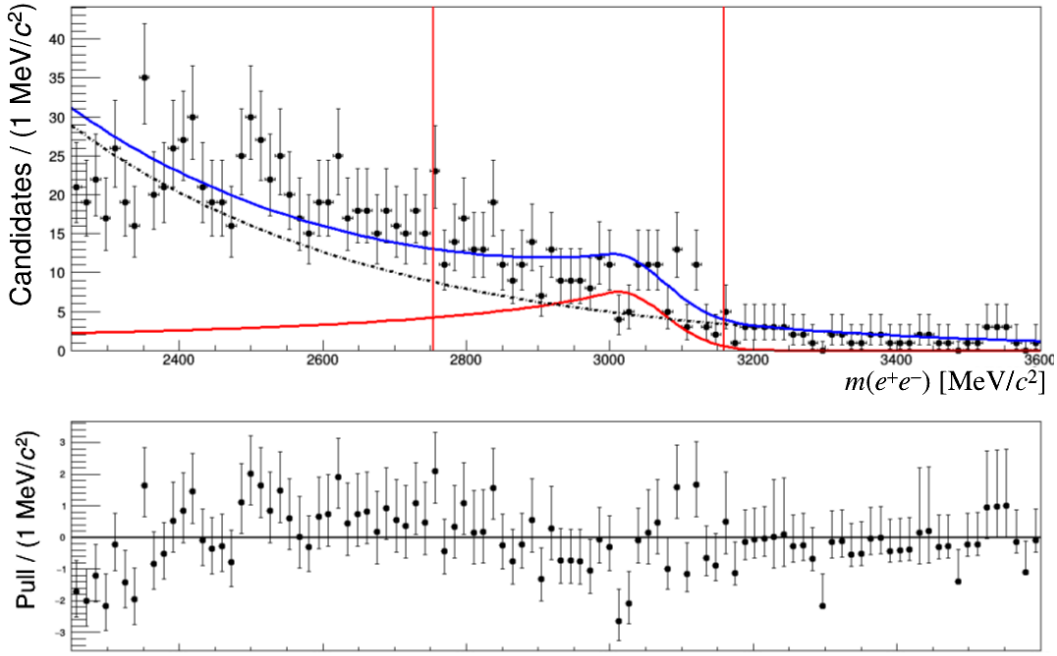


Figure B.13: Fit to the e^+e^- mass for bin 5, no bremsstrahlung added, PID$e < 0$.

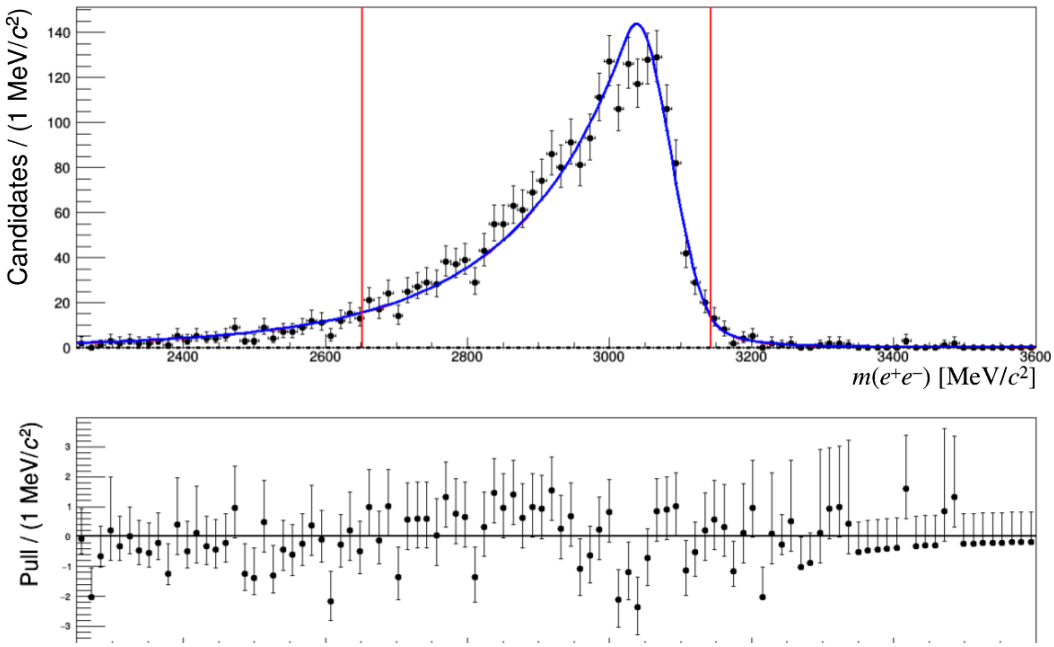


Figure B.14: Fit to the e^+e^- mass for bin 6, no bremsstrahlung added, PID$e > 0$.

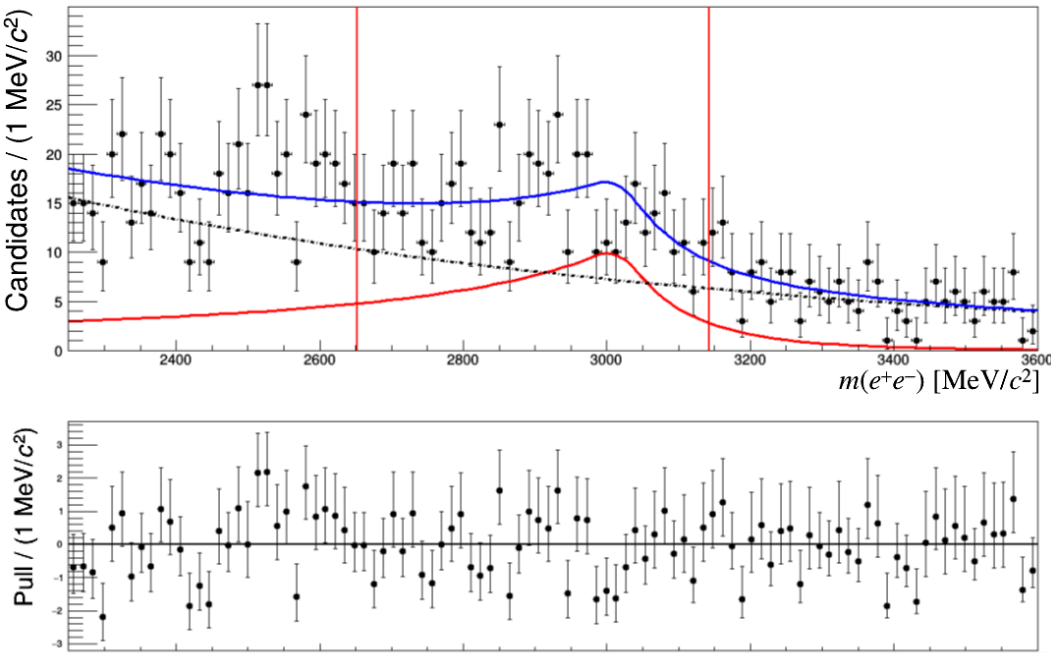


Figure B.15: Fit to the e^+e^- mass for bin 6, no bremsstrahlung added, PID$e < 0$.

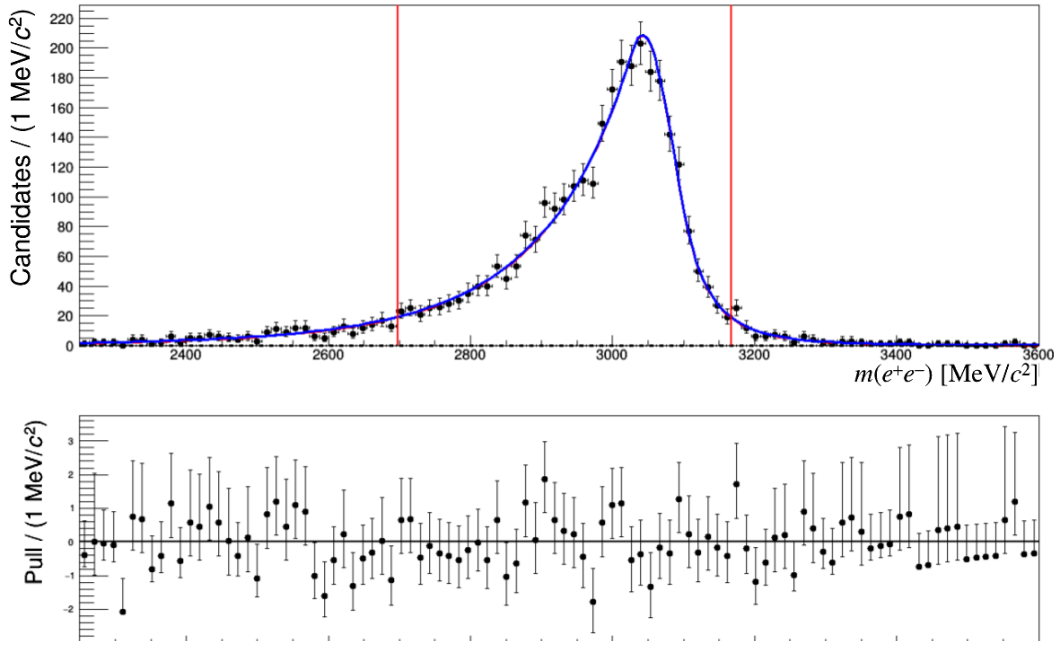


Figure B.16: Fit to the e^+e^- mass for bin 7, no bremsstrahlung added, PID$e > 0$.

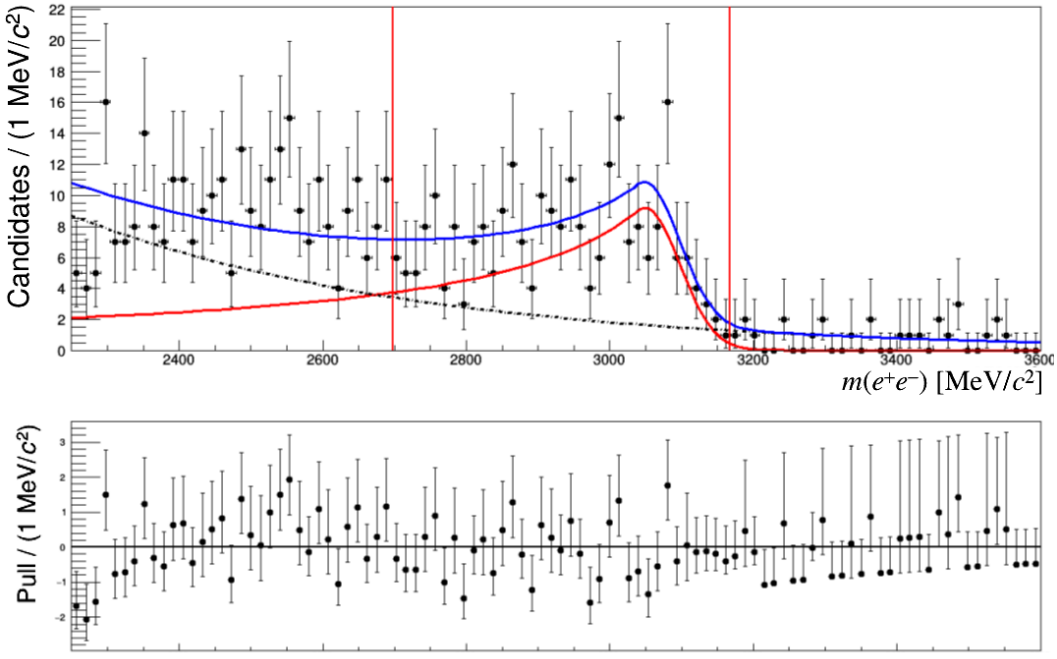


Figure B.17: Fit to the e^+e^- mass for bin 7, no bremsstrahlung added, PID$e < 0$.

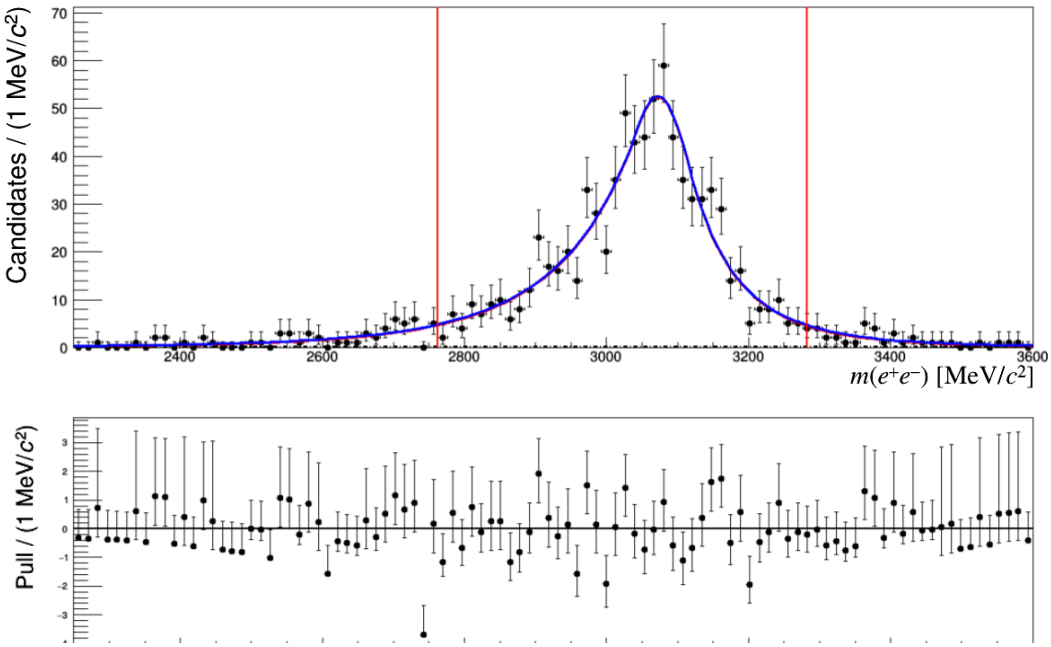


Figure B.18: Fit to the e^+e^- mass for bin 1, bremsstrahlung added, PID$e > 0$.

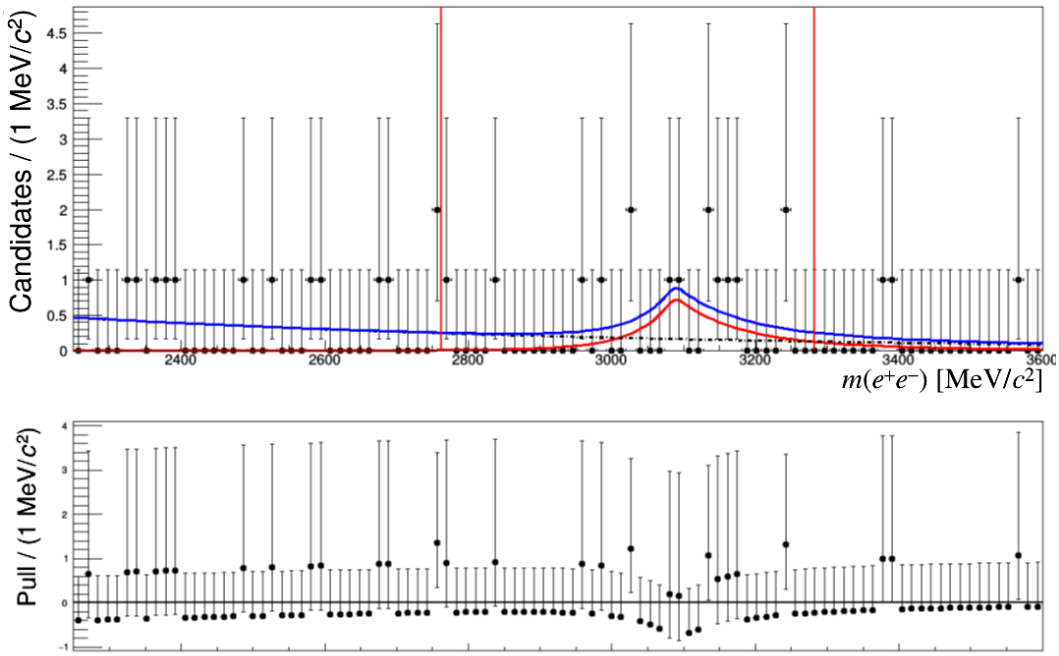


Figure B.19: Fit to the e^+e^- mass for bin 1, bremsstrahlung added, PID$e < 0$.

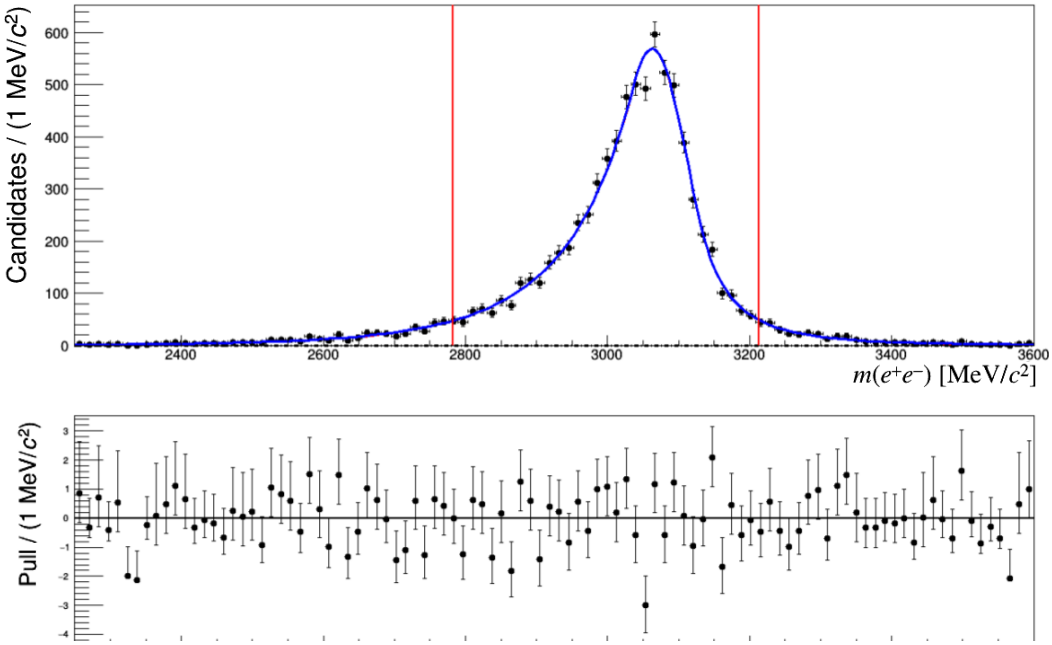


Figure B.20: Fit to the e^+e^- mass for bin 2, bremsstrahlung added, PID$e > 0$.

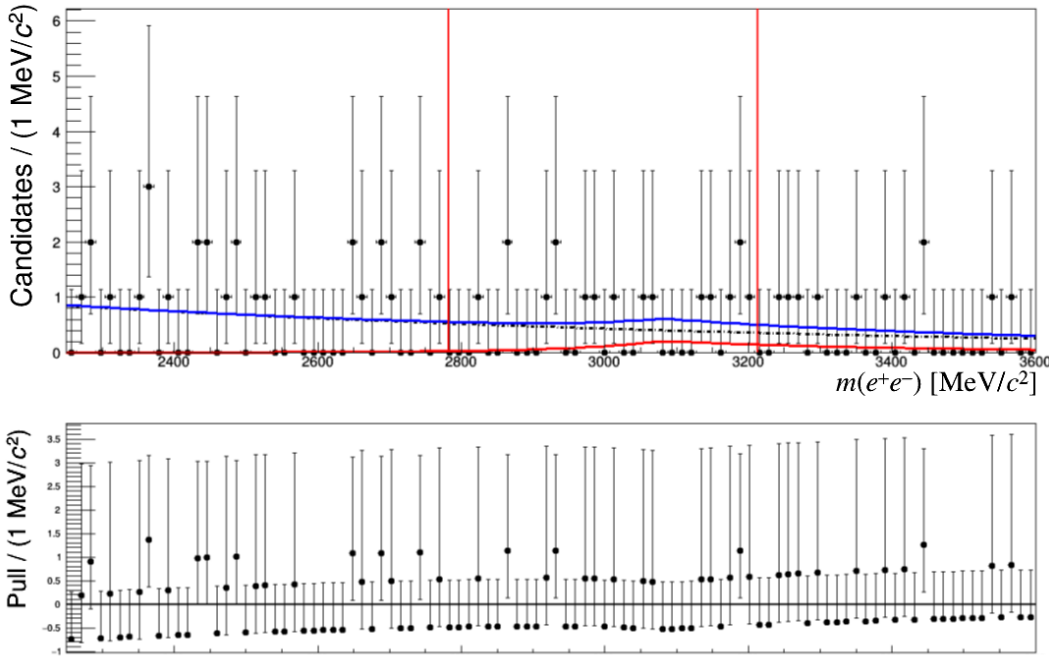


Figure B.21: Fit to the e^+e^- mass for bin 2, bremsstrahlung added, PID$e < 0$.

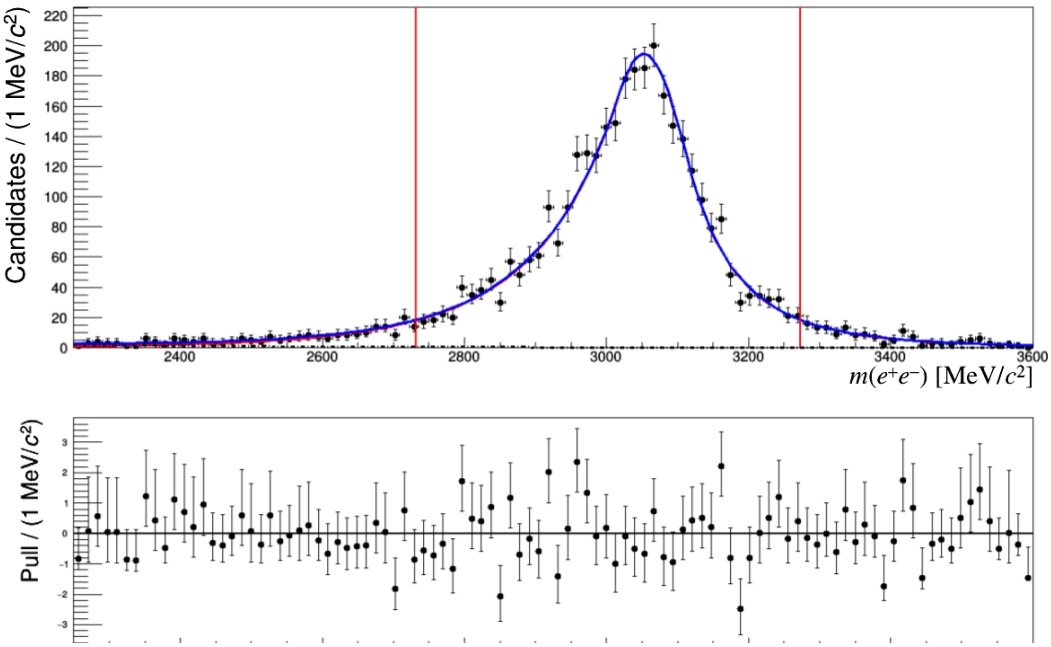


Figure B.22: Fit to the e^+e^- mass for bin 3, bremsstrahlung added, PID$e > 0$.

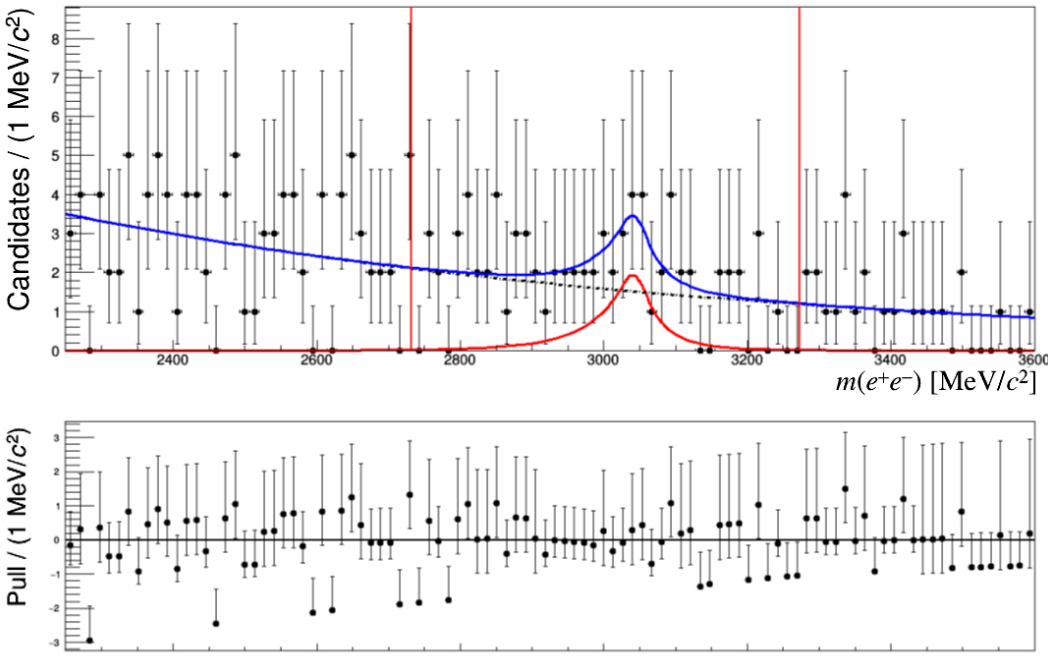


Figure B.23: Fit to the e^+e^- mass for bin 3, bremsstrahlung added, PID$e < 0$.

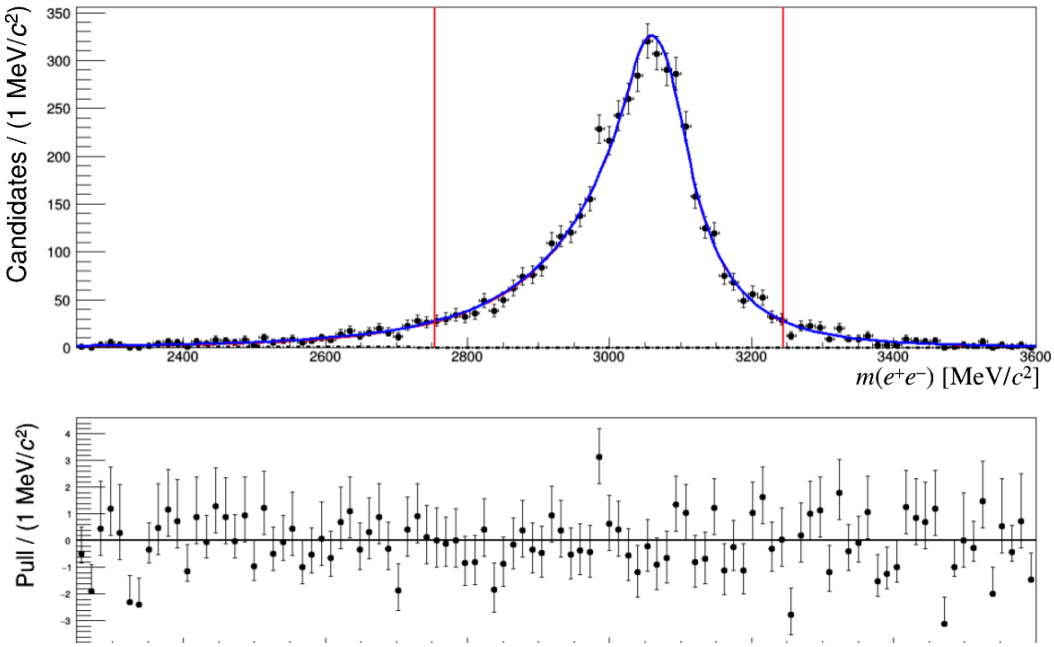


Figure B.24: Fit to the e^+e^- mass for bin 4, bremsstrahlung added, PID$e > 0$.

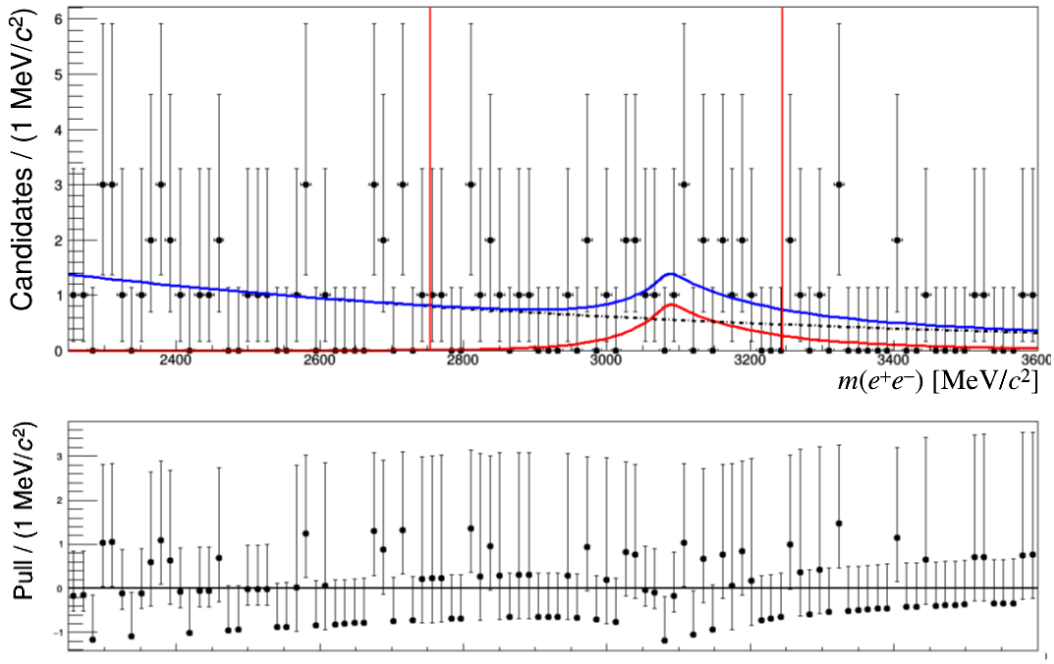


Figure B.25: Fit to the e^+e^- mass for bin 4, bremsstrahlung added, PID$e < 0$.

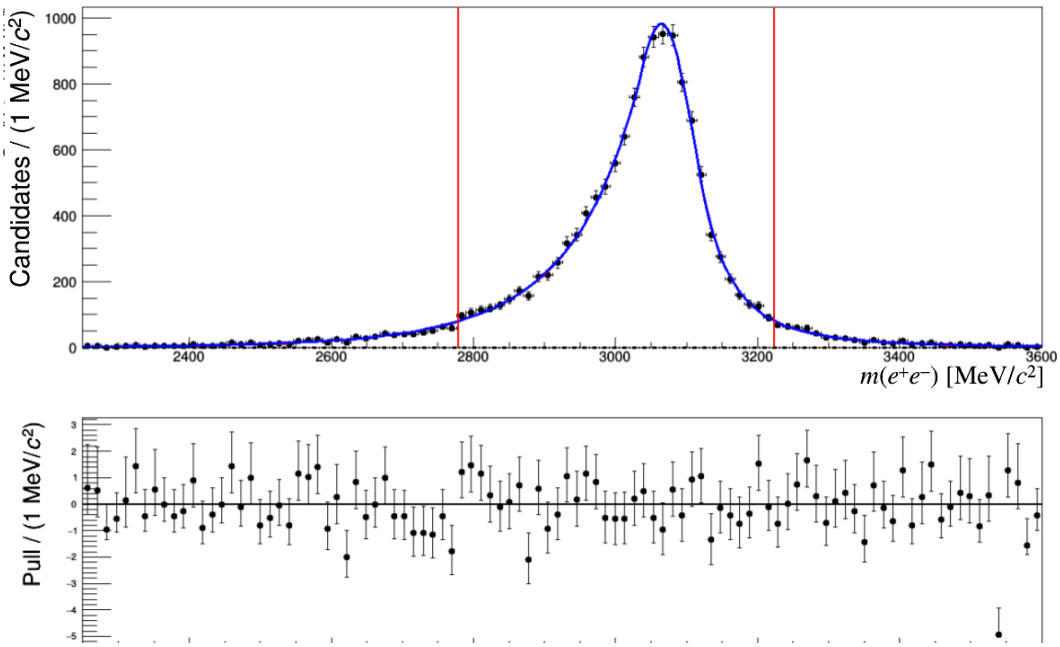


Figure B.26: Fit to the e^+e^- mass for bin 5, bremsstrahlung added, PID$e > 0$.

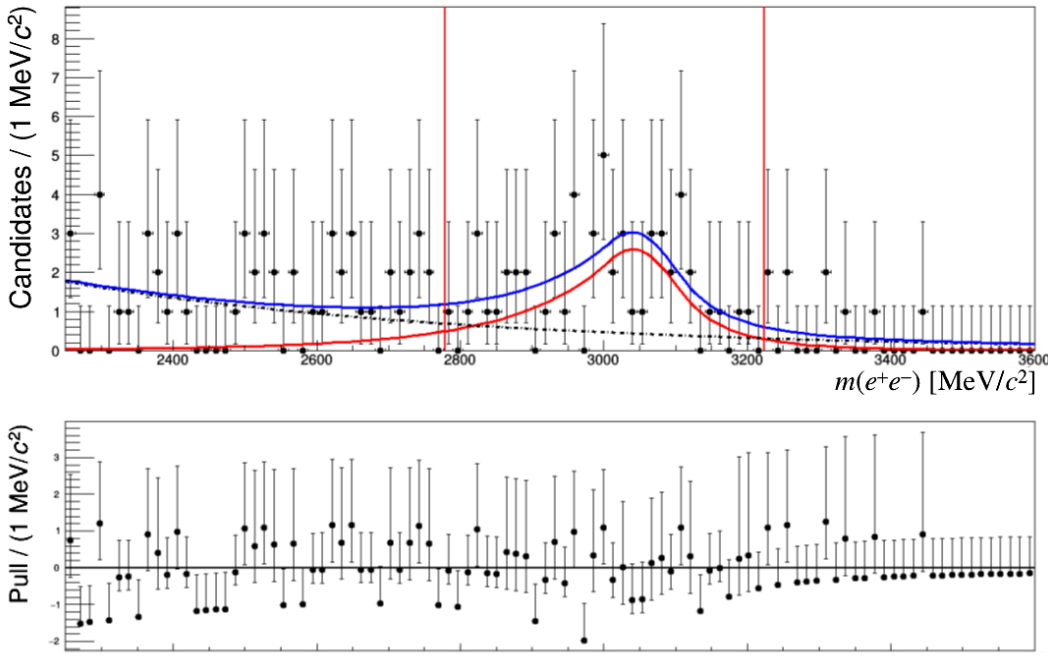


Figure B.27: Fit to the e^+e^- mass for bin 5, bremsstrahlung added, PID$e < 0$.

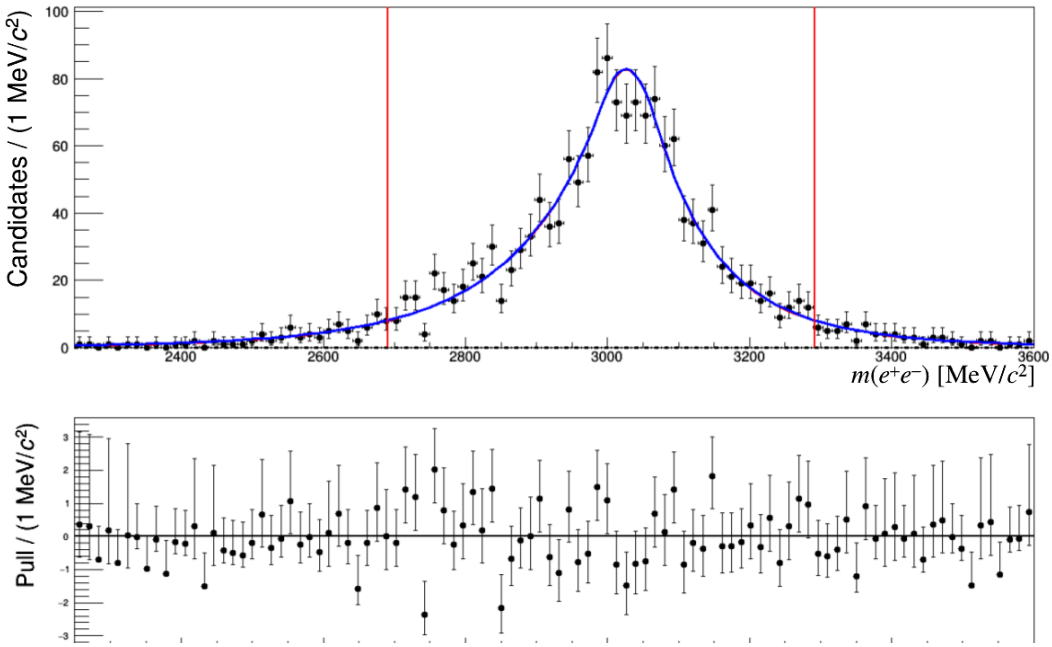


Figure B.28: Fit to the e^+e^- mass for bin 6, bremsstrahlung added, PID$e > 0$.

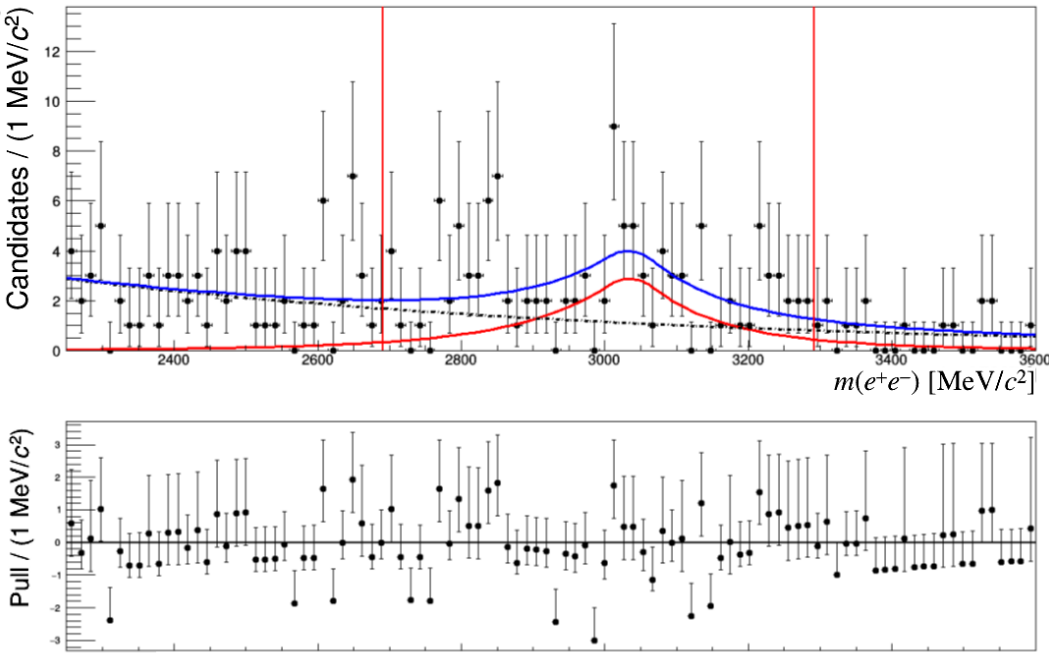


Figure B.29: Fit to the e^+e^- mass for bin 6, bremsstrahlung added, PID$e < 0$.

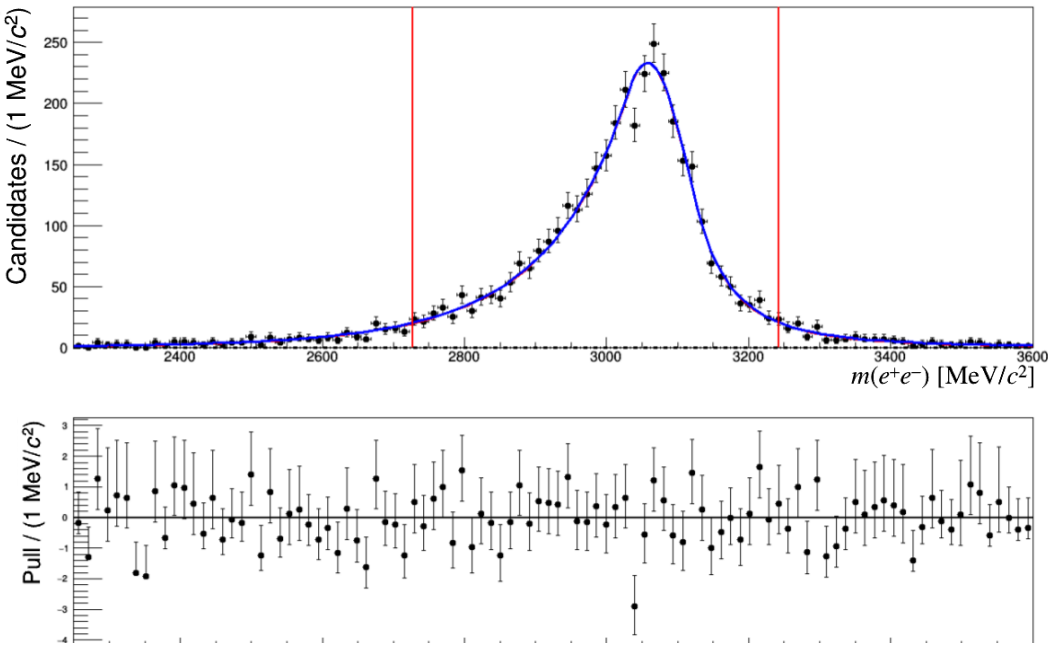


Figure B.30: Fit to the e^+e^- mass for bin 7, bremsstrahlung added, PID$e > 0$.

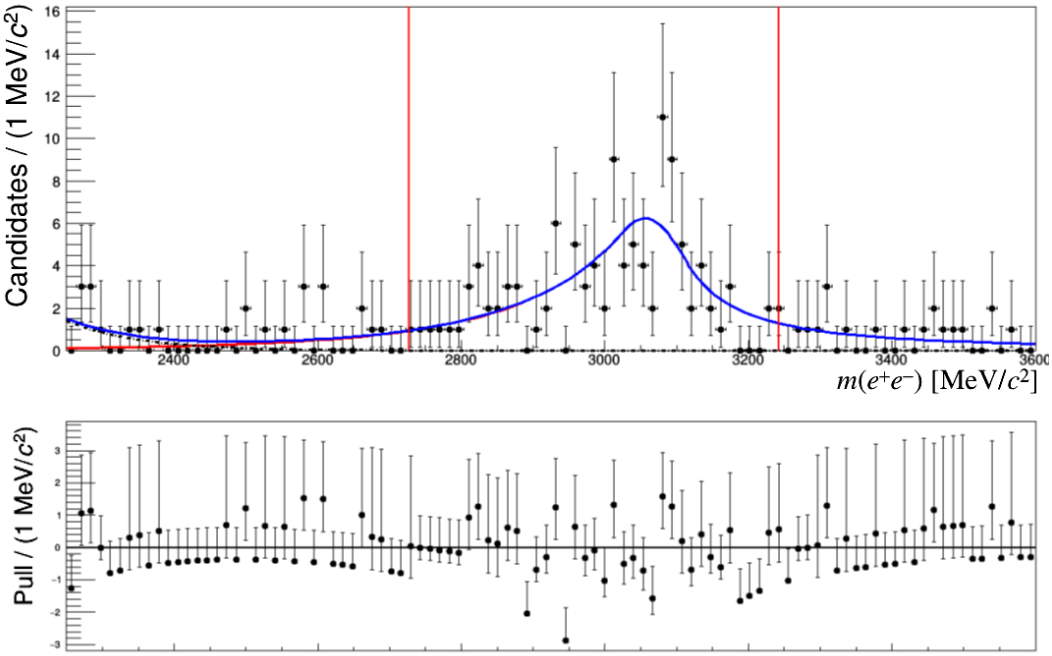


Figure B.31: Fit to the e^+e^- mass for bin 7, bremsstrahlung added, PID$e < 0$.

Bibliography

[1] LHCb collaboration, R. Aaij *et al.*, *Observation of the resonant character of the $Z(4430)^-$ state*, Phys. Rev. Lett. **112** (2014), no. 22 222002, arXiv:1404.1903.

[2] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays*, Phys. Rev. Lett. **115** (2015) 072001, arXiv:1507.03414.

[3] M. Gell-Mann, *Symmetries of baryons and mesons*, Phys. Rev. **125** (1962) 1067.

[4] M. Borsato, *Study of the $B^0 \to K^{*0} e^+ e^-$ decay with the LHCb detector and development of a novel concept of PID detector: the Focusing DIRC*, PhD thesis, Santiago de Compostela U., 2015.

[5] E. Fermi, *Tentativo di una teoria dei raggi β*, Il Nuovo Cimento (1924-1942) **11** (2008) 1.

[6] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.

[7] C. N. Yang and R. L. Mills, *Conservation of isotopic spin and isotopic gauge invariance*, Phys. Rev. **96** (1954) 191.

[8] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961), no. 4 579 .

[9] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[10] A. Salam and J. C. Ward, *Weak and electromagnetic interactions*, Il Nuovo Cimento (1955-1965) **11** (1959) 568.

[11] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[12] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

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

[14] M. Kobayashi and T. Maskawa, *Cp-violation in the renormalizable theory of weak interaction*, Progress of Theoretical Physics **49** (1973), no. 2 652.

[15] Super-Kamiokande collaboration, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562.

[16] SNO collaboration, Q. R. Ahmad *et al.*, *Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002) 011301.

[17] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, Progress of Theoretical Physics **28** (1962), no. 5 870.

[18] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π decay of the k_2^0 meson*, Phys. Rev. Lett. **13** (1964) 138.

[19] *List of publications*, `http://inspirehep.net/search?p=collaboration:%27PAMELA%27`.

[20] *List of publications*, `http://inspirehep.net/search?p=collaboration:%27AMS%27`.

[21] WMAP, C. L. Bennett *et al.*, *First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Preliminary maps and basic results*, Astrophys. J. Suppl. **148** (2003) 1, `arXiv:astro-ph/0302207`.

[22] A. D. Sakharov, *Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32, [Usp. Fiz. Nauk161,61(1991)].

[23] L. Canetti, M. Drewes, and M. Shaposhnikov, *Matter and Antimatter in the Universe*, New J. Phys. **14** (2012) 095012, `arXiv:1204.4186`.

[24] D. V. Perepelitsa, *Sakharov Conditions for Baryogenesis*, `http://phys.columbia.edu/~dvp/dvp-sakharov.pdf`, 2008.

[25] F. R. Klinkhamer and N. S. Manton, *A saddle-point solution in the weinberg-salam theory*, Phys. Rev. D **30** (1984) 2212.

[26] C. Jarlskog, *Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal* CP *nonconservation*, Phys. Rev. Lett. **55** (1985) 1039.

[27] M. B. Gavela, P. Hernandez, J. Orloff, and O. Pene, *Standard model CP violation and baryon asymmetry*, Mod. Phys. Lett. **A9** (1994) 795, arXiv:hep-ph/9312215.

[28] R. D. Peccei, *The Strong CP problem and axions*, Lect. Notes Phys. **741** (2008) 3, arXiv:hep-ph/0607268, [3(2006)].

[29] S. Yu. Khlebnikov and M. E. Shaposhnikov, *Extra Space-time Dimensions: Towards a Solution to the Strong CP Problem*, Phys. Lett. **B203** (1988) 121.

[30] D. R. Nelson, G. T. Fleming, and G. W. Kilcup, *Is strong CP due to a massless up quark?*, Phys. Rev. Lett. **90** (2003) 021601, arXiv:hep-lat/0112029.

[31] R. G. Leigh, *The Strong CP problem, string theory and the Nelson-Barr mechanism*, in *International Workshop on Recent Advances in the Superworld Woodlands, Texas, April 13-16, 1993*, 1993. arXiv:hep-ph/9307214.

[32] R. N. M. Zurab G.B̃erezhiani and G. Senjanovic, *Planck scale physics and solutions to the strong CP problem without axion*, Phys. Rev. **D47** (1993) 5565, arXiv:hep-ph/9212318.

[33] A. A. Anselm and A. A. Johansen, *Can electroweak theta term be observable?*, Nucl. Phys. **B412** (1994) 553, arXiv:hep-ph/9305271.

[34] *Search for the CP-violating strong decays $\eta \to \pi^+\pi^-$ and $\eta'(958) \to \pi^+\pi^-$*, Physics Letters B **764** (2017) 233 .

[35] J. Kuckei *et al.*, *Strong CP violation and the neutron electric dipole form-factor*, Phys. Atom. Nucl. **70** (2007) 349, arXiv:hep-ph/0510116.

[36] P. Krekora, Q. Su, and R. Grobe, *Electric-field-induced relativistic larmor-frequency reduction*, Phys. Rev. A **66** (2002) 013405.

[37] P. J. Mohr, D. B. Newell, and B. N. Taylor, *CODATA Recommended Values of the Fundamental Physical Constants: 2014*, Rev. Mod. Phys. **88** (2016), no. 3 035009, arXiv:1507.07956.

[38] J. M. Pendlebury *et al.*, *Revised experimental upper limit on the electric dipole moment of the neutron*, Phys. Rev. **D92** (2015), no. 9 092003, arXiv:1509.04411.

[39] C. Jarlskog and E. Shabalin, *How large are the rates of the CP violating $\eta, \eta' \to \pi\pi$ decays?*, Phys. Rev. **D52** (1995) 248.

[40] KLOE collaboration, F. Ambrosino *et al.*, *Upper limit on the $\eta \to \pi^+\pi^-$ branching ratio with the KLOE detector*, Phys. Lett. **B606** (2005) 276, arXiv:hep-ex/0411030.

[41] BESIII collaboration, M. Ablikim *et al.*, *Search for CP and P violating pseudoscalar decays into $\pi\pi$*, Phys. Rev. **D84** (2011) 032006, arXiv:1106.5118.

[42] T. K. Collaboration *et al.*, *"$|v_{us}|$ and lepton universality from kaon decays with the kloe detector"*, Journal of High Energy Physics **2008** (2008), no. 04 059.

[43] NA62, E. Goudzovski, *Lepton flavour universality test at the CERN NA62 experiment*, Nucl. Phys. Proc. Suppl. **210-211** (2011) 163, arXiv:1008.1219.

[44] PEN, D. Počanić *et al.*, *PEN: a low energy test of lepton universality*, PoS **HQL2016** (2017) 042, arXiv:1701.05254.

[45] BESIII Collaboration, M. Ablikim *et al.*, *Precision measurement of the mass of the τ lepton*, Phys. Rev. D **90** (2014) 012001.

[46] G. Ciezarek *et al.*, *A Challenge to Lepton Universality in B Meson Decays*, Nature **546** (2017) 227, arXiv:1703.01766.

[47] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. **C40** (2016), no. 10 100001.

[48] LHCb collaboration, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetries of $B \to K^{(*)}\mu^+\mu^-$ decays*, JHEP **06** (2014) 133, arXiv:1403.8044.

[49] LHCb collaboration, R. Aaij *et al.*, *Angular analysis and differential branching fraction of the decay $B_s^0 \to \phi\mu^+\mu^-$*, JHEP **09** (2015) 179, arXiv:1506.08777.

[50] LHCb collaboration, R. Aaij *et al.*, *Measurements of the S-wave fraction in $B^0 \to K^+\pi^-\mu^+\mu^-$ decays and the $B^0 \to K^*(892)^0\mu^+\mu^-$ differential branching fraction*, JHEP **11** (2016) 047, arXiv:1606.04731, [Erratum: JHEP04,142(2017)].

[51] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays*, JHEP **06** (2015) 115, arXiv:1503.07138.

[52] W. Altmannshofer *et al.*, *Symmetries and Asymmetries of $B \to K^*\mu^+\mu^-$ Decays in the Standard Model and Beyond*, JHEP **01** (2009) 019, arXiv:0811.1214.

[53] LHCb collaboration, R. Aaij *et al.*, *Measurement of Form-Factor-Independent Observables in the Decay $B^0 \to K^{*0}\mu^+\mu^-$*, Phys. Rev. Lett. **111** (2013) 191801, arXiv:1308.1707.

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

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

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

[57] J. F. Gunion and H. E. Haber, *Higgs bosons in supersymmetric models (I)*, Nuclear Physics B **272** (1986), no. 1 1 .

[58] S. M. Boucenna *et al.*, *Phenomenology of an SU(2) × SU(2) × U(1) model with lepton-flavour non-universality*, Journal of High Energy Physics **2016** (2016) 59.

[59] A. Abada *et al.*, *Tree-level lepton universality violation in the presence of sterile neutrinos: impact for R_K and R_π*, Journal of High Energy Physics **2013** (2013) 48.

[60] P. Langacker, *The physics of heavy $Z^{'}$ gauge bosons*, Rev. Mod. Phys. **81** (2009) 1199.

[61] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, *Quark flavor transitions in $L_\mu - L_\tau$ models*, Phys. Rev. D **89** (2014) 095033.

[62] R. Gauld, F. Goertz, and U. Haisch, *An explicit Z' boson explanation of the $B \to K^* \mu^+ \mu^-$ anomaly*, JHEP **01** (2014) 069, arXiv:1310.1082.

[63] D. Becirevic, S. Fajfer, N. Kosnik, and O. Sumensari, *Leptoquark model to explain the B-physics anomalies, R_K and R_D*, Phys. Rev. **D94** (2016), no. 11 115021, arXiv:1608.08501.

[64] G. Hiller and M. Schmaltz, *R_K and future $b \to s\ell\ell$ physics beyond the standard model opportunities*, Phys. Rev. D **90** (2014) 054014.

[65] A. Crivellin, D. Muller, and T. Ota, *Simultaneous explanation of $R(D^{(*)})$ and $b \to s\mu^+\mu^-$: the last scalar leptoquarks standing*, JHEP **09** (2017) 040, arXiv:1703.09226.

[66] BABAR Collaboration, J. P. Lees *et al.*, *Evidence for an Excess of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ Decays*, Phys. Rev. Lett. **109** (2012) 101802.

[67] Belle, M. Huschle *et al.*, *Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*, Phys. Rev. **D92** (2015), no. 7 072014, arXiv:1507.03233.

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

[69] Belle, S. Hirose *et al.*, *Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$*, Phys. Rev. Lett. **118** (2017), no. 21 211801, arXiv:1612.00529.

[70] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$*, Phys. Rev. Lett. **115** (2015), no. 11 111803, arXiv:1506.08614, [Erratum: Phys. Rev. Lett.115,no.15,159901(2015)].

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

[72] Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, The European Physical Journal C **77** (2017) 895.

[73] A. Davis. Private communication with S. Fajfer, 2016.

[74] O. S. Brüning *et al.*, *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, Geneva, 2004.

[75] L. Capriotti, *Search for the decay X(3872)→ $J/\psi\omega$ in the LHCb experiment at CERN*, Master's thesis, Sapienza University of Rome, 2014.

[76] R. Scrivens *et al.*, *Overview of the status and developments on primary ion sources at CERN*, .

[77] LHCb collaboration, R. Aaij *et al.*, *Precision luminosity measurements at LHCb*, JINST **9** (2014), no. 12 P12005, arXiv:1410.0149.

[78] F. Marcastel, *CERN's Accelerator Complex. La chaîne des accélérateurs du CERN*, , General Photo.

[79] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030. LHCb-TDR-009.

[80] R. Lindner, *LHCb layout. LHCb schema*, LHCb Collection., Feb, 2008.

[81] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[82] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011. LHCb-TDR-005.

[83] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) 09007, arXiv:1405.7808.

[84] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037. LHCb-TDR-003.

[85] LHCb RICH Group, M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759.

[86] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029. LHCb-TDR-008.

[87] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024. LHCb-TDR-006.

[88] G. A. Cowan, *Performance of the LHCb silicon tracker*, Nucl. Instrum. Methods Phys. Res. A **699** (2013) 156 , Proceedings of the 8th International "Hiroshima" Symposium on the Development and Application of Semiconductor Tracking Detectors.

[89] LHCb Outer Tracker Group, R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014), no. 01 P01002, arXiv:1311.3893.

[90] J. Gassner, M. Needham, and O. Steinkamp, *Layout and Expected Performance of the LHCb TT Station*, Tech. Rep. LHCb-2003-140, CERN, Geneva, Apr, 2004.

[91] J. R. Harrison, *Radiation damage studies in the LHCb VELO detector and searches for lepton flavour and baryon number violating tau decays*, PhD thesis, The University of Manchester, 2014.

[92] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007. LHCb-TDR-001.

[93] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036. LHCb-TDR-002.

[94] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010. LHCb-TDR-004.

[95] A. A. Alves, Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, arXiv:1211.1346.

[96] LHCb collaboration, *LHCb trigger system: Technical Design Report*, CERN-LHCC-2003-031. LHCb-TDR-010.

[97] LHCb collaboration, *Trigger schemes*, https://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html.

[98] A. Puig Navarro, *The LHCb trigger in 2011 and 2012*, Tech. Rep. LHCb-PUB-2014-046. CERN-LHCb-PUB-2014-046, CERN, Geneva, Nov, 2014.

[99] A. Puig Navarro, *The LHCb Turbo stream*, Nucl. Instrum. Methods Phys. Res. , A **824** (2016) 38.

[100] R. Forty, *RICH pattern recognition for LHCb*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **433** (1999), no. 1 257 .

[101] H. Bethe, *Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie*, Annalen der Physik **397** (1930), no. 3 325.

[102] LHCb collaboration, M. Fontana, *Study on the performance of the Particle Identification Detectors at LHCb after the LHC First Long Shutdown (LS1)*, PoS **ICHEP2016** (2016) 295.

[103] G. Barrand *et al.*, *GAUDI - The software architecture and framework for building LHCb data processing applications*, in *Proceedings, 11th International Conference on Computing in High-Energy and Nuclear Physics (CHEP 2000): Padua, Italy, February 7-11, 2000*, pp. 92–95, 2000.

[104] LHCb collaboration, *The Gauss Project*, `http://lhcbdoc.web.cern.ch/lhcbdoc/gauss/`.

[105] T. Sjöstrand *et al.*, *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, `arXiv:1410.3012`.

[106] A. Ryd *et al.*, *EvtGen: A Monte Carlo Generator for B-Physics*, `https://evtgen.hepforge.org/doc/EvtGenGuide.pdf`, 2004.

[107] N. Davidson, T. Przedzinski, and Z. Was, *PHOTOS interface in C++: Technical and Physics Documentation*, Comput. Phys. Commun. **199** (2016) 86, `arXiv:1011.0937`.

[108] S. Agostinelli *et al.*, *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003), no. 3 250 .

[109] J. Allison *et al.*, *Recent developments in Geant4*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **835** (2016) 186 .

[110] LHCb collaboration, *The Boole Project*, `http://lhcbdoc.web.cern.ch/lhcbdoc/boole/`.

[111] LHCb collaboration, *The Brunel Project*, `http://lhcbdoc.web.cern.ch/lhcbdoc/brunel/`.

[112] LHCb collaboration, *The Moore Project*, `http://lhcbdoc.web.cern.ch/lhcbdoc/moore/`.

[113] LHCb collaboration, *The DaVinci Project*, `http://lhcbdoc.web.cern.ch/lhcbdoc/davinci/`.

[114] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81.

[115] A. L. Read, *Presentation of search results: The CL_s technique*, J. Phys. G: Nucl. Part. Phys. **28** (2002) 2693.

[116] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, `arXiv:physics/0703039`.

[117] A. P. Bradley, *The use of the area under the ROC curve in the evaluation of machine learning algorithms*, Pattern Recognition **30** (1997), no. 7 1145 .

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

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

[120] WASA-at-COSY collaboration, P. Adlarson *et al.*, *Exclusive measurement of the $\eta \to \pi^+\pi^-\gamma$ decay*, Phys. Lett. **B707** (2012) 243, `arXiv:1107.5277`.

[121] KLOE, D. Babusci *et al.*, *Measurement of $\Gamma(\eta \to \pi^+\pi^-\gamma)/\Gamma(\eta \to \pi^+\pi^-\pi^0)$ with the KLOE Detector*, Phys. Lett. **B718** (2013) 910, `arXiv:1209.4611`.

[122] CLEO collaboration, A. Lopez *et al.*, *Measurement of Prominent η-Decay Branching Fractions*, Phys. Rev. Lett. **99** (2007) 122001.

[123] K. G. Wilson, *Non-lagrangian models of current algebra*, Phys. Rev. **179** (1969) 1499.

[124] J. Barranco, D. Delepine, V. Gonzalez Macias, and L. Lopez-Lozano, *Two Higgs doublet model and leptoquarks constraints from D meson decays*, `arXiv:1404.0454`.

[125] S. Fajfer, I. Nisandzic, and U. Rojec, *Discerning new physics in charm meson leptonic and semileptonic decays*, Phys. Rev. **D91** (2015), no. 9 094009, `arXiv:1502.07488`.

[126] W. E. Johns, *Measurements of the semileptonic decay of the neutral charmed meson $D^0 \to K^- \mu^+$ muon-neutrino*, PhD thesis, Colorado U., 1995.

[127] R. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Computer Physics Communications **77** (1993), no. 2 219 .

[128] L. Anderlini *et al.*, *The PIDCalib package*, Tech. Rep. LHCb-PUB-2016-021. CERN-LHCb-PUB-2016-021, CERN, Geneva, Jul, 2016.

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