Test of Lepton Flavour Universality using the B0 to D* tau nu decays at LHCb

Dawid Gerstel

CPPM-T-2020-02

AIX-MARSEILLE UNIVERSITÉ
Ecole Doctorale 352 : Physique et Sciences de la Matière
Faculté Des Sciences de Luminy
Centre de Physique des Particules de Marseille

Thèse présentée pour obtenir le grade universitaire de docteur

Discipline : Physique et Sciences de la Matière
Spécialité : Physiques des Particules et Astroparticules

Dawid GERSTEL

Test of Lepton Flavour Universality using the $B^0 \to D^{*-}\tau^+\nu_\tau$ decays at LHCb

Soutenue le 30/11/2020 devant le jury :

Karim TRABELSI	IJCLab, Orsay, France	Rapporteur
Isabelle RIPP-BAUDOT	IPHC, Strasbourg, France	Rapporteur
Cristinel DIACONU	CPPM, Marseille, France	Examinateur
Olivier LEROY	CPPM, Marseille, France	Directeur de thèse

To my parents / Rodzicom

“Perceive that which cannot be seen with the eye”
— Miyamoto Musashi, A Book of Five Rings

”Practice any art, music, singing, dancing, acting, drawing, painting, sculpting, poetry, fiction, essays, reportage, no matter how well or badly, not to get money and fame, but to experience becoming, to find out what’s inside you, to make your soul grow.“
— Kurt Vonnegut Jr.

Contents

Abstract

This thesis presents the measurement of the $R(D^*) \equiv \frac{\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)}$ ratio with $2\,\text{fb}^{-1}$ of pp collisions collected at $\sqrt{s} = 13\,\text{TeV}$ by LHCb during 2015-2016 using 3-prong tau decays. The study comprises a test of the Lepton Flavour Universality in $b \to c\ell\nu_\ell$ decays to help resolve the tension between the Standard Model $R(D^*)$ estimation and the experimental results from the B-factories and LHCb. The analysis builds upon a previous LHCb measurement [1, 2] with a new dataset and improved techniques. Most importantly, a novel fast simulation technique, ReDecay, is used to generate large simulated samples and Multivariate Analysis techniques are exploited in signal selection. Since the analysis has yet to undergo an internal LHCb review, and several systematic uncertainties must be computed, the $R(D^*)$ value is blinded. Nonetheless, nearly complete documentation of the analysis is presented in this thesis. The current relative statistical uncertainty on the $R(D^*)$ is 5.56%.

Synthèse

Cette thèse présente la mesure du rapport $R(D^*) \equiv \frac{\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)}$ avec les données accumulées par l'expérience LHCb en 2015 et 2016, en utilisant les désintégrations du tau en trois pions. L'étude comprend un test de l'universalité de la saveur leptonique dans les désintégrations $b \to c\ell\nu_\ell$ pour aider à résoudre la tension entre l'estimation de $R(D^*)$ dans le Modèle Standard et les résultats expérimentaux des usines à B et de LHCb. L'analyse s'appuie sur une mesure précédente LHCb [1, 2] avec un nouvel ensemble de données et des techniques améliorées. En particulier, une nouvelle technique de simulation rapide, ReDecay, est utilisée pour générer de nombreux échantillons simulés ainsi que les techniques d'analyses multivariées sont exploitées dans la sélection des candidats. Étant donné que l'analyse n'a pas encore fait l'objet d'une revue interne dans la collaboration LHCb et que plusieurs incertitudes systématiques doivent être calculées, la valeur du $R(D^*)$ est cachée. Néanmoins, une documentation presque complète de l'analyse est présentée dans cette thèse. L'incertitude statistique relative sur $R(D^*)$ est égale actuellement à 5,56%.

Acknowledgements

First, I would like to thank my supervisor, Olivier Leroy, for entrusting me with this challenging analysis within the LHCb Collaboration. He helped me see the big picture of the $R(D^*)$ analysis whenever I got stuck in the technicalities. (On the other hand, being stuck taught me valuable lessons sometimes!)

I am grateful to Adam Morris for the daily collaboration, patience and ambition to develop good quality software along the guidelines of reproducible research. I have learnt a lot thanks to his countless spot-on suggestions.

I would like to thank all the $R(D^*)$ colleagues, in particular Guy Wormser for the guidance over the project, his clever remarks as well as enthusiasm and optimism despite the obstacles encountered. Also I appreciate friendly help of Antonio Romero Vidal, Victor Renaudin and Davide Fazzini.

I thank Resmi P.K. for the kind and tenacious teamwork in the stressful but productive last months of my contract and I wish her good luck in the review process of the analysis.

I would like to express my appreciation of the support from the CPPM-LHCb group. I thank Julien and Giampi for reading the manuscript; Anton for friendly and informative discussions. Also, I appreciate the help received from Aoife Bharucha from CPT in preparation of the theory chapter.

I thank the LHCb Semileptonic Working Group convenors: Greg, Lucia, Michel and Marcello for enabling me to be a liaison and their feedback for my LHCb and FPCP talks.

A big thank you goes to my PhD friends: Nihel, Cédric, Joan, Nghia, Grig, Rima, Anna, Hichem, Marie, Sylvain, Giovanni, Khan, Jacopo and Aishik for mutual support and conversations helping digest the student canteen lunch menu.

I would like to thank my martial arts instructors: sifu David Delannoy, sifu Eric Sanchez, sifu Greg, sifu Bou; and friends: Adam, Kévin, Reinier, Max, Mélanie, for letting me tread the warrior path and helping me not to put on too many kilograms ;-)

Last but not least, I would like to express my gratitude for the support I have received from my family and girlfriend, Titin.

This thesis has received support from Excellence Initiative of Aix-Marseille Université - A*MIDEX, a French "Investissements d'Avenir programme", AMX-INT-18-020.

Introduction

The approach to explain the matter and its interactions at the most fundamental level is the subject of particle physics. Its best theoretical framework is the Standard Model (SM), formulated in the 70's. The theory has been extraordinarily successful, withstanding the most stringent experimental tests in particle accelerators. Its pinnacle was the discovery of the Higgs boson by ATLAS and CMS in 2012 [3, 4]. Nonetheless, the SM remains an effective theory and fails to explain several fundamental questions. Therefore, quests for New Physics effects have been ongoing for the past decades by either direct searches of new particles or indirectly contradicting its predictions.

One of the assumptions of the SM is the Lepton Flavour Universality (LFU). It states that the electroweak couplings are independent of the lepton flavour, apart from its mass. Consequently, the decay rates of hadrons into (semi)leptonic final states should differ only due to the invariant masses of the leptons involved. Should a greater discrepancy be found, the LFU must be violated.

The LFU has been intensely tested over the recent years. Several tensions with the SM have been found in $b \to c\ell\nu_\ell$ and $b \to s\ell\ell$ channels as measured by the B-factories and LHCb. The usual observables investigated are the ratios of branching fractions of such processes with different charged lepton in the final state, $\ell = e, \mu, \tau$.

This thesis documents the $R(D^*) \equiv \frac{\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)}$ measurement performed with the 2015-2016 LHCb dataset. The project builds upon the LHCb analysis of the 2011-2012 data with three-prong tau decays [1, 2], where the τ^+ is reconstructed in two modes: $\tau^+ \to \pi^+\pi^-\pi^+\nu_\tau$ and $\tau^+ \to \pi^+\pi^-\pi^+\pi^0\nu_\tau$. The measured observable is $\mathcal{K}(D^*)$, defined as the ratio of branching fractions of $B^0 \to D^{*-}\tau^+\nu_\tau$ over a normalisation channel, $B^0 \to D^{*-}3\pi^\pm$, $\mathcal{K}(D^*) \equiv \frac{\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{*-}\pi^+\pi^-\pi^+)}$. $R(D^*)$ is derived from $\mathcal{K}(D^*)$ and the branching fractions of $B^0 \to D^{*-}\mu^+\nu_\mu$ and $B^0 \to D^{*-}3\pi^\pm$, known from other measurements (PDG world averages): $R(D^*) = \mathcal{K}(D^{*-})\frac{\mathcal{B}(B^0 \to D^{*-}\pi^+\pi^-\pi^+)}{\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)}$.

Since this analysis has yet to undergo the LHCb internal review, and several systematic uncertainties have to be computed, the $R(D^*)$ result remains blinded. Nonetheless, nearly complete documentation of the project is laid out in this thesis.

Chapter 1 introduces the Standard Model of particle physics. An emphasis is put on the Lepton Flavour Universality from both phenomenological and experimental perspectives. Chapter 2 describes the LHCb detector and software. Chapter 3 lays out strategy of the $R(D^*)$ measurement. In Chapter 4 the data samples are described. Chapter 5 documents the selection steps for the main decay modes of the analysis. Chapter 6 explains the use of the control samples. The main result is reported in Chapter 7. Systematic uncertainties are discussed in Chapter 8 before the conclusions and future prospects.

Chapter 1

Phenomenology of $R(D^*)$

This chapter introduces the theoretical framework behind the measurements described in this thesis. The underlying theory, the Standard Model of particle physics is treated briefly in Section 1.1. Then, Section 1.2 summarises phenomenology of the Lepton Flavour Universality searches, driven by an interplay between effective field theory approaches and experimental results. The two genres of anomalies, concerning the $b \to c\ell\nu_\ell$ and $b \to s\ell\ell$ processes, notably the target of this analysis, the $R(D^*)$ ratio, are explained there. Finally an overview of possible New Physics models awaiting experimental or phenomenological verification is laid out. Throughout, prospects of improving current results in the near future are mentioned.

1.1 The Standard Model of particle physics

The Glashow-Weinberg-Salam Model [5, 6, 7], or the Standard Model (SM), is a Quantum Field Theory (QFT) describing fundamental fields (particles) and interactions (forces). It comprises a synthesis of the Electroweak and Strong interactions with the Higgs mechanism. The following sections shortly highlight its basic properties.

1.1.1 Particles and their interactions

The Standard Model (SM) of particle physics groups the 12 elementary particles and anti-particles, referred to as fermions, into three generations as depicted below.

I	II	III
$\begin{pmatrix} u \\ d \end{pmatrix}$	$\begin{pmatrix} c \\ s \end{pmatrix}$	$\begin{pmatrix} t \\ b \end{pmatrix}$
$\begin{pmatrix} e \\ \nu_e \end{pmatrix}$	$\begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}$	$\begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}$

The second and third generation particles replicate the properties of those from the first one, except they are increasingly heavier: *e.g.* the muon is heavier than the electron, whereas the tau is heavier than the muon. Similar relationship occurs among quarks: the mass of up- and down-type quarks goes up with generation. This rule, however, does not hold for neutrinos: their absolute masses are not known and the Standard Model assumes they are massless particles. This has been, nonetheless, contradicted historically in solar neutrino oscillations [8, 9] and the long baseline experiments [10], governed by the PMNS matrix [11].

The matter all around us is built exclusively from the first generation. The second and third generation fermions are unstable, so they cannot build matter.

There are three distinct interaction types or 'forces' accounted for in the SM: electromagnetic, strong and weak. The gravity is ignored in the model as it is too feeble in particle interactions. Nonetheless, it would be important to include it in a more complete theory. Apart from it, the electromagnetism governs most of the phenomena that we perceive in our lives.

The strong force is short ranged and bonds the nucleons together. Without it atoms would not exist. The weak nuclear force is responsible for nuclear decays (*e.g.* the beta decay) and is essential in the nuclear fusion fuelling the Sun.

Each of these interactions is mediated by a boson exchange as listed in Table 1.1. Besides, there exists a Higgs field that gives mass to fermions and weak interaction mediators.

Table 1.1 – Fundamental interactions and their force carriers. Gravity is included for completeness and, due to lack of its QFT, the graviton remains hypothetical.

Interaction	Mediator
Electromagnetic	photon
Strong	gluons
Weak	$W^{\pm}$ and Z^0 bosons
Gravity	graviton?

The SM is a quantum field theory and its particles are described in the language of fields whereby an excitation of a field manifests itself in a particle being produced.

The three charges

At the core of each interaction is a specific charge.

The charged leptons carry the *electric charge* $Q = -1$, whereas their corresponding neutrinos, none, disqualifying them from electromagnetic interactions. The up-type quarks are of $Q = +2/3$ charge, while the down-type ones have $Q = -1/3$.

The weak interaction stems from the *isospin*, usually expressed by its third component T_3. The charged currents exhibit the vector minus axial vector ($V - A$) structure, making only left-handed particles sensitive to this interaction. The left-handed quarks

and leptons are ordered in isospin doublets ($T_3 = \pm 1/2$), whereas the right-handed particles form singlets ($T_3 = 0$). The neutrinos are only left-handed in the SM and they couple only to $V - A$ structures. This implies their zero mass, which assumption has, nonetheless, been proven wrong by the observation of neutrino oscillations [8, 9]. The existence of massive neutrinos necessitates the Physics Beyond the Standard Model (BSM).

Only quarks carry the *colour charge* that results in the strong interaction. It can have three distinct positive values (red, green and blue). Quarks hadronise into 'colourless' mesons (colour-anticolour) or baryons (three different colours).

Weak interaction

The weak interaction is the sole one that violates parity. This stems from the $V - A$ structure of the charged current interactions (*i.e.* mediated by $W^{\pm}$ bosons), that is required to preserve the Lorentz invariance of the weak interaction. Table 1.2 lists bilinear terms for various types of currents. Consequently, the charged current weak interactions, such as in the $b \to c\ell\nu_\ell$ processes, imply a vertex factor:

$$\frac{-ig_W}{\sqrt{2}} 1/2\gamma^\mu(1-\gamma^5), \tag{1.1}$$

where g_W denotes the weak coupling constant and γ are the Dirac matrices.

Table 1.2 – Lorentz-invariant bilinear currents for 2 spinors: ψ and ϕ [12].

Type	Form	Boson spin
Scalar	$\overline{\psi}\phi$	0
Pseudoscalar	$\overline{\psi}\gamma^5\phi$	0
Vector	$\overline{\psi}\gamma^\mu\phi$	1
Axial vector	$\overline{\psi}\gamma^\mu\gamma^5\phi$	1
Tensor	$\overline{\psi}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)\phi$	2

Only left-handed particles and right-handed antiparticles are subject to the weak charged current.

1.1.2 Gauge Bosons in the SM formalism

The SM describes the particles and their interaction in a language of fields. It is based on the Lagrangian density $\mathcal{L}$ dependent on the fields $\psi(x)$ and their derivatives $\partial\psi(x)$. The quantities that cannot be observed reflect theory invariances under symmetries. For instance, in electromagnetism one cannot measure the absolute phase of the field (that depends on x), which choice remains arbitrary by applying a *local gauge transformation i.e.* multiplying the field by $e^{i\theta}$. Such operations belong to the $U(1)$ symmetry group under which the $\mathcal{L}$ is invariant. In this case, we change the field accordingly:

$$\psi(x) \to e^{i\theta(x)}\psi(x). \tag{1.2}$$

This, however, affects the derivative as follows:

$$\partial\psi(x) \to e^{i\theta(x)}\partial\psi(x) + ie^{i\theta(x)}\psi(x)\partial\theta(x), \tag{1.3}$$

which renders the lagrangian density $\mathcal{L}$ non-invariant. To remedy this problem, the covariant derivative is used that corresponds to a new so-called gauge field. This way the new lagrangian density $\mathcal{L}$ is obtained which is invariant and carries a term for the kinetic energy of the gauge field. Each of the fundamental interactions has at least one such gauge field identified by corresponding gauge boson(s):

- In **electromagnetism** the gauge boson is the photon γ;
- In **weak interaction** there are three gauge bosons: $W^{\pm}$ and Z^0;
- In **strong interaction** there are eight gluons.

1.1.3 Electroweak interactions

The Cabibbo-Kobayashi-Maskawa (CKM) matrix [13] describes the strength of flavour-changing weak interaction. Formally, it relates the weak eigenstates of quarks to their mass eigenstates. Historically, the CKM matrix is a generalisation of the 2×2 Cabibbo matrix, relating u, d, s quarks transitions, to all six quarks. The magnitude of each matrix element, $|V_{ij}|$, represents an amplitude of the transition of the q_i quark into q_j quark (or vice versa). Its square, $|V_{ij}|^2$, is the probability of that transition.

By allowing the matrix elements to be complex numbers, CP violation can be introduced resulting in different decay rates of particles and antiparticles. The CKM matrix is written as

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

and with the matrix elements expanded in Wolfenstein parametrisation as

$$V_{\text{CKM}} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4),$$

where [14]

$$\lambda = 0.22453 \pm 0.00044, \qquad A = 0.836 \pm 0.015 \tag{1.4}$$

$$\bar{\rho} \equiv \rho(1 - \lambda^2/2 + \dots) = 0.122^{+0.018}_{-0.017} \qquad \bar{\eta} \equiv \eta(1 - \lambda^2/2 + \dots) = 0.355^{+0.012}_{-0.011}. \tag{1.5}$$

The CKM matrix unitarity implies: $\sum_k |V_{ik}|^2 = \sum_i |V_{ik}|^2 = 1$ and $\sum_k V_{ik}V_{jk}^* = 0$.

The relationships between the CKM matrix elements or parameters can be represented graphically using the 'unitarity' triangle shown in Figure 1.1 (left). The two sides of the triangle, expressed in terms of the CKM matrix elements, as well as the vertex $(\bar{\rho}, \bar{\eta})$, can be obtained experimentally to find out if they indeed form the triangle, testing the CKM

matrix unitarity. Figure 1.1 (right) depicts how constraints from various measurements average out providing the $\bar{\rho}$ and $\bar{\eta}$ values.

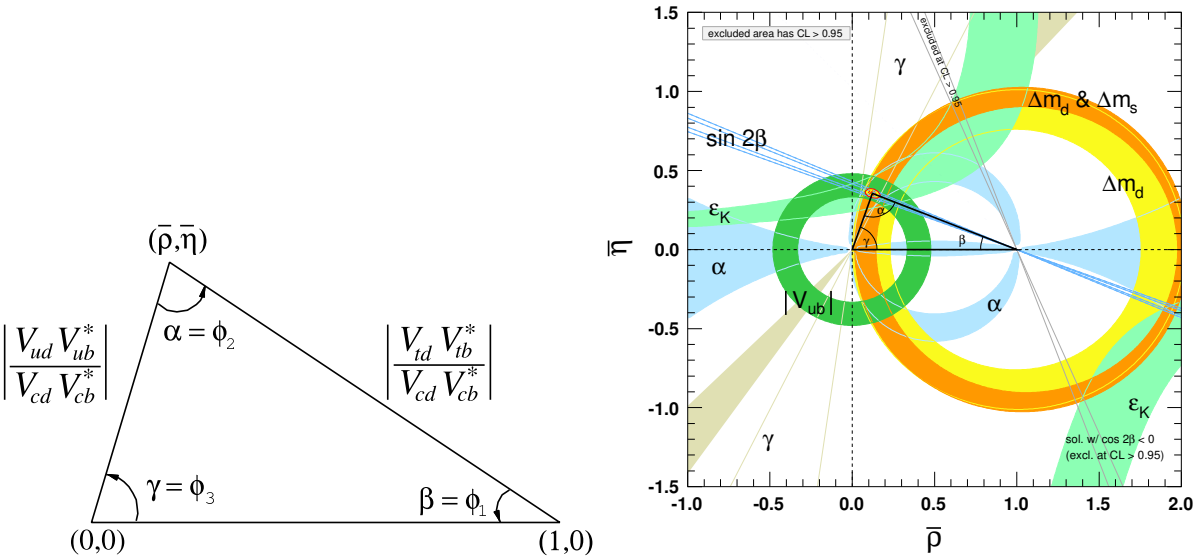


Figure 1.1 – (Left) the unitarity triangle with its two sides measurable experimentally; (right) experimental constraints on the $(\bar{\rho}, \bar{\eta})$ plane [14].

The non-zero off-diagonal elements allow for quark mixing between the generations, which in case of flavour-changing charged currents (FCCC), like for the $b \to c\ell\nu_\ell$ decays, can occur at tree level. On the other hand, the flavour-changing neutral currents (FCNC), such as the $b \to s\ell\ell$ transitions are forbidden at the tree level by the GIM mechanism [15], hence they comprise loop processes.

1.1.4 Higgs mechanism

To accommodate the electroweak symmetry breaking, that makes the $W^{\pm}$ and Z^0 bosons massive, a spinless quantum field is introduced [16, 17, 18, 19], usually called the Higgs field. The field comprises a complex scalar isospin doublet with the components ϕ_1 and ϕ_2. Its generic potential $V(\phi)$, invariant under rotation around the axes perpendicular to ϕ_1 and ϕ_2, depends on two parameters, m and g:

$$V(\phi) = m^2(\phi_1^2 + \phi_2^2) + g(\phi_1^2 + \phi_2^2)^2. \tag{1.6}$$

For positive m^2, the potential has a paraboloid shape with an equilibrium at the bottom, with two degenerate values of m. When m^2 is negative, the potential resembles the Mexican hat and its equilibrium is in the hat's trough. Then the ϕ_1 and ϕ_2 terms designate two different particles: a massive Higgs boson and a massless Goldstone boson. The latter has a screening effect rendering the interaction short-ranged. From the uncertainty principle, short-range interaction allows for massive bosons. Hence, the weak interaction bosons may obtain their masses.

The existence of the Higgs-boson-like particle has been confirmed in 2012 independently by ATLAS and CMS experiments [3, 4].

The Higgs mechanism also gives rise to the masses of fermions.

1.1.5 Strong interactions

Strong interaction is governed by Quantum Chromodynamics (QCD). It has $SU(3)_C$ representation in the SM. All particles carrying colour charge are subject to it. The $SU(3)_C$ group has 8 generators spawning 8 types of gluons that carry this interaction. As they themselves carry the colour charge, they can couple one to another. The invariance under $SU(3)_C$ results in confinement that requires hadrons to have a zero net colour by forming usually mesons or baryons. Moreover, states with 4 or 5 quarks have been observed [20, 21, 22]. Quarks, consequently, do not exist as free particles. An important feature of the Strong Interaction is that the strong coupling increases with distance. Once the energy of the interaction is high enough, a quark-antiquark pair is produced in a process called hadronisation. The thus made hadrons are observable in particle detectors.

In the hadronic effects in the non-perturbative region are factorised into so-called form factors. They reflect the spin-structure of the hadrons involved and impact the effective Hamiltonian operators, which are sensitive to New Physics, as discussed in Section 1.2.3. The form factors parameters are obtained from computationally costly lattice QCD calculations combined with experimental studies.

The branching fractions of the $B^0 \to D^{*-}\ell\nu_\ell$ transitions, studied in this thesis, have large uncertainties due to the hadronic effects. It is much more precise to determine the ratios of the branching fractions of such processes with different charged lepton, ℓ in the final state. Therefore, the Lepton Flavour Universality tests, briefly covered in Section 1.2, study the ratios, such as $R(D^*)$.

1.1.6 Limitations of the Standard Model

Even though the Standard Model is a highly successful theory, it has to be complemented [12]. First of all, there are 26 free parameters whose numbers are fitted experimentally. These are:

- the masses of the 12 fermions;
- the 3 coupling constants of the electromagnetic, weak and strong interactions: α, G_F and α_S;
- the 2 parameters describing the Higgs field, μ and λ or, equivalently the vacuum expectation value and mass of the Higgs field, v and m_H;
- the PMNS [1] and CKM matrices parameters: θ_{12}, θ_{13}, θ_{23}, δ, and λ, A, ρ and η;

[1] The PMNS matrix parameters are not included in the SM. Nevertheless, their detachment from the SM can be considered a shortcoming of the theory.

- Possibly the QCD Lagrangian can have non-zero phase leading to CP violation in the strong interaction, which, nonetheless, is measured to be extremely small, $\theta_{CP} \approx 0$.

The large number of these free parameters suggests that the Standard Model is merely an effective theory of some more complete and higher energy one. Moreover, there are a few unresolved questions expected to be addressed by a more complete theory:

- How to fully explain the matter-antimatter asymmetry?
- What is the nature of dark matter?
- What is the dark energy?
- Where does the hierarchy of fermion masses stem from (see Figure 1.2)?
- How can gravity be incorporated within a more complete theory?

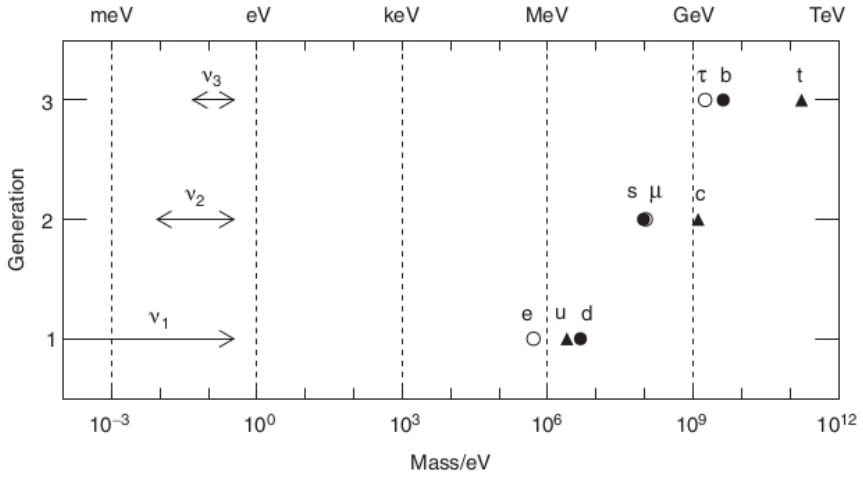


Figure 1.2 – Mass hierarchy of the fermions arranged by their generation [12]. It is noteworthy how different masses of charged fermions of different generations are whilst being quite similar within a generation.

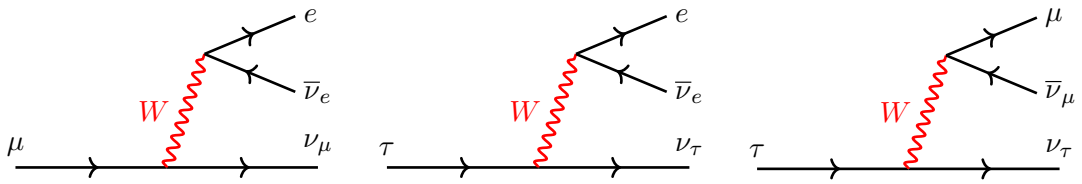
The searches to complement the Standard Model are generically denoted by the so-called New Physics. Perhaps the leptons and quarks are remnants of degenerate states named leptoquarks that would exist at higher energies?

1.2 Lepton Flavour Universality

The aim of this project is to test Lepton Flavour Universality (LFU). This SM principle is laid out in Section 1.2.1, followed by brief overview of the experimental LFU tests in Section 1.2.2 and introduction of the Effective Field Theory (EFT) formalism, relevant for $R(D^*)$ prediction, both within and beyond the SM, in Section 1.2.3. Finally, possible New Physics models to address these anomalies are discussed in Section 1.2.4.

1.2.1 LFU in the Standard Model

Let us consider the charged-current leptonic decays of the charged leptons, $\mu \to e\overline{\nu}_e\nu_\mu$, $\tau \to e\overline{\nu}_e\nu_\tau$ and $\tau \to \mu\overline{\nu}_\mu\nu_\tau$, as presented in Fig. 1.3 [12]. These are well-studied processes that occur in the Standard Model via the W boson emission. The SM assumes that the coupling at each $W\ell\nu_\ell$ vertex, where $\ell = e, \mu, \tau$, is the same and equals the Fermi constant, G_F. Let us, however, abandon this assumption momentarily, allowing for breaking the LFU in these processes. Hence, each vertex implies now a non-universal G_F^ℓ.



(a) $\mu \to e\overline{\nu}_e\nu_\mu$ (b) $\tau \to e\overline{\nu}_e\nu_\tau$ (c) $\tau \to \mu\overline{\nu}_\mu\nu_\tau$

Figure 1.3 – Feynman diagrams of the leptonic decays of the charged leptons.

To the next-to-leading order precision, the decay rate of $\mu \to e\overline{\nu}_e\nu_\mu$ is

$$\Gamma(\mu \to e\overline{\nu}_e\nu_\mu) = \frac{G_F^e G_F^\mu m_\mu^5}{192\pi^3} \times f(m(e)^2/m(\mu)^2), \tag{1.7}$$

and the one of the $\tau \to \ell\overline{\nu_\ell}\nu_\tau$ reads

$$\Gamma(\tau \to \ell\overline{\nu_\ell}\nu_\mu) = \frac{G_F^\ell G_F^\tau m_\tau^5}{192\pi^3} \times f(m(\ell)^2/m(\tau)^2), \tag{1.8}$$

where $f(x) = 1 - 8x + 8x^3 - x^4 - 12x^2 \log x$ is the phase-space correction factor [23, 24].

Because $\mu \to e\overline{\nu}_e\nu_\mu$ is the sole decay of the μ, its decay rate is the inverse of the muon lifetime, τ_μ

$$\Gamma(\mu \to e\overline{\nu}_e\nu_\mu) = \Gamma_{\text{tot}}(\mu \to X) = \frac{1}{\tau_\mu}. \tag{1.9}$$

The decay rate of $\tau \to \ell\overline{\nu_\ell}\nu_\tau$ similarly depends on its lifetime, but also on its branching fraction to allow for alternative decay modes

$$\mathcal{B}(\tau \to \ell\overline{\nu_\ell}\nu_\tau) = \frac{\Gamma(\tau \to \ell\overline{\nu_\ell}\nu_\tau)}{\Gamma_{\text{tot}}(\tau \to X)} = \Gamma(\tau \to \ell\overline{\nu_\ell}\nu_\tau) \times \tau_\tau, \tag{1.10}$$

hence, combining 1.8 and 1.10 yields

$$\tau_\tau = \mathcal{B}(\tau \to \ell\overline{\nu_\ell}\nu_\tau)\frac{192\pi^3}{G_F^\ell G_F^\tau m_\tau^5} \times 1/f(m(\ell)^2/m(\tau)^2). \tag{1.11}$$

By comparing the μ and τ lifetimes, we obtain the ratio of non-universal Fermi constants

$$\frac{G_F^\tau G_F^\ell}{G_F^e G_F^\mu} = \frac{\tau_\mu m_\mu^5}{\tau_\tau m_\tau^5}\mathcal{B}(\tau \to \ell\overline{\nu_\ell}\nu_\tau) \times \frac{f(m(e)^2/m(\mu)^2)}{f(m(\ell)^2/m(\tau)^2)}. \tag{1.12}$$

Substituting $\ell = e$ and putting values for the observables reported in Tab.1.3, we obtain

$$\frac{G_F^\tau}{G_F^\mu} = \frac{\tau_\mu m_\mu^5}{\tau_\tau m_\tau^5}\mathcal{B}(\tau \to e\overline{\nu}_e\nu_\tau) \times \frac{f(m(e)^2/m(\mu)^2)}{f(m(e)^2/m(\tau)^2)} = 1.002 \pm 0.003. \tag{1.13}$$

Combining 1.10 and 1.12 for $\ell = e, \mu$ one can obtain

$$\frac{G_F^e}{G_F^\mu} = \frac{\mathcal{B}(\tau \to e\overline{\nu}_e\nu_\tau)}{\mathcal{B}(\tau \to \mu\overline{\nu}_\mu\nu_\tau)} = 0.997 \pm 0.003. \tag{1.14}$$

Therefore, to next-to-leading order corrections, within the uncertainties, the experimental results are consistent with $G_F^e = G_F^\mu = G_F^\tau = G_F$ in the weak charged current decays of the charged leptons.

This section considered purely leptonic decays of the charged leptons. The next section introduces semileptonic decays of b-hadrons, where the LFU is, nonetheless, questioned.

Table 1.3 – Input values to the LFU derivation in leptonic decays of the charged leptons [14].

Quantity	Value
$m(e)$	$(0.5109989461 \pm 0.0000000031)\,\mathrm{MeV}/c^2$
$m(\mu)$	$(105.6583745 \pm 0.0000024)\,\mathrm{MeV}/c^2$
$m(\tau)$	$(1776.86 \pm 0.12)\,\mathrm{MeV}/c^2$
$\tau(\mu)$	$(2.1969811 \pm 0.0000022) \times 10^{-6}\,\mathrm{s}$
$\tau(\tau)$	$(2.903 \pm 0.005) \times 10^{-13}\,\mathrm{s}$
$\mathcal{B}(\tau \to e\overline{\nu}_e\nu_\tau)$	$(17.82 \pm 0.04)\%$
$\mathcal{B}(\tau \to \mu\overline{\nu}_\mu\nu_\tau)$	$(17.39 \pm 0.04)\%$

1.2.2 Tests of Lepton Flavour Universality

There are many observables that can verify LFU. In the following, an emphasis is put on the ones that are relevant to either the flavour changing charged current (FCCC) processes, like $b \to c\ell\nu_\ell$ or the flavour changing neutral current (FCNC) ones, like $b \to s\ell\ell$.

The first one occurs at the tree level, while the second one happens at the loop level. Both of them exhibit tensions between the SM prediction and experimental results. The next two Sections overview the recent measurements in these two modes. A selected few other tensions are covered afterwards.

1.2.2.1 Overview of the $b \to c\ell\nu_\ell$ measurements

The LFU in $b \to c\ell\nu_\ell$ decays has been recently studied in quite a few measurements of the $R(D^*)$ and $R(D)$ ratios defined as:

$$R(D^*) \equiv \frac{\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)} \qquad (1.15) \qquad R(D) \equiv \frac{\mathcal{B}(B^0 \to D^{-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{-}\mu^+\nu_\mu)} \qquad (1.16)$$

where $\ell = \mu^+, e^+$.

$R(D^*)$ in the SM was originally estimated in 2012 as [25]

$$R(D^*)_{\mathrm{SM}} = 0.252 \pm 0.003. \qquad (1.17)$$

Then, several other estimations followed, most importantly, using improved form-factor estimations [26, 27, 28] based on the BGL parametrisation, derived from the Belle paper [29], resulting in the current arithmetic average, provided by HFLAV [30]:

$$R(D^*)_{\mathrm{SM;wa}} = 0.258 \pm 0.005. \qquad (1.18)$$

In 2012 BaBar measured $R(D^*)$, with $\ell = \mu, e$ in the denominator, as [31, 32] $0.332 \pm 0.024\,(\mathrm{stat}) \pm 0.018\,(\mathrm{syst})$, *i.e.* 2.7σ away from the then SM prediction. It spawned a follow-up measurement by Belle (with hadronic tag and leptonic τ reconstruction) [33] and LHCb (with muonic tau reconstruction) [34] published in 2015. The Belle result was in-between the SM prediction and the BaBar result, while LHCb confirmed the BaBar result. It resulted in further checks: in 2016 by Belle with semileptonic tag [35]; in 2017 by Belle using a hadronic tag ($\tau \to \pi\nu_\tau$ and $\tau \to \rho\nu_\tau$) [36, 37]; in 2018 by LHCb with hadronic tau reconstruction [38, 39]; and in 2019 Belle re-analysed the full dataset using semileptonic tag and measured both $R(D^*)$ and $R(D)$ [40] (this supersedes the result in [35]). Figure 1.4 summarises the theoretical and experimental $R(D^*)$ results. Currently the theoretical-experimental tension, based on world averages as published by HFLAV [30] exhibits $\approx 2.5\sigma$ tension as depicted in Figure 1.5, where $R(D^*)$ is plotted with respect to $R(D)$. The latter shows 1.4σ deviation from the SM. Exploiting the -0.38 correlation between the two observables, the combined world average manifests approximately 3.1σ tension with the SM. However, a recent measurement of $\overline{B}{}^0_s \to D_s^{+(*)}$ form factors [41] increases the tension to 3.8σ.

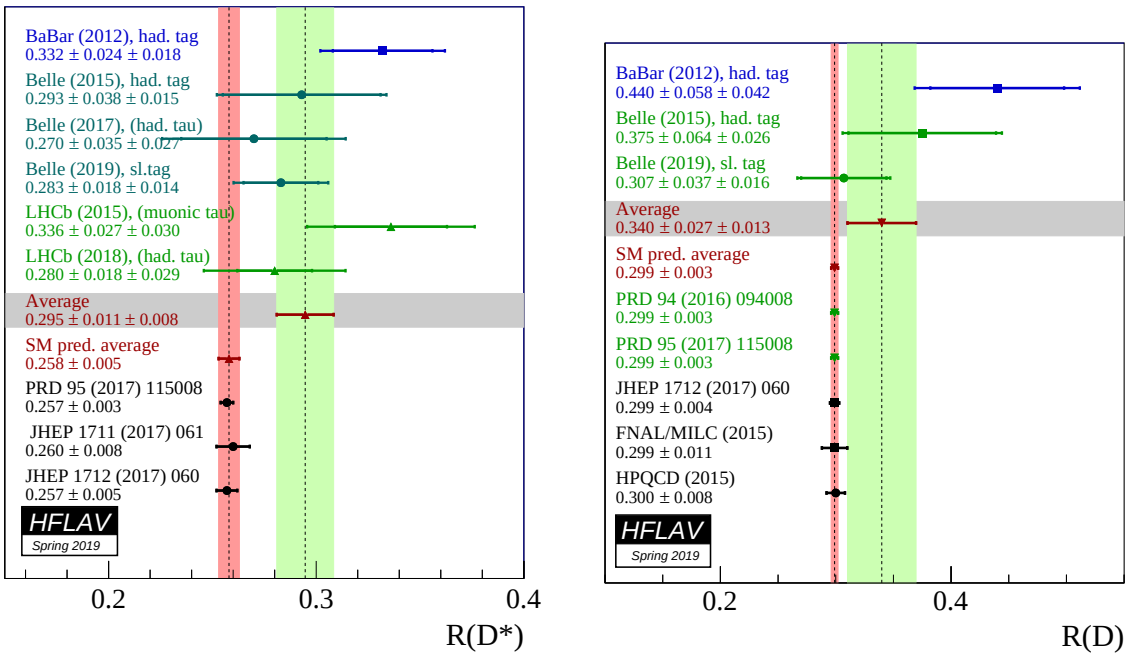


Figure 1.4 – Overview of world measurements of $R(D^*)$ and $R(D)$ by HFLAV [30].

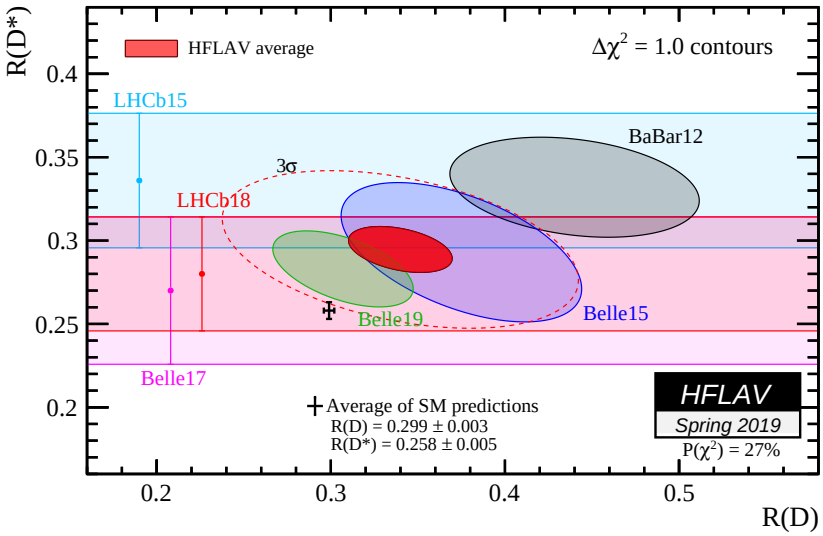


Figure 1.5 – Overview of combined world results on $R(D)$-$R(D^*)$ by HFLAV [30].

This 3.1σ tension is intriguing because the process considered is at the tree level and that most of possible New Physics explanations require new states up to one TeV [42].

Therefore it is crucial to further investigate $R(D^*)$ and similar ratios in the $b \to c\ell\nu$ family.

Another $b\to c\ell\nu_\ell$ observable is $R(J/\psi) \equiv \frac{\mathcal{B}(B_c^+\to J/\psi^-\tau^+\nu_\tau)}{\mathcal{B}(B_c^+\to J/\psi^-\mu^+\nu_\mu)}$ [43]. Its result $R(J/\psi) = 0.71\pm 0.17\,(\text{stat}) \pm 0.17\,(\text{syst})$ exhibits $\approx 2\sigma$ tension with the SM prediction of $R(J/\psi)_{\text{SM}} = 0.2601 \pm 0.0036$ [44].

It is curious that all the results of $R(D^*)$, $R(D)$ and $R(J/\psi)$ deviate systematically above the SM estimations. To complement these searches, several other observables are currently studied in $b\to c\ell\nu_\ell$ modes, for example, $R(D_s^{(*)}), R(\Lambda_c^{(*)})$, with different or more spectator quarks than in $R(D^*)$ and possibly with an excited charm hadron in the final state.

1.2.2.2 Overview of the $b\to s\ell\ell$ measurements

One of the most striking anomalies are in the ratios of branching fractions of $B\to K\ell\ell$, where $\ell = \mu, e$ and $K = K^+$ or $K = K^{*0}$, denoted by $R(K) \equiv \frac{\mathcal{B}(B^+\to K^+\mu^+\mu^-)}{\mathcal{B}(B^+\to K^+e^+e^-)}$ [45, 46, 47, 48] and $R(K^\star) \equiv \frac{\mathcal{B}(B^0\to K^{*0}\mu^+\mu^-)}{\mathcal{B}(B^0\to K^{*0}e^+e^-)}$ [49, 47, 48]. Despite their prior measurements by BaBar and Belle that agreed with the SM ($< 1\sigma$ tensions), the recent, more precise, results from LHCb show 2.5σ and $2.2-2.4\sigma$ deviation from the SM for $R(K)$ and $R(K^\star)$ respectively. To complement, LHCb extended the investigation to the baryon decays by studying the ratio $R_{pK}^{-1} \equiv \frac{\mathcal{B}(\Lambda_b^0\to pK^-e^+e^-)}{\mathcal{B}(\Lambda_b^0\to pK^-\mu^+\mu^-)}$ [50], which, nonetheless, turned out to be consistent with the SM. Another searches for New Physics are the measurements of angular observables such as P_5' studied in the $B^0\to K^{*0}\mu^+\mu^-$ decay [51]. Itself the variable is $2.5-2.9\sigma$ away from the SM prediction in selected q^2 regions, but combined with other angular variables, the global discrepancy reaches 3.3σ. Curiously, the effect seems to be well explained by altered only one vector coupling $Re(C_9)$ [2] that was found to be 1.04 ± 0.25 below the SM value [51]. All these measurements are enlisted in Table 1.4 and summarised in Figure 1.6.

[2] C_9 is a Wilson coefficient of the Effective Field Theory, which is briefly introduced in Section 1.2.3.

Table 1.4 – Selected observables in $b \to s\ell\ell$ decays from LHCb, BaBar and Belle.

Parameter	Value	$q^2[GeV^2/c^4]$	Ref.	Experiment	SM tension
$R(K)$ superseded by [46]	$0.745^{+0.090}_{-0.074} \pm 0.036$	$[1.0, 6.0]$	[45]	LHCb	2.6σ
$R(K)$	$0.846^{+0.060\ +0.016}_{-0.054\ -0.014}$	$[1.1, 6.0]$	[46]	LHCb	2.5σ
$R(K)$	$0.74^{+0.40}_{-0.31} \pm 0.06$	$[0.10, 8.12]$	[47]	BaBar	$< 1\sigma$
$R(K)$	$1.43^{+0.65}_{-0.44} \pm 0.12$	> 10.11	[47]	BaBar	$< 1\sigma$
$R(K)$	$1.01^{+0.28}_{-0.25} \pm 0.02$	$[0.1, 4.0]$	[52]	Belle	$< 1\sigma$
$R(K)$	$0.85^{+0.30}_{-0.24} \pm 0.01$	$[4, 8.12]$	[52]	Belle	$< 1\sigma$
$R(K)$	$1.03^{+0.28}_{-0.24} \pm 0.01$	$[1, 6]$	[52]	Belle	$< 1\sigma$
$R(K)$	$1.97^{+1.03}_{-0.89} \pm 0.02$	$[10.2, 12.8]$	[52]	Belle	$< 1.1\sigma$
$R(K)$	$1.16^{+0.30}_{-0.27} \pm 0.01$	> 14.18	[52]	Belle	$< 1\sigma$
$R(K)$	$1.10^{+0.16}_{-0.15} \pm 0.02$	whole range	[52]	Belle	$< 1\sigma$
$R(K^\star)$	$0.66^{+0.11}_{-0.07} \pm 0.03$	$[0.045, 1.1]$	[49]	LHCb	2.2σ
$R(K^\star)$	$0.69^{+0.11}_{-0.07} \pm 0.05$	$[1.1, 6.0]$	[49]	LHCb	2.4σ
$R(K^\star)$	$1.06^{+0.48}_{-0.33} \pm 0.08$	$[0.10, 8.12]$	[47]	BaBar	$< 1\sigma$
$R(K^\star)$	$1.18^{+0.55}_{-0.37} \pm 0.11$	> 10.11	[47]	BaBar	$< 1\sigma$
$R(K^\star)$	$0.52^{+0.36}_{-0.26} \pm 0.06$	$[0.045, 1.1]$	[53]	Belle	$\approx 1\sigma$
$R(K^\star)$	$0.96^{+0.45}_{-0.29} \pm 0.11$	$[1.1, 6]$	[53]	Belle	$< 1\sigma$
$R(K^\star)$	$0.90^{+0.27}_{-0.21} \pm 0.10$	$[0.1, 8]$	[53]	Belle	$< 1\sigma$
$R(K^\star)$	$1.18^{+0.52}_{-0.32} \pm 0.11$	$[15, 19]$	[53]	Belle	$< 1\sigma$
$R(K^\star)$	$0.94^{+0.17}_{-0.14} \pm 0.08$	> 0.045	[53]	Belle	$< 1\sigma$
R_{pK}^{-1}	$1.17^{+0.18}_{-0.16} \pm 0.07$	$[0.1, 6.0]$	[50]	LHCb	1σ
P_5'	*cf.* Figure 1.6	[4,6]	[51]	LHCb	2.5σ
P_5'	*cf.* Figure 1.6	[6,8]	[51]	LHCb	2.9σ
all angular observables	n/a	$\approx [0, 19]$	[51]	LHCb	3.3σ

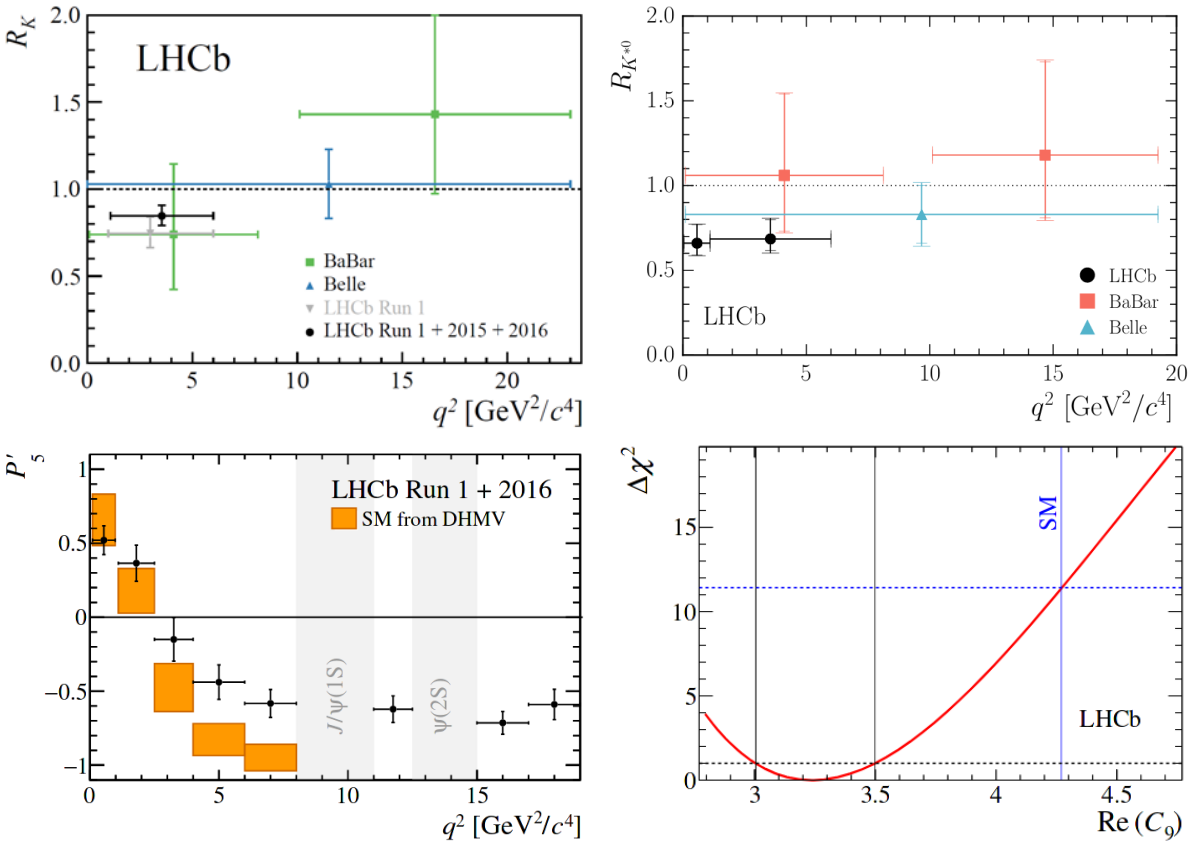


Figure 1.6 – (Top left) $R(K)$ results from LHCb [45, 46], BaBar [47] and Belle [48]; (Top right) $R(K^\star)$ LHCb result compared to various theoretical predictions [49]; (Bottom left) LHCb P'_5 result [51]; (Bottom right) $\Delta\chi^2$ distribution of various $Re(C_9)$ values with its minimum 1.04 ± 0.25 below the SM prediction [54].

1.2.3 Effective Field Theory computation of $R(D^*)$

The branching fractions in the $R(D^*)$ ratio are theoretically computed in the Effective Field Theory (EFT) formalism introduced in the following.

The EFT [55, 56] is a common approach in tackling computations of branching fractions of weak decays. It allows to separate out effects from different energy scales, as *e.g.* around a b-hadron mass, $\mathcal{O}(m_{X_b})$, as opposed to around the W-boson mass, $\mathcal{O}(m_W)$. Various contributions are summed up in a series as shown in Eq. 1.19, which comprises the SM effective Hamiltonian. Another advantage of this formalism is that it accommodates the potential New Physics effects by adding extra operators whose contributions are normalised by their coefficients. Some ongoing analyses exploit this feature by 'injecting' New Physics candidates there.

The EFT ignores the W boson in the diagram Fig. 1.7 and assumes a four-fermion vertex resulting in the four-vector operator, $\mathcal{O}_{V\ell}$, as depicted in Fig. 1.8.

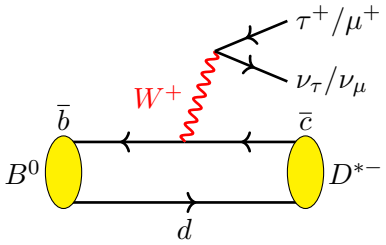


Figure 1.7 – The leading order SM Feynman diagram of the decays in the $R(D^*)$ numerator, $B^0 \to D^{*-}\tau^+\nu_\tau$, and denominator, $B^0 \to D^{*-}\mu^+\nu_\mu$, respectively.

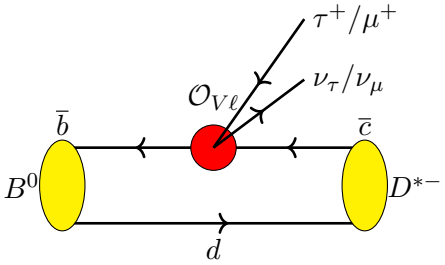


Figure 1.8 – The leading order SM Feynman diagram of the decays in the $R(D^*)$ numerator, $B^0 \to D^{*-}\tau^+\nu_\tau$, and denominator, $B^0 \to D^{*-}\mu^+\nu_\mu$, respectively as in the effective Hamiltonian approach.

The effective Hamiltonian in the $b \to c\ell^-\overline{\nu}_\ell$ decays is as follows [57]:

$$\mathcal{H}_{\text{eff}}(b \to c\ell^-\overline{\nu}_\ell) = \frac{4G_F}{\sqrt{2}} V_{cb} \sum_i \mathcal{C}_i \mathcal{O}_i, \tag{1.19}$$

where the index i runs over the various 4-fermion operators.

In fact, due to the universality of the lepton couplings, the operator product expansion in Eq. 1.19 boils down to the dominant SM operator, $\mathcal{O}_{V\ell}$, as follows:

$$\mathcal{C}_{V\ell}\mathcal{O}_{V\ell} = (\bar{c}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu P_L \nu_\ell), \tag{1.20}$$

where P_L is a function of the Dirac γ_5 matrix, $P_L = (1-\gamma_5)/2$ and $\mathcal{C}_{V\ell}$ is normalised to unity. Potential New Physics contributions would manifest themselves in extra terms is the series in Eq. 1.19.

The estimation of bilinears $\bar{c}\Gamma b$, as in eq. 1.20 (or in a beyond-the-SM form), where Γ is a 4×4 matrix resulting from the product of the Dirac γ-matrices, relies on form factors that depend on the spin of the decaying hadron as shown in Tab. 1.5. In the case of the $B^0 \to D^{*-}\tau^+\nu_\tau$, the $A_{0,1,2}$ and V form factors have to be known.

Table 1.5 – Relevant form-factors in $B \to H$ transition as a function of the spin of the H hadron and the bilinear $\bar{c}\Gamma b$ [57].

$J^P(H)$	Γ	Form factors
0^-	γ_μ	f_0, f_+
0^-	$\sigma_{\mu\nu}$	f_T
1^-	γ_μ	A_0, A_1, A_2
1^-	$\gamma_\mu\gamma_5$	V
1^-	$\sigma_{\mu\nu}$	T_2, T_3
1^-	$\sigma_{\mu\nu}\gamma_5$	T_1

The differential decay rate of the $R(D^*)$ numerator or denominator ($B^0 \to D^{*-}\ell^+\nu_\ell$, where $\ell = \tau^+, \mu^+$ respectively), and the $R(D^*)$ itself neglecting the mass of the light lepton, can be expressed in terms of hadronic helicity amplitudes H_{mn} [25], which depend on the form factors:

$$\frac{d\Gamma_\ell}{dq^2} = \frac{G_F^2|V_{cb}|^2|p|q^2}{96\pi^3 m_B^2}\left(1-\frac{m_\ell^2}{q^2}\right)^2 \times \left[(|H_{++}|^2+|H_{--}^2|+|H_{00}|^2)(1+\frac{m_\ell^2}{2q^2}) + \frac{3m_\ell^2}{2q^2}|H_{0t}|^2\right],$$
$$R(D^*) \equiv \frac{d\Gamma_\tau/dq^2}{d\Gamma_\ell/dq^2} \overset{m_\ell=0}{\approx} \left(1-\frac{m_\tau^2}{q^2}\right)^2 \left[(1+\frac{m_\tau^2}{2q^2}) + \frac{3m_\tau^2}{2q^2}\frac{|H_{0t}|^2}{|H_{++}|^2+|H_{--}^2|+|H_{00}|^2}\right], \tag{1.21}$$

(1.22)

where q^2 is the four-momentum transfer to the lepton system, $q^2 = (p_B - p_{D^*})^2$, $|p|$ is the 3-momentum of D^* in B rest frame, $|p| = \frac{\sqrt{\lambda(m_B^2, m_{D^*}^2, q^2)}}{2m_B}$, $\lambda(a,b,c) = a^2+b^2+c^2-2(ab+bc+ca)$. The helicity amplitudes are q^2 functions dependent on the $A_{0,1,2}$ and V form-factors:

$$H_{\pm\pm}^{SM}(q^2) = (m_B + m_{D^*})A_1(q^2) \mp \frac{2m_B}{m_B+m_{D^*}}|p|V(q^2), \tag{1.23}$$
$$H_{00}^{SM}(q^2) = \frac{1}{2m_{D^*}\sqrt{q^2}}\left[(m_B^2 - m_{D^*}^2)A_1(q^2) - \frac{4m_B^2|p|^2}{m_B+m_{D^*}}A_2(q^2)\right], \tag{1.24}$$
$$H_{0t}^{SM}(q^2) = \frac{2m_B|p|}{\sqrt{q^2}}A_0(q^2). \tag{1.25}$$

These form factors are further expressed in terms of the universal form factor, $h_{A_1}(w)$,

and the ratios $R_{0,1,2}(w)$ as follows:

$$h_{A_1}(w) = A_1(q^2)\frac{1}{R_m}\frac{2}{w+1}, \tag{1.26}$$

$$A_0(q^2) = \frac{R_0(w)}{R_m}h_{A_1}(w), \tag{1.27}$$

$$A_2(q^2) = \frac{R_2(w)}{R_m}h_{A_1}(w), \tag{1.28}$$

$$V(q^2) = \frac{R_1(w)}{R_m}h_{A_1}(w), \tag{1.29}$$

where $R_m = 2\sqrt{m_B m_{D^*}}/(m_B + m_{D^*})$ and $w = v_B \cdot v_{D^*} = (m_B^2 + m_{D^*}^2 - q^2)/(2m_B m_{D^*})$, with $m_{B(D^*)}$ being the mass of the B (D^*) meson. The minimum value, $w = 1$, corresponds to zero recoil of the D^* meson in the B rest frame, *i.e.*, the largest kinematically allowed value of q^2.

They can be obtained from lattice QCD, light-cone sum rules, heavy-quark expansion theory and others. The following section briefly introduces the two common form factor parametrisations: CLN and BGL.

Form factor parametrisations

The Caprini-Lellouch-Neubert (CLN) parametrisation [58] relies on the heavy-quark spin symmetry in the $B^{(*)}$ and $D^{(*)}$ mesons), a dispersion technique and introduces the short-distance and $1/m_Q$ (where Q denotes both the b and c quark) corrections near zero recoil ($w = 1$). The form factor results are valid within approximately 2% in the semileptonic region. The four above-mentioned form factors are written as [58]

$$h_{A_1}(w) = h_{A_1}(1)\left[1 - 8\rho^2 z + (53\rho^2 - 15)z^2 - (231\rho^2 - 91)z^3\right], \tag{1.30}$$

$$R_0(w) = R_0(1) - 0.11(w-1) + 0.01(w-1)^2, \tag{1.31}$$

$$R_1(w) = R_1(1) - 0.12(w-1) + 0.05(w-1)^2, \tag{1.32}$$

$$R_2(w) = R_2(1) - 0.11(w-1) - 0.06(w-1)^2, \tag{1.33}$$

where $z \equiv \frac{\sqrt{w+1}-\sqrt{2}}{\sqrt{w+1}+\sqrt{2}}$.

These depend only on five parameters: ρ^2, $R_0(1)$, $R_1(1)$, $R_2(1)$ and $h_{A_1}(1)$ (the latter is often referred to as $\mathcal{F}(1)$). They are all, except $R_0(1)$, extracted from the differential spectrum of the $B^0 \to D^{*-}\ell^+\nu_\ell$ decays, and have been measured most precisely by BaBar and Belle. $R_0(1)$ is helicity suppressed in Equation 1.21, where it is incorporated in the H_{0t} term (see Equations 1.25 and 1.27), so it cannot be measured similarly, however it still is important for the $R(D^*)$ theoretical result.

Another approach to the form factors is the Boyd-Grinstein-Lebed (BGL) parametrization [59, 60, 61], which is more general and relies on a series expansion with arbitrary precision. Recently, there has been some doubts concerning validity of the CLN assumptions at the current precision: *e.g.* as argued in [27] the residual uncertainty is not negligible; some precise lattice QCD calculations should be considered. Therefore,

many measurements try to compare results of, for instance, V_{cb} and $R(D^*)$ as obtained with the both methods. The form factors relevant to $B \to D^* \ell \nu$ decays depend on three functions, $f(w)$, $g(w)$ and $\mathcal{F}_1(w)$:

$$h_{A_1}(w) = \frac{f(w)}{\sqrt{m_B m_{D^*}}(1+w)}\,, \tag{1.34}$$

$$R_1(w) = (w+1) m_B m_{D^*} \frac{g(w)}{f(w)}\,, \tag{1.35}$$

$$R_2(w) = \frac{w-r}{w-1} - \frac{\mathcal{F}_1(w)}{m_B(w-1)f(w)}\,, \tag{1.36}$$

which are expressed in terms of convergent power series of z as:

$$f(z) = \frac{1}{P_{1^+}(z)\phi_f(z)} \sum_{n=0}^{\infty} b_n z^n\,, \tag{1.37}$$

$$g(z) = \frac{1}{P_{1^-}(z)\phi_g(z)} \sum_{n=0}^{\infty} a_n z^n\,, \tag{1.38}$$

$$\mathcal{F}_1(z) = \frac{1}{P_{1^+}(z)\phi_{\mathcal{F}_1}(z)} \sum_{n=0}^{\infty} c_n z^n\,. \tag{1.39}$$

The terms $P_{1^\pm}(z)$ are Blaschke factors and $\phi_{f,g,\mathcal{F}_1}(z)$ are 'outer functions', both detailed *e.g.* in [62].

The Table 1.6 gives the world averages of the form factor parameters as of Summer 2019 HFLAV summary [30]. The $h_{A_1}(1)$ parameter is written (i) as obtained in an HFLAV experimental world average, with the electroweak corrections $\eta_{\rm EW}$ and multiplied by $|V_{cb}|$; (ii) as calculated from lattice QCD by Fermilab/MILC or HPQCD. It is noteworthy that $h_{A_1}(1)$ is the only form factor that remains in the zero-recoil ($w = 1$) limit in the BGL parametrisation.

Table 1.6 – HFLAV world average of the CLN form factor parameters as of the Summer 2019 [30].

Parameter	HFLAV world average
ρ^2	$1.122 \pm 0.015\,(\text{stat}) \pm 0.019\,(\text{syst})$
$R_1(1)$	1.270 ± 0.026
$R_2(1)$	0.852 ± 0.018
$\eta_{\rm EW} h_{A_1}(1)\lvert V_{cb}\rvert$	$35.27 \pm 0.11\,(\text{stat}) \pm 0.36\,(\text{syst})$
	Lattice QCD [63]
$h_{A_1}(1)$ Fermilab/MILC 14 [64]	$0.906 \pm 0.004\,(\text{stat}) \pm 0.012\,(\text{syst})$
$h_{A_1}(1)$ HPQCD 17B [65]	$0.895 \pm 0.010\,(\text{stat}) \pm 0.024\,(\text{syst})$

The $R_0(1)$ in the HQET limit equals $[R_0(1)]_{\rm HQET} = 1$, but was more precisely estimated in [25] as $R_0(1) = 1.14 \pm 0.11$, and later in [26] as $R_0(1) = 1.17 \pm 0.02$. These form factor parameters are subject to future updates from lattice computations, especially beyond the zero-recoil, as well as experimental studies. Preliminary results by the JLQCD

Collaboration [66] and by Fermilab/MILC [67] at the Lattice2019 Symposium already demonstrate the recoil-dependent form factors of $B \to D^{(*)}\ell\nu_\ell$ decays. Arriving at a consensus between the CLN vs. BGL parametrisation (*cf.* the long-standing discrepancy between the inclusive/exclusive determination of $|V_{cb}|$ may stem from differences in the CLN/BGL parametrisations) would be desirable. Last but not least, Beyond-the-SM form factors, combined from experiment and LQCD, would shed light on *raison d'être* of possible New Physics models.

1.2.4 New Physics models

Recent years have brought harvest as much to the experimental as the theoretical results on Lepton Flavour Universality. With the appearance of new measurements, some models were encouraged or dismissed.

There are three [68] typical candidates to account for the $R(D^*)$ and $R(D)$ anomalies:

- Two-Higgs-doublet models [69]
- Heavy vector bosons, *e.g.* W' [70]
- Leptoquarks [71, 72]

Charged Higgses lead to too large B_c^+ lifetime and result in disfavoured q^2 distributions [68]. The W' imply also Z' of unnaturally large widths.

Leptoquarks (LQs) are postulated as particles that carry both lepton and baryon numbers. They appear in grand unification theories. LQs might be more fundamental particles decaying into a lepton and a quark (usually from the same generation) at lower energies. Their quantum numbers like spin, electric charge and weak isospin depend on a specific model as it is represented in Table 1.9.

Spin	$3B+L$	$SU(3)_c$	$SU(2)_W$	$U(1)_Y$	Allowed coupling
0	-2	$\bar{3}$	1	1/3	$\bar{q}_L^c \ell_L$ or $\bar{u}_R^c e_R$
0	-2	$\bar{3}$	1	4/3	$\bar{d}_R^c e_R$
0	-2	$\bar{3}$	3	1/3	$\bar{q}_L^c \ell_L$
1	-2	$\bar{3}$	2	5/6	$\bar{q}_L^c \gamma^\mu e_R$ or $\bar{d}_R^c \gamma^\mu \ell_L$
1	-2	$\bar{3}$	2	$-1/6$	$\bar{u}_R^c \gamma^\mu \ell_L$
0	0	3	2	7/6	$\bar{q}_L e_R$ or $\bar{u}_R \ell_L$
0	0	3	2	1/6	$\bar{d}_R \ell_L$
1	0	3	1	2/3	$\bar{q}_L \gamma^\mu \ell_L$ or $\bar{d}_R \gamma^\mu e_R$
1	0	3	1	5/3	$\bar{u}_R \gamma^\mu e_R$
1	0	3	3	2/3	$\bar{q}_L \gamma^\mu \ell_L$

Figure 1.9 – Possible LQs with respect to their baryon (B) and lepton (L) numbers, QCD and weak isospin representation and the weak hypercharge. Their allowed coupling to the SM fermions is also shown [14].

Essentially, leptoquarks are classified as either scalar (spin 0) or vector (spin 1) type

particles. LQs might help explain why there are exactly 3 generations of quarks and leptons, and why this number is equal for these two groups of particles. Moreover, their existence would allow for transitions between leptons and quarks.

ATLAS and CMS searches of LQs of the first and seconds generations [73, 74, 75] and the third generation [76, 77, 78, 79, 80] excluded such particles below the mass of 1.3-1.5 TeV and 740-1002 GeV, respectively. A CMS study allowing for a mixed generation leptoquark [81] excluded such a LQ candidate below the mass of 1420 GeV.

Attempts [82] at resolving both $R(D^*)$ and $R(K^\star)$ puzzles simultaneously rules out the possibility of a single scalar LQ of mass around 1 TeV, but allow for a vector LQ. This LQ poses, however, a challenge of UV completion that has to be specified to compute the loop effects. The LQ proposed is the weak singlet vector U_1 of mass 1-2 TeV. Prospective constraints might be obtained from LHCb and Belle II from $\mathcal{B}(B \to K\mu\tau)$ upper limits.

The paper [68] argues that any two of the three anomalies: $R(D^*)$, $R(K)$ and μ magnetic moment, a_μ (exhibiting 3.3σ tension with the SM [14]), can be explained simultaneously by a $C_9 = -C_{10}$ contribution. The authors propose a model with two LQs of the same mass and coupling to quarks and leptons, with one $SU(2)$ singlet and the other SU(2) triplet.

Another attempt at explaining the anomalies with two scalars [83] originates from two $SU(5)$ operators and can accommodate all observed B meson anomalies.

The most recent model-independent analysis [54] interestingly explains both $b \to s\ell\ell$ and $b \to c\ell\nu_\ell$ anomalies by introducing lepton flavour universal NP C_9^U besides lepton flavour violating $C_{9\mu}^V = -C_{10\mu}^V$ ('Scenario-8' in the paper). Figure 1.10 shows fits to the $(C_{9\mu}^V = -C_{10\mu}^V, C_9^U)$ plane for (left) only $b \to s\ell\ell$ anomalies and (right) also $R(D^{(\star)})$. Inclusion of the latter one, however, increases the pull of that fit to 7.0σ.

In [84] the authors discuss the angular observables in $B^0 \to D^{(*)}\ell\overline{\nu}_\ell$ that can pinpoint the probable New Physics structures even if the $R(D^*)$ tension with the SM disappears. The angular observables can be obtained in an unbiased way despite the missing neutrinos, as proposed in [85].

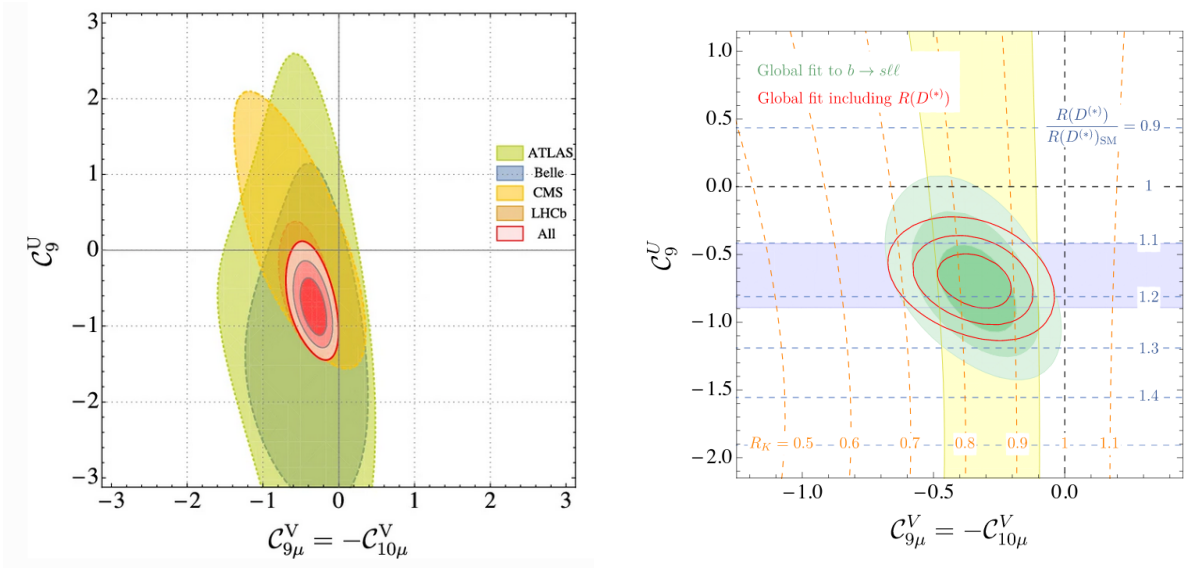


Figure 1.10 – Fits to the $(C_{9\mu}^V = -C_{10\mu}^V, C_9^U)$ plane based on (left) $b \to s\ell\ell$ from various experiments and (right) global $b \to s\ell\ell$ and $R(D^{(\star)})$ results [54].

Chapter 2

The LHCb experiment

The Large Hadron Collider Beauty (LHCb) experiment was initially designed to search for New Physics by studying decays of the beauty hadrons. Its physics programme has, however, been much extended and now covers the charm, strange and top quark physics, heavy ions, electro-weak physics and the Higgs physics.

So far , the LHC produced $\mathcal{O}(10^{12})$ $b\bar{b}$ pairs in the acceptance of the LHCb detector [86]. This is a record number in a b-hadron physics experiment.

The LHCb experiment is situated at the Large Hadron Collider briefly described in the following section. Next, the LHCb detector and its software will be covered with an emphasis put on their importance for the $R(D^*)$ measurement.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [87] comprises a two-ring hadron accelerator installed in the 27 km toric tunnel bored in 1984-1989 for the LEP machine. It has 8 straight segments and 8 arcs between 45 m and 170 m underground and two 2.5 km transfer tunnels that inject particles from the CERN accelerator complex shown in Figure 2.1. The proton (or heavy ions) beams rotate in opposite directions in the two rings with ultra-high vacuum. They are accelerated by radio-frequency cavities and curved by superconducting dipole magnets that operate at extremely low temperatures. The protons collide at four points, which are equipped with very specialised and complex detectors: ATLAS, ALICE, CMS and LHCb.

ATLAS and CMS are general-purpose experiments, whose programme includes searches for the Higgs boson, t-quark physics and supersymmetry. ALICE studies quark-gluon plasma properties at heavy ion collisions.

Before the particles enter the LHC, they undergo many stages illustrated in Figure 2.1. Protons are extracted from a hydrogen ion source and accelerated in a linear accelerator LINAC-2 up to an energy of 50 MeV. Then they enter three increasingly larger circular

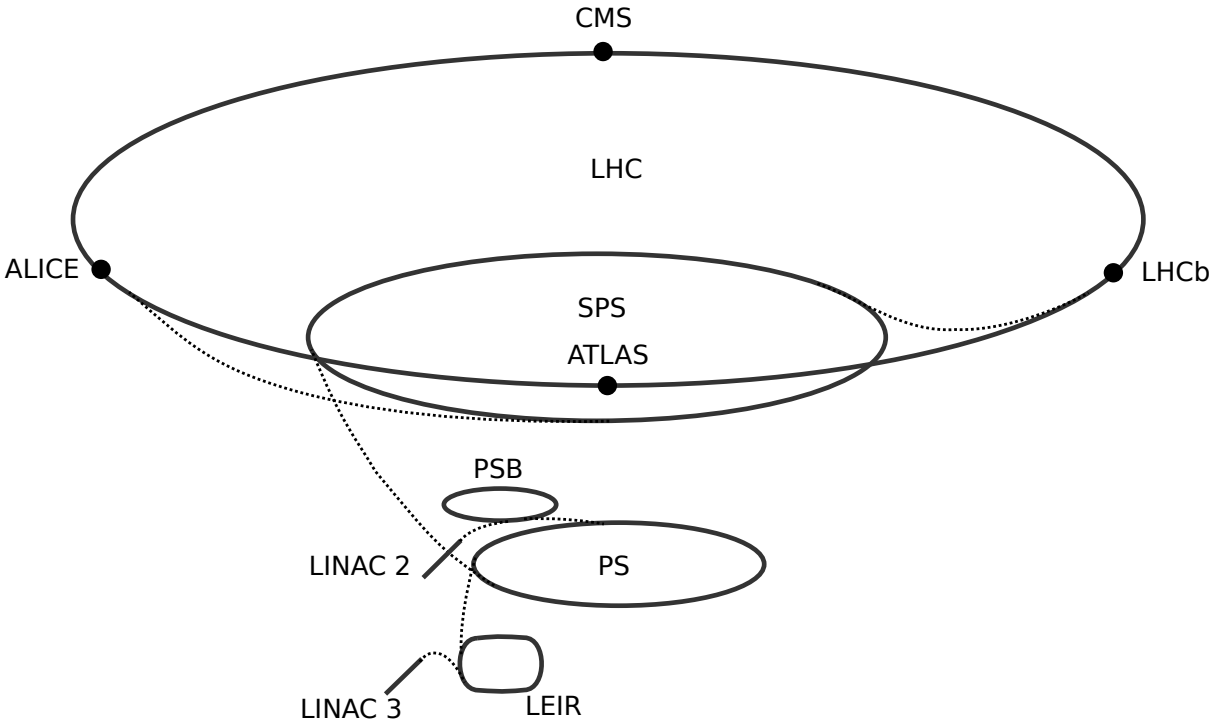


Figure 2.1 – CERN accelerator complex [88] with the proton acceleration route from LINAC-2 to the LHC.

accelerators: Proton Synchrotron Booster (PSB) of 1.4 GeV, Proton Synchrotron (PS) of 28 GeV and the Super Proton Synchrotron (SPS) of 450 GeV. Finally, the protons enter the LHC. The route of heavy ions is similar, except that it starts at LINAC-3, then goes through the Low Energy Ion Ring (LEIR), before proceeding to the PS.

The beams are accelerated using radio-frequency (RF) cavities whereby they get some energy from the electric field. Their trajectory is curved at the 8 arcs with superconducting dipole magnets of 8.3 Tesla and tightened at the collision points with quadrupole magnets.

In head-on collisions of protons of energy E_p each, the centre-of-mass energy equals $\sqrt{s} = 2E_p$. Therefore, it grows linearly with the proton energy.

To date there has been two data-taking periods at LHC: Run1 in 2011-2012 and Run2 in 2015-2018. In Run1 the LHC operated at centre-of-mass energy of $\sqrt{s} = 7 - 8\,\mathrm{TeV}$ while in Run2 at $\sqrt{s} = 13\,\mathrm{TeV}$.

The particles in LHC circulate at 40 MHz frequency grouped in bunches. There are about 3000 bunches per beam with $\sim 10^{11}$ protons in each at the beginning of each around 20-hour collision period, so-called fill.

Number of events, N_{events}, produced at an LHC collision depends on the event cross-section σ_{event} and the machine luminosity $\mathcal{L}$:

$$N_{\text{events}} = \sigma_{\text{event}} \int \mathcal{L} dt, \tag{2.1}$$

where

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \varepsilon_n \beta^*} F, \tag{2.2}$$

with N_b particles per bunch, n_b bunches per beam, revolution frequency f_{rev}, relativistic gamma factor γ_r, the normalised transverse beam emittance ϵ_n, the beta function β^* and factor F accounting for collision geometry.

The LHC peak instantaneous luminosity at the beginning of a fill is $L = 10^{34}\,\text{cm}^{-2}\text{s}^{-1}$ during Run2.

2.2 The LHCb Detector

The $b\bar{b}$ pairs produced in pp collision at LHC are heavily boosted either in the forward or background direction of the beam (z) axis. Therefore, LHCb being a forward detector, can observe almost half of the $b\bar{b}$ decay products in its pseudorapidity range $\eta \in [2, 5]$.

The LHCb [89] geometrical acceptance is shown in Figure 2.2 (left) and also there (right) it is compared with the one of ATLAS and CMS.

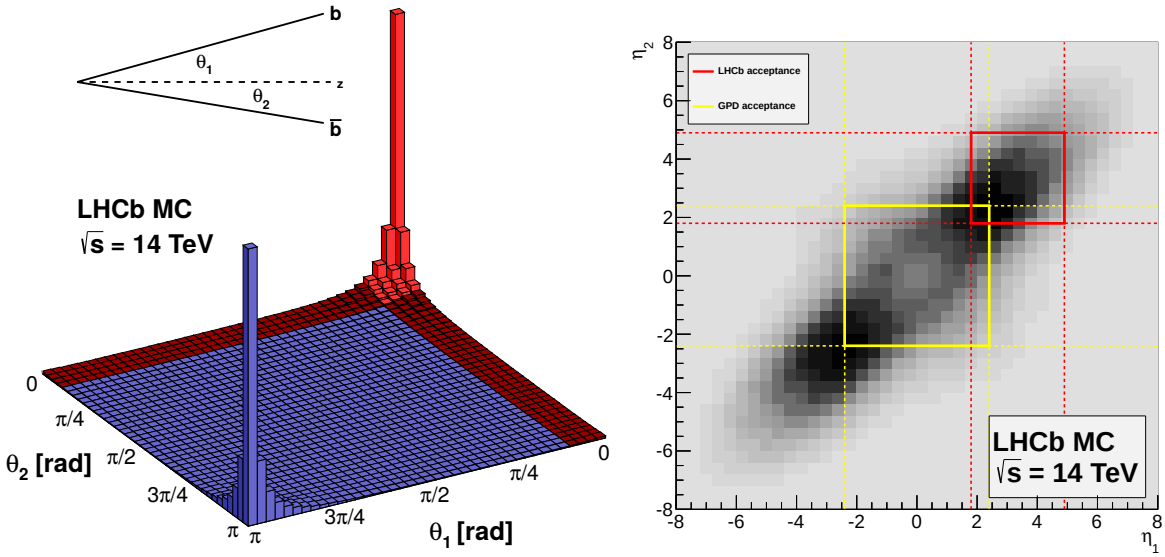


Figure 2.2 – Left: Distribution of simulated $b\bar{b}$ pairs of quarks as a function of their polar angles θ_1 and θ_2. In red is shown angular acceptance of the LHCb detector. Right: Pseudorapidity distribution of the pairs of b quarks. The red region lies within the LHCb acceptance, whereas the yellow marks the one of ATLAS or CMS.

LHCb operates at $\mathcal{L} = 4 \times 10^{32}\,\text{cm}^{-2}\text{s}^{-1}$, in order to reduce event occupancy to the desired average of $\mu = 1.1$ interactions per bunch crossing. This luminosity is decreased with respect to the LHC one by the luminosity levelling procedure, whereby the two beams are offset such that constant luminosity is maintained as long as possible. The

instantaneous luminosity within a typical LHCb fill is depicted in 2.3 (left), while the integrated luminosities gathered in between 2010-2016 are shown in 2.3 (right).

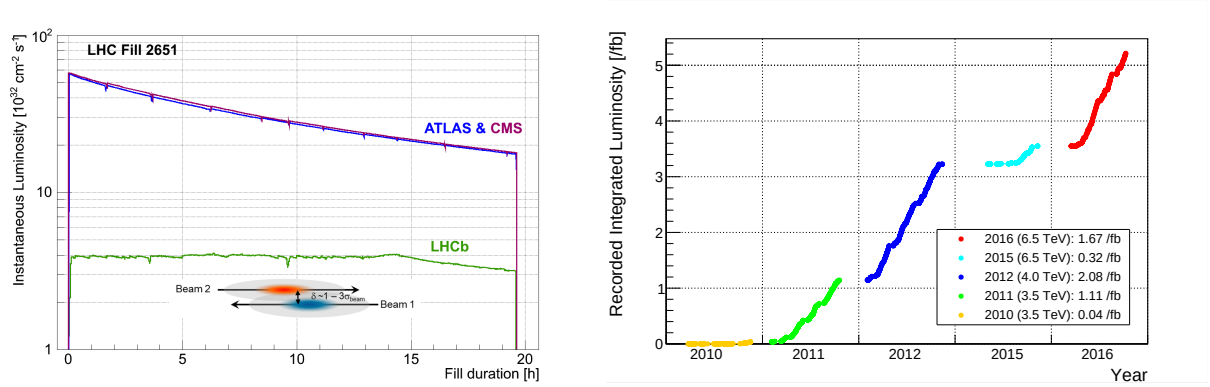


Figure 2.3 – Left: Instantaneous luminosity as a function of time compared between ATLAS, CMS and LHCb. The LHCb maintains approximately constant value thanks to tuning the distance between the two beams. Right: Recorded integrated luminosity at LHCb for different years with their corresponding energies per single proton.

The analysis presented in this thesis is performed using the 2015-2016 dataset with its $\sim 2\,\text{fb}^{-1}$ gathered at the centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ per proton-proton (pp) collision. The previous $R(D^*)$ analysis with hadronic τ reconstruction was based on the $\sim 3.2\,\text{fb}^{-1}$ from the 2011-2012 (Run1) dataset. The running conditions for the two collected datasets are compared in Table 2.1.

Table 2.1 – Comparison of the integrated luminosity, centre-of-mass energy per pp collision, $\sqrt{s}$, and the resulting number of $b\bar{b}$ pairs produced in the LHCb acceptance. This number reflects the *à priori* relative dataset yields, with no trigger nor selection taken into account, for the Run1 (2011-2012) and this analysis (2015-2016).

Period	Integrated luminosity	$\sqrt{s}$	Number of $b\bar{b}$
2011-2012	$3.2\,\text{fb}^{-1}$	7-8 TeV	2.5×10^{11}
2015-2016	$2.0\,\text{fb}^{-1}$	13 TeV	2.9×10^{11}

The three major ingredients of most LHCb analyses are: tracking, particle identification (PID) and trigger.

The tracking allows to reconstruct 'tracks' representing the trajectories of the charged particles passing through the LHCb detector (see Section 2.2.1, while the PID allows to infer the nature of those particles (see Section 2.2.2). The trigger selects the events that are kept for further treatment (see Section 5.2).

These procedures are intertwined as partial reconstruction is sufficient to pass or not early trigger conditions and then run more complete reconstruction before the final trigger steps.

The following sections introduce these indispensable components, where the relevant LHCb subdetectors, depicted in Figure 2.4, are introduced.

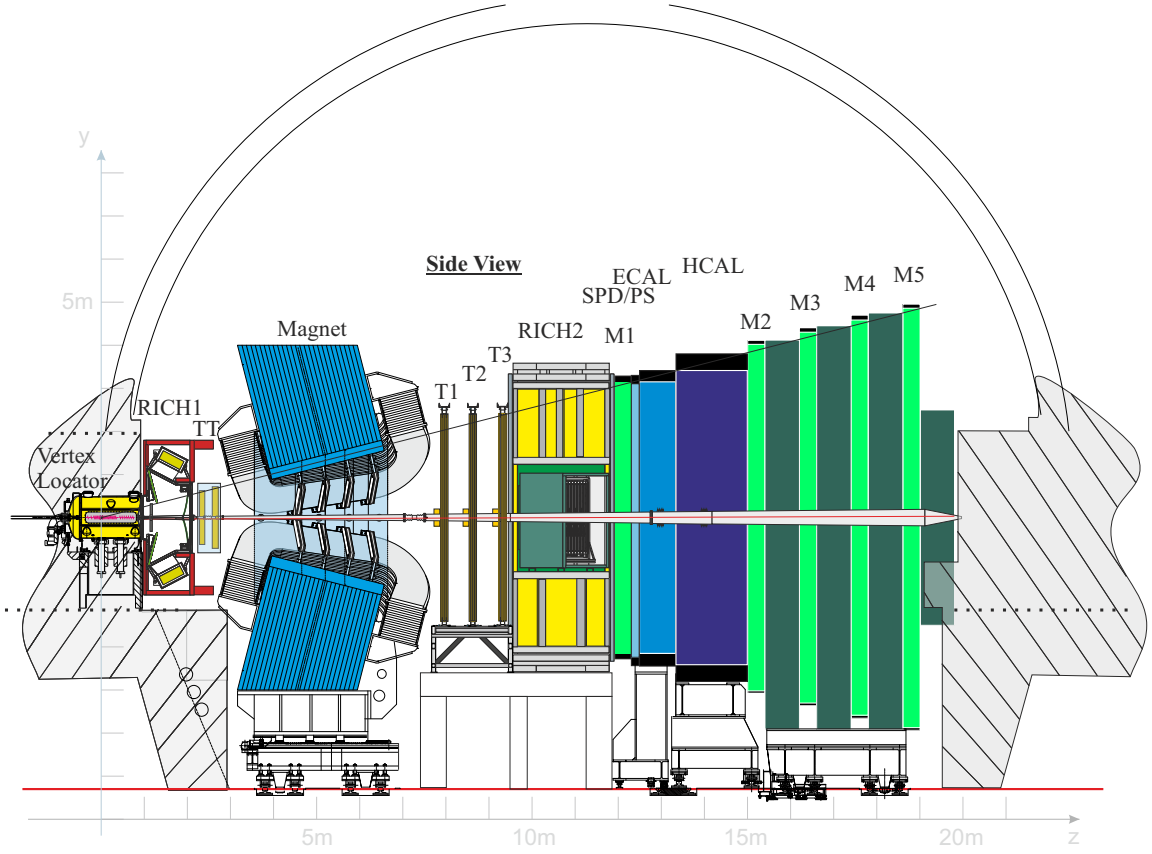


Figure 2.4 – A profile view of the LHCb detector in the non-bending $y - z$ plane with its subdetectors along the beam (z) axis.

2.2.1 Tracking

The tracking system reconstructs trajectories of charged particles. It uses information from points where they interacted with the detector spawning 'hits'. The tracking system is composed of a vertex detector surrounding the interaction point, and tracking stations, which are located on either side of a dipole magnet. Particle momentum is deduced from the deflection angle in the magnetic field. The following sections briefly describe these components as well as the track reconstruction techniques.

2.2.1.1 The vertex detector

The Vertex Locator (VELO) is a silicon micro-strip detector surrounding the interaction point at a minimum distance of 8 mm. Its geometrical acceptance is in the range $1.6 <$

$\eta < 4.9$. Its role is to measure the trajectories of particles that originate at the primary vertex (PV) and subsequently reconstruct the PV. It also reconstructs the secondary vertices (SV) where heavy quark hadrons decay (*e.g.* B^0 in the signal of this analysis).

The VELO has 42 modules along the beam axis, split vertically into two halves either side of it, each enveloped in an aluminium foil. The foil allows for close proximity of the VELO's silicon sensors to the beam, while protecting the LHC primary vacuum from the secondary vacuum at the detector.

The modules along the beam axis are shown in Figure 2.5 (right). Each module has two types of sensors: one measuring the radius, r, where the strips are concentric, 38-102 μm apart; and one measuring the azimuthal angle, ϕ, where the strips are oriented radially with denser segmentation in the inner region due to higher occupancy. The VELO sensors are depicted in Figure 2.5 (left).

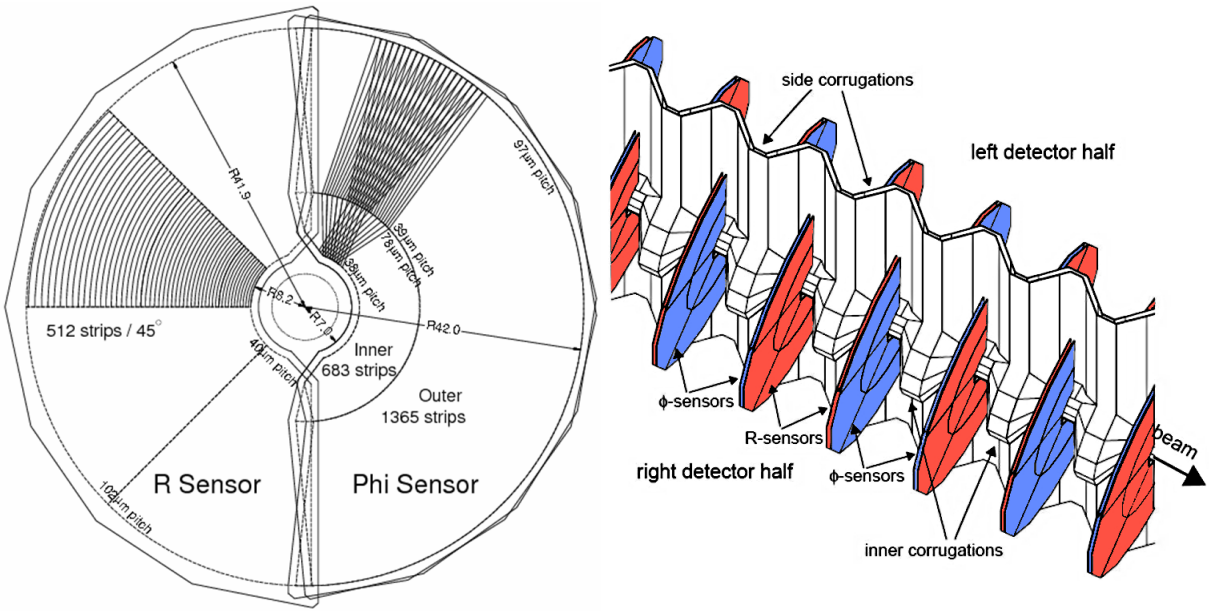


Figure 2.5 – Left: The VELO r and ϕ sensors. Right: The VELO modules alongside the beam axis.

2.2.1.2 The tracking stations

There is one tracking station upstream of the magnet (Tracker Turicensis, TT) and three downstream of it (*cf.* Figure 2.4). The Tracker Turicensis reconstructs tracks coming from particles decaying outside the VELO.It consists of four layers of silicon microstrip sensors of 183 μm pitch. In the first and last layers the strips are vertical while, to improve the reconstruction precision, the strips are rotated by a stereo angle of -5° and 5° in the second and third layers, respectively.

The three upstream tracking stations (T1-T3) are implemented with two technologies: in the inner region close to the beam pipe, the Inner Tracker (IT) is a silicon tracker

similar to the TT, while, in the outer region as the required granularity is lower due to the smaller track density, the Outer Tracker (OT) is made of straw-tubes detectors.

2.2.1.3 The magnet

The warm dipole magnet, illustrated in Figure 2.6, is used to determine particles momenta. To achieve a relative momentum resolution at 5 GeV/c of 0.4%, the magnet provides an integrated magnetic field of 4 Tm between the VELO and the most upstream tracking station. The relative resolution obtained is 4×10^{-3} that requires an integrated magnetic field of 4 Tm between the VELO and the three downstream tracking stations.

The magnet is composed of two coils with 5.85 kA current installed in a rectangular yoke with a window inside, spanning the detector's acceptance.

The magnet polarity is flipped regularly to record approximately equal amounts of data at each polarity to mitigate potential bias due to the detector asymmetry.

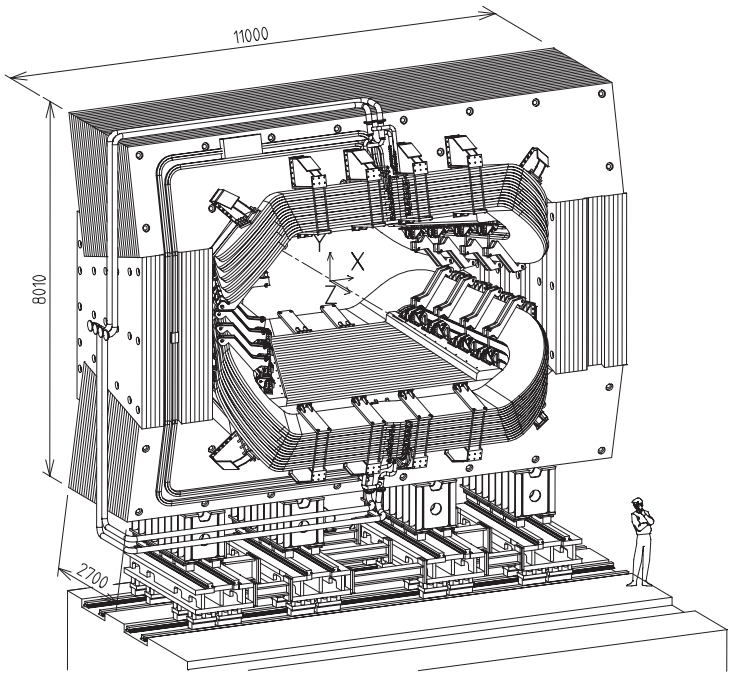


Figure 2.6 – A schematic of the bending dipole magnet projected onto the $x - y$ plane.

2.2.1.4 The tracking algorithm

The trajectories of the charged particles are reconstructed from all hits collected in the tracking system. The tracks are classified according to the set of subdetectors their corresponding particles pass through, as shown in Figure 2.7. In this analysis only 'long' tracks, which pass through all the tracking subdetectors, are considered. Their reconstruction starts with finding a straight line segment in the VELO. Then two independent algorithms follow: the 'forward tracking' and the 'track matching'.

The forward tracking appends a single hit in one of the T1-T3 stations to a VELO segment. Then additional hits are searched in the remaining T1-T3 stations, in a window around the initial trajectory. Thus reconstructed track needs to then pass certain quality cuts.

The track matching, instead, attempts to combine the VELO tracks with those from T1-T3 stations by extrapolating the both sets through the magnet. The tracks from the both algorithms are considered, duplicates are removed and relevant hits from the TT are added to improve the momentum resolution.

Finally, a Kalman filter is applied to the tracks, accounting for multiple scattering and energy losses. Track quality is measured with the χ^2 per degree of freedom metric.

'Ghost' tracks consist in trajectories that do not correspond to a real particle passage. They mostly stem from incorrect matching of the VELO and T1-T3 tracks. In representative pp inelastic collisions they amount to typically 6.5% tracks per event, but can comprise up to 20% events in case of large-multiplicity. The 'ghost' tracks are removed with a neural network that allows to choose a working point with 60% 'ghost' track rejection while keeping 99% of genuine tracks [90].

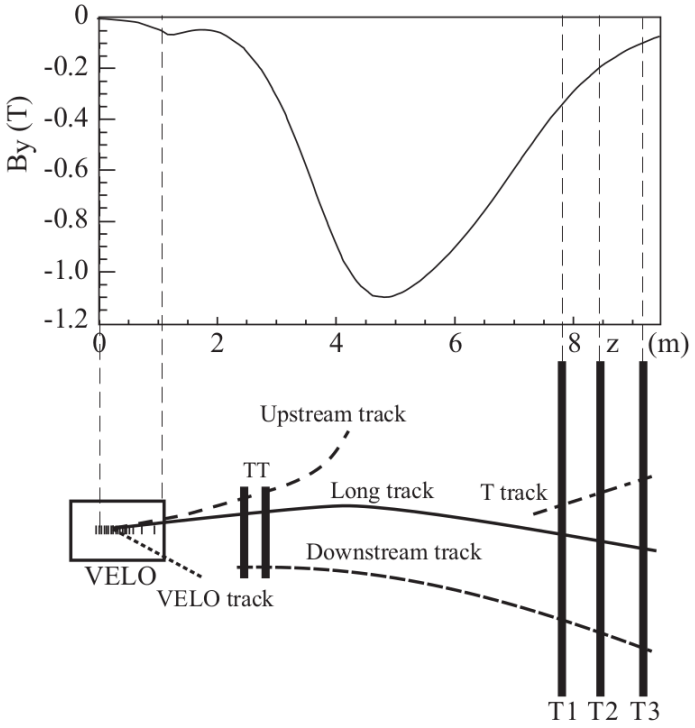


Figure 2.7 – Track types depending on the subdetectors where hits occurred overlaid with the magnetic field distribution along the beampipe.

2.2.1.5 The tracking performance

The tracking system allows for primary vertex (PV) reconstruction with a resolution up to 10-40 μm in the x and y coordinates (perpendicular to the beam) and 50-300 μm

along the beam z coordinate. The lower boundaries are achieved for large number of tracks in PV, as depicted in Figure 2.8.

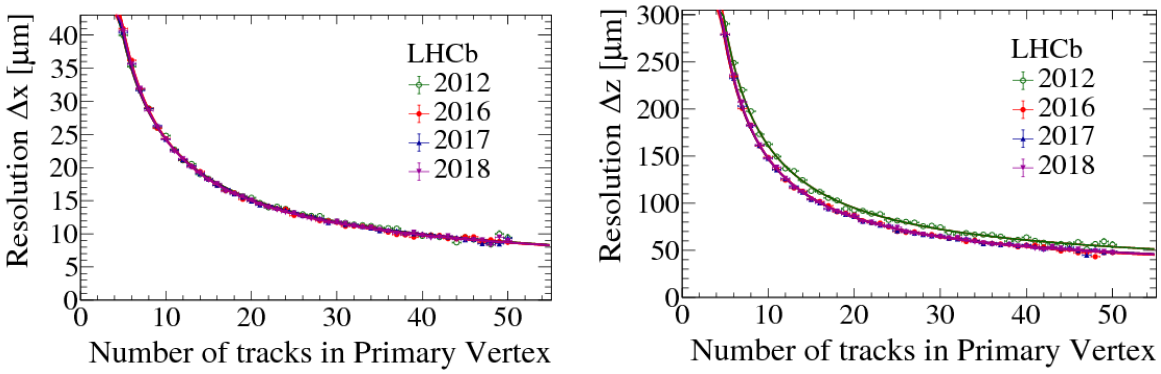


Figure 2.8 – Primary vertex (PV) resolution in the x (left) and z (right) coordinate as a function of number of tracks in the PV [91].

One of the most important geometrical variables in tracking is impact parameter (IP), defined as the distance between the PV and the track direction, measured perpendicularly to the track, in the plane containing the PV and the track.

Individual tracks have IP resolution in the x and y axis down to around 12 μm for high momentum particles as depicted in Figure 2.9.

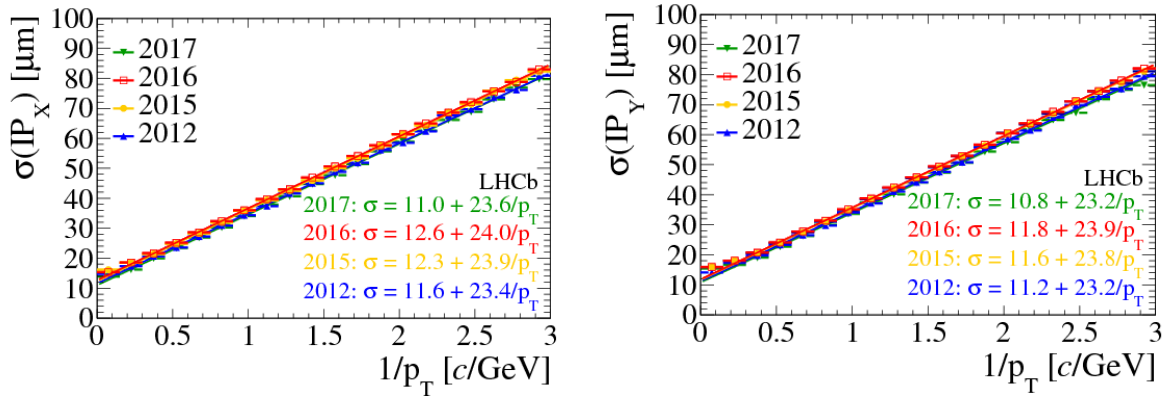


Figure 2.9 – IP x (left) and y (right) coordinate dependence on the transverse momentum p_{T} [91].

The tracking efficiency is 96% as measured for the muons that pass through all the tracking stations. The relative momentum resolution for charged tracks is around 0.5% for tracks whose corresponding particles have momentum below 20 GeV/c and around 0.8% for ones with momentum around 100 GeV/c. See Figures 2.10 for the tracking efficiency and momentum resolution as a function of particle's momentum.

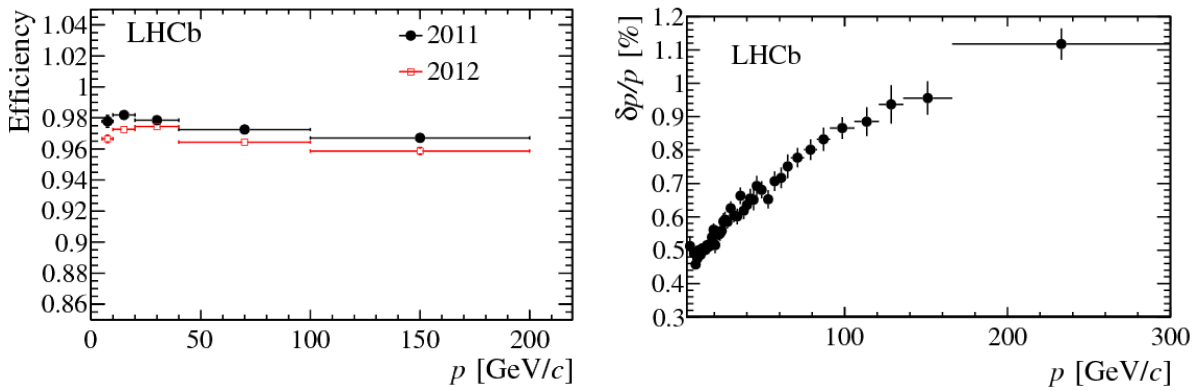


Figure 2.10 – Tracking efficiency as a function of track momentum (left) and momentum resolution as a function of momentum [92] (right) as measured in Run1.

2.2.2 Particle Identification

Particle identification (PID) is crucial in this analysis. It allows to identify the pions and the kaons that form the candidates for the signal and normalisation modes. This section describes the PID system of the LHCb detector. It is composed of the three elements:

- Ring-Imaging Cherenkov detectors (RICH); they identify long-lived charged hadrons: pions, kaons and protons;
- electromagnetic and hadronic calorimeters (ECAL and HCAL) that measure energy of photons, electrons and hadrons;
- muon stations that identify and measure momentum of the muons in the first trigger level.

2.2.2.1 The Ring Imaging Cherenkov Detectors

There are two Ring Imaging Cherenkov detectors (RICH) that are designed to distinguish proton, charged pions and charged kaons. They rely on the Cherenkov effect whereby a radiation is emitted when a charged particle passes through a dielectric of refractive index n with a speed, β, greater than the speed of light in that medium, $1/n$ (in units where $c \equiv 1$). Consequently, the light is emitted at the Cherenkov angle, θ_C, thus forming a cone of an opening angle relative to the particle direction given by the formula:

$$cos\theta_C = \frac{1}{n\beta}. \tag{2.3}$$

Once the cone of light arrives at either RICH detector, it produces a ring whose radius depends on θ_C. Hence the radii of the Cherenkov rings are used to deduce the speed of the particle. Once matched with tracks of known momenta, the particle mass can be obtained. Figure 2.11 shows momentum distributions for different particle species.

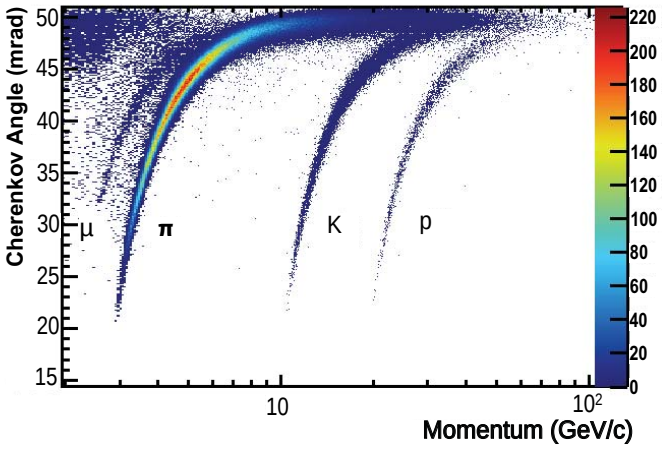


Figure 2.11 – Cherenkov angle distribution as a function of momentum for various particle species as reconstructed in one of the RICH detectors.

The two RICH detectors are between the VELO and TT, and after the tracking system. They have different angular and momentum coverages: the first one, RICH1, covers low momentum particles, around 1-60 GeV/c. It covers the full LHCb acceptance of $\pm 25 - \pm 300\,\mathrm{mrad}$ horizontally and $\pm 250\,\mathrm{mrad}$ vertically. the RICH2 allows for a momentum range from about 15 GeV/c to over 100 GeV/c. Its acceptance is lowered to $\pm 120\,\mathrm{mrad}$ horizontally and $\pm 100\,\mathrm{mrad}$ vertically, as high-momentum particles are boosted forwards.

RICH1 uses aerogel and fluorobutane ($\mathrm{C_4F_{10}}$) and RICH2 uses $\mathrm{CF_4}$. Both RICH detectors are shown in Figure 2.12

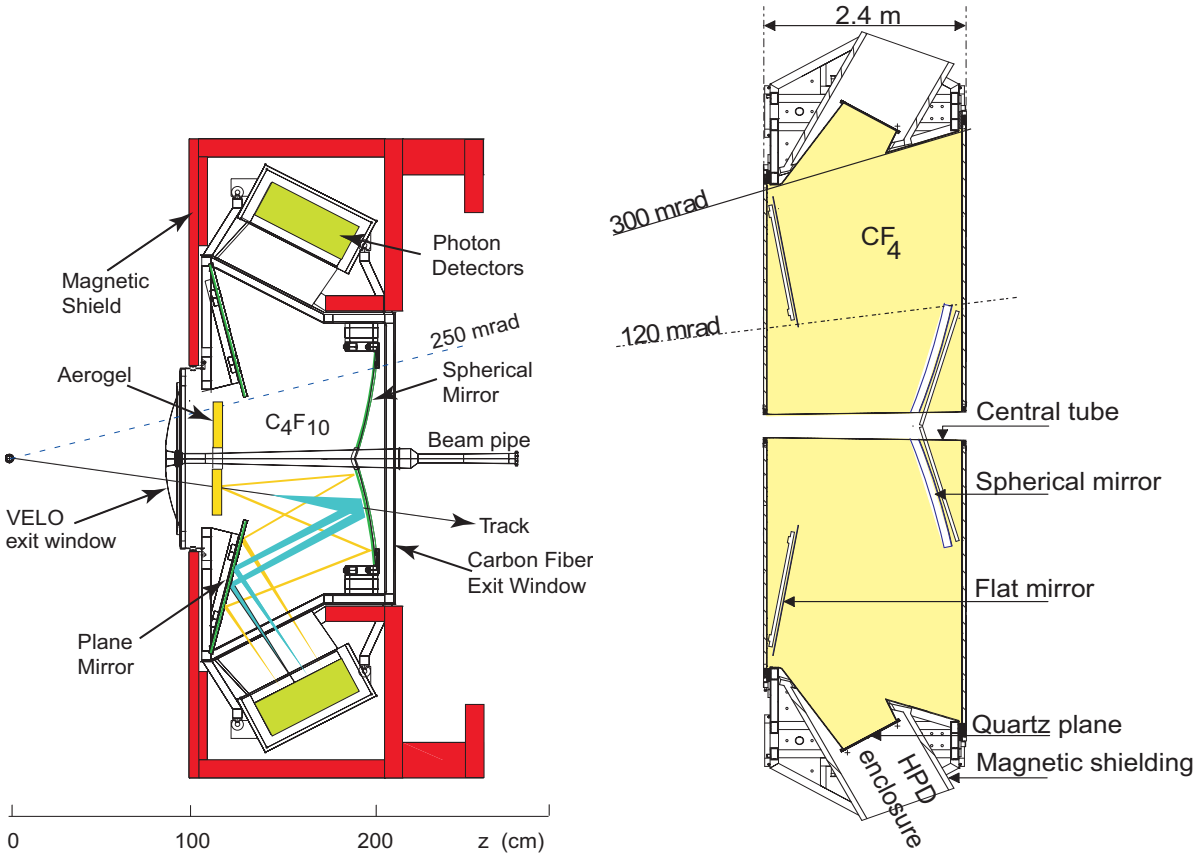


Figure 2.12 – Left: Side view of RICH1. Right: Top-down view of RICH2.

The Cherenkov light emitted in any of the RICH detectors is detected by the Hybrid Photon Detectors (HPDs). They are vacuum phototubes that convert photons to electrons that are detected at the silicon sensors at the anode end.

2.2.2.2 The calorimeters

The calorimeters distinguish electrons, photons and hadrons by measuring their energies and positions. Also, they provide transverse energy measurement that is used by the first level trigger, described in the following Section 5.2. They need to, therefore, provide relevant information fast. This requirement drives the calorimeter system design.

The calorimetry system is situated between the first and remaining muon stations. It consists of two main detectors: the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL). The ECAL is more upstream and is preceded with two additional detectors: the scintillator Pad Detector (SPD) and the Preshower Detector (PS). They detect backgrounds for the electron trigger such as pions. To contain high energy photons, the ECAL thickness is 25 radiation lengths, after 2.5 of the SPD and PS. HCAL contains only 5.6 interaction length, because the trigger requirements on the HCAL resolution do not impose a stringent shower confinement condition. The

calorimeters angular coverage matches that of the RICH1. Their segmentation varies, similarly to other subdetectors, to account for higher occupancy near the beam.

The ECAL and HCAL have alternating layers of absorber and detector. The absorber is lead (ECAL) or iron (HCAL), whereas the detector is a plastic scintillator. Particles interact with the absorber generating showers of secondary particles that are detected by the scintillators which generate light read out by photomultipliers.

The SPD and PS detectors are made of two planes of scintillating pads separated with a lead layer. The SPD detects charged particles and with the ECAL they distinguish *e.g.* electrons from pions. The PS helps discriminate between electrons and charged hadrons based on smaller energy deposit from the hadrons.

2.2.2.3 The muon detector

There are five muon stations (μ) as shown in Figure 2.13 (left). The first one is between the RICH2 and the calorimeters and helps improve the p_{T} measurement for the first level trigger.

The four remaining stations are interleaved with a muon shield, which is comprised of the electromagnetic and hadronic calorimeters. It has three iron filters and has a total absorber thickness of 20 interaction lengths. They measure positions of charged particles passage. With a fine segmentation, the first three stations allow for transverse momentum, p_{T}, reconstruction with about 20% resolution. The other stations are less precise and they identify highly penetrating muons. The muon stations are composed of logical pads with binary readout each, and whose segmentation is more fine closer to the beam as depicted in Figure 2.13 (right). Most pads use multi-wire proportional counters. Only an inner region of M1 uses triple gas electron multipliers (GEMs) due to high flux and considerable ageing.

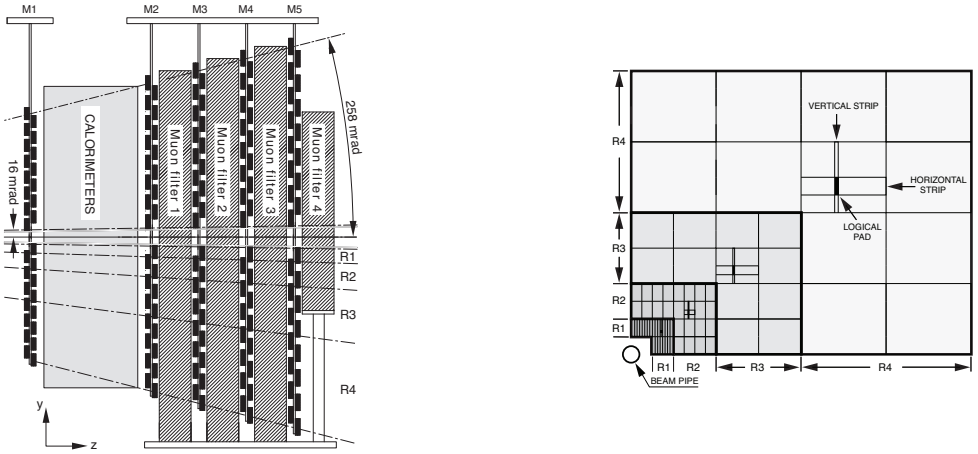


Figure 2.13 – Left: Side view of the five muon stations along the beam axis. Right: A front view of a muon station quadrant. Varying segmentation for four regions, R1-4, of roughly equal occupancy is shown.

2.2.2.4 PID variables and performance

The PID variables allow to assign a mass hypothesis to a given track. They rely on likelihood functions $\mathcal{L}$ for each particle type, computed based on the information from the RICH detectors and the calorimeters. Initially all tracks are assumed to be pions. Each track is sequentially assumed to be the electron, muon, kaon and proton, and the corresponding likelihood functions are computed. Each hypothesis is used to compute the 'delta-log-likelihood' (DLL) function, where the difference is with respect to the previously computed log-likelihood of the pion hypothesis, as follows:

$$DLL(x) = \log L(x) - \log L(\pi) \tag{2.4}$$

LHCb features a good separation between the kaons and pions. Figure 2.14 demonstrates the efficiency of two illustrative cuts on $\Delta \log L(K-\pi)$ for genuine kaons (red) and for pions that pass the DLL cut and are hence mis-identified as kaons (black). The momentum range up to 100 GeV/c covers almost entirely the phase-space of the kaons and pions in this project. For a loose cut of $DLL > 0$ The kaon efficiency on average is roughly 95%, while a chance of identifying a pion as kaon is around 10%. For a harsher cut, of $DLL > 5$ these values shift to around 80% and 3%, respectively.

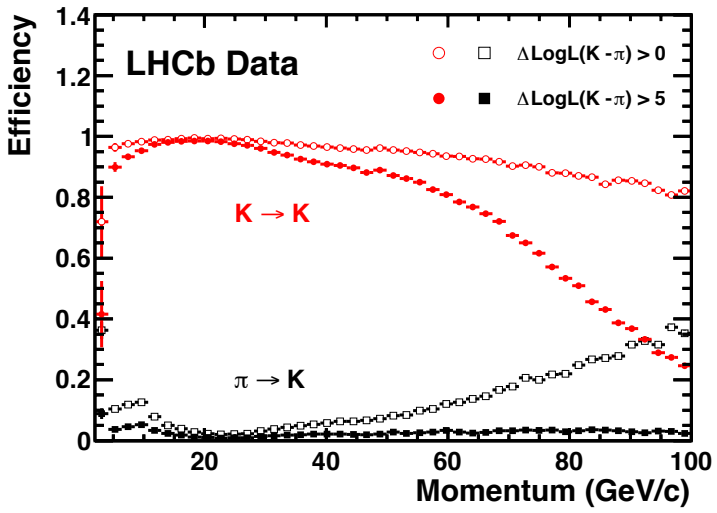


Figure 2.14 – $\Delta \log L(K-\pi)$ cut efficiency in data as a function of momentum for genuine kaons identified as such (red) and for pions misidentified as kaons (black). Two different cuts are shown (empty vs. filled in shapes) [92].

Figure 2.15 provides another perspective on the kaons-pion separation depending on the number of tracks or primary vertices in the event. With increased track multiplicity or number of PV's, the kaon identification efficiency drops while pion mis-identification becomes more likely.

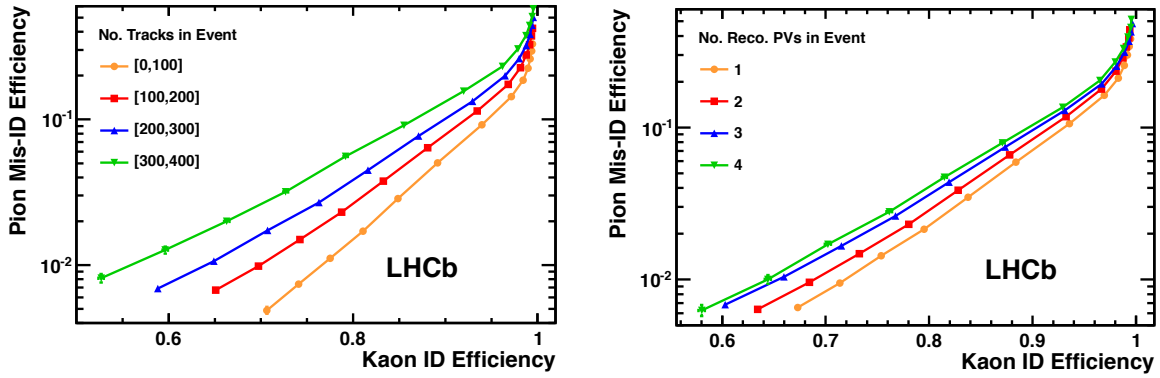


Figure 2.15 – Pion misidentification fraction as a function of kaon efficiency as obtained by cutting on the $\Delta \log L(K-\pi)$ depending on the number of tracks in the event (left) or number of reconstructed primary vertices (right). Specific working point on the curves is subject to the cut value.

Another set of variables, `ProbNNx`, is based on a neural network combining information from the RICH, calorimeters, muon stations and the tracking system [93]. It includes and complements the DLL variables, enabling a more complete PID determination. The `ProbNNx` variables are in range $\in [0,1]$ and can be interpreted as probabilities of a given charged track (*i.e.* π, K, μ, p, e). Each variable stems from a separate neural network. The neural network is implemented as a Multi Layer Perceptron (MLP) with the TMVA package integrated in ROOT [94] and trained on simulated events. Its performance is measured with calibration samples from data. In this analysis, two such variables are used in selection, `probNNpi` and `probNNk`, to select or reject pions and kaons respectively. They are calibrated globally at LHCb using pions and kaons extracted, for high momentum tracks, from $D^0 \to K^-\pi^+$, where D^0 originates from $D^{*+} \to D^0\pi^+$. The samples are extracted using the *sPlot* technique [95].

To facilitate the measurement of the PID efficiencies, the LHCb software has a dedicated Particle Identification Calibration (PIDCalib) package [96]. It provides a parametrisation of track PID efficiency based on its phase-space *e.g.* (transverse) momentum, η, multiplicity. By comparison of the phase-space between the LHCb calibration samples and the simulated ones from this analysis, a weighted average is computed resulting in the estimated PID efficiency.

In this analysis, loose cuts are applied on the DLL variables (*e.g.* `PIDK`) in preselection and, more stringent ones, using the `ProbNNpi` and `ProbNNk` variables in further selection as described in Chapter 5.

2.2.3 The trigger

Trigger [97] is a system that decides if a given event (*i.e.* ensemble of detector signals in a time window) should be recorded in a mass storage. Recording all events is technically impossible: the LHC operates at 40 MHz bunch-crossing rate while the LHCb

detector processes signals at 10 MHz rate. Therefore it is important to rapidly decide if the current event is likely to be of physics interest to LHCb. This is a complex task performed in a pipeline of three subsequent steps:

- **Level-0 Trigger (L0)**: Hardware-based; operates at the bunch crossing rate of 40 MHz, reduces the event rate to around 1 MHz;
- **High Level Trigger 1 (HLT1)**: software-based; inclusive filter of events, reduces the event rate to 150 kHz;
- **High Level Trigger 2 (HLT2)**: software-based; combines inclusive and exclusive selections of b- and c-hadron decays reducing the event rate to 12.5 kHz.

A schematic overview of the trigger is shown in Figure 2.16. Each step verifies if certain

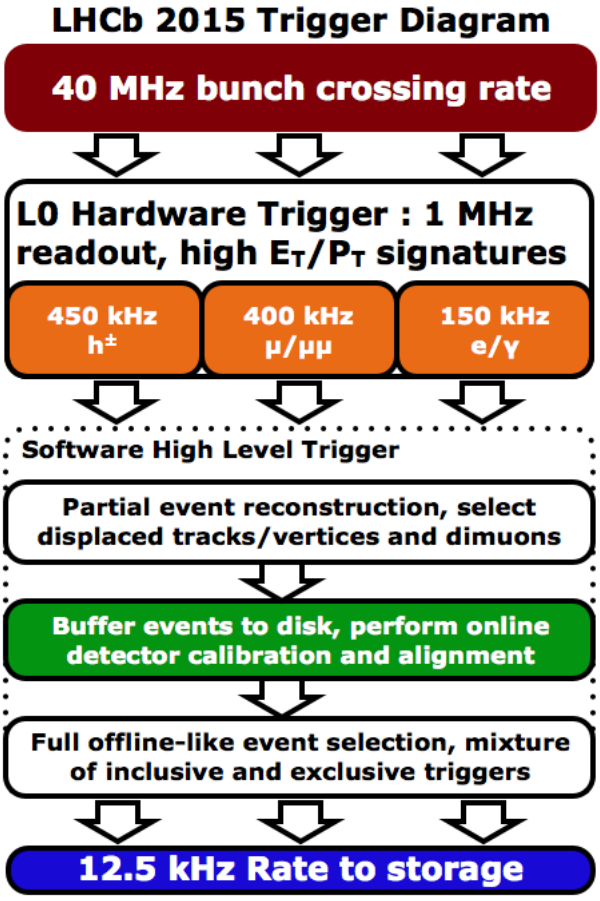


Figure 2.16 – An overview of the trigger flow used in Run2 (2015-2018).

conditions, referred to as 'lines' are satisfied. The following sections detail the three stages of the LHCb trigger.

2.2.3.1 The hardware trigger

The L0 is made with custom electronics and use only fast available sub-detector information to be able to operate at the LHC bunch crossing (40 MHz) rate. It performs a

loose selection, which reduces the event rate down to a maximum of 1 MHz at which the full detector information can be read-out and treated by the next trigger stage.

The b- and c-hadrons, due to their high masses decay into products that have high transverse momenta, p_{T}, and energies, E_{T}. These properties are computed by two components of the L0: the calorimeter trigger, L0CALO, and the muon trigger, L0MUON. This information is combined in the L0 decision unit to compute the decision to accept an event or not.

L0CALO uses information mostly from the electromagnetic and hadron calorimeters, ECAL and HCAL. SPD and PS are also used to distinguish hadrons, photons and electrons. Consequently its decisions is a logical 'or' of its three subsystems: L0Hadron, L0Photon and L0Electron (this split is shown in Figure 2.16).

L0Hadron records the maximum transverse energy, E_{T}, deposited in a cluster of 2×2 HCAL cells. In case there is an energy deposit in the ECAL, just in front of the HCAL cell, both energies are summed up.

L0Photon measures the same quantity in the ECAL, requires 1-2 hits in the PS and none in the SPD. L0Electron is similar to L0Photon except there has to be at least one SPD hit.

L0MUON considers straight tracks in the 5 muon stations. Using their directions, and assuming their origin at the interaction point, their p_{T}, are computed with a resolution of about 20%. There are two relevant lines: L0Muon and L0DiMuon. The former verifies if the maximum p_{T} of the set of tracks exceeds certain threshold. The latter takes into account the product of the two highest p_{T} tracks. Typical thresholds are reported in Table 2.2.

An event is kept if any of the three subsystems has E_{T} greater than certain threshold values, as reported in the Table 2.2, where typical trigger conditions in Run-2 are given. If there are many SPD hits, the event becomes too occupied and slow to reconstruct and is, therefore, rejected with no significant loss in typical signal efficiencies.

Table 2.2 – Selection values used for the L0 hardware trigger in Run-2 collision data [91] for given data-taking years. Dimuon line is a product of the highest and second highest transverse momenta, p_{T}.

L0 trigger	$E_{\mathrm{T}}/p_{\mathrm{T}}$ threshold		SPD threshold
	2015	2016	
Hadron	$> 3.6\,\mathrm{GeV}$	$> 3.7\,\mathrm{GeV}$	< 450
Photon	$> 2.7\,\mathrm{GeV}$	$> 2.78\,\mathrm{GeV}$	< 450
Electron	$> 2.7\,\mathrm{GeV}$	$> 2.4\,\mathrm{GeV}$	< 450
Muon	$> 2.8\,\mathrm{GeV}$	$> 1.8\,\mathrm{GeV}$	< 450
Muon high p_{T}	$> 6.0\,\mathrm{GeV}/c$	$> 6.0\,\mathrm{GeV}/c$	none
Dimuon	$> 1.69\,\mathrm{GeV}/c^2$	$> 2.25\,\mathrm{GeV}/c^2$	< 900

2.2.3.2 The software triggers

With a maximum rate of events of 1 MHz, the next trigger steps are processed in the Event Filter Farm (EFF). These are split into High-Level-Trigger-1 (HLT1) and High-Level-Trigger-2 (HLT2). HLT1 performs an inclusive selection looking at 1-2-track signatures or muon/dimuon track(s) displaced from the PV. Its output is written to disk to allow processing in-between the fills and re-calibrate the detector. HLT2 fully reconstructs events using a combination of inclusive and exclusive lines.

2.2.3.3 Trigger for $R(D^*)$

Trigger lines used in this analysis are classified as 'trigger on signal' (TOS) or 'trigger independently of signal' (TIS). The TOS events are triggered independently of the presence of the rest of the event, while the TIS events are recorded independently of the signal part within the event.

The trigger conditions are reported in Table 2.3. They consist in a logical 'or' of the topological B^0 requirements and the exclusive $B^0 \to D^{*-}$ decays with $D^{*-} \to \pi^- D^0 (\to K^+\pi^-)$. The topological lines allow to select 2-4 body decays of the B^0 meson that ensures that both signal ($B^0 \to D^{*-}\tau^+\nu_\tau$) and normalisation ($B^0 \to D^{*-}3\pi^\pm$) modes should be selected.

Table 2.3 – Trigger requirement in this analysis is a logical 'or' between 'B^0 topology' and '$B^0 \to D^{*-}$ exclusive'. The logical 'or' is implied between the lines in each cell for a given trigger step (*i.e.* L0, HLT1 and HLT2) and the logical 'and' is implied between the three trigger steps of a given column.

Level	B^0 topology	$B^0 \to D^{*-}$ exclusive
L0	B0_L0Global_TIS	B0_L0Global_TIS
	B0_L0HadronDecision_TOS	Dst_L0HadronDecision_TOS
HLT1	B0_Hlt1TrackMVADecision_TOS	B0_Hlt1TrackMVADecision_TOS
	B0_Hlt1TwoTrackMVADecision_TOS	B0_Hlt1TwoTrackMVADecision_TOS
HLT2	B0_Hlt2Topo2BodyDecision_TOS	D0_Hlt2RareCharmD02KPiDecision_Dec
	B0_Hlt2Topo3BodyDecision_TOS	D0_Hlt2CharmHadDstp2D0Pip\
	B0_Hlt2Topo4BodyDecision_TOS	_D02KmPipTurboDecision_Dec

2.3 Software

The LHCb software is based on the GAUDI framework [98]. Figure 2.17 represents a schematic overview of the LHCb data flow.

In case of collision data, the first software tool used is MOORE that runs the software triggers, HLT1 and HLT2 described in Section 2.2.3.2. Then, the events are reconstructed with BRUNEL and undergo preselection with DAVINCI, where additional useful

information about the events is computed with custom algorithms and some cuts are applied. The preselection is the first step specific to each analysis. This results in 'ntuples' in the ROOT format, ready for a specific analysis.

Simulation is an indispensable part of many analyses and this one in particular. It is conducted by the GAUSS [99] that runs three steps: pp collision generation (PYTHIA [100]), decay of resulting particles (EVTGEN [101]) and their propagation through the detector (GEANT4 [102]). Next, the GEANT4 output is digitised with BOOLE mimicking resolution of the real detector. The subsequent steps of the simulated samples are identical to those of the collision data.

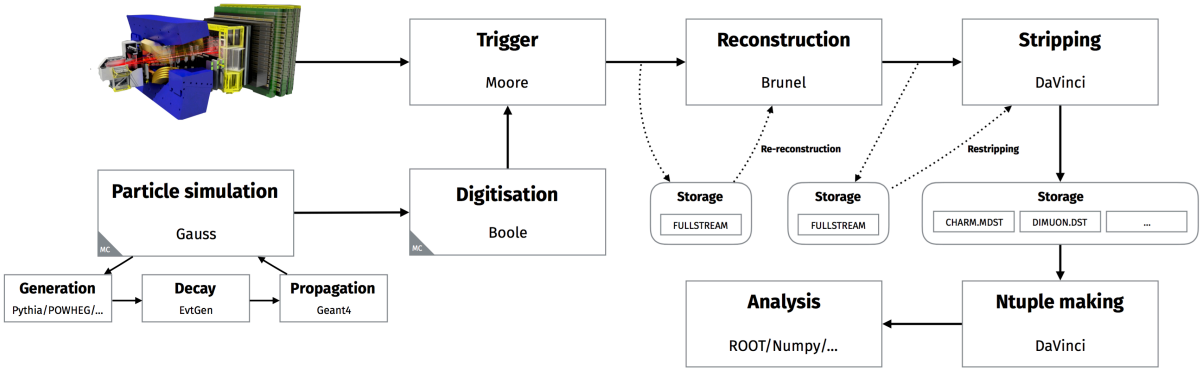


Figure 2.17 – The LHCb data and simulation flow with steps and software tools used shown.

Chapter 3

Methodology of the $R(D^*)$ measurement

This chapter provides an overview of the $R(D^*)$ measurements. Section 3.1 compares the various methods used to measure $R(D^*)$ at LHCb and the B-factories. Section 3.2 defines the method used to measure $R(D^*)$ with three-prong tau decays. Section 3.3 discusses the backgrounds involved in this measurement. Section 3.4 describes the reconstruction techniques of the decay kinematics. Finally, an overview of the analysis is covered in Section 3.5.

3.1 Measuring $R(D^*)$ at LHC

$R(D^*)$ has been measured by the BaBar, Belle and LHCb experiments.

BaBar and Belle, detecting e^+e^- collisions, produce the $\Upsilon(4S)$ resonance spawning two B mesons. One B meson is signal, reconstructed from its decay to a charm meson (D or D^*). The other one, B_{tag}, reconstructed from its either hadronic or semileptonic decay, enables to constrain the kinematics of the total BB_{tag} system and hence measure the missing mass due to neutrinos. In the $R(D^*)$ measurements at the B factories, the τ^+ is reconstructed through its leptonic modes, $\tau^+ \to \ell\nu_\tau\bar{\nu}_\ell$, where $\ell = e^+, \mu^+$ (the charged conjugate is implied throughout the text) or through the modes with a hadron, $\tau \to \pi\nu_\tau$ or $\tau \to \rho\nu_\tau$. At the $R(D^*)$ numerator, the $B^0 \to D^{(*)}\tau^+\nu_\tau$ mode, spawns 2-3 neutrinos (depending on the mode), while at the denominator, $B^0 \to D^{(*)}\ell\nu_\ell$ spawns only 1 neutrino, since the ℓ is reconstructed directly. The resulting missing mass can be obtained from $M^2_{\text{miss}} = (p_{e^+e^-} - p_{B_{\text{tag}}} - p_{D^{(*)}} - p_\ell)^2$ (in units of $c \equiv 1$), where $p_{e^+e^-}$, $p_{B_{\text{tag}}}$, $p_{D^{(*)}}$ and p_ℓ are the 4-momenta of the beam or the corresponding particles. For the signal events, the M_{miss} distribution peaks above zero, due to missing 2-3 neutrinos, while it peaks at zero in case of $B^0 \to D^{(*)}\ell\nu_\ell$, where only 1 neutrino is missing, making it possible to distinguish the two. Beside the missing mass, another variable is used in the fit, for example the magnitude of the three-momentum of the lepton, measured in the B rest frame, $|\mathbf{p}^*_\ell|$ [31, 32] or an output of a neural network [33].

The LHC, in its pp collisions has no information about the energy of the interacting partons. Therefore the energy of the $b\bar{b}$ produced is unknown. Even if the $b\bar{b}$ system decays entirely in the LHCb acceptance, one cannot constrain the kinematics of its hadronisation. Furthermore, hadron colliders have higher multiplicities making the events significantly more contaminated. Consequently, the semileptonic decays in the $R(D^*)$ ratio can be only partially reconstructed (*cf.* Section 3.4). Moreover, to measure $R(D^*)$, the analysts rely on large simulation samples of events or 'cocktails' thereof (*cf.* Chapter 4) to obtain their templates, as well as on complex multivariate selection (*cf.* Chapter 5).

There are two branches of the $R(D^*)$ measurements at LHCb: with the muonic τ reconstruction, $\tau \to \mu\overline{\nu}_\mu\nu_\tau$, used in Ref. [34], and the 3-prong decay, $\tau^+ \to \pi^+\pi^-\pi^+(\pi^0)\overline{\nu}_\tau$ [1], used in Refs. [38, 39] and the analysis presented in this thesis. The hadronic mode allows for precise reconstruction of the τ^+ vertex thanks to the intersection of the three charged pions trajectories.

3.2 The signal, normalisation and $\mathcal{K}(D^*)$

Let us return to the $R(D^*)$ definition:

$$R(D^*) \equiv \frac{\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)}. \tag{3.1}$$

The mode in the numerator of $R(D^*)$, $B^0 \to D^{*-}\tau^+\nu_\tau$, will be referred to throughout as the signal. Let us discuss how its decay products can be observed. The D^{*-} instantly decays mostly[2] ($(67.7 \pm 0.5)\%$) into $\overline{D}^0\pi^-$ and then $\overline{D}^0$ decays into K^+ and π^- ($(3.950 \pm 0.031)\%$). The two decay modes of the τ considered for $R(D^*)$ at LHCb are the muonic one, $\tau^+ \to \mu^+\nu_\mu\overline{\nu}_\tau$ ($\approx 17.4\%$), and the hadronic one, $\tau^+ \to \pi^+\pi^-\pi^+(\pi^0)\overline{\nu}_\tau$ ($\approx 13.5\%$). The latter one is chosen in this analysis. Neutrinos are not detected, making it impossible to reconstruct a clear B^0 mass peak for the signal mode. There is, however, a technique to cope with it as discussed in Section 3.4. Thus, the visible final state of the signal mode consists of $5\pi^\pm$ and one K^+. The signal mode topology is shown in Figure 3.1 (left).

[1] The branching fractions of the two modes are: $\mathcal{B}(\tau \to \mu\overline{\nu}_\mu\nu_\tau) \approx 17.4\%$ and $\mathcal{B}(\tau^+ \to \pi^+\pi^-\pi^+(\pi^0)\overline{\nu}_\tau) \approx 13.5\%$.

[2] See Table A.1 in Appendix A for a review of the relevant branching fractions.

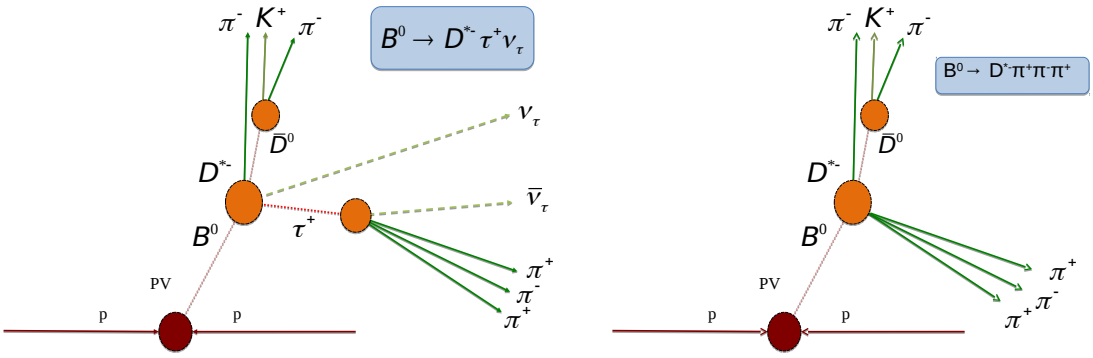


Figure 3.1 – Schematic view of the $B^0 \to D^{*-}\tau^+\nu_\tau$ signal decay (left) and the $B^0 \to D^{*-}3\pi^\pm$ normalisation mode (right).

The three-prong decays allow for the τ vertex reconstruction. Also, that decay mode enables angular analysis studies, which go beyond the scope of the work presented in this thesis, that would be sensitive to New Physics parameters. This latter feature is currently used in ongoing analyses at LHCb.

The denominator of Eq. 3.1, has been measured by B-factories and LHCb [103] as: $\mathcal{B}(B^0 \to D^{*-}\ell\nu_\mu) = (5.06 \pm 0.02 \pm 0.12)\%$, where $\ell = e^+, \mu^+$.

At this point we have all the ingredients to measure $R(D^*)$ hadronically. Nonetheless, introducing the normalisation mode, $B^0 \to D^{*-}3\pi^\pm$, of the same visible final state as the signal, whose topology is depicted in Figure 3.1 (right), can improve the experimental precision, because it cancels out systematic detector and reconstruction effects. Then, the target measurement, denoted by $\mathcal{K}(D^*)$, becomes:

$$\mathcal{K}(D^*) = \frac{\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{*-}\pi^+\pi^-\pi^+)}\,, \tag{3.2}$$

which can be expressed in terms of the relative yields of the signal and normalisation modes, N_{sig}, N_{norm}, obtained from fitting the data, divided by their respective efficiencies, $\varepsilon_{\rm sig}$ and $\varepsilon_{\rm norm}$ obtained mostly from MC studies, divided by the sum of the branching fractions of the τ^+ decay modes considered:

$$\mathcal{K}(D^{*-}) \equiv \frac{\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{*-}3\pi)} = \frac{N_{\rm sig}}{N_{\rm norm}}\frac{\varepsilon_{\rm norm}}{\varepsilon_{\rm sig}}\frac{1}{\mathcal{B}(\tau^+ \to 3\pi\overline{\nu}_\tau) + \mathcal{B}(\tau^+ \to 3\pi\pi^0\overline{\nu}_\tau)}\,. \tag{3.3}$$

Finally, $R(D^*)$ is obtained as:

$$R(D^*) = \mathcal{K}(D^{*-})\frac{\mathcal{B}(B^0 \to D^{*-}\pi^+\pi^-\pi^+)}{\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)}. \tag{3.4}$$

The normalisation mode has been measured by LHCb [104], BaBar [105] and Belle [106] with a weighted average: $(7.21 \pm 0.28) \times 10^{-3}$.

3.3 Backgrounds

The most important backgrounds are:

- 3π directly from B decays; so called *prompt* backgrounds
- $B^0 \to D^{*-}DX$ decays where $D \to 3\pi X$; so-called *double charm* backgrounds
- Feed-down from excited D^{*-} mesons; denoted by D^{**}
- Non-physical wrong-sign (WS) decays $D^{*\pm}\tau^{\pm}$, where the D^* and τ candidates have the same charge
- Combinatorics, where particles from different decays are accidentally combined in one signal candidate

3.4 Reconstruction of decay kinematics

In the signal decay the two neutrinos are not reconstructed. However, thanks to precisely measured decay vertices of the $\overline{D}^0$, $3\pi^{\pm}$ system and B^0 (this is enabled by the VELO detector), and exploiting some vectorial algebra, it is possible to mitigate that shortcoming. Also, similar techniques are used in reconstructing backgrounds with missing neutral particles.

The following sections describe the reconstruction techniques in the signal and $B \to D^{*-}D_s^+(X)$, with $D_s^+ \to 3\pi N$ hypotheses. These methods were described in the $R(D^*)$ Run-1 analysis [107] and are summarised in the following two sections.

The resulting momenta are exploited by the MVA described in Section 5.8. They also allow for derivation of the q^2 and lifetime of the tau-candidate, τ. These are all three variables used in the signal yield fit in Section 7.

3.4.1 Reconstruction in the signal hypothesis

Despite the two missing neutrinos, the well-measured vertices of B^0 and τ^+ allow for reconstruction of their flight directions (where the B^0 one is a line from the primary vertex to the B^0 decay vertex; the τ^+ one is a line between the B^0 and τ^+ decay vertices). Using their known masses allows to derive the magnitude of the momenta up to 2-fold ambiguities in the laboratory frame, in units where $c = 1$.

From the energy conservation in the decay $\tau \to 3\pi\nu_\tau$ one can derive the magnitude of the τ momentum as a function of the angle between the τ and 3π vectors, $\theta_{\tau,3\pi}$, to two-fold ambiguity:

$$|\vec{p}_\tau| = \frac{(m_{3\pi}^2 + m_\tau^2)|\vec{p}_{3\pi}|\cos\theta_{\tau,3\pi} \pm E_{3\pi}\sqrt{(m_\tau^2 - m_{3\pi}^2)^2 - 4m_\tau^2|\vec{p}_{3\pi}|^2\sin^2\theta_{\tau,3\pi}}}{2(E_{3\pi}^2 - |\vec{p}_{3\pi}|^2\cos^2\theta_{\tau,3\pi})}, \tag{3.5}$$

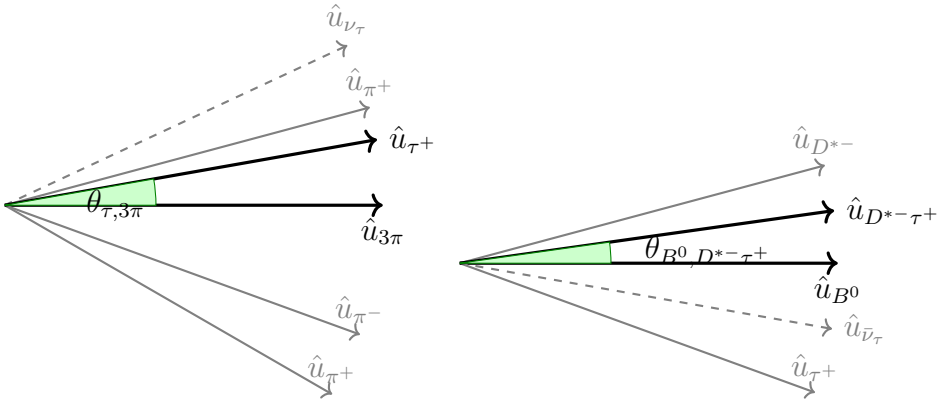


Figure 3.2 – Illustration of the angle between the $3\pi\nu$ system and the τ (left) and between the $D^{*-}\tau^+$ system and the B^0 (right).

where $\theta_{\tau,3\pi}$ is the angle between the 3π system three-momentum and the τ line of flight, illustrated in Figure 3.2 (left); $m_{3\pi}$, $|\vec{p}_{3\pi}|$ and $E_{3\pi}$ are the mass, three-momentum and energy of the 3π system, respectively; and m_τ is the known τ mass.

Using an approximation $\theta_{\tau,3\pi} = \theta^{\max}$ that gives

$$\theta^{\max}_{\tau,3\pi} = \arcsin\left(\frac{m_\tau^2 - m_{3\pi}^2}{2m_\tau|\vec{p}_{3\pi}|}\right), \tag{3.6}$$

making the square root in Eq. 3.5 vanish, yielding a non-ambiguous estimate of the τ momentum.

The B^0 momentum is obtained similarly:

$$|\vec{p}_{B^0}| = \frac{(m_Y^2 + m_{B^0}^2)|\vec{p}_Y|\cos\theta_{B^0,Y} \pm E_Y\sqrt{(m_{B^0}^2 - m_Y^2)^2 - 4m_{B^0}^2|\vec{p}_Y|^2\sin^2\theta_{B^0,Y}}}{2(E_Y^2 - |\vec{p}_Y|^2\cos^2\theta_{B^0,Y})} \tag{3.7}$$

with

$$\theta^{\max}_{B^0,Y} = \arcsin\left(\frac{m_{B^0}^2 - m_Y^2}{2m_{B^0}|\vec{p}_Y|}\right), \tag{3.8}$$

where Y represents the $D^{*-}\tau$ system. The angle between the $D^{*-}\tau$ system and B^0 is illustrated in Figure 3.2 (right). The three-momentum and energy of the $D^{*-}\tau$ system, using the estimated τ momentum are:

$$\vec{p}_Y = \vec{p}_{D^{*-}} + \vec{p}_\tau, \quad E_Y = E_{D^{*-}} + E_\tau, \tag{3.9}$$

where $\vec{p}_{D^{*-}}$ and $\vec{p}_\tau$ are the three-momenta of the D^{*-} and the τ candidates, and $E_{D^{*-}}$ and E_τ their energies.

This method is applied to obtain the rest frame variables $q^2 \equiv (p_{B^0} - p_{D^{*-}})^2 = (p_\tau + p_{\nu_\tau})^2$ and the τ decay time, t_τ, with sufficient precision to discriminate the signal from the $D^{*-}D\,X$ backgrounds in the signal yield fit in Section 7.

Figure 3.3 compares the reconstructed and truth-matched distribution of q^2 for the simulated signal $B^0 \to D^{*-}\tau^+\nu_\tau$ sample after initial selection(Section 5.4). The agreement is satisfactory, although one can observe a small bias in the lower values of the q^2 distribution. This bias should be addressed in systematics.

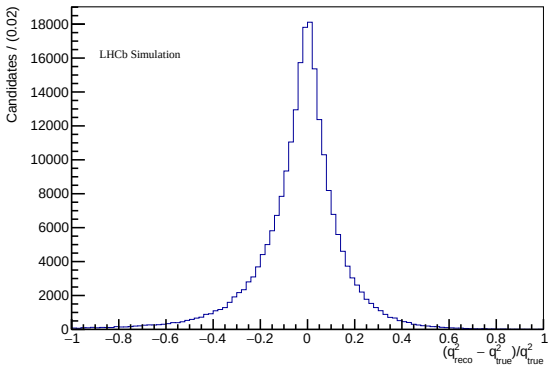


Figure 3.3 – Comparison of reconstructed and true q^2 for the simulated signal $B^0 \to D^{*-}\tau^+\nu_\tau$ sample after initial selection(Section 5.4).

3.4.2 Reconstruction assuming a double-charm origin for the candidate

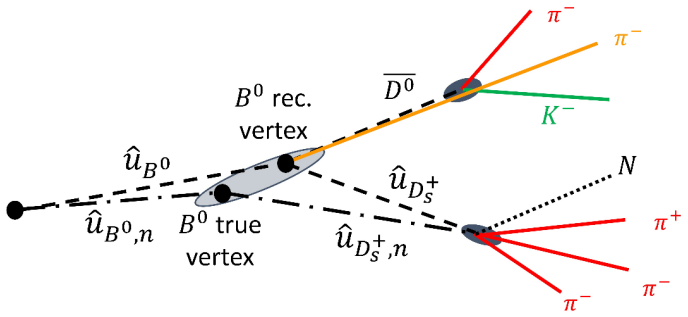


Figure 3.4 – Flight directions of the B^0 and D_s^+ in the decay $B^0 \to D^{*-}D_s^+(\to \pi^+\pi^-\pi^+N)$. The original B^0 flight direction $\hat{u}_{B^0}$ is compared to the improved one by accounting for neutral particles in the D_s^+ decay, $\hat{u}_{B^0,n}$.

The dominant backgrounds after the selection comprise the $B^0 \to D^{*-}DX$ decays, especially $B^0 \to D^{*-}D_s^+(\to \pi^+\pi^-\pi^+N)$ where N denotes a system of unreconstructed neutral particles. The anti-D_s^+ MVA in Section 5.8 requires kinematic reconstruction of this mode described here.

Knowing the flight directions of the B^0 and D_s^+, the D^{*-} vertex and using the known B^0 mass, allows us to write the momentum conservation here as:

$$|\vec{p}_B|\hat{u}_B = |\vec{p}_{D_s^+}|\hat{u}_{D_s^+} + \vec{p}_{D^{*-}}. \tag{3.10}$$

The solutions can be obtained through a vectorial or scalar product techniques as:

$$|\vec{P_{B,v}}| = \frac{|\vec{p}_{D^{*-}} \times \hat{u}_{D_s^+}|}{|\hat{u}_B \times \hat{u}_{D_s^+}|}, \tag{3.11a}$$

$$|\vec{P_{B,s}}| = \frac{\vec{p}_{D^{*-}} \cdot \hat{u}_B - (\vec{p}_{D^{*-}} \cdot \hat{u}_{D_s^+})(\hat{u}_B \cdot \hat{u}_{D_s^+})}{1 - (\hat{u}_B \cdot \hat{u}_{D_s^+})^2}, \tag{3.11b}$$

for the B^0 momentum, and

$$|\vec{P_{D_s,v}}| = \frac{|\vec{p}_{D^{*-}} \times \hat{u}_B|}{|\hat{u}_{D_s^+} \times \hat{u}_B|}, \tag{3.12a}$$

$$|\vec{P_{D_s,s}}| = \frac{(\vec{p}_{D^{*-}} \cdot \hat{u}_B)(\hat{u}_B \cdot \hat{u}_{D_s^+}) - \vec{p}_{D^{*-}} \cdot \hat{u}_{D_s^+}}{1 - (\hat{u}_B \cdot \hat{u}_{D_s^+})^2}, \tag{3.12b}$$

for the D_s^+ momentum.

These methods are equivalent in the absence of extra particles, but not otherwise, therefore the both results are used in the anti-D_s^+ MVA. Since this partial reconstruction works without imposing a mass to the $3\pi N$ system, the reconstructed $3\pi N$ mass can be used as a discriminating variable.

The reconstructed B^0 momenta, $|\vec{P_{B,v}}|$ and $|\vec{P_{B,s}}|$, can be further improved by accounting for the presence of neutral particles in the D_s^+ decay, denoted by $|\vec{P_{B,vn}}|$ and $|\vec{P_{B,sn}}|$.

At this point one might be tempted to think that this technique can discriminate the $B^0 \to D^{*-}D_s^+$ backgrounds from the signal, which has two missing neutrinos. Unfortunately other backgrounds, such as $B \to D^{*-}D^0(X)$ and $B \to D^{*-}D^+(X)$ decays, also have missing energy (due to two unreconstructed kaons from the B^0 and $3\pi^\pm$ vertices), making them similar to the signal. Therefore, the just described reconstruction must be exploited by a precise MVA described in the next chapter.

3.5 Workflow

The workflow of this analysis is quite involved. To facilitate reading the following chapter, Figure 3.5 represents a simplified workflow of the project. It can be described in the following sequence:

1. **Dataset:** Simulation software is tailored to produce all the samples required. Here Generator cuts and Filtering scripts are made or adapted to spare CPU usage and/or disk space for the huge Monte Carlo (MC) samples to be produced. Filtering (Stripping) is applied as a means of pre-selection. Also physics checks are made for both (i) the correct decay chains in the decay-describing simulation inputs (so-called DEC files) and (ii) first ReDecay validation. Preparing, requesting and producing simulation took us around one year of work.
2. **Selection:** This step involves reprocessing the 'raw' simulation and collision data to the analysis-ready ntuple format using the DaVinci software. Initial cut-based

selection is applied there. Then MC is reweighted (kinematics and, for signal only, form-factors) to agree better with the expected collision data. Lastly, the data and simulation undergo cut-based and multivariate cuts, which differ for the signal, normalisation and control samples modes.

3. **Obtaining $\mathcal{K}(D^*)$:** here the normalisation and signal (final fit) yields are extracted, efficiencies are computed and $\mathcal{K}(D^*)$ is inferred.
4. **Result:** combining the $\mathcal{K}(D^*)$ value and the external branching fractions results in $R(D^*)$. Systematics are estimated to account for the uncertainty on the result.

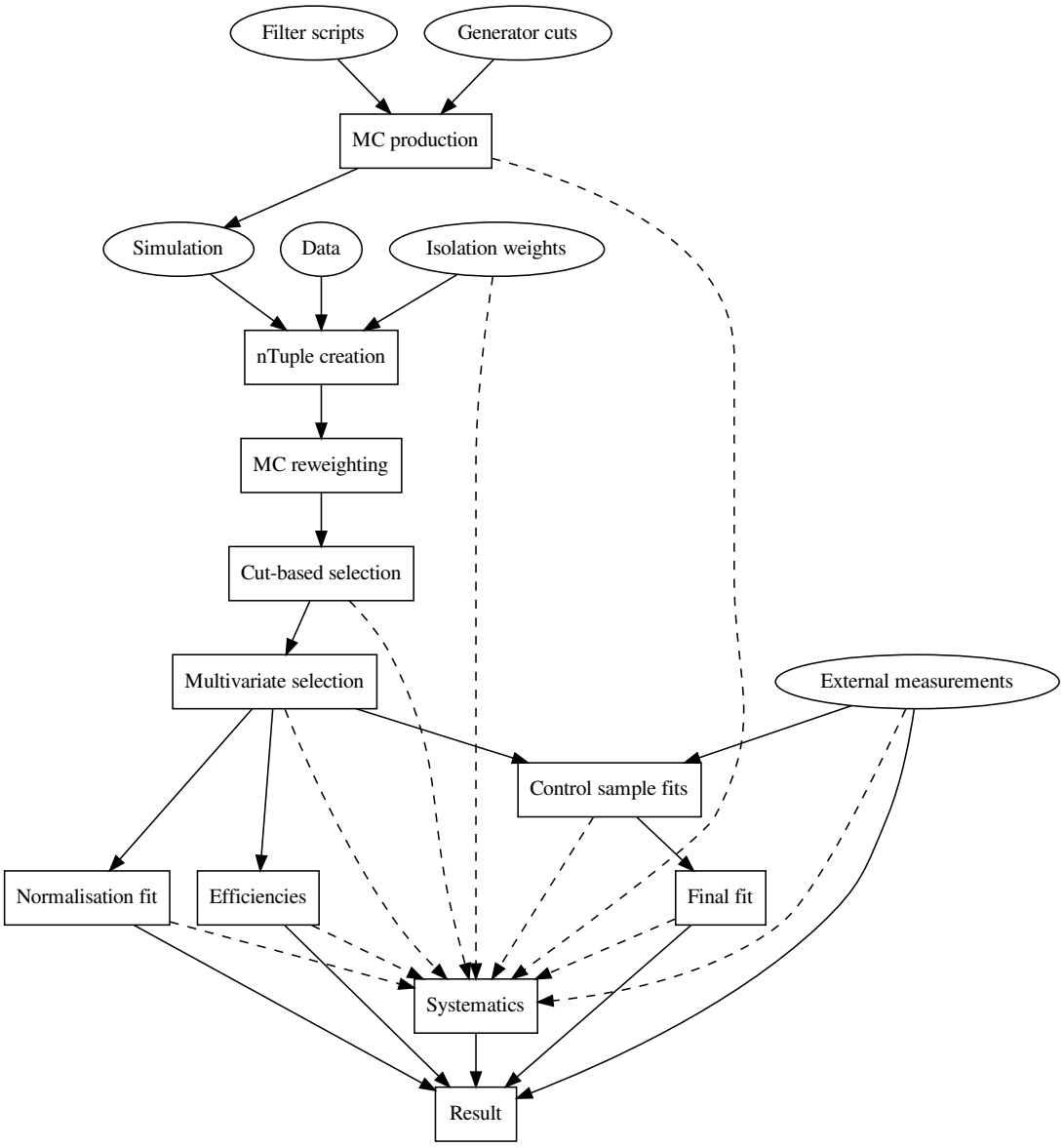


Figure 3.5 – Analysis workflow from the simulation setup, through the selection, to the signal and normalisation fits, efficiencies computation, systematics and final $R(D^*)$ result.

The next chapter describes the collision and simulation datasets, comprising the first step in the list above. A special emphasis is put on the validation of the fast simulation algorithm ReDecay.

Chapter 4

Dataset

This chapter presents the collision and simulated dataset used for the project in Section 4.1. Then, Section 4.2 discusses various fast simulation techniques available at LHCb and validates the chosen one, ReDecay.

4.1 Collision and simulated datasets

In this section are given the real data and MC samples used to perform our analysis.

4.1.1 Real data samples

The real data are the full datasets recorded by LHCb during the year 2015 and 2016, i.e. $2\,\text{fb}^{-1}$ of data. The centre-of-mass energy for this period was $13\,\text{TeV}$.

4.1.2 MC samples

This analysis requires very large MC samples, to model the shape of the background used in the final signal fit. We generated 2.5 billions of simulated events, during the year 2018. To achieve that, we pioneered the use of the ReDecay software [108], within the LHCb semileptonic group. Table 4.1 summarised the MC samples produced. The signal modes are simulated using the TAUOLA package [109].

Table 4.1 – Monte Carlo samples produced for this analysis.

Decay	No. events produced [M]
$B^0 \to D^{*-}\tau^+(\to \pi^+\pi^-\pi^+\overline{\nu}_\tau)\nu_\tau$	29
$B^0 \to D^{*-}\tau^+(\to \pi^+\pi^-\pi^+\pi^0\overline{\nu}_\tau)\nu_\tau$	49
$B^0 \to D^{*-}\pi^+\pi^-\pi^+$	10
$B^0 \to D^{**-}\tau^+(\to \pi^+\pi^-\pi^+\overline{\nu}_\tau)\nu_\tau$	20
$B_s^0 \to D^{*-}D_s^+X$	23
$B^+ \to D^{**0}D_s^+X$	13
$B^0 \to D^{*-}D_s^+X$	12
$B^0 \to D^{**-}D_s^+X$	1
$b\bar{b} \to D^{*-}\pi^+\pi^-\pi^+X$	34
$b\bar{b} \to D^{*-}D^{\{0,+\}}X$	3
total	194

In $B^0 \to D^{**}\tau^+(\to 3\pi^\pm\overline{\nu}_\tau)\nu_\tau$, D^{**} denotes any of the states: $D_1(2420)^{0,-}, D_0^*(2300)^{0,-}, D_1'(2430)^{0,+}$ or $D_2(2460)^{0,+}$. In $B^+ \to D^{**0}D_s^+X$ and $B^0 \to D^{**-}D_s^+X$, D^{**} denotes: $D_1(2420)^{0,-}, D_1'(2430)^{0,+}$ or $D_2(2460)^{0,+}$. Table 4.2 reports the branching fractions and the visible contributions of various D^{**} species as simulated in the inclusive $b\bar{b} \to D^{*-}3\pi^\pm X$ sample.

Table 4.2 – List of branching fractions of various $D\hat{}^{**}$ decays in the inclusive $b\bar{b} \to D^{*-}3\pi^\pm X$ MC.

D^{**} state	$\mathcal{B}(B \to D^{**}\tau\nu_\tau)$ [%]	$\mathcal{B}(D^{**} \to D^{*-})$ [%]	visible contribution [%]
$D_1'(2430)^0$	0.2	0.66	0.132
$D_1(2420)^0$	0.13	0.67	0.09
$D_2^*(2460)^0$	0.2	0.21	0.042
$D_0(2400)^0$	0.2	0	0
Total from B^+			0.264
$D_1'(2430)^+$	0.2	0.33	0.066
$D_1(2420)^+$	0.13	0.33	0.04
$D_2^*(2460)^+$	0.2	0.1	0.02
$D_0(2400)^+$	0.2	0	0
Total from B^0			0.13

4.2 Fast simulation with ReDecay

A necessary condition to make this measurement competitive is to improve its systematic precision with respect to the previously published Run-1 $R(D^*)$ with hadronic τ reconstruction [39, 38]. The major factor contributing to the systematic uncertainty was the sample size of the simulated events. To lower that uncertainty such that it

becomes equal to the statistical one, required approximately 4-6 times increased sample size for most of the event types. Discussions with the Run-1 analysis proponents resulted in finally producing the simulated samples as reported in Table 4.1.

Given the demanded size of the simulation, typical production times (queues within LHCb computing resources) and the analysts' time constraints (PhD, post-doc contracts duration), it was clear from the beginning that a much faster simulation approach was needed.

Therefore, one of the first tasks in the project was to research the fast simulation techniques available within LHCb, including validating them specifically for this project.

In LHCb there exist several such techniques. Some of them reduce the simulated event to one decay mode (so-called signal, but it may refer to a background event) by rejecting all 'underlying' tracks, which do not come from the same 'mother' hadron. We could not, however, choose such strategy, because we need information from the full event to develop charged track isolation algorithms.

Another way to speed up the simulation is used by the ReDecay algorithm [110]. The idea is to decay the mother hadron multiple times (usually 100 times) while sharing the underlying part of the event (*i.e.* all tracks that do not originate from the mother hadron) for all decays. This way the original event is fully simulated once, but spawns 99 extra events, produced at low computing cost. Two facts are noteworthy:

- The 'mother' hadron combined with the remainder of the event ensures energy and momentum conservation;
- Each of 100 decays of the 'mother' hadron decays generally at a different point, following the exponential lifetime distribution. Hence, each decay 'could have happened' in a full simulation.

4.2.1 ReDecay rationale

ReDecay is designed to shorten the simulation time by factor $\mathcal{O}(10)$. If it were just for the sake of speed, the Particle Gun or Tracker-only techniques could have been used. The former simulates only a decay of one of the modes listed in Table 4.1, while the rest of the event, vital for the charged isolation study, is ignored, disqualifying possibility of using this method. The Tracker-only simulation misses information from the RICH detectors, implying no PID variables such as `ProbNNpi` and `ProbNNk`. ReDecay preserves the full event information and is compatible with simulating the full detector, therefore was the only solution, readily available in 2018, up to the task.

4.2.2 ReDecay validation

Before launching the simulation jobs, the ReDecay applicability had to be carefully checked. Figure 4.1 illustrates the concept of ReDecay in the signal mode.

- First, a single event is fully simulated (*e.g.* Figure 4.1 left).
- The B^0 at its *origin* (not decay) vertex is saved with its momentum.
- The underlying (*i.e.* non-signal) tracks are also saved.
- The following $N-1$ (here $100-1$) events occur as a merge of (i) the tracks/vertices from a new decay of the stored B^0 and (ii) the tracks from the preserved underlying event. One such event is depicted on the RHS in Figure 4.1.
- The above procedure constitutes one ReDecay block. There is as many blocks as number of events requested divided by the size of a single block, N. Consequently each block corresponds to one 'original' complete event.

The main concern with ReDecay was potential correlations within its blocks. Recycling the underlying event and the B^0 origin vertex might reduce event-to-event variation resulting in a dataset of smaller statistical power than a fully-simulated one of the same size. Whether, or to what extend, that was the case had to be found out by inspecting distributions of the most important variables from small-size validation samples produced.

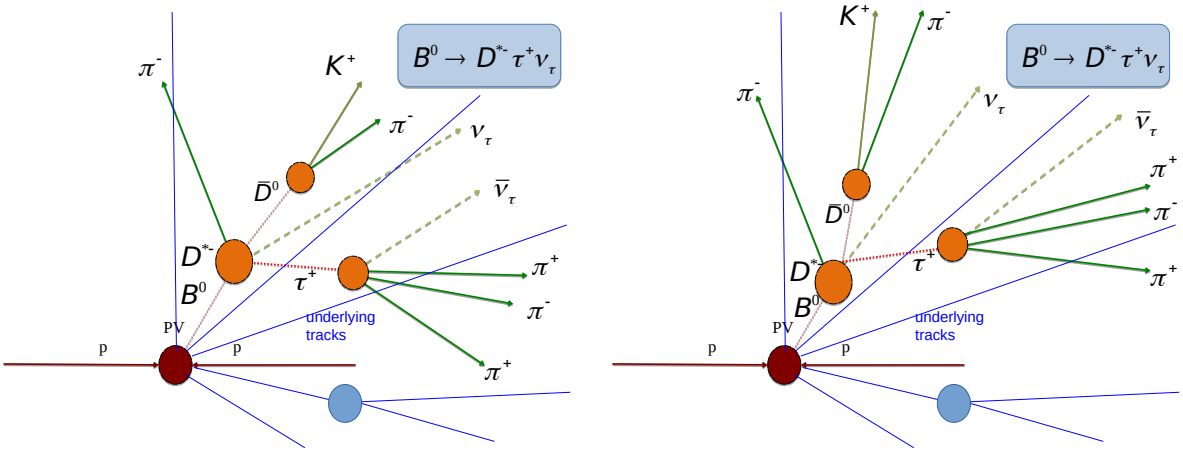


Figure 4.1 – Illustration of two $B^0 \to D^{*-}\tau^+\nu_\tau$ ($\tau^+ \to 3\pi^\pm\overline{\nu}_\tau$) events as simulated in one ReDecay block. Note that the vertices and tracks of the signal mode differ, while the underlying event remains the same.

Figure 4.2 (left) shows agreement of $\eta(B^0)$ between ReDecay and full simulation. This variable has the worst, but still acceptable, agreement. Other variables are reported in Appendix B. Figure 4.2 (right) depicts bin-to-bin correlation matrix due to using ReDecay. The off-diagonal datapoints point to correlation of events within one block. Therefore, the statistical uncertainty of a histogram of any quantity does not follow the Poisson distribution, but has to be obtained otherwise, for example by bootstrapping the full ReDecay blocks. See Appendix B for details.

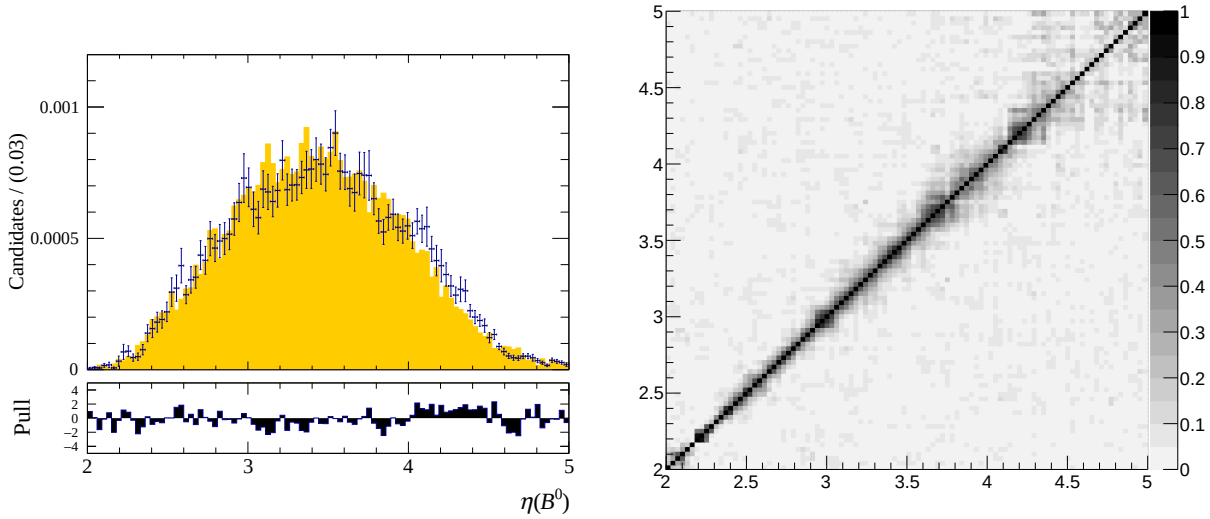


Figure 4.2 – Left: $\eta(B^0)$ distribution for the 28k $B^0 \to D^{*-}\tau^+\nu_\tau$ ($\tau^+ \to 3\pi^\pm\overline{\nu}_\tau$) mode events after stripping; full simulation is in yellow and ReDecay in bars with uncertainty. Right: Bin-to-bin correlation of $\eta(B^0)$.

Another check concerns the uncertainties in bin counts for any observables in the samples obtained with ReDecay. Due to the possibility of having more than one event in the ReDecay block, the bin content uncertainty might differ from the Poissonian one (*i.e.* $\sqrt{N}$ for N entries in a bin). Figures 4.3 and 4.4 compare the bin contents and uncertainties between the usual approach, block- and event-bootstrapped samples for the signal yield fit variables after full selection, respectively for signal and $b\bar{b} \to D^{*-}3\pi^\pm X$ inclusive background. The top panel shows the distribution of a given variable for reference. The middle one demonstrates that the bin content is the same in all three methods and is meant as a technical check of the bootstrapping implementation. The bottom one compares the bin uncertainties, where the bootstrapped cases usually, but not always, exhibit larger uncertainties than as calculated with the $\sqrt{N}$ method. The difference is usually below 2% and frequently below 1%. This observation motivates investigating the impact of ReDecay on the signal yield templates uncertainties and possibly using the bootstrapped templates in future updates of the analysis.

The conclusion from this chapter is that ReDecay is perfectly applicable for the analysis, however one has to be careful assessing its bin-content uncertainties, where the safest approach is block-bootstrapping.

Having both the collision and simulated datasets ready, the next step is to apply selection as explained in the next chapter.

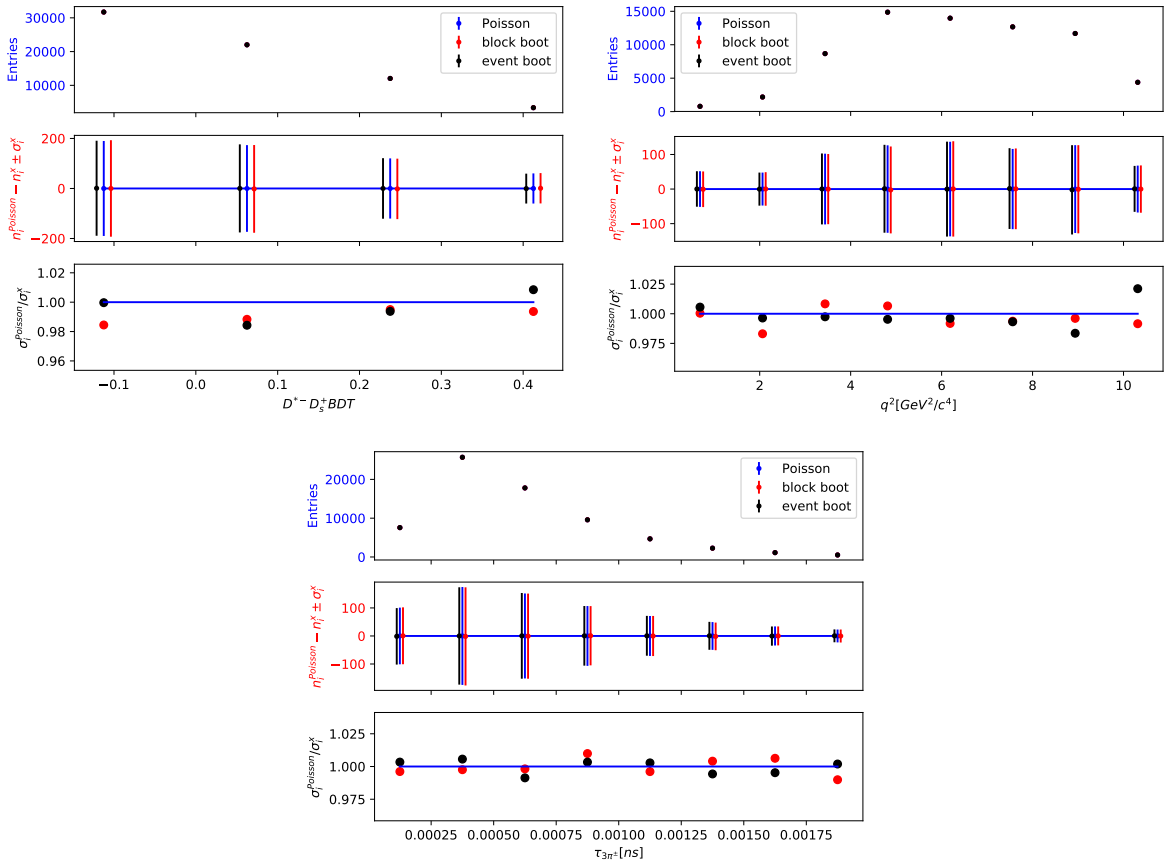


Figure 4.3 – Signal yield fit variables from the $B^0 \to D^{*-}\tau^+\nu_\tau$ MC. For each variable top panel shows the distribution itself; the middle panel shows the difference between the default bin content (blue) and those obtained with block (red) or event (black) bootstrapping; the bottom panel shows the uncertainty ratio of the Poisson method (blue) and block (red) or event (black) bootstrapping.

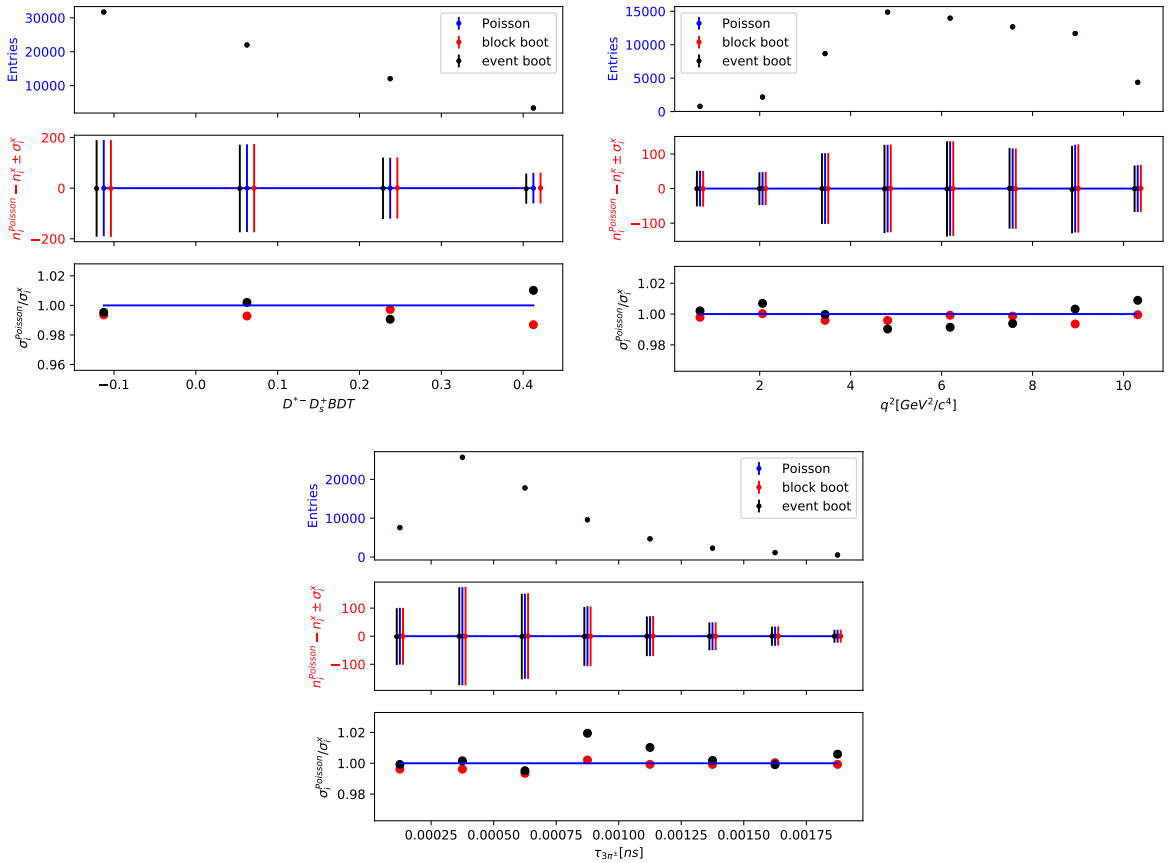


Figure 4.4 – Signal yield fit variables from the $b\bar{b} \to D^{*-}\pi^+\pi^-\pi^+X$ MC. For each variable top panel shows the distribution itself; the middle panel shows the difference between the default bin content (blue) and those obtained with block (red) or event (black) bootstrapping; the bottom panel shows the uncertainty ratio of the Poisson method (blue) and block (red) or event (black) bootstrapping.

Chapter 5

Selection of signal and normalisation modes

This chapter discusses how the signal, normalisation and control samples (*cf.* Chapter 4) are selected. Most of the selection workflow is common for all the samples until the last steps of the offline selection. An exception is made for the combinatorial background which is studied with Wrong Sign (WS) [1] candidates, that need to have a distinct selection already at the first stage, the preselection ('stripping' in LHCb jargon). Then they follow the same steps as for the other modes.

Data and MC selection is a multi-step process that can be split in the following categories:

- 'Online' selection, which consists of all the steps before the data is saved in the ROOT ntuples format.
- 'Offline' selection, which consists of all the steps after the data is saved in the ROOT ntuples format.

The 'online' selection includes the geometrical acceptance of the LHCb detector (Section 5.1), its trigger (Section 5.2) and preselection (Section 5.3).

The 'offline' selection starts by applying initial cuts that are common to all samples (Section 5.4). Next, the Multivariate Analysis techniques are treated in Sections 5.5, 5.6, 5.7 and 5.8. The distinction between the signal and normalisation modes is made in Sections 5.9 and 5.10. Finally the efficiencies are summarised in Section 5.11, where the normalisation to signal mode efficiency ratio is computed. This ratio is the first intermediate result obtained that enters the $R(D^*)$ formula.

[1] Where the $D^{*\pm}$ are required to be of the same charge as $\tau^{\pm}$ as a proxy for combinatorial B^0 decays whereby the $D^{*\pm}$ and $\tau^{\pm}$ come from two different b-hadrons.

5.1 Geometrical acceptance

Due to the geometrical coverage of the LHCb detector, only a part of the events of interest are recorded. This effect impacts efficencies of the signal and normalisation modes and is studied for the simulated samples with EvtGen. The geometrical acceptance obtained is reported in Table 5.9.

5.2 Trigger

The trigger system, as described in Chapter 2, saves on disk events of potential interest to the LHCb physics programme in general. In this analysis it is fundamental to precisely measure the ratio of the normalisation and signal efficiencies. This is studied with simulation and requires knowing the yields before and after the trigger cuts. Therefore, in simulation trigger cuts are postponed to a later stage in the pipeline: after the ROOT ntuples have been made. Obviously, that changes the selection order with respect to the real collision data, but does not impact the total efficiencies being the product of efficiencies of all the steps. Furthermore, due to the altered order, the trigger efficiencies as measured for the simulated samples do not reflect those of real data. The trigger efficiencies as measured after the geometrical acceptance and preselection are reported in Table 5.10.

5.3 Preselection

After the geometrical acceptance and trigger cuts have been applied to data, the next step is preselection, referred to at LHCb as 'stripping'. It comprises generic selections ('lines') used usually by more than one analysis. In this project we rely on the two 'lines' written by the proponents of the previous $R(D^*)$ measurement [38, 39]:

- `StrippingB0d2DstarTauNuForB2XTauNuAllLines`,
- `StrippingB0d2DstarTauNuWSForB2XTauNuAllLines`.

For most parts of the project the selection with the first line is used. The second one is needed only to study combinatorial background, often denoted by 'Wrong Sign' (WS).

The first one, `StrippingB0d2DstarTauNuForB2XTauNuAllLines`, forms a B^0 candidate by combining a $D^{*-} \to \overline{D}^0(\to K^+\pi^-)\pi^-$ candidate with a $\tau^+ \to 3\pi^\pm\overline{\nu}_\tau$ candidate.

The second one, `StrippingB0d2DstarTauNuWSForB2XTauNuAllLines`, forms a B^0 candidate from a same-sign combination of a $D^{*\pm}$ and a $\tau^\pm$.

Apart from these sign differences, the cuts are the same in the two lines and are summarized in Table 5.1, where

- DIRA(particle, PV) is the cosine of the angle between the particle momentum and the line of flight from the best primary vertex to the particle decay vertex (see Figure 5.1);
- DOCA(particle$_i$, particle$_j$) is the distance of closest approach between the flight directions of the particle$_i$ and particle$_j$.

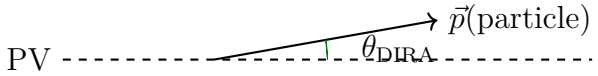


Figure 5.1 – Illustration of DIRA(particle, PV).

The stripping versions used are 24r1 for the 2015 data and 28r1 for the 2016 data.

First, the $K^-\pi^+$ system is formed by requirements concerning $K^-\pi^+$ track momenta (p and/or p_{T}), vertex quality χ^2/ndf, track ghost probability (*i.e.* probability that a given track does not correspond to a real particle passage), sufficient distance from the PV (the variables χ^2 and χ^2_{IP} relative to the PV), and particle identity (PIDK is a DLL variable describing a probability that a particle is a kaon; the DLL variables are discussed in Section 2.2.2.4). Then, a D^0 candidate is built by constraining the absolute difference from the known D^0 mass. Moreover, the K^- and π^+ from the D^0 need to satisfy DOCA$(K^-, \pi^+) < 0.5\,\text{mm}$ and form a vertex of a χ^2-quality of $\chi^2_{\text{DOCA}(K^-,\pi^+)} < 15$. Next, similar track quality requirements as for K^- and π^+ are applied to the π^- from D^{*-}. Then, the D^{*-} candidate is constructed based on the thus obtained π^+D^0 system by the requirements of $\chi^2_{\text{vtx}}/\text{ndf} < 25$ on the $D^0 - \pi^-$ vertex, a requirement of $135 < m(D^{*-}) - m(D^0) < 160\,\text{MeV}/c^2$, $p_{\text{T}} > 1250\,\text{MeV}/c$ and $m(D^{*-})$ within 50 MeV/c^2 of the known D^{*-} mass.

The $\tau^+ \to 3\pi^\pm\overline{\nu}_\tau$ candidate is formed by cuts on its reconstructed mass $m(3\pi^\pm) \in [400, 3500]\,\text{MeV}/c^2$, track quality $\chi^2_{\text{track}}/\text{ndf} < 25$, DIRA($\tau^+$, PV) > 0.99, max[DOCA($\pi^\pm{}_i, \pi^\pm{}_j$)] $< 0.15\,\text{mm}$ (where $\pi^\pm{}_i$ and $\pi^\pm{}_j$ are the $\pi^\pm$ candidates in the $3\pi^\pm$ system), min$[m(\pi^+\pi^-)] < 1670\,\text{MeV}/c^2$, at most one daughter with $p_{\text{T}} < 300\,\text{MeV}/c$ and at least one daughter with min[χ^2_{IP}] [2] w.r.t. any PV > 5.

Furthermore, each $\pi^\pm$ candidate needs to satisfy $p_{\text{T}} > 250\,\text{MeV}/c$, min[$\chi^2_{\text{IP}}$] w.r.t. any PV > 4, track quality $\chi^2_{\text{track}}/\text{ndf} < 4$ and PIDK < 8.

[2]Impact parameter chi squared is the difference between the χ^2 of the PV reconstructed with and without the given track; its minimum, for all PV's is denoted by min$[\chi^2_{\text{IP}}]$.

Table 5.1 – List of Stripping cuts.

variable	cut
D^{*-}	
vertex quality χ^2_{track}/ndf	< 25
$m(D^{*-}) - m(\bar{D}^0)$	$\in [135, 160]\,\text{MeV}/c^2$
p_{T}	$> 1250\,\text{MeV}/c$
$\lvert m(D^0\pi^-) - m(D^{*-}(\text{PDG}))\rvert$	$< 50\,\text{MeV}/c^2$
π^- from D^{*-}	
p_{T}	$> 50\,\text{MeV}/c$
vertex quality χ^2_{track}/ndf	< 30
Track ghost probability	< 0.6
D^0	
p_{T}	$> 1.2\,\text{GeV}/c$
$\lvert m(K^-\pi^+) - m(D^0(\text{PDG}))\rvert$	$< 40\,\text{MeV}/c^2$
DIRA(D^0,PV)	> 0.995
χ^2 separation from related PV,	> 36
vertex quality χ^2_{track}/ndf	< 10
DOCA(K^-, π^+)	$< 0.5\,\text{mm}$
$\chi^2_{\text{DOCA}(K^-,\pi^+)}$	< 15
π^- from D^0	
p	$> 2\,\text{GeV}/c$
p_{T}	$> 250\,\text{MeV}/c$
vertex quality χ^2_{track}/ndf	< 3
Track ghost probability	< 0.4
PIDK	< 50
Min χ^2_{IP} w.r.t. any PV	> 10
K^+ from D^0	
p	$> 2\,\text{GeV}/c$
p_{T}	$> 250\,\text{MeV}/c$
vertex quality χ^2_{track}/ndf	< 30
Track ghost probability	< 0.4
PIDK	> -3
Min χ^2_{IP} w.r.t. any PV	> 10
τ^+	
$m(3\pi^\pm)$	$\in [400, 3500]\,\text{MeV}/c^2$
vertex quality χ^2_{track}/ndf	< 25
DIRA(τ^+, PV)	> 0.99
max[DOCA($\pi^\pm{}_i, \pi^\pm{}_j$)]	$< 0.15\,\text{mm}$
min[$m(\pi^+\pi^-)$]	$< 1670\,\text{MeV}/c^2$
At most 1 daughter with p_{T}	$< 300\,\text{MeV}/c$
At least 1 daughter with min χ^2_{IP} w.r.t. any PV	> 5
$\pi^\pm$ from τ^+	
p_{T}	$> 250\,\text{MeV}/c$
min[χ^2_{IP}] w.r.t. any PV	> 4
vertex quality χ^2_{track}/ndf	< 4
PIDK < 8	
B^0	
$m(D^{*-}3\pi)$	$\in [m(B^0) - 2579\,\text{MeV}/c^2,\, m(B^0) + 300\,\text{MeV}/c^2]$
DIRA(B^0, PV)	> 0.995
DOCA($D^{*-}, 3\pi^\pm$)	$< 1\,\text{mm}$

The B^0 candidate is formed by combining the aforementioned D^{*-} and τ^+ candidates, and the following cuts are made on the combination: the difference between $m(D^{*-}\tau^+)$ and the known B^0 mass is required to be in $[-2579, 300]\,\text{MeV}/c^2$, the DIRA($B^0$, PV) is required to be greater than 0.995 and the DOCA($D^{*-}, 3\pi^\pm$) between the D^{*-} and τ^+ candidates is required to be less than 15 mm.

These 'stripping' cuts are applied in the DaVinci framework. In this analysis, to speed up the computing workflow, we have put several highly performant cuts already at this stage. These extra cuts are reported in Table 5.2. They comprise: transverse momentum of the 'slow' pion from D^{*-} to be $p_{\text{T}}(\pi^-{}_{\text{from } D^{*-}}) > 110\,\text{MeV}/c$; good-quality τ^+ vertex of $\text{vtx}(\chi^2/\text{ndof})_{\tau^+} < 10$; the radial distance of τ^+ from the primary ver-

tex of τ^+ radial distance from PV $\in [0.2, 5.0]\,\mathrm{mm}$ to avoid spurious pions from the LHC beampipe; the significance of the impact parameter of the D^0 and the $\pi^\pm$ from the τ^+ with respect to the primary vertex to be respectively $\chi^2[IP_{PV}(\overline{D}^0)] > 10$ and $\chi^2[IP_{PV}(\pi^\pm)]$, $\pi^\pm$ from $\tau^+ > 15$; either the flight distance significance of the τ^+ from the B^0 vertex to be `tau_FDCHI2_ORIVX` to pre-select the signal or the mass of the B^0 to be $m(B^0) > 5\,\mathrm{GeV}/c^2$ to pre-select the normalisation mode, which is fully reconstructed.

Table 5.2 – Initial cut-based selection for all the modes applied in DaVinci.

variable	cut	background targeted
$p_\mathrm{T}(\pi^-_{\text{from } D^{*-}})$	$> 110\,\mathrm{MeV}/c$	combinatorial
$\mathrm{vtx}(\chi^2/\mathrm{ndof})_{\tau^+}$	< 10	combinatorial
τ^+ radial distance from PV	$\in [0.2, 5.0]\,\mathrm{mm}$	spurious
$\chi^2[IP_{PV}(\overline{D}^0)]$	> 10	prompt
$\chi^2[IP_{PV}(\pi^\pm)]$, $\pi^\pm$ from τ^+	> 15	combinatorial
`tau_FDCHI2_ORIVX`	> 4	prompt
or $m(B^0)$	$> 5\,\mathrm{GeV}/c^2$	partially reconstructed

The 'stripping' cuts merged with these few other requirements form our custom preliminary selection whose efficiency is reported in Table 5.9. The following sections tackle the 'offline' selection defined as occurring after the ROOT samples have been produced.

5.4 Initial offline selection cuts

The initial selection applied 'offline' is common to all the modes and is reported in Table 5.3. It consists of: removal of the combinatorial backgrounds by requiring the $\overline{D}^0$ share the primary vertex with the τ^+; accepting events with only one $B^0 \to D^{*-}\tau^+\nu_\tau$ candidate, `totCandidates` $= 1$ to avoid candidate multiplicity; for the same reason number of hits in the SPD detector, reflective of the event occupancy, is limited to below 450. Finally, either signal or normalisation events are 'promoted' by requiring either $[\mathrm{vtx}_z(\tau^+) - \mathrm{vtx}_z(B^0)]/\mathrm{error} > 2$ to remove the prompt background in the signal case or, in case of normalisation mode, the $\overline{D}^0$ vertex is required to be 4 significance units downstream of the τ^+ one, to select the $3\pi^\pm$ system coming directly from the B^0, and the B^0 mass must be $m(B^0) > 5150\,\mathrm{MeV}/c^2$. This last logical 'or' is motivated merely by computing reasons and is reported for consistency with the efficiencies measured. The efficiencies of the initial selection are measured line-by-line and reported in Table 5.10.

Figure 5.2 shows the importance of the $[\mathrm{vtx}_z(\tau^+) - \mathrm{vtx}_z(B^0)]/\mathrm{error} > 2$ cut after all the initial cuts except the last three lines of Table 5.3. It is evident how this cut removes the 'prompt' background (in grey).

At this stage the prompt background, whereby the $3\pi^\pm$ system comes directly from the B^0 is the most dominant, as can be seen in the simulated inclusive $b\bar{b} \to D^{*-}3\pi^\pm X$ sample, whose composition will be discussed later and is shown in Figure 5.27. The second largest background is the $B \to D^{*-}D_s^+$. Figure 5.3 shows the invariant mass

Table 5.3 – Initial cut-based selection for all the modes applied after preselection (*cf.* Section 5.3). The last three cuts comprise merely a computing trick to keep events that pass a logical 'or' of signal or normalisation conditions. These cuts are reported here for consistency with the efficiencies computation in Table 5.10.

variable	cut	background targeted
$PV(\overline{D}^0)$	$= PV(\tau^+)$	combinatorial
`totCandidates`	$= 1$	track multiplicity
$[\mathrm{vtx}_z(\tau^+) - \mathrm{vtx}_z(PV)]/\mathrm{error}$	> 10	prompt
`nSPDHits`	< 450	track multiplicity
$[\mathrm{vtx}_z(\tau^+) - \mathrm{vtx}_z(B^0)]/\mathrm{error}$	> 2	remove prompt (for signal)
OR		
$[\mathrm{vtx}_z(\overline{D}^0) - \mathrm{vtx}_z(\tau^+)]/\mathrm{error}$	> 4	require prompt (for normalisation)
AND $m(B^0)$	$> 5150\,\mathrm{MeV}/c^2$	partially reconstructed (for normalisation)

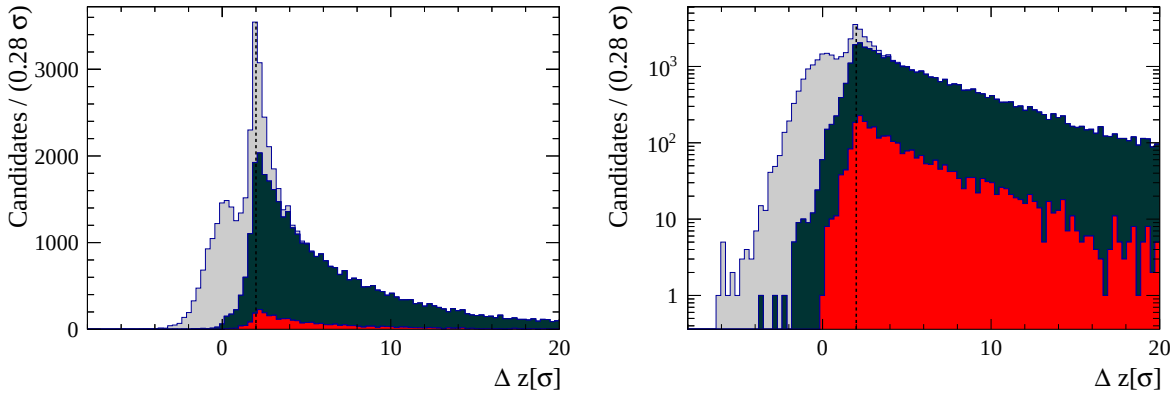


Figure 5.2 – Δz/uncertainty distribution in linear (left) and logarithmic (right) scale of the simulated signal (red), double charm background (black) and prompt background (grey), after the initial cuts except the last three lines of the Table 5.3. A cut at 2σ is shown.

distribution of the $3\pi^{\pm}$ system in the collision data, after the initial selection. The D_s^+ peak is clearly visible.

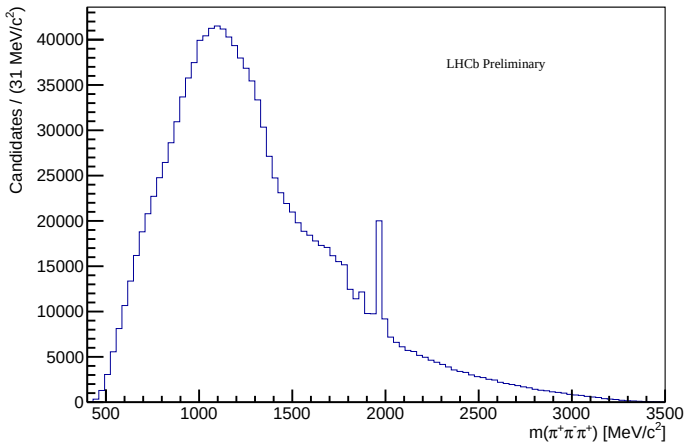


Figure 5.3 – $m(\pi^+\pi^-\pi^+)$ after the initial selection in the collision 2015-2016 data. D_s^+ and D^+ peaks are visible.

5.5 Anti combinatorial background BDT

The objective of the first MVA algorithm is to remove the combinatorial backgrounds events, where the τ^+ and D^{*-} systems do not originate at the same B^0 vertex. The following variables are used for the anti-combinatorial-background BDT with the TMVA package integrated in ROOT [94]:

- $p_{\mathrm{T}}(D^{*-})$ and $p_{\mathrm{T}}(\tau^+)$
- $\eta(D^{*-})$ and $\eta(\tau^+)$
- $\log(\chi^2[\mathrm{IP}_{\mathrm{PV}}(B^0)])$ and $\log(\chi^2[\mathrm{IP}_{\mathrm{PV}}(D^{*-})])$, logarithm of the quality (χ^2) of the impact parameter of a given particle w.r.t. the best PV of that particle
- $\log\left[\mathrm{vtx}(\chi^2/\mathrm{ndof})_{B^0}\right]$, logarithm of the B^0 decay vertex quality
- $\log(\mathrm{acos}(\mathrm{DIRA}(B^0, \mathrm{PV})))$, logarithm of the angle between the B^0 momentum and the line of flight from the PV to the B^0 decay vertex
- $\log(\mathrm{acos}(|\mathrm{DIRA}(D^*, \mathrm{ORIVX}(D^*))))$, logarithm of the acos of the magnitude of the angle between the D^{*-} momentum and the line of flight from the D^{*-} origin to decay vertex
- $\log(\mathrm{acos}(\mathrm{DIRA}(\tau, \mathrm{ORIVX}(\tau))))$, logarithm of the acos of the magnitude of the angle between the τ^+ momentum and the line of flight from the τ^+ origin to decay vertex
- BPVVDR(τ), radial distance between the τ decay vertex and the PV

The signal training samples are $B^0 \to D^{*-}\tau^+\nu_\tau$ MC and for background the wrong-sign data is used, where $D^{*\pm}$ and $\tau^\pm$ have the same charge and come from different

$b-$hadrons (see Section 5.3). The training samples are selected with the initial selection (see Section 5.4) and a loose detachment cut $[\text{vtx}_z(\tau^+) - \text{vtx}_z(B^0)]/\text{error} > 2$.

The Figures 5.4 and 5.5 show the discrimination power of the input variables for the signal and background as defined above.

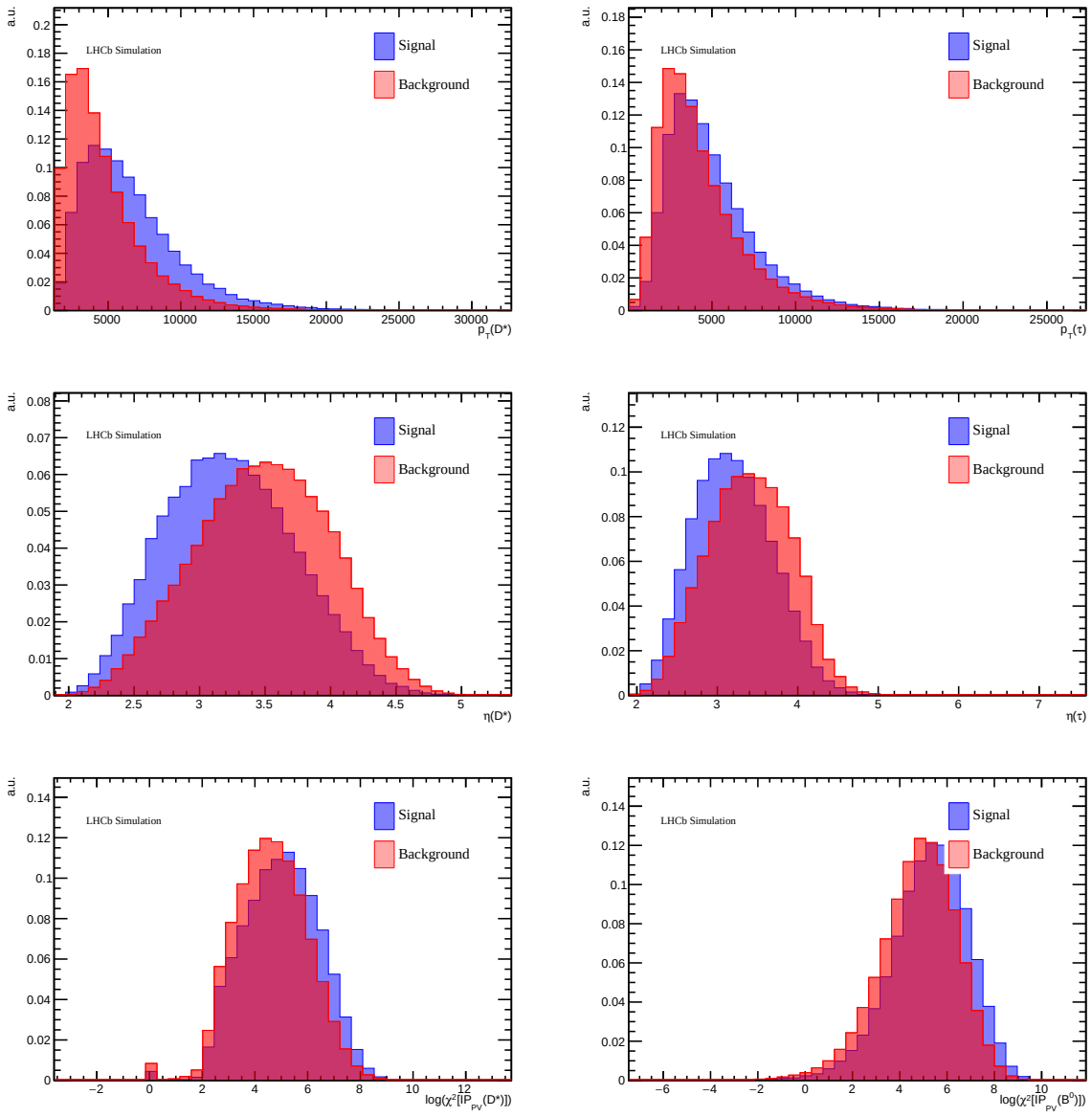


Figure 5.4 – Distributions of input variables for the combinatorial BDT for the signal (blue) and background (red).

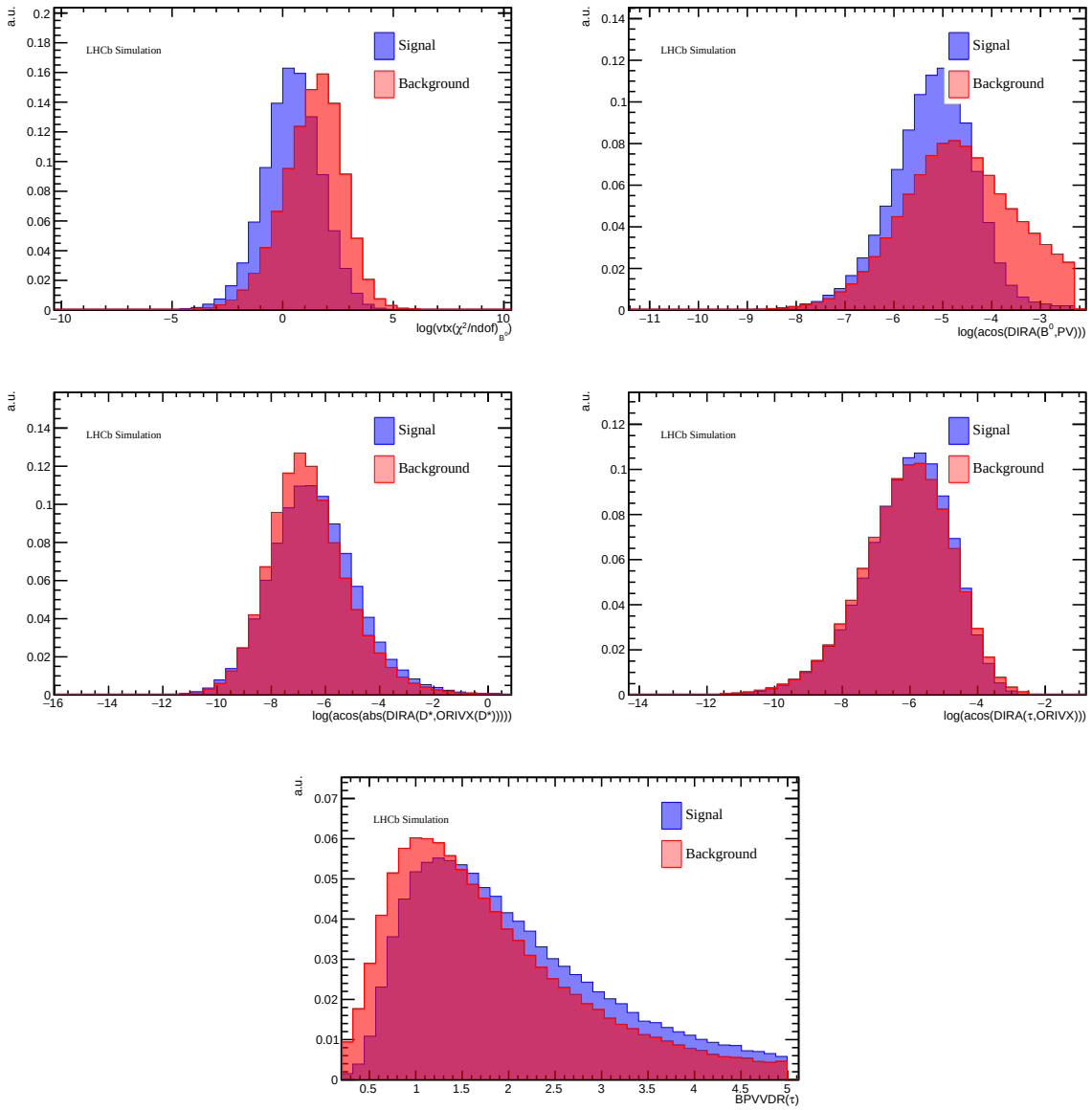
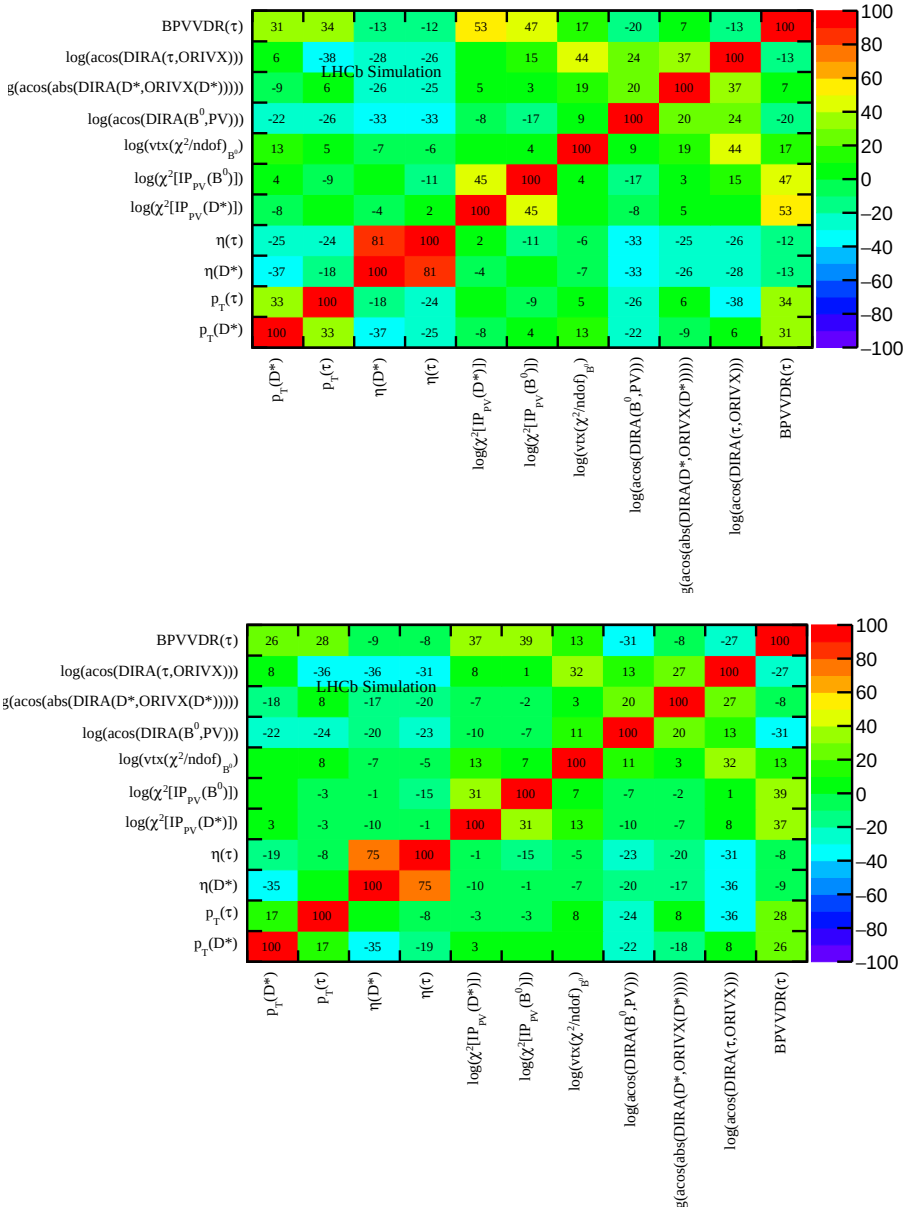


Figure 5.5 – Distributions of input variables for the combinatorial BDT for the signal (blue) and background (red) training samples.

Figure 5.6 depicts correlations between the input variables of the combinatorial BDT. The variables exhibit relative correlations below 50-60%, except $\eta(\tau)$ vs. $\eta(D^*)$ (correlation 81% (75%) for signal (background)).

Figure 5.6 – Correlation matrices for the input variables to the combinatorial BDT for signal (top) and background (bottom) training samples.



The BDT performance is inspected on testing samples that are statistically independent of the ones used in the training. Figure 5.7 (right) shows the background rejection as a function of signal efficiency for three BDT algorithms considered: Ada Boost (BDT), Ada Boost with decorrelated variables (BDTD) and Gradient Boost (BDTG). The most performant (with curve closest to the top-right corner), BDTD is chosen as the default one. Figure 5.7 (left) shows the combinatorial BDTD distribution for the training signal and background samples. A cut is applied in the middle of the anti-combinatorial BDT range, *i.e.* at 0, as a preliminary working point. This preserves approximately 85% of the signal and rejects around 70% of the combinatorial background.

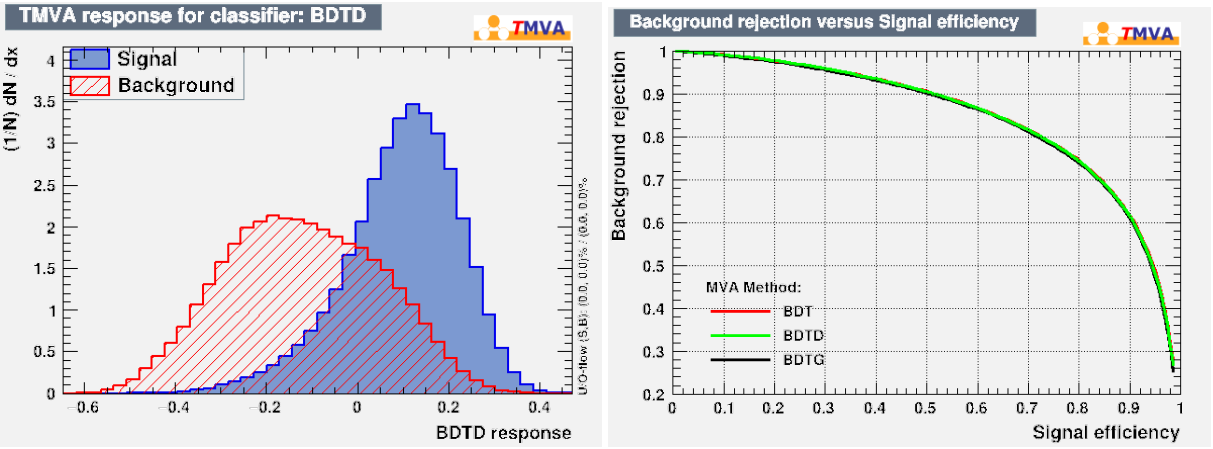


Figure 5.7 – Left: Distribution of the combinatorial BDT for the testing samples, for signal and background. Right: Signal efficiency as a function of background rejection for the combinatorial BDT for the testing samples.

Figure 5.8 compares the MC/data distribution of the combinatorial BDT for the fully reconstructed $B^0 \to D^{*-}3\pi^{\pm}$ mode. The data sample used has been s-weighted [95] with respect to the B^0 mass to select approximately pure $B^0 \to D^{*-}3\pi^{\pm}$ sample.

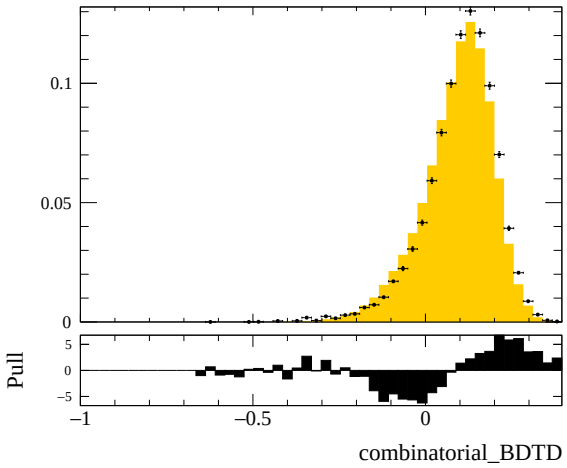


Figure 5.8 – Data (black points) / MC (coloured bins) comparison for the combinatorial BDTD distribution as for the $B^0 \to D^{*-}3\pi^{\pm}$ mode. See Appendix D for the distributions of the input variables.

Despite a small shift towards higher BDTD values for data w.r.t. MC, the agreement is rather satisfactory. This variables is used only once, in the selection of all the samples. Therefore, a correction should be applied by cutting at a slightly smaller value of the combinatorial BDTD. Nonetheless, as long as the optimisation remains coarse, there is no need for such a correction.

5.6 Charged isolation

In order to reject background events that have additional charged tracks besides the signal candidates, a charged isolation technique is used. It detects charged tracks forming good vertices with the signal tracks as shown in Figure 5.9, where the decay $B^0 \to D^{*-}D^0(\to K^-\pi^+\pi^-\pi^+)K^+$ has two extra charged tracks (K^+ and K^- marked magenta) on top of those compatible with the $B^0 \to D^{*-}\tau^+\nu_\tau$ tracks (green).

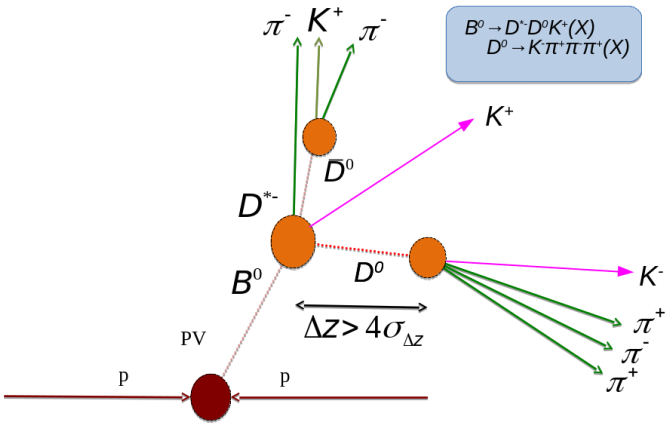


Figure 5.9 – The $B^0 \to D^{*-}D^0(\to K^-\pi^+\pi^-\pi^+)K^+$ decay, where the green tracks are compatible with the $B^0 \to D^{*-}\tau^+\nu_\tau$. The charged isolation detects the non-isolating (magenta) tracks.

The isolation algorithms use track-based and vertex-based isolation variables. The former, for a given signal-candidate track, reflect how close it is with the non-signal (underlying) tracks in the event. Therefore each out of 6 charged tracks in $B^0 \to D^{*-}\tau^+\nu_\tau$ has such a variable. The vertex based ones, however, scan all the tracks usually constrained to a given cone around the flight direction of the particle whose vertex is considered.

Regardless of the decay mode considered (*i.e.* signal, normalisation or backgrounds), the isolation variables are, in this analysis, always computed relative to the signal candidate.

The variables considered are based on the LHCb packages `TupleToolIsoGeneric` and `TupleToolVtxIsolnPlus` and are enlisted here:

- 6 track isolation variables from a customised `TupleToolIsoGeneric`: IsoSumBDT(<particle>), where <particle> is any of the 6 final states charged tracks, *i.e.*, $\pi^\pm$ from the D^{*-}, D^0 and τ^+ vertices, and a K^+ from the $\overline{D}^0$ vertex.
- 4 vertex isolation variables from (or derived from) `TupleToolVtxIsolnPlus`:

- `tau_isoPlus_nGood` – number of charged tracks forming 'good' vertex with the τ^+;
- `B0_isoPlus_nGood` – number of charged tracks forming 'good' vertex with the B^0;
- `nTauIso0` – number of charged tracks compatible with the τ^+ vertex after removing the tracks from the D^{*-}
- `nB0Iso25sig` – number of charged tracks compatible with the B^0 vertex that worsen the B^0 vertex χ^2 by not more than 25, after removing the tracks from the D^{*-}

The pairs of variables `tau_isoPlus_nGood` and `nTauIso0` as well as `B0_isoPlus_nGood` and `nB0Iso25sig` are very similar. They indeed exhibit 55% and 94% correlations, respectively, for the signal events, as shown in Figure 5.12 (top) and 69% and 92%, respectively, for the isolation backgrounds Figure 5.12 (bottom). Nonetheless, to exploit maximum information, all four variables are considered.

Instead of relying on rectangular cuts on a these *track/vertex-based* variables, a custom BDT is used, incorporating them into a single *event-based* one.

The training is performed on MC samples extracted from the $b\bar{b} \to D^{*-}3\pi^{\pm}X$ after the initial selections as described in Section 5.4 with additional cuts as reported in Table 5.4. There, the common selection for the two modes comprises loose vertex detachment requirement and the B^0 mass cut, to focus on more signal-like phase-space region (*i.e.* the signal mode will eventually be required to meet the vertex detachment and lower B^0 reconstructed mass criteria). The distinction between 'signal' and 'background' is made using the MC-truth information by requiring no extra charged tracks ($\pi^{\pm}$ or $K^{\pm}$) in the signal decay tree and at least one extra track (a $\pi^{\pm}$ or $K^{\pm}$) for the background. The isolation background is required to have either more than one $K^{\pm}$ or more than five $\pi^{\pm}$ excluding those from K^0_S decays (that 69% of the time spawn $\pi^+\pi^-$ pair) that decay outside VELO thus being 'well-isolated' from the signal candidate.

Table 5.4 – Selection of the training samples for the charged isolation BDT applied on top of the initial cuts discussed in Section 5.4.

signal	background
1 $K^{\pm}$ and 5 $\pi^{\pm}$ in decay	> 1 $K^{\pm}$ or > 5 $\pi^{\pm}$not from K^0_S
$[\mathrm{vtx}_z(\tau^+) - \mathrm{vtx}_z(B^0)]/\mathrm{error} > 2$	
$m(B^0) < 5225\,\mathrm{MeV}/c^2$	

The Figures 5.10 and 5.11 show the discrimination power of these input variables for the signal and background as defined above.

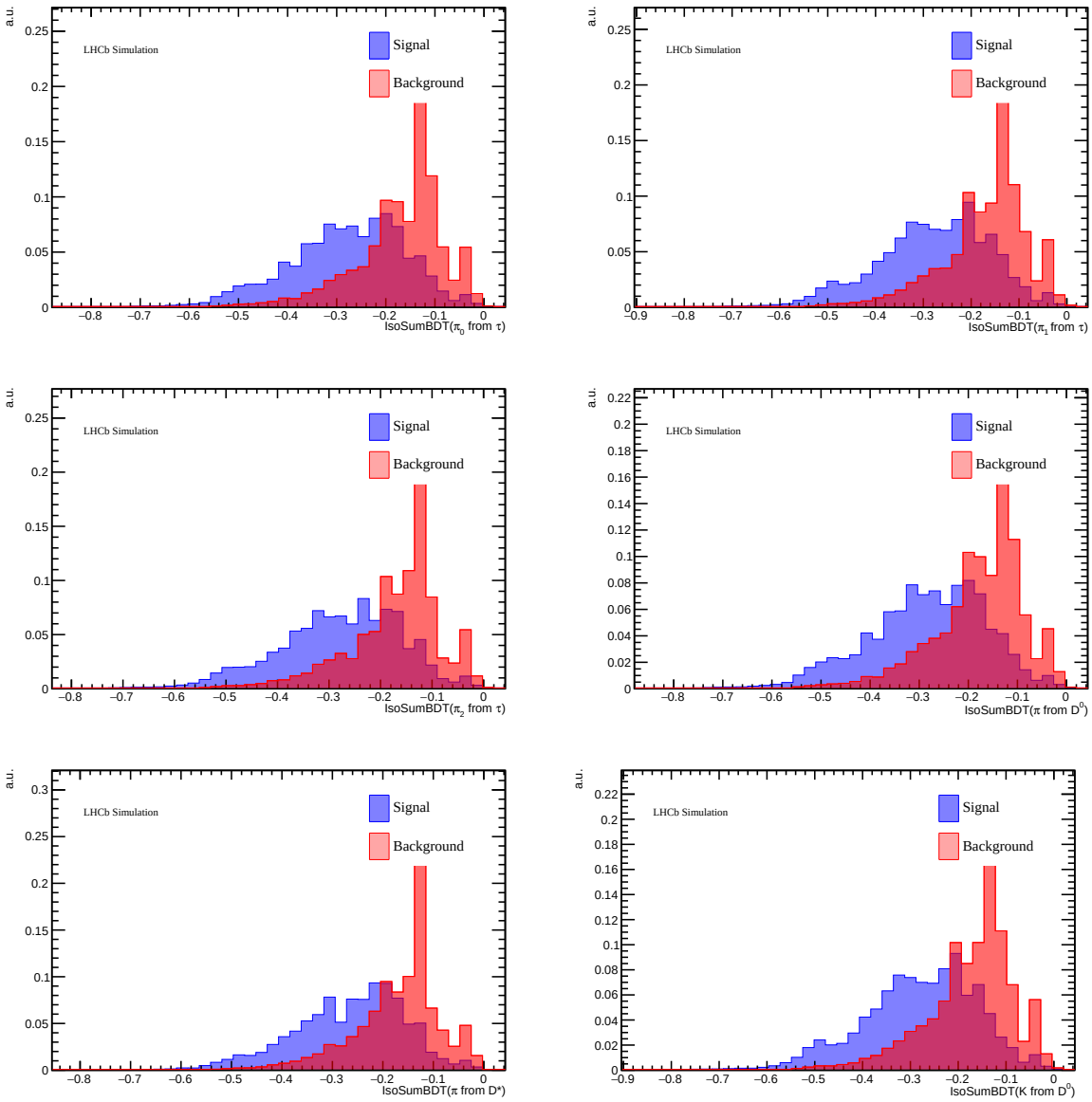


Figure 5.10 – Distributions of input variables for the charged isolation BDT for the isolation signal (blue) and background (red).

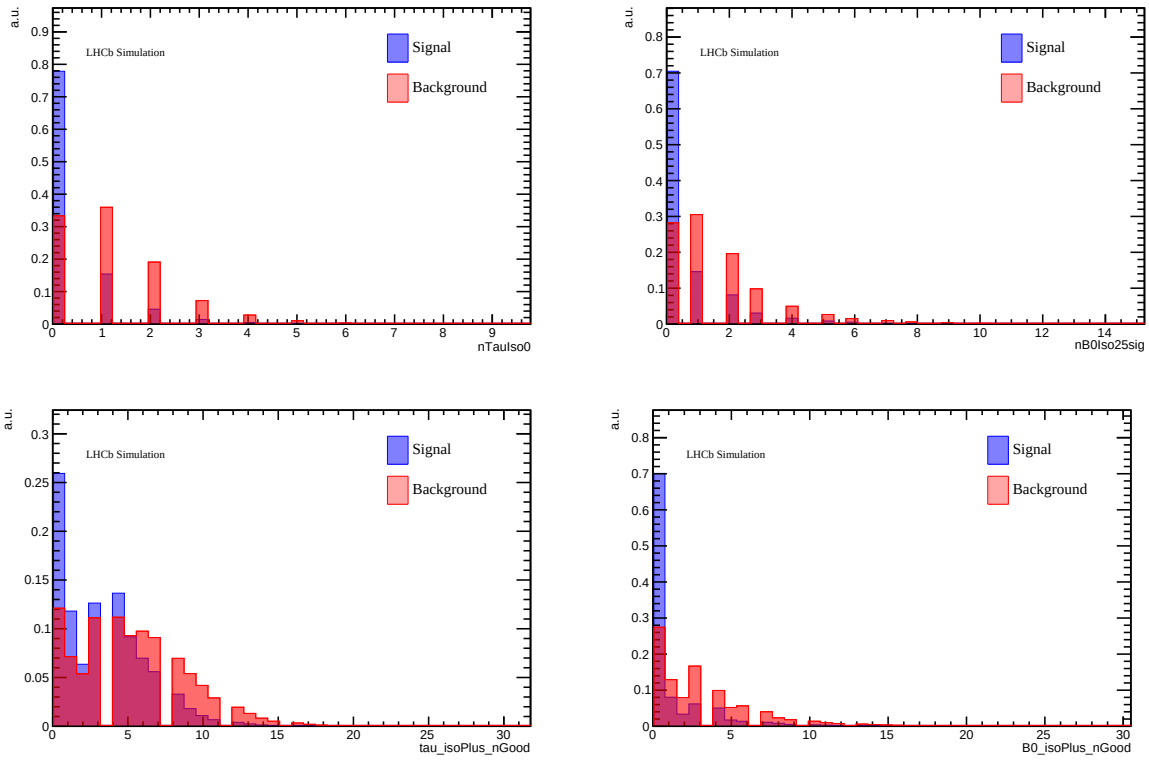


Figure 5.11 – Distributions of input variables for the charged isolation BDT for the isolation signal (blue) and background (red).

Figure 5.12 – Correlation matrices for the input variables to the `metaIsoBDT` for signal (top) and background (bottom) training samples.

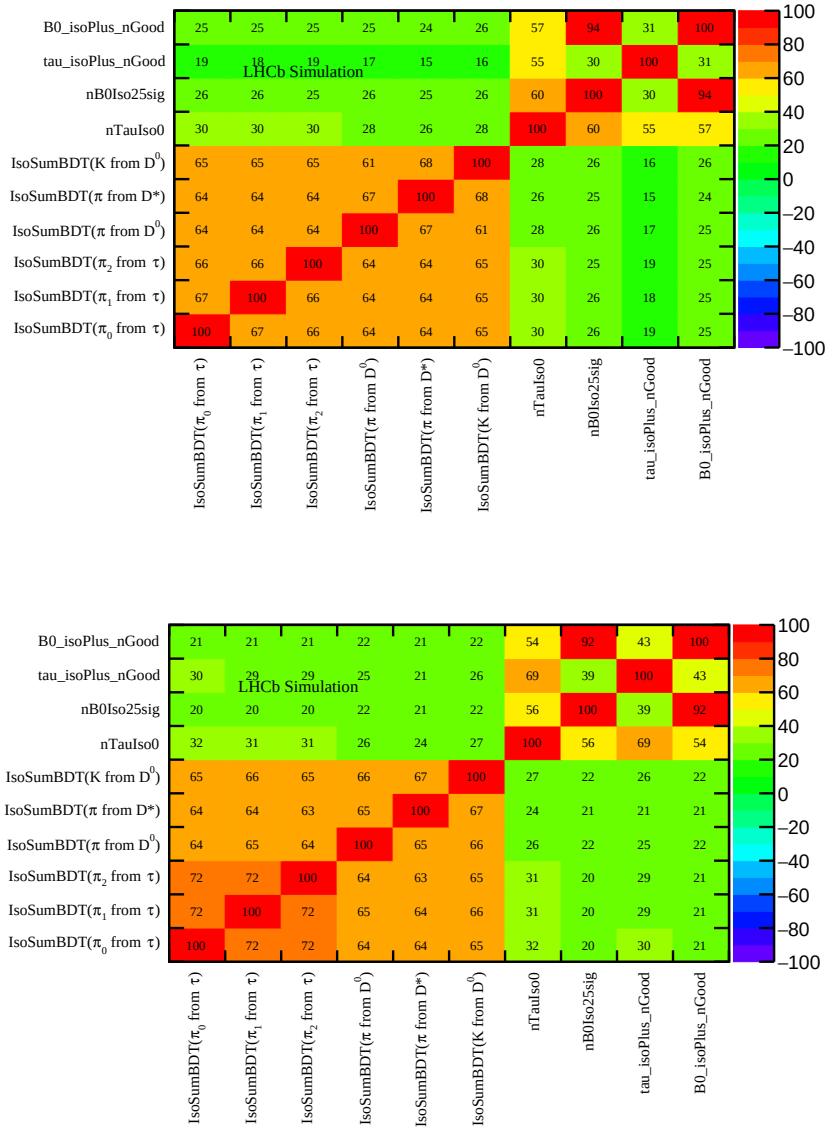


Figure 5.13 shows the `metaIsoBDT` distribution for the signal and background and its background rejection as a function of signal efficiency. The AdaBoost algorithm (denoted by BDT) is the most performant and has, therefore, been chosen as the default one.

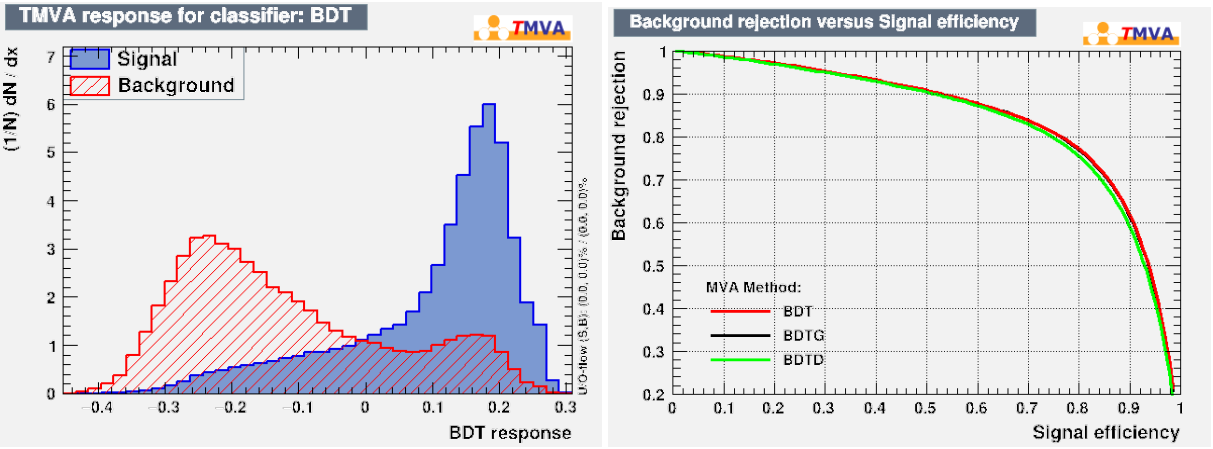


Figure 5.13 – Left: Distribution of the `metaIsoBDT` for the testing samples, for signal and background. Right: Signal efficiency as a function of background rejection for `metaIsoBDT` for the testing samples.

Figure 5.14 compares the MC/data distribution of the isolation BDT for the $B^0 \to D^{*-}3\pi^{\pm}$ mode. The data sample used has been s-weighted with respect to the B^0 mass to select approximately pure $B^0 \to D^{*-}3\pi^{\pm}$ sample.

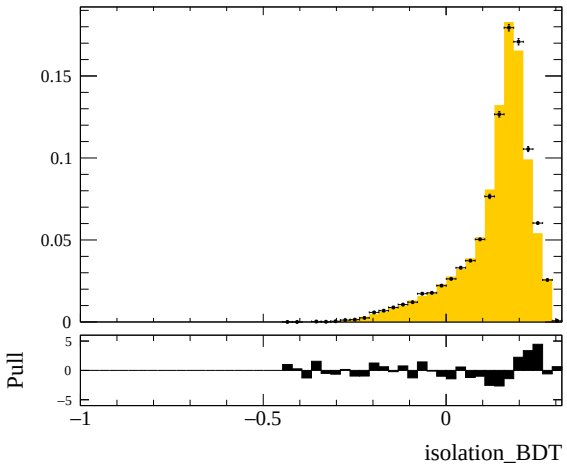


Figure 5.14 – Data (black points) / MC (coloured bins) comparison for the isolation BDT distribution as for the $B^0 \to D^{*-}3\pi^{\pm}$ mode. See Appendix D for the distributions of the input variables.

The agreement is satisfactory, especially around the cut region at isolation BDT $= 0.0$.

5.7 Detachment BDT

To remove the background where the $3\pi^{\pm}$ system comes directly from the B^0, a vertex detachment BDT is used. It separates the $3\pi^{\pm}$ vertex from the B^0 one to match the

signal topology.

The input variables are as follows:

- $\log([\text{vtx}_z(\tau) - \text{vtx}_z(B^0)]/error)$, $z-$coordinate of the distance between the τ^+ and B^0 decay vertices
- $\log([r(\tau) - r(B^0)]/error)$, radial coordinate of the distance between the τ^+ and B^0 decay vertices
- $[\text{vtx}_z(D^0) - \text{vtx}_z(\tau)]/error)$, $z-$coordinate of the distance between the D^0 and τ decay vertices
- $\log(\text{acos}(\text{DIRA}(\tau, \text{ORIVX}(\tau))))$, logarithm of the angle between the τ^+ momentum and the line of flight from the τ^+ origin to decay vertex
- $\log(\chi^2[\tau(\tau)])$, τ^+ decay vertex χ^2
- $\log(\chi^2[ORIVX(\tau)])$, χ^2 of the τ^+ production vertex

The training samples are MC $B^0 \to D^{*-}\tau^+\nu_\tau$ for the detachment BDT signal and the inclusive $b\bar{b} \to D^{*-}3\pi^\pm X$ sample for the background, selected as specified in Table 5.5. Both samples need to meet the loose vertex detachment requirement, $[\text{vtx}_z(\tau^+) - \text{vtx}_z(B^0)]/\text{error} > 2$ (that eventually will be imposed for the final selection) and have the τ^+ candidates further from the beampipe than the B^0 decay vertex, $r(\tau^+) - r(B^0) > 0$. In addition, the signal is truth-matched with the $B^0 \to D^{*-}\tau^+\nu_\tau$ decay, whereas the background must not originate from the τ^+ nor the charm mesons: D^0, D^+, D_s^+. This way, the prompt $3\pi^\pm$, coming directly from the B^0 candidate are selected.

Table 5.5 – Selection of the training samples for the detachment BDT applied on top of the initial cuts discussed in Section 5.4.

signal	background
$[\text{vtx}_z(\tau^+) - \text{vtx}_z(B^0)]/\text{error} > 2$	
$r(\tau^+) - r(B^0) > 0$	
Truth-matched signal	$3\pi^\pm$ not from τ^+ nor D^0,D^+,D_s^+

Figure 5.15 shows the discrimination power of these input variables for the signal and background as defined above.

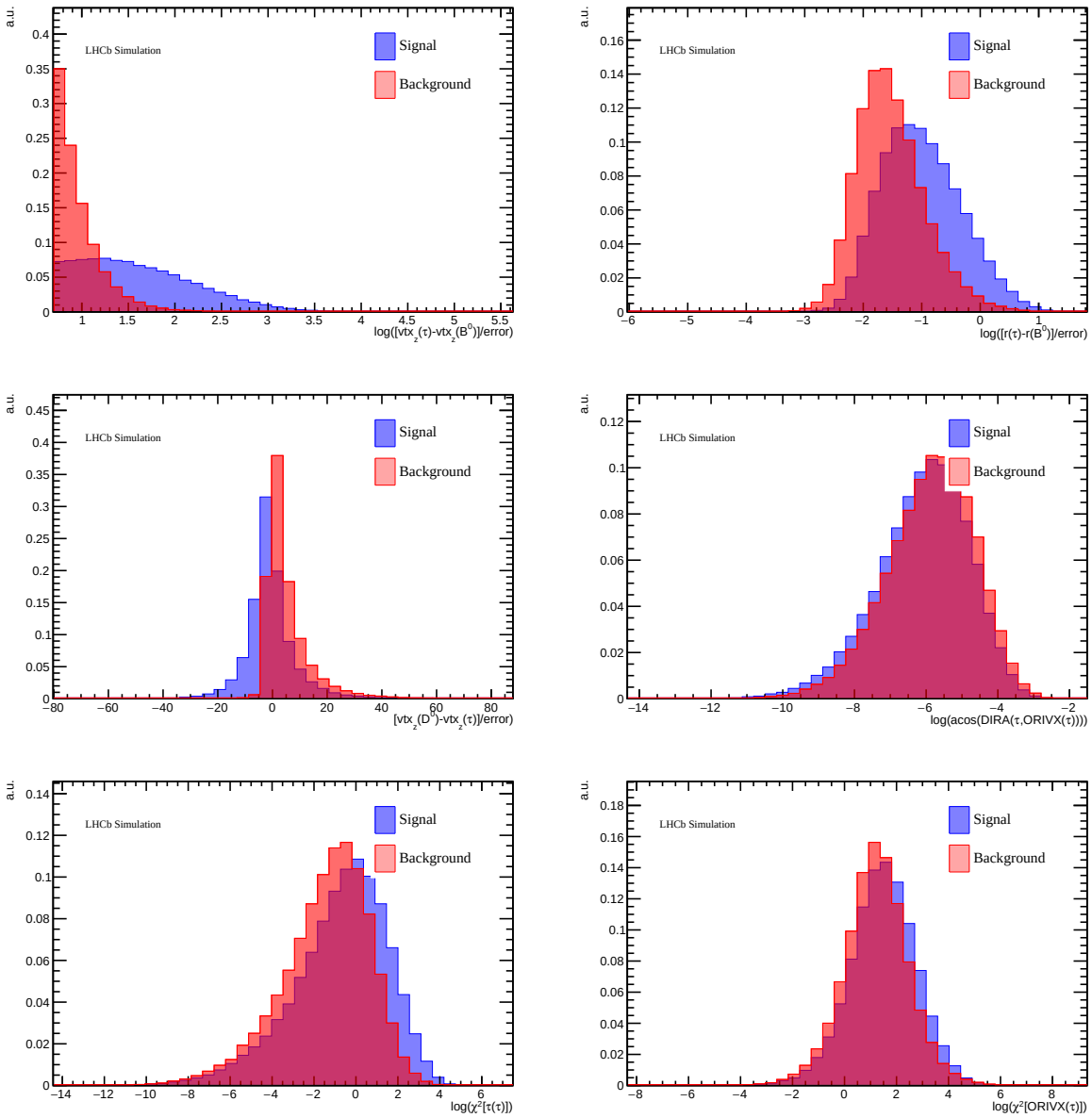


Figure 5.15 – Distributions of input variables for the detachment BDT for the signal (blue) and background (red).

Figure 5.16 – Correlation matrices for the input variables to the detachment BDT for signal (top) and background (bottom) training samples.

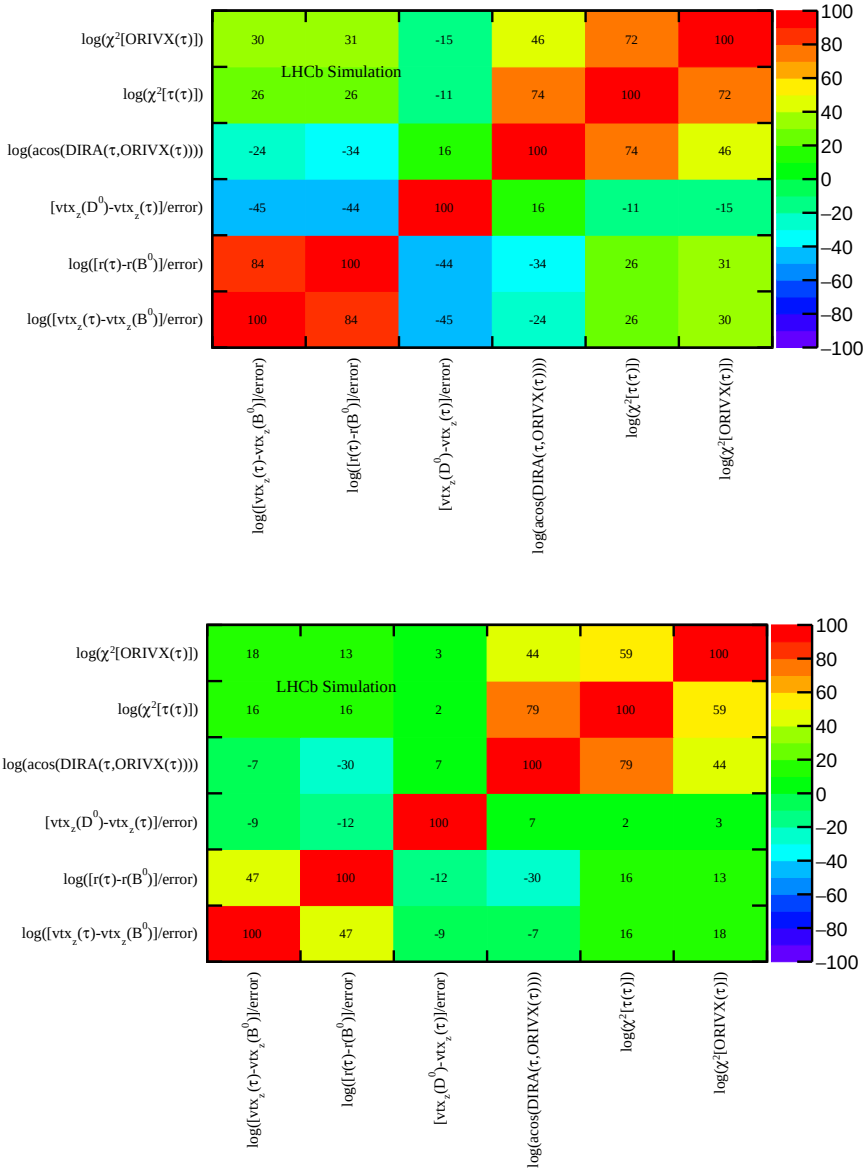


Figure 5.17 shows the detachment BDT distribution for the signal and background and its background rejection as a function of signal efficiency. The Gradient Boost BDT (BDTG) is the most performant and has, therefore, been chosen as the default one.

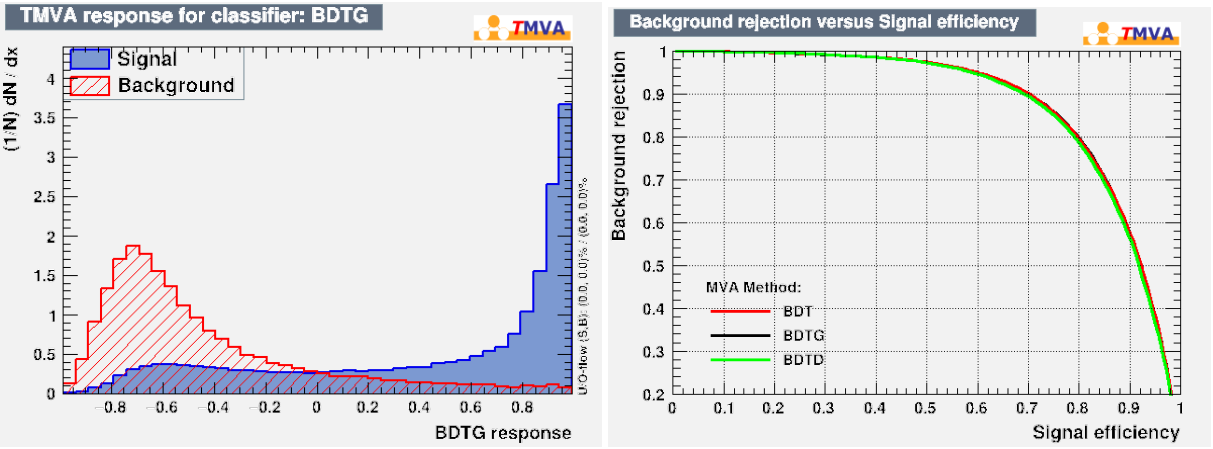


Figure 5.17 – Left: Distribution of the detachment BDT for the testing samples, for signal and background. Right: Signal efficiency as a function of background rejection for detachment BDT for the testing samples.

A cut is applied at `detachment_BDTG` > 0.2, which keeps 70% of the signal mode, relative to the yield after the isolation cut, while rejecting around 90% of the prompt background. The cut value is a preliminary working point and is subject to further optimisation. Nonetheless, the rather flat distributions of both signal and background in the region $[-0.2, 0.2]$ of the detachment BDT, suggests the optimal working point should not be very sensitive to exact cut value.

Figure 5.18 compares the MC/data distribution of the detachment BDT for the $B^0 \to D^{*-}3\pi^{\pm}$ mode. The data sample used has been s-weighted with respect to the B^0 mass to select approximately pure $B^0 \to D^{*-}3\pi^{\pm}$ sample.

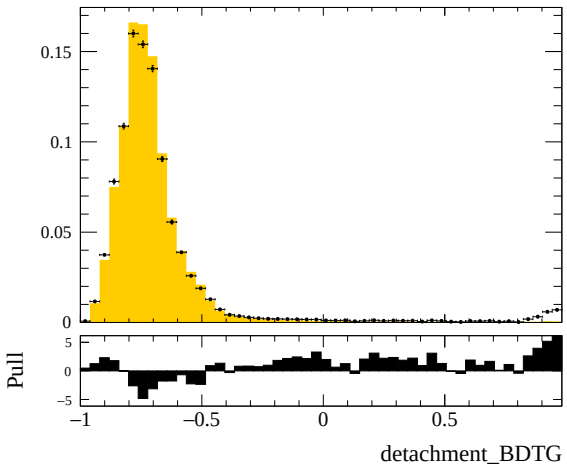


Figure 5.18 – Data (black points) / MC (coloured bins) comparison for the detachment BDT distribution as for the $B^0 \to D^{*-}3\pi^{\pm}$ mode. See Appendix D for the distributions of the input variables.

The small peak at high detachment BDTG score for data is due to $B^0 \to D^{*-}D_s^+$ ($\to 3\pi^{\pm}$)

events, as depicted in Figure 5.19, where the $m(\pi^+\pi^-\pi^+)$ is shown for the detachment BDTG > 0.8 in the relevant data sample.

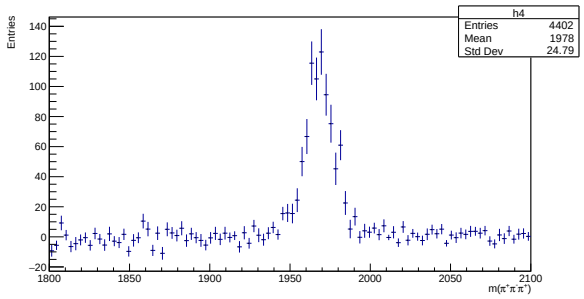


Figure 5.19 – Distribution of $m(\pi^+\pi^-\pi^+)$ for the detachment BDTG > 0.8 in the data sample used in the data/MC comparison. The D_s^+ mass peak is visible.

Despite a small shift between data and MC distributions, the agreement is satisfactory. It is, nonetheless, possible to apply a correction to MC while optimising the cut. As long as the optimisation remains coarse, there is no need for such a correction.

5.8 Anti D_s^+ BDT

The most important background category after rejecting the prompt background are events with $D^{*-}D_s^+$ X in the final state. Figure 5.27 shows that this background is the most abundant after the full selection. To control the $D^{*-}D_s^+$ X contributions a dedicated BDT is made with the following 17 input variables:

- Partial reconstruction variables
  - Reconstruction as a signal event (*cf.* 3.4.1)
    - $\diamond$ $|\vec{p}(B^0)|-|\vec{p}(\tau)|-|\vec{p}(D^*)|$, where $|\vec{p}(B^0)|$ and $|\vec{p}(\tau)|$ come from the partial reconstruction under the signal hypothesis (*cf.* Section 3.4.1)
  - Reconstruction as a double charm background event (*cf.* 3.4.2)
    - $\diamond$ $|\vec{P}_{B,sn}|$: the B^0 momentum reconstructed using the scalar product method, using the corrected B^0 decay vertex accounting for the neutral tracks from the D_s^+ candidate
    - $\diamond$ $\log\left(\left||\vec{P}_{B,v}|/\vec{P}(B^0)|\right|\right)$: the reconstructed B^0 momentum with the vector method, normalised to the one reconstructed based on the charged tracks only
    - $\diamond$ $\log\left(\left||\vec{P}_{B,vn}|/|\vec{P}(B^0)|\right|\right)$: the reconstructed B^0 momentum with the vector method, using the corrected B^0 decay vertex accounting for the

neutral tracks from the D_s^+ candidate, normalised to the reconstructed based on the charged tracks only

- $\diamond$ $\log\left(\left||\vec{P}_{B,sn}|-|\vec{P}_{B,vn}|\right)/|\vec{P}_{B,vn}|\right|\Big)$: the normalized difference of B^0 momenta between the scalar and vector method, using the corrected B^0 decay vertex for neutral tracks from the D_s^+ candidate, normalised to the one obtained with the vector method
- $\diamond$ $m(N)_v^2 = m(B^0)^2 + m(B^0)^2_{\text{PDG}} + 2(\vec{p}_{B,v}\cdot\vec{p}_{D^*D_s} - E(B^0)_v E(B^0))$: the squared mass of the reconstructed neutral vector, N, in $D_s^+ \to 3\pi N$
- $\diamond$ $\log(\sqrt{|(m(D_s)^2_{vn}|})$, where $m(D_s)^2_{vn}$ is the reconstructed mass squared of the $D_{(s)}^{(*,**)}$ system

- neutral and charged isolation variables

  Background events where the 3π system is coming from D_s^+ decays are often accompanied by a large neutral energy coming from the rest of the D_s^+ decay. This neutral energy is searched in cones around the 3π system.

  The background events with $D_s^+ \to 3\pi^\pm\ X$ often carry significant energy in the neutral particles that is deposited in the LHCb electromagnetic calorimeter. The neutral energy around the $3\pi^\pm$ system is measured with the `EWTupleIsolation` tool by the three variables:

  - `tau_0_40_nc_mult`: the multiplicity of neutral objects in a cone of 0.4 opening (defined in the $\Delta\phi$, $\Delta\eta$ reference frame) centred around the 3π vector;
  - `tau_0_40_nc_PZ`: the sum of the neutral energy contained in the cone of 0.4 opening around the 3π vector;
  - `tau_0_30_nc_PZ`: the sum of the neutral energy in a cone of 0.3 opening around the 3π vector.

  Despite that the charged isolation BDT is already applied at an earlier stage, two charged track isolation variables are included in the BDT:

  - `tau_0_20_cc_mult`: the multiplicity of charged tracks in the cone of 0.2 opening angle around the $3\pi^\pm$ system
  - `tau_0_20_cc_PZ`: the energy due to charged tracks in the cone of 0.2 opening angle around the $3\pi^\pm$ system

- 2 variables related to the $\pi^+\pi^-$ dynamics: min$[m(\pi^+\pi^-)]$ and max$[m(\pi^+\pi^-)]$
  The signal decays through $\tau^+ \to 3\pi^\pm\nu_\tau$ that occurs purely via the a_1 mode that decays to $\rho\pi$. Consequently the $\pi^+\pi^-$ reconstructed mass should peak around the ρ mass. On the other hand, the D_s^+ decays mostly through η and η': $\eta \to \pi^+\pi^-\pi^0$ and $\eta' \to \eta\pi^+\pi^-$. This puts the upper boundary on the $\pi^+\pi^-$ mass of around 400 MeV. These phase-space differences justify using the minimum and maximum masses of the $\pi^+\pi^-$ system in the BDT and can be seen in Figure 5.21.

- kinematic variables

  - BPVVDR(B^0): the radial distance between the primary vertex and the B^0 decay vertex
  - $\log(E(3\pi))$: where $E(3\pi)$ is the energy of the 3π system
  - $m(B^0)$: the mass of the B^0 candidate computed on the six-track system. This variable distinguishes fully reconstructed $D_s^+ \to 3\pi^\pm$ from the partially reconstructed signal in $\tau^+ \to 3\pi^\pm\nu_\tau$. Also, it discriminates against the $3\pi^\pm$ system where the pions do not share the same vertex. Its importance is shown in Figure 5.20.

The BDT is trained using as signal the MC signal sample where the τ^+ decays into 3 pions, requiring the events to be true $B^0 \to D^{*-}\tau^+\nu_\tau$ decays (MC truth). For the background, MC-truth-matched $B^0 \to D^{*-}D_s^+X$ events are used. Table 5.6 reports all cuts used for the training sample applied in addition to the initial cut-based selection (*cf.* Section 5.4).

Table 5.6 – Anti-D_s^+ BDT cuts for signal and background applied on top of the initial cut-based selection (*cf.* Section 5.4).

variable	cut
ranges	
$\lvert\vec{P}_{B,s}\rvert$	> 0
$m(N)_v^2$	reject if unreconstructed
$\lvert\vec{p}(B^0)\rvert - \lvert\vec{p}(\tau)\rvert - \lvert\vec{p}(D^*)\rvert$	reject if unreconstructed
detachment	
$[\mathrm{vtx}_z(\tau^+) - \mathrm{vtx}_z(B^0)]/\mathrm{error}$	> 2
signal only	
Signal MC truth	
background only	
$B^0 \to D^*\ D\ X$ MC truth	
(τ) candidate true ID	$= D_s^+$

Figures 5.20, 5.21, 5.22 show the distribution of the input variables for the training samples. Correlations between the variables are reported in Figure 5.23, where $\log(\mathrm{abs}(|\vec{\mathrm{P}}_{\mathrm{B,v}}|/\vec{\mathrm{P}}(\mathrm{B}^0)|))$ exhibits 96% and 93% correlation with $\log(\mathrm{abs}(|\vec{\mathrm{P}}_{\mathrm{B,vn}}|/|\vec{\mathrm{P}}(\mathrm{B}^0)|))$ for signal and background respectively. Despite the very high correlation both variables are kept to maximally exploit them.

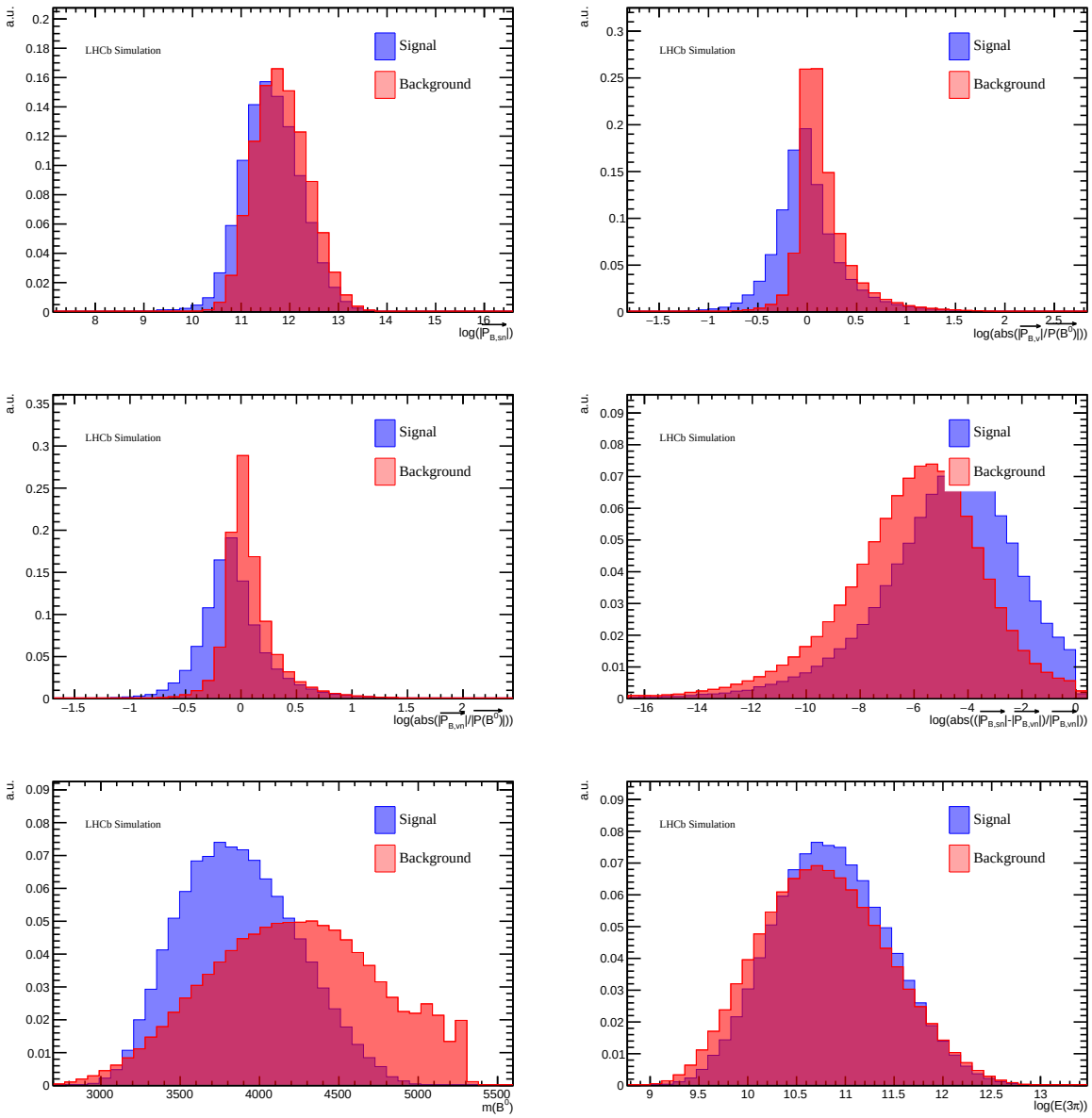


Figure 5.20 – Input variables distributions for the anti-D_s^+ BDT for signal (blue) and background (red).

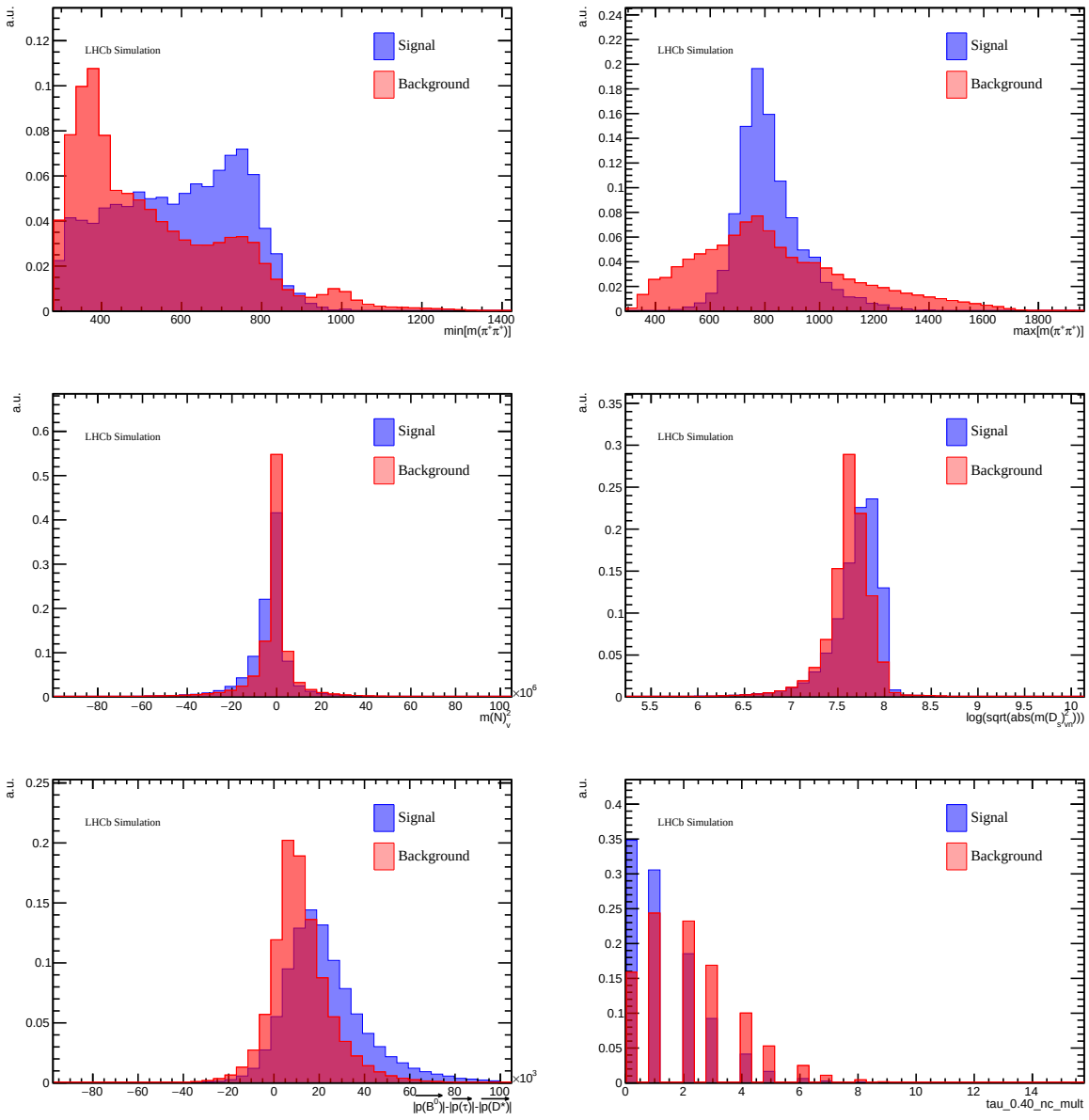


Figure 5.21 – Input variables distributions for the anti-D_s^+ BDT for signal (blue) and background (red).

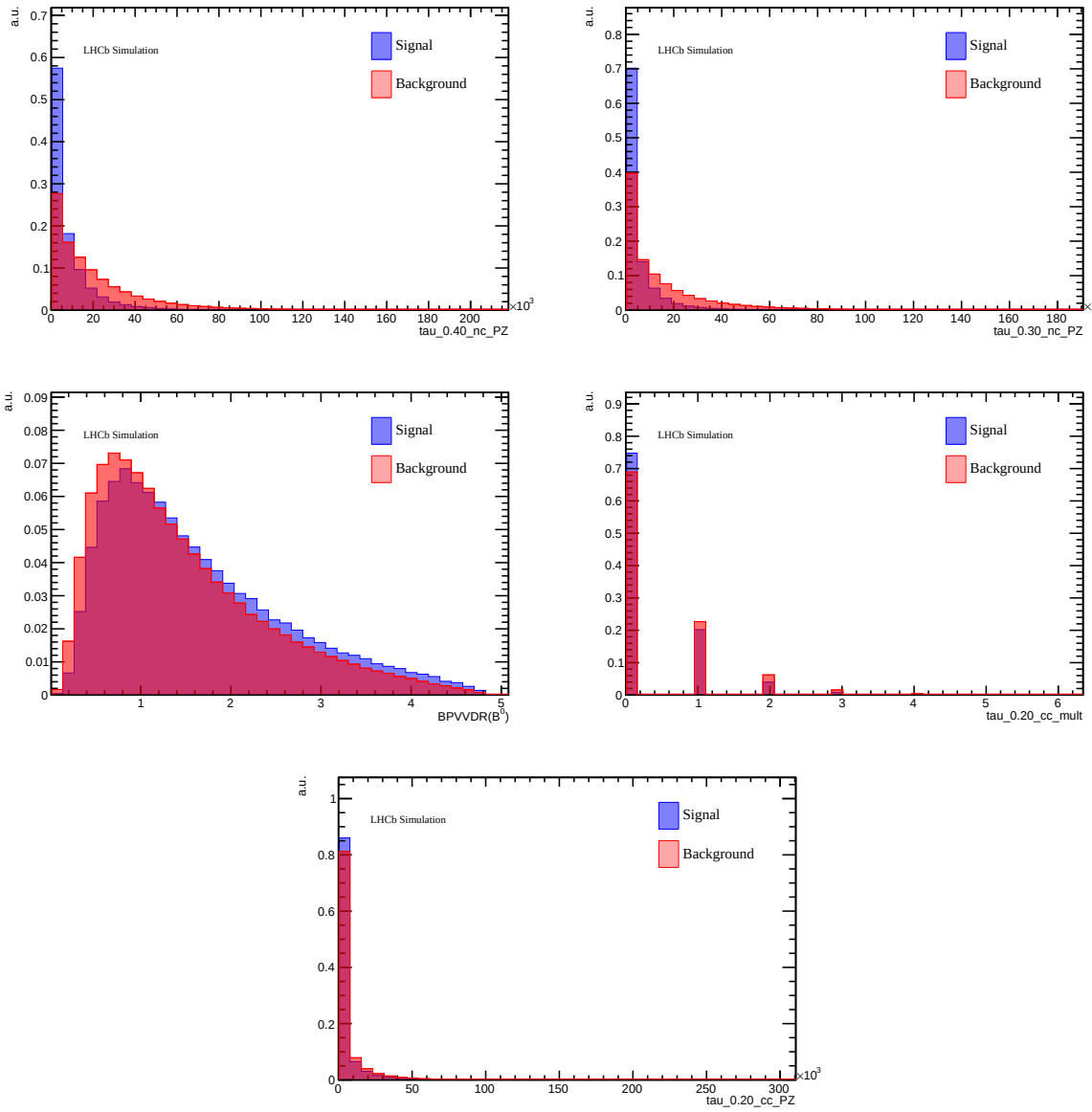


Figure 5.22 – Input variables distributions for the anti-D_s^+ BDT for signal (blue) and background (red).

Figure 5.23 – Input variables correlation matrix for the signal (top) and background (bottom) training samples of the anti-D_s^+ BDT.

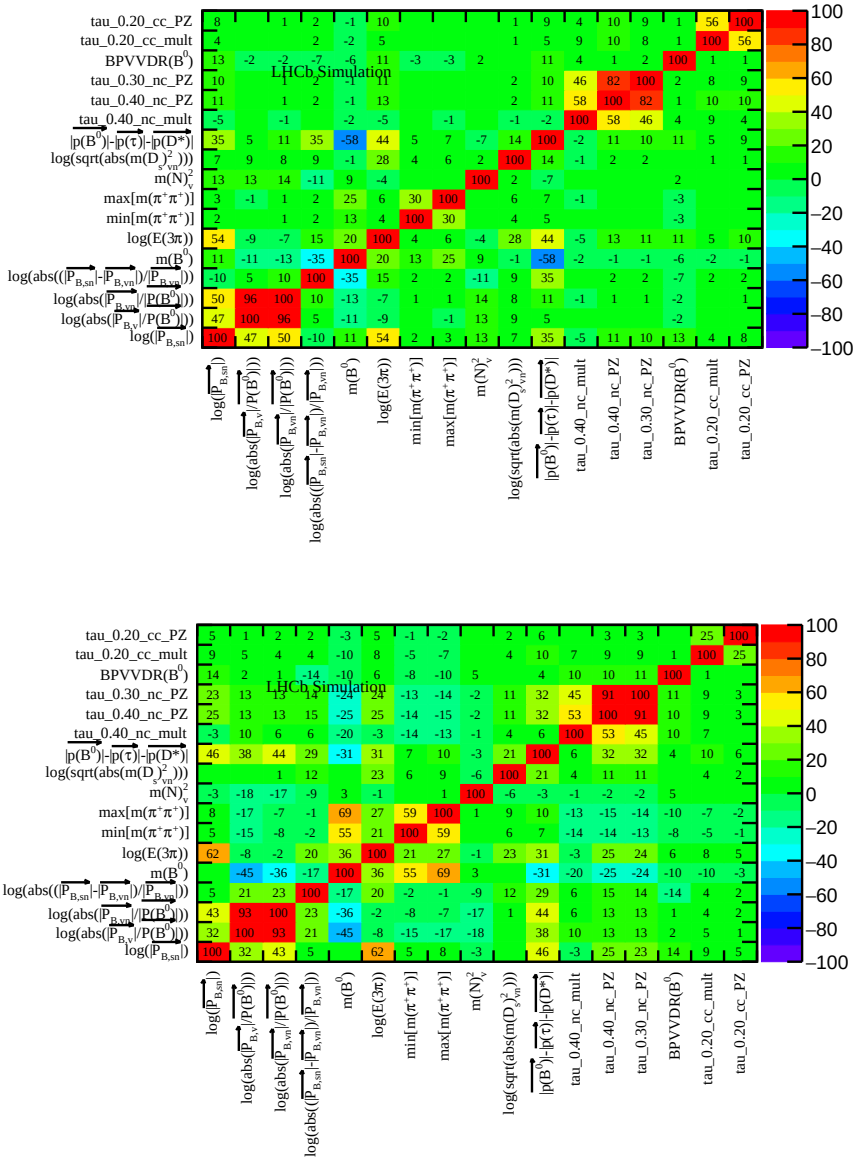


Figure 5.24 depicts (left) the anti-D_s^+ BDT distribution for the training samples and (right) the background rejection as a function of signal efficiency. The AdaBoost BDT (BDT) is the most performant and has, therefore, been chosen as the default one.

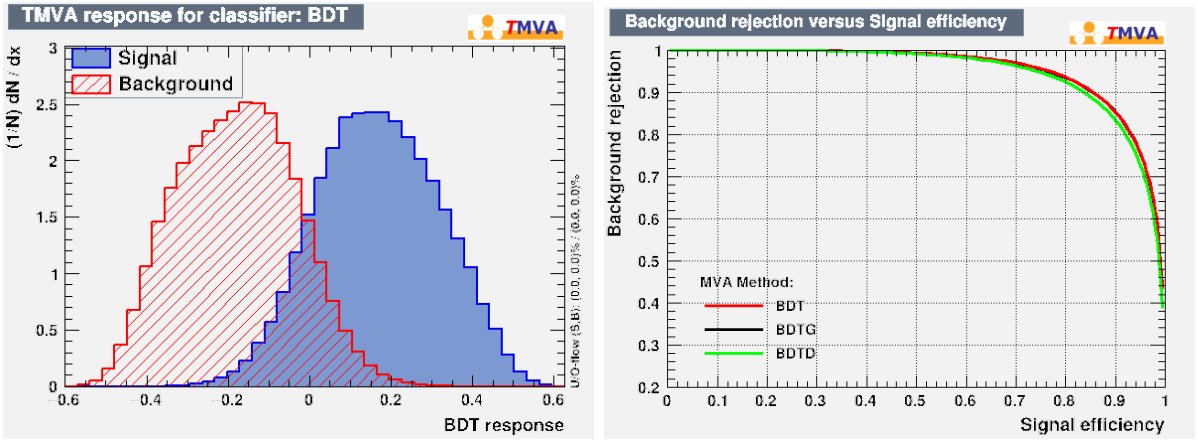


Figure 5.24 – Left: BDT distribution for the signal (blue) and background (red) testing samples. Right: background rejection as a function of signal efficiency for the testing samples.

The cut applied to the anti-D_s^+ BDT is anti-D_s^+ BDT > -0.2, which keeps almost all the signal (signal efficiency: 99.7%) and rejects around 31% of the background. The cut is very loose since this variable is used for the signal yield fit. Its primary objective is to distinguish the signal from the double-charm (mostly with D_s^+) components. Therefore, the cut rejects only the most obvious $B^0 \to D^{*-} D_s^+ \; X$ events, without sacrificing the signal statistics for the final fit in Chapter 7.

This BDT most importantly should discriminate the signal from the double-charm backgrounds. Since it relies on the resonant structure of the $3\pi^\pm$ system, it is more meaningful to see its data/MC agreement for the double-charm modes rather than for the prompt pions in the normalisation mode. Therefore, the validation shown in this section does not follow the scheme for the three preceding BDTs.

Let us examine the BDT distributions for the data control samples. Because the $B \to D^{*-} D_s^+ X$ control sample is selected using that BDT, we cannot use it in this exercise, unless we modify the BDT. We can, nonetheless, inspect the data/MC agreement in the other double-charm modes, *i.e.* the $B \to D^* D^0 \; X$ and $B \to D^* D^+ \; X$ control samples, which are selected with no cuts correlated with the BDT. Their selection is documented in Chapter 6.

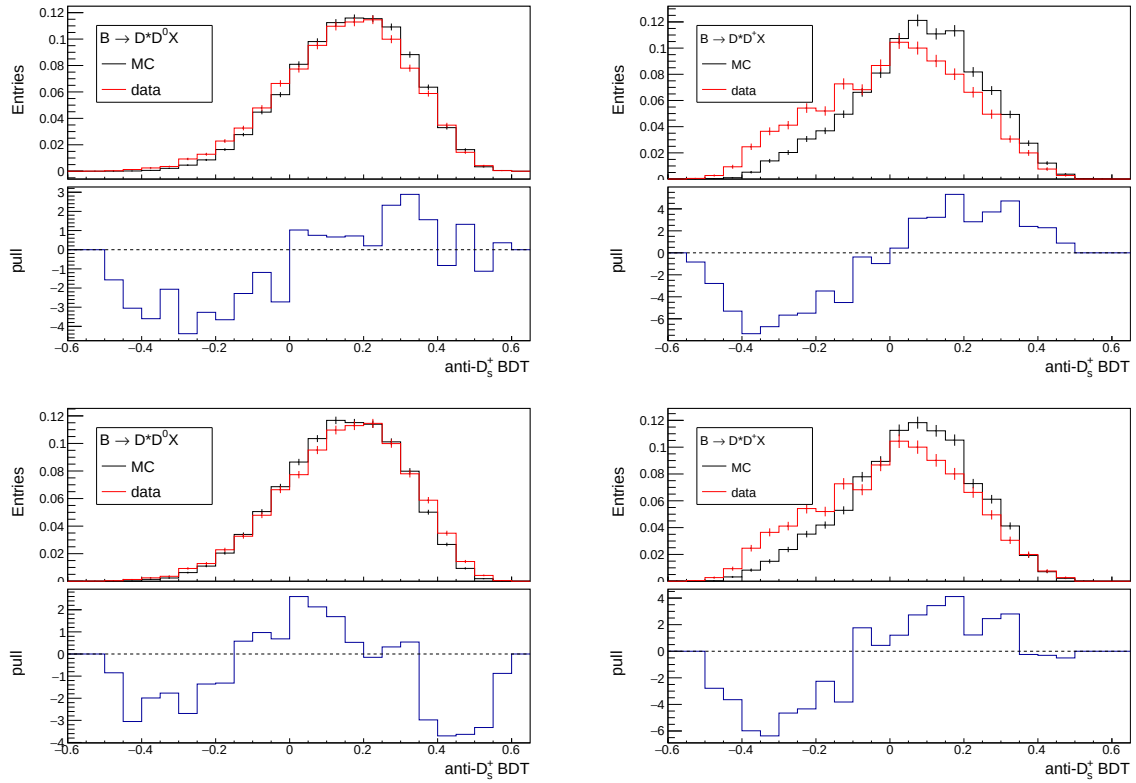


Figure 5.25 – Top: Anti D_s^+ BDT comparison for data and MC samples of the $B \to D^* D^0\ X$ (left) and $B D^* D^+\ X$ (right) modes; pull distributions are shown below each subplot. Bottom: similar distributions after shifting the MC distribution by -0.02 BDT score.

Figure 5.25 (top) shows the data/MC comparison for the anti D_s^+ BDT distributions of the $B \to D^* D^0\ X$ (left) and $B \to D^* D^+\ X$ (right) modes. In both cases there is a little shift towards higher BDT scores in MC (with usually negative (positive) pulls below (above) the anti-D_s^+ BDT$= 0$). Figure 5.25 (bottom) depicts similar graphs with a shift of -0.02 BDT score applied to MC that slightly improves the agreement for the $B^0 \to D^* D^+\ X$ modes (bottom right). However, it results in another 'dip' in the pull distribution of the $B \to D^* D^0\ X$ mode around the BDT score $0.4 - 0.5$ (bottom left). Therefore, is is not straightforward to mitigate the data/MC discrepancy simultaneously for the two modes. Nonetheless, given that the background structure is different in data and the inclusive MC, the agreement is rather acceptable, especially in the fit region (*i.e.* anti-D_s^+ BDT> -0.2). The comparison of the data/MC distributions of the input variables to the BDT is shown in the Appendix D. Discrepancies are found in the $B \to D^{*-} D^0\ X$ case and they should be investigated in the next steps of the analysis. The comparison with the $B \to D^{*-} D^+\ X$ samples, gives rather satisfactory agreement.

5.9 Remaining signal cuts

To select the signal mode, further cuts are applied, as reported in Table 5.7. The τ^+ decay vertex lies 2σ downstream of the B^0 decay vertex to remove the $3\pi^\pm$ from B^0 (referred to as 'prompt' background). The D^{*-} mass is constrained to satisfy $m(D^{*-}) - m(\overline{D}^0) \in [143, 148]\,\mathrm{MeV}/c^2$. This way the $D^*(2010)^-$ meson is selected. To avoid the double charm backgrounds, an upper boundary is put on the τ^+ mass: $m(\tau^+) < 1825\,\mathrm{MeV}/c^2$. Due to the partial reconstruction of the decay, an upper boundary is also put on the B^0 mass: $m(B^0) < 5350\,\mathrm{MeV}/c^2$. The reconstructed q^2 has to be in the range $[0, 11]\,\mathrm{GeV}^2/c^4$ to remove the combinatorial backgrounds. As explained in the previous sections, the four MVA cuts are applied. Finally, the PID requirements are put on: the pion from D^{*-}, `ProbNNpi` above 0.1; the $3\pi^\pm$ from the τ^+ candidate, `ProbNNpi` above 0.6; for the opposite-sign pion `ProbNNk` less than 0.1.

Table 5.7 – Remaining cuts for the signal modes. See text for further explanation.

variable	cut	background targeted
$[\mathrm{vtx}_z(\tau^+) - \mathrm{vtx}_z(B^0)]/\mathrm{error}$	> 2	prompt
$m(K^-\pi^+)$	$\in [1840, 1890]\,\mathrm{MeV}/c^2$	combinatorial D^0
$m(D^{*-}) - m(K^-\pi^+)$	$\in [143, 148]\,\mathrm{MeV}/c^2$	combinatorial D^{*-}
$m(\tau^+)$	$< 1600\,\mathrm{MeV}/c^2$	double-charm
$m(B^0)$	$< 5100\,\mathrm{MeV}/c^2$	combinatorial
q^2	$\in [0, 11]\,\mathrm{GeV}^2/c^4$	combinatorial
ProbNNpi π^- from D^{*-}	> 0.1	misidentification
ProbNNpi $\pi^\pm$ from τ^+	> 0.6	misidentification
ProbNNk π^- from τ^+	< 0.1	misidentification
anti D_s^+ BDT	> -0.2	$D^{*-}D_s^+\ X$
isolation BDT	> 0.0	double-charm
combinatorial BDTD	> 0.0	combinatorial
detachment BDTG	> 0.2	prompt

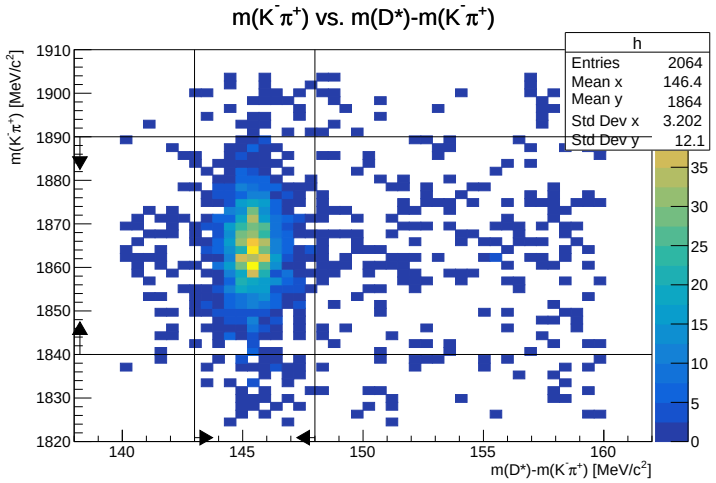


Figure 5.26 – $m(K^-\pi^+)$ as a function of $m(D^{*-}) - m(K^-\pi^+)$ for the data sample after the full selection except the cuts on these variables, which are marked by arrows.

To assure the selection achieves its objectives, composition of the inclusive $b\bar{b} \rightarrow D^{*-}3\pi^{\pm}X$ MC sample is inspected after the initial selection (Figure 5.27 top) and after the full selection (Figure 5.27 middle). Only D^{*-} candidates with matched tracks are considered, since combinatorial D^0 and D^{*-} events are removed with a data-driven technique, as discussed in Section 7.1.1.2. Moreover, the composition is depicted with a more stringent anti-D_s^+ BDT cut at anti-D_s^+ BDT > 0 (Figure 5.27 bottom), which shows an increased signal purity (obtained at the price of reducing the signal statistics by around 20%). Future datasets will therefore enable to exploit higher BDT regions.

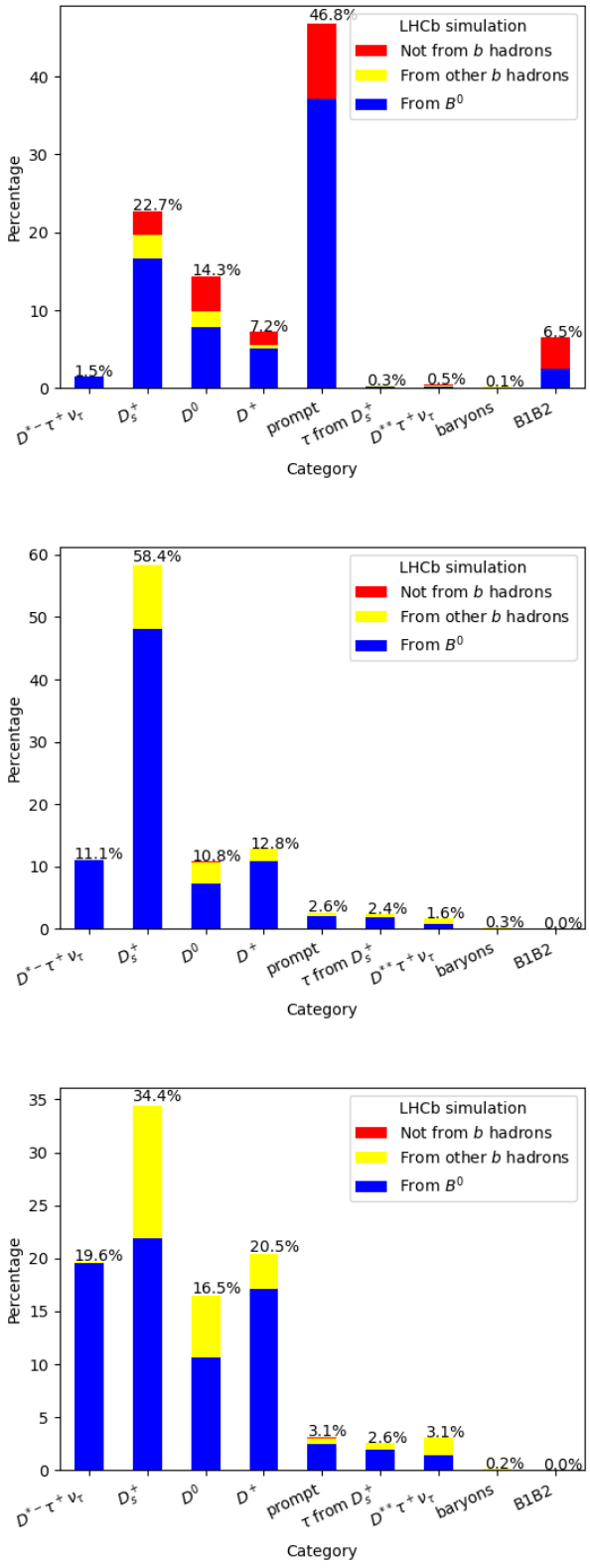


Figure 5.27 – Inclusive MC sample composition after the initial selection (top), after the full signal selection (middle) and with a harsher anti-D_s^+ BDT cut at anti-D_s^+ BDT > 0. The 'prompt' category encompasses events where the $3\pi^{\pm}$ system originates at the b-hadron decay vertex.

After the full selection, almost all candidates originate from a b-hadron and a vast majority of them come from the B^0 meson. Most importantly, the 'prompt' background is reduced and the dominant backgrounds are the three double charm decay modes, especially $D^{*-}D_s^+X$. The latter one is controlled by the anti-D_s^+ BDT in the signal-yield fit, hence the loose cut on this variable.

The selection reported here is reasonably optimised. Even though it is possible to still fine-tune it, the PhD constraints require the author to quickly converge on this part of the project.

5.10 Selection for the normalisation mode

The normalisation mode selection should exploit the unique characteristics of this mode: the fully reconstructed $B^0 \to D^{*-}\ 3\pi^{\pm}$ decay and the $3\pi^{\pm}$ system coming directly from the B^0. At the same time, it should be as similar to the signal selection as possible, to cancel out biases in the $R(D^*)$ ratio.

The first requirement is met by constraining the B^0 mass spectrum to the range $[5150, 5400]\,\mathrm{MeV}$. The second criterion is satisfied by demanding a different $3\pi^{\pm}$ vertex topology. This is ensured by demanding that the $\overline{D}^0$ decay vertex lie further downstream of the $3\pi^{\pm}$ one: $[\mathrm{vtx}_z(\overline{D}^0) - \mathrm{vtx}_z(\tau^+)]/\mathrm{error} > 4$, instead of the vertex detachment as in the signal mode.

Regarding the MVA selection, the normalisation mode does not use the detachment BDT, since the $B^0 \to D^{*-}3\pi^{\pm}$ decay has the $3\pi^{\pm}$ 'attached' to the B^0 decay vertex. Furthermore, the anti-D_s^+ BDT, relying on the intermediate resonant states in the $3\pi^{\pm}$ system would be irrelevant for the normalisation mode. The other two BDTs, the combinatorial and isolation ones, are kept with the same cuts as for the signal mode.

The PID requirements are the same as for signal. Even though they might be reported in the initial selection, they are reported here for consistency with the efficiencies computation.

The other cuts are kept at the same values to reduce bias in the relative signal-to-normalisation yields.

The selection of the normalisation mode is reported in Table 5.8.

Table 5.8 – Remaining cuts for the normalisation mode. See text for further explanation.

variable	cut	targeted background
$m(B^0)$	$\in [5150, 5400]\,\text{MeV}$	combinatorial
$m(D^{*-}) - m(\overline{D}^0)$	$\in [143, 148]\,\text{MeV}/c^2$	combinatorial D^{*-}
$[\text{vtx}_z(\overline{D}^0) - \text{vtx}_z(\tau^+)]/\text{error}$	> 4	non-prompt
ProbNNpi π^- from D^{*-}	> 0.1	misidentification
ProbNNpi $\pi^\pm$ from τ^+	> 0.6	misidentification
ProbNNk π^- from τ^+	< 0.1	misidentification
isolation BDT	> 0.0	double-charm
combinatorial BDTD	> 0.0	combinatorial

Figure 5.28 shows the invariant mass spectrum of the $D^{*-}\ 3\pi^\pm$ system after the full normalisation-mode selection. The peak around the B^0 mass is clearly visible. The lower mass tail stems from the $B^0 \to D^{*-}\ 3\pi^\pm\ X$ events. Moreover, the peak may contain feeddown from D_s^+, *i.e.* $D_s^+ \to 3\pi^\pm$, which events are subtracted as described in Chapter 7.

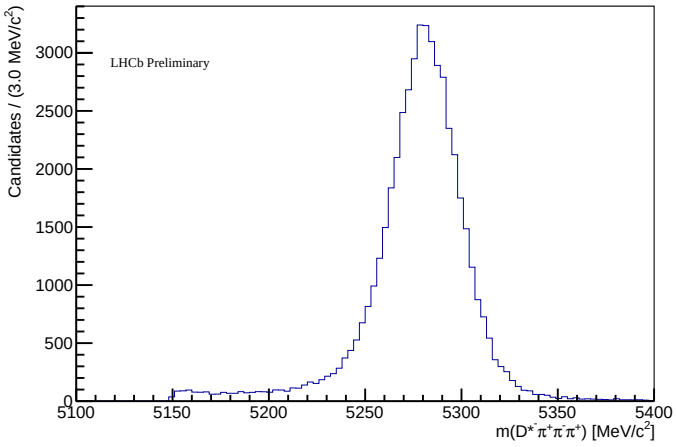


Figure 5.28 – $m(D^{*-}\pi^+\pi^-\pi^+)$ in data after full normalisation-mode selection

5.11 Efficiencies

The total efficiency to select any sample is computed as a product of the 'online' and 'offline' efficiencies,

$$\varepsilon_{\text{tot}} = \varepsilon_{\text{online}}\varepsilon_{\text{offline}}, \tag{5.1}$$

which are factorised as follows:

$$\varepsilon_{\text{'online'}} = \varepsilon_{\text{geom. acceptance}}\varepsilon_{\text{preselection}}, \tag{5.2}$$

$$\varepsilon_{\text{offline}} = \varepsilon_{\text{initial}}\varepsilon_{\text{mode-specific}}\varepsilon_{\text{PID}}, \tag{5.3}$$

where the $\varepsilon_{\text{mode-specific}}$ stands for the signal or normalisation efficiencies.

The 'online' efficiencies for the two signal modes and the normalisation mode are reported in Table 5.9.

Table 5.10 reports the 'offline' (*i.e.* starting from the ROOT ntuples saved on disk ending at the final selection) selection efficiencies per step of the initial selection, common to all the modes, and then for the signal and normalisation modes. The signal selection is applied to all MC samples, therefore the $b\bar{b} \to D^{*-}3\pi^{\pm}X$ efficiencies are shown whenever possible. The absolute and cumulative efficiencies are shown in the second and third columns. In the initial selection, the first three lines are the trigger cuts. They are applied also for the remaining initial cuts, hence the indentation. The signal and normalisation efficiencies are quoted as measured on the samples with the initial cuts already applied, hence the discontinuity in the cumulative efficiencies between the initial and signal cuts, and between the initial and normalisation cuts. The PID efficiencies are computed with a data-driven approach (see 2.2.2.4) that requires all the other cuts be already applied. Therefore only the PID absolute efficiencies are measured.

Table 5.9 – 'Online' efficiencies (in %) for the two signal modes and the normalisation mode split into the geometrical acceptance and preselection steps.

Selection Step	$3\pi\overline{\nu}_{\tau}$	$3\pi\pi^0\overline{\nu}_{\tau}$	$B^0 \to D^{*-}3\pi^{\pm}$
geometrical acceptance	16.44 ± 0.19	15.86 ± 0.15	15.56 ± 0.14
preselection	0.449 ± 0.001	0.370 ± 0.004	1.183 ± 0.003

Table 5.10 – 'Offline' selection absolute and cumulative efficiencies (in %) as a function of the selection stage for the simulated $b\bar{b} \to D^{*-}3\pi^{\pm}X$, $B^0 \to D^{*-}\tau^+\nu_\tau$ and $B^0 \to D^{*-}3\pi^{\pm}$ modes. The cells group cuts into the selection stages, where each cell assumes all the cuts from the previous stage have been applied. Uncertainties are of the order of last digit shown. See Table C.1 for more details.

Cut	**Absolute efficiencies**			**Cumulative efficiencies**		
	$B^0 \to D^{*-}\tau^+\nu_\tau$		$B^0 \to D^{*-}3\pi^{\pm}$	$B^0 \to D^{*-}\tau^+\nu_\tau$		$B^0 \to D^{*-}3\pi^{\pm}$
	$3\pi\overline{\nu}_\tau$	$3\pi\pi^0\overline{\nu}_\tau$		$3\pi\overline{\nu}_\tau$	$3\pi\pi^0\overline{\nu}_\tau$	
Initial selection						
`L0`	89.51	86.60	89.08	89.51	86.60	89.08
`Hlt1`	89.76	87.32	90.92	87.14	83.88	88.02
`Hlt2`	79.90	77.31	90.33	73.25	69.02	85.10
$PV(\overline{D}^0) = PV(\tau^+)$	69.76	65.73	79.94	69.76	65.73	79.94
`totCandidates` $= 1$	60.89	52.22	71.97	58.06	49.87	67.75
$[\mathrm{vtx}_z(\tau^+) - \mathrm{vtx}_z(PV)]/\mathrm{error} > 10$	71.66	66.59	78.60	57.01	48.29	62.64
`nSPDHits` < 450	72.24	67.78	83.97	56.37	47.56	61.99
Signal selection						
$m(D^{*-}) - m(K^-\pi^+) \in [143, 148]\,\mathrm{MeV}/c^2$	94.63	93.98	-	94.63	93.98	-
$m(K^-\pi^+) \in [1840, 1890]\,\mathrm{MeV}/c^2$	97.36	97.39	-	92.28	91.70	-
$m(3\pi) < 1825\,\mathrm{MeV}/c^2$	98.24	98.77	-	90.73	90.68	-
$m(B^0) < 5100\,\mathrm{MeV}/c^2$	99.29	99.03	-	90.46	90.27	-
$q^2 \in [0, 12]\,\mathrm{GeV}^2/c^4$	97.52	97.22	-	88.74	88.53	-
combinatorial BDTD > 0	80.37	76.71	-	74.72	71.89	-
$[\mathrm{vtx}_z(\tau^+) - \mathrm{vtx}_z(B^0)]/\mathrm{error} > 2$	99.81	99.78	-	74.72	71.89	-
isolation BDT > 0	87.85	83.86	-	67.42	62.41	-
anti D_s^+ BDT > -0.2	98.30	86.10	-	67.12	54.87	-
PID	76.23	78.86	-	-	-	-
Normalisation selection						
$[\mathrm{vtx}_z(\overline{D}^0) - \mathrm{vtx}_z(\tau^+)]/\mathrm{error} > 4$	-	-	94.30	-	-	94.30
$m(D^*3\pi^{\pm}) \in [5150, 5400]\,\mathrm{MeV}$	-	-	97.87	-	-	93.32
$m(D^{*-}) - m(\overline{D}^0) \in [143, 148]\,\mathrm{MeV}$	-	-	94.97	-	-	89.04
combinatorial BDTD > 0	-	-	81.37	-	-	74.19
isolation BDT > 0	-	-	88.33	-	-	66.94
PID	-	-	73.96	-	-	-

Table 5.11 – Summary of total efficiencies of the signal and normalisation modes. The signal efficiency is a weighted average of the $3\pi^{\pm}\overline{\nu}_\tau$ and $3\pi^{\pm}\pi^0\overline{\nu}_\tau$ modes where the weights are the branching fractions of these modes.

quantity	value
$\varepsilon_{B^0 \to D^{*-}\tau^+\nu_\tau(3\pi\overline{\nu}_\tau)}$	$(1.2 \pm 0.03) \times 10^{-4}$
$\varepsilon_{B^0 \to D^{*-}\tau^+\nu_\tau(3\pi\pi^0\overline{\nu}_\tau)}$	$(7.0 \pm 0.2) \times 10^{-5}$
$\varepsilon_{B^0 \to D^{*-}\tau^+\nu_\tau}$	$(1.03 \pm 0.01) \times 10^{-4}$
$\varepsilon_{B^0 \to D^{*-}3\pi^{\pm}}$	$(2.91 \pm 0.03) \times 10^{-4}$
$\varepsilon_{B^0 \to D^{*-}3\pi^{\pm}} / \varepsilon_{B^0 \to D^{*-}\tau^+\nu_\tau}$	2.81 ± 0.04

Finally, the ratio of the total efficiencies of the normalisation to the signal modes is

computed and shown in Table 5.11. This ratio enters the $\mathcal{K}(D^*)$ formula.

The next chapter describes the data-driven control samples fits that help constrain the final signal and normalisation fits. Their selections are described there for convenience.

Chapter 6

Control samples fits

The previous chapter described the full selection, in particular, for the signal mode. Before the signal yield fit is performed, it is useful to apply data-driven corrections to the simulated samples of the most important backgrounds. The data samples used to this end are referred to as the control samples. They are described in this chapter.

First of all, the resonant structure of the D_s^+ decay is corrected as explained in Section 6.1. Next, in Section 6.2 the composition of the $B \to D^{*-} D_s^+ X$ inclusive mode is fitted, providing constraints for the final fit. Finally, in Section 6.3 the $B \to D^{*-} D^+ X$ and $B \to D^{*-} D^0 X$ samples are used to correct the q^2 distribution of the respective decay modes in the final fit templates.

6.1 D_s^+ decay model

The anti-D_s^+ BDT used in the final signal yield fit relies on the modelling of the $3\pi^\pm X$ system as in the decays of, most importantly, the $\tau \to 3\pi^\pm \nu_\tau$ (signal) and $D_s^+ \to 3\pi^\pm X$ (background). The τ^+ decays usually through the $a_1(1260)^+$ state, which in turn goes into $\rho^0\pi^+$. The D_s^+ may also decay to the $a_1(1260)^+$, but the branching fraction is unknown. Nonetheless, the branching fraction of the D_s^+ decay to ρ^0 through the η' state has been measured as $\mathcal{B}(D_s^+ \to \eta' X) = (10.3 \pm 1.4)\%$, where η' decays to $\rho^0\gamma$ with the branching fraction $\mathcal{B}(\eta' \to \rho^0\gamma) = (29.5 \pm 0.4)\%$. See Table A.2 for the relevant D_s^+ branching fractions. Consequently, the ρ^0 contribution from the D_s^+ decay can be mistakenly attributed to the one from the τ^+.

Therefore it is crucial that contributions from various resonances, especially the η', be correctly normalised in the simulation to reflect the real data as closely as possible. Moreover, the branching fractions of certain decay modes are not well-known. To select the data sample enriched in the $D_s^+ \to 3\pi^\pm X$ decays, the signal-mode selection is applied (*cf.* Tab. 5.7) with the reverse requirement on the anti-D_s^+ BDT: anti-D_s^+ BDT < -0.2. A simultaneous maximum likelihood binned fit is performed to the four variables: min$[m(\pi^+\pi^-)]$, max$[m(\pi^+\pi^-)]$, $m(\pi^+\pi^+)$ and $m(3\pi^\pm)$. The variables min$[m(\pi^+\pi^-)]$,

max$[m(\pi^+\pi^-)]$ and $m(\pi^+\pi^+)$ are reconstructed masses of all the combinations of two-pion systems. Therefore, they enable to pinpoint the invariant masses of the resonances spawning two pions. The last variable, $m(3\pi^\pm)$ allows to distinguish the exclusive $D_s^+ \to \pi^+\pi^-\pi^+$ from events where an extra energy is carried by additional neutral particles, *e.g.* $D_s^+ \to \tau^+$ ($\to \pi^+\pi^-\pi^+\overline{\nu}_\tau$) ν_τ or $D_s^+ \to 3\pi^\pm X$, where X escapes detection.

Besides, the fitting range of the four variables is constrained by phase-space *i.e.* the lower ranges are around the double or triple of the $\pi^\pm$ invariant mass, whereas the upper limits are usually around the D_s^+ invariant mass, except the min$[m(\pi^+\pi^-)]$ with the limit of 1200 MeV/c^2 as heavier resonances producing $\pi^+\pi^-$ occur very rarely in D_s^+ decays. Table 6.1 shows differences of the D_s^+ decay model selection with respect to the signal cuts. This control sample comprises 8821 events of the collision data.

Table 6.1 – Selection of the D_s^+ decay model control sample

variable	cut
anti-D_s^+ BDT	< -0.2
$m(\pi^+\pi^+)$	$\in [240, 1600]$ MeV
max$[m(\pi^+\pi^-)]$	$\in [280, 1700]$ MeV
min$[m(\pi^+\pi^-)]$	$\in [280, 1200]$ MeV
$m(\pi^+\pi^-\pi^+)$	$\in [440, 1840]$ MeV

Figure 6.1 shows projections of the simulated $D_s^+ \to 3\pi^\pm X$ templates to the four fitted variables: min$[m(\pi^+\pi^-)]$, max$[m(\pi^+\pi^-)]$, $m(\pi^+\pi^+)$ and $m(\pi^+\pi^-\pi^+)$. The components are: $D_s^+ \to \eta'\pi^+(\pi^0)$ (green); $D_s^+ \to \eta\pi^+(\pi^0)$ (red); $D_s^+ \to \omega\pi^+(\pi^0)$ and $D_s^+ \to \phi\pi^+(\pi^0)$ (orange); other D_s^+ modes (yellow); non-D_s^+ background(blue).

The components are broken down into:

- decays with at least one pion from η: $D_s^+ \to \eta\pi^+(\pi^0)$
- decays with at least one pion from η': $D_s^+ \to \eta'\pi^+(\pi^0)$
- decays with at least one pion from ω or ϕ: $D_s^+ \to \omega\pi^+(\pi^0)$ or $D_s^+ \to \phi\pi^+(\pi^0)$
- decays where none of the pions originates from an intermediate resonance: $D_s^+ \to 3\pi^\pm X$ modes
- Non-D_s^+ background (*e.g.* $D_s^+ \to \tau^+\nu_\tau$)

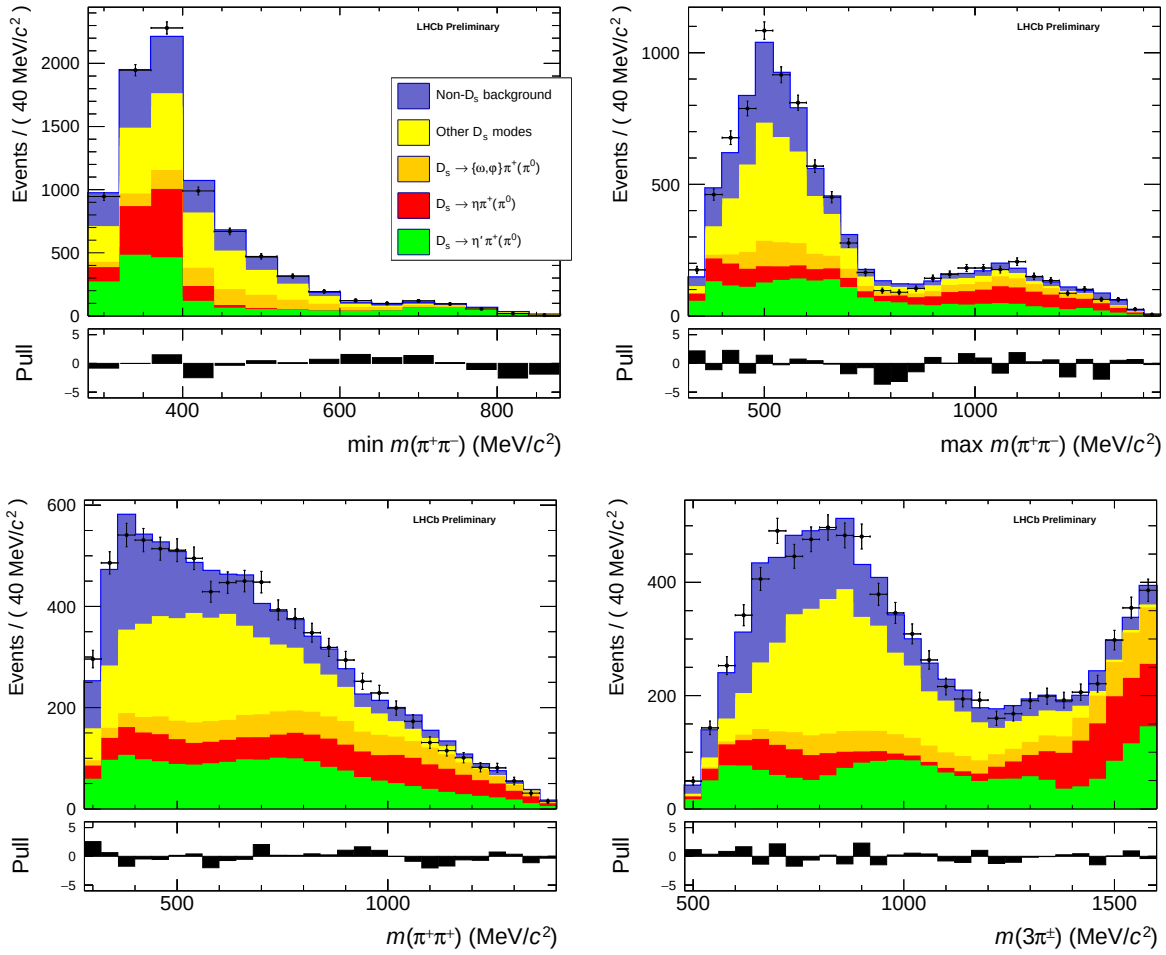


Figure 6.1 – Projections of the $D_s^+ \to 3\pi^\pm\ X$ components for the variables: $\min[m(\pi^+\pi^-)]$, $\max[m(\pi^+\pi^-)]$, $m(\pi^+\pi^+)$ and $m(\pi^+\pi^-\pi^+)$ in the fit to the control data samples. The components are: $D_s^+ \to \eta'\pi^+(\pi^0)$ (green); $D_s^+ \to \eta\pi^+(\pi^0)$ (red); $D_s^+ \to \omega\pi^+(\pi^0)$ and $D_s^+ \to \phi\pi^+(\pi^0)$ (orange); other D_s^+ modes (yellow); non-D_s^+ background(blue).

These modes are sub-divided more finely as reported in Table 6.2, where their relative fractions (normalised to all $D_s^+ \to 3\pi^\pm\ X$ decays), obtained from the fit, are shown. Importantly, the estimated fractions of the $D_s^+ \to a_1\ X$ are reported. The last column represents a multiplicative correction factor to be applied to the MC templates in order to correctly normalise the modes to match the fitted fractions. That correction is applied to the relevant MC samples to the signal-fit templates, discussed in Chapter 7, after accounting for the efficiency differences between the D_s^+ and signal-mode selections.

Table 6.2 – Relative fractions of various $D_s^+ \to 3\pi^\pm X$ templates obtained from the fit to the collision data D_s^+ control sample (see Figure 6.1) and the corresponding corrections factors applied to the simulation.

Template	$D_s^+ \to 3\pi^\pm$ X fraction	MC Correction factor
$\eta 3\pi^\pm$	0.069 ± 0.019	1.26 ± 0.35
$\eta' 3\pi^\pm$	0.034 ± 0.023	0.71 ± 0.48
$\omega 3\pi^\pm$	0.187 ± 0.033	6.0 ± 1.1
$\phi 3\pi^\pm$	0.0300 ± 0.0070	0.97 ± 0.22
$K^0 3\pi^\pm$	0.0120 ± 0.0030	0.91 ± 0.20
$D_s^+ \to \tau^+ \nu_\tau$	0.0160 ± 0.0040	0.94 ± 0.21
$X_{\rm nr} 3\pi^\pm$	0.051 ± 0.061	3.8 ± 4.6
$a_1 \eta$	0.0080 ± 0.0020	1.34 ± 0.37
$a_1 \eta'$	0.0040 ± 0.0030	0.66 ± 0.45
$a_1 \omega$	0.0210 ± 0.0040	3.36 ± 0.59
$a_1 \phi$	0.0030 ± 0.0010	0.76 ± 0.17
$a_1 K^0$	0.00 ± 0.00	1.37 ± 0.31
$\omega\rho^+$ or $\phi\rho^+$	0.071 ± 0.019	0.38 ± 0.10
$\omega\pi^+$ or $\phi\pi^+$	0.0500 ± 0.0090	1.29 ± 0.24
$\eta\rho^+$	0.162 ± 0.019	0.94 ± 0.11
$\eta\pi^+$	0.014 ± 0.010	0.65 ± 0.46
$\eta'\rho^+$	0.189 ± 0.015	0.690 ± 0.056
$\eta'\pi^+$	0.079 ± 0.018	1.02 ± 0.23

6.2 $B \to D^{*-} D_s^+(X)$ control mode

The $B \to D^{*-} D_s^+(X)$ modes form an important family of backgrounds in the signal yield fit. Their relative abundances in the double-charm cocktail simulation samples are corrected by performing a fit to a data sample enriched in $B \to D^* D_s^+(\to 3\pi^\pm)X$ events. The selection of the control sample differs from the nominal selection by requiring the 3π invariant mass to be within 20 MeV/c^2 of the D_s^+ mass and skipping the anti-D_s^+ BDT cut, as well as the B^0 and τ^+ mass constraints.

Consequently, only the exclusive $D_s^+ \to 3\pi^\pm$ modes are considered. Nonetheless, the relative fractions of the $B \to D^{*-} D_s^+(X)$ decays are valid for the inclusive $D_s^+ \to 3\pi^\pm$ X decays, as it is the case in the signal yield fit [1].

This control sample comprises 2878 events collision data events.

The fit is performed to $m(D^{*-}\pi^+\pi^-\pi^+)$ shifted by subtracting $m(K^-\pi^+)_{D^0}$ and $m(3\pi^\pm)$ for better resolution, and to $q^2 \equiv (p_{B^0} - p_{D^{*-}})^2$. The probability density function used

[1] The $D_s^+ \to 3\pi^\pm$ X decay model corrections, discussed in the previous section, are not applicable here, because they concern the inclusive $D_s^+ \to 3\pi^\pm$ X modes, while here we discuss the exclusive $D_s^+ \to 3\pi^\pm$ events only.

in the fit is

$$\mathcal{P} = f_{\text{c.b.}} \mathcal{P}_{\text{c.b.}} + \frac{(1 - f_{\text{c.b.}})}{k} \sum_i f_i \mathcal{P}_i, \tag{6.1}$$

where $f_{\text{c.b.}}$ and $\mathcal{P}_{\text{c.b.}}$ are fractions and probability density function of the combinatorial background, modelled with the $D^{*\pm} 3\pi^{\pm}$ system, where the $D^{*\pm}$ has the same charge as the $3\pi^{\pm}$; f_i are floating-in-the-fit fractions of the $B \to D^{*-} D_s^+ X$ components as defined in Table 6.4, relative to the most abundant $B^0 \to D^{*-} D_s^{*+}$ decays; $k = 1 + f_{D_s^+} + f_{D_{s0}^{*+}} + f_{D_{s1}^{\prime +}} + f_{D^{**} D_s X} + f_{B_s \to D^* D_s^+ X}$ and $f_{D_s^{*+}} = 1$ by definition. The shapes of each component are taken from binned templates from simulation.

The fit projections on the fitted variables and also onto t_τ and anti-D_s^+ BDT are depicted in Figure 6.2.

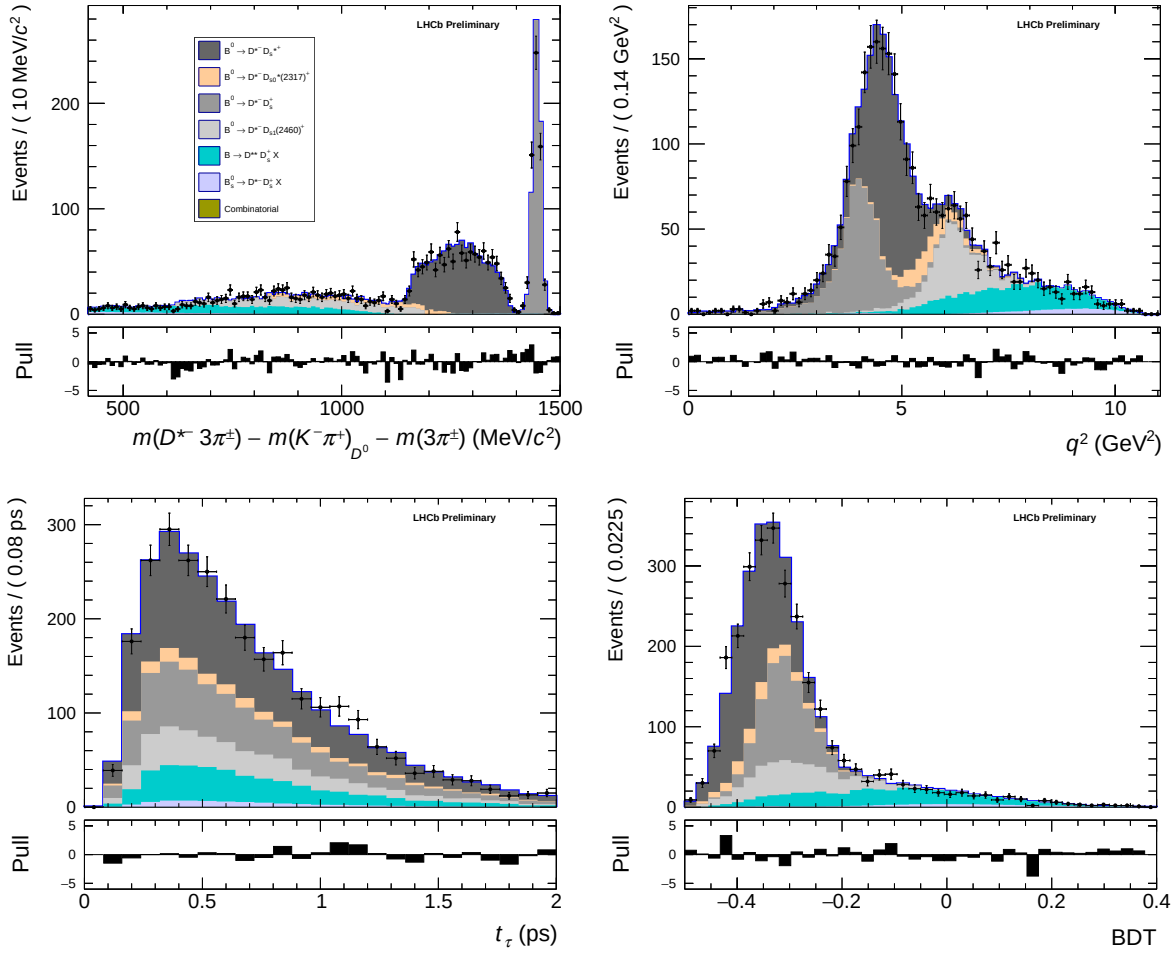


Figure 6.2 – Data-fit projections of the $B \to D^{*-} D_s^+ X$ components for the variables: $m(D^{*-}\pi^+\pi^-\pi^+) - m(K^-\pi^+)_{D^0} - m(3\pi^{\pm})$, q^2, t_τ and anti-D_s^+ BDT. The first two variables are used in the fit.

The $B \to D^{*-} D_s^+(X)$ modes can be grouped into the exclusive $B^0 \to D^{*-} D_s^{+(*)}$ modes, and the inclusive $B \to D^{**-} D_s^+ X$ and $B_s^0 \to D^{*-} D_s^+ X$ modes.

The exclusive modes include the exclusive $B^0 \to D^{*-} D_s^+$ mode and those with excited states of the D_s^+, denoted by D_s^{*+}, D_{s0}^{*+} and $D_{s1}^{\prime +}$. Their $q^2 = (p_{B^0} - p_{D^{*-}})^2$ peaks at the mass of the given D_s^+ states as can be seen in the top-right of Figure 6.2 for the mass squared of the D_s^+ species reported in Table 6.3.

Table 6.3 – Invariant mass and its square for the D_s^+ states of interest [14]. Uncertainties are smaller than the last digit shown.

D_s^+ state	mass$[\,\mathrm{MeV}/c^2\,]$	mass$^2[\,\mathrm{GeV}^2/c^4\,]$
D_s^+	1968	3.9
D_s^{*+}	2112	4.5
D_{s0}^{*+}	2318	5.4
$D_{s1}^{\prime +}$	2460	6.1

The inclusive $B \to D^{*-} D_s^+ X$ modes often have at least one extra particle, possibly from a feed-down from excited D^{*-} and excited D_s^+ states. These additional particles carry momentum that contributes to the q^2, hence shifting this distribution to higher values as can be seen in top-right of Figure 6.2 for the modes $B^0 \to D^{**-} D_s^+ X$ and $B_s^0 \to D^{*-} D_s^+ X$.

Table 6.5 reports the $B^0 \to D^{*-} D_s^+$ and combinatorial background yields obtained in the fit to the simulation sample and the fractions of the other modes.

Table 6.4 – List of components in the $D^* D_s^+ X$ control fit and how they are normalised.

Component	Normalisation
$B^0 \to D^{*-} D_s^+$	$N_{D_s^+} \times f_{D_s^+}/k$
$B^0 \to D^{*-} D_s^{*+}$	$N_{D_s^+} \times 1/k$
$B^0 \to D^{*-} D_{s0}^{*+}$	$N_{D_s^+} \times f_{D_{s0}^{*+}}/k$
$B^0 \to D^{*-} D_{s1}^{\prime +}$	$N_{D_s^+} \times f_{D_{s1}^{\prime +}}/k$
$B \to D^{**-} D_s^+ X$	$N_{D_s^+} \times f_{D^{**} D_s X}/k$
$B_s^0 \to D^{*-} D_s^+ X$	$N_{D_s^+} \times f_{B_s \to D^* D_s^+ X}/k$
Background	N_{bkg}

Table 6.5 – $D^* D_s^+ X$ control fit parameters. The fractions are normalised relative to $B^0 \to D^{*-}\ D_s^{*+}$, as defined in Table 6.4.

Parameter	Fit result
N_{D_s}	2757 ± 37
f_{D_s}	0.58 ± 0.03
$f_{D_{s0}^{*+}}$	0.12 ± 0.03
$f_{D_{s1}^{\prime +}}$	0.35 ± 0.03
$f_{D^{**} D_s X}$	0.34 ± 0.04
$f_{B_s \to D^* D_s X}$	0.06 ± 0.02
N_{bkg}	0 ± 3

The various $B \to D^{*-}D_s^+X$ components contribute significantly in the signal fit, therefore their relative fractions derived in this section are used as constraints in the signal yield fit (*cf.* Chapter 7), leaving their normalisation, $N_{D_s^+}$, unconstrained.

6.3 $B \to D^{*-}D^0(X)$ and $B \to D^{*-}D^+(X)$ control modes

$B \to D^{*-}D^+(X)$ and $B \to D^{*-}D^0(X)$ are the other major double-charm backgrounds present in the signal fit. It is important to inspect and correct their q^2 distribution.

They are selected similarly to signal except that no MVA is used to keep high statistics and no requirements are put on the B^0 and τ^+ masses.

The $D^0 \to 3\pi^\pm\, X$ in $B \to D^{*-}D^0X$ is dominated by the decay $D^0 \to K^-\pi^+\pi^-\pi^+$ whose branching fraction is $\mathcal{B}(D^0 \to K^-\pi^+\pi^-\pi^+) = (8.23 \pm 0.14)\%$. The isolation tool scans the $3\pi^\pm$ vertex in search of the extra kaon. The kaon mass is assigned to it and the invariant mass of the thus created $K^-\pi^+\pi^-\pi^+$ system has to be within $\pm 50\,\mathrm{MeV}/c^2$ of the D^0 mass.

To select the $B \to D^{*-}D^+X$ control sample, the decay $D^+ \to \pi^+K^-\pi^+$ ($\mathcal{B}(D^+ \to \pi^+K^-\pi^+) = (9.38 \pm 0.16)\%$) is used. The opposite-sign pion in the $\pi^+\pi^-\pi^+$ system, the π^-, is required to be a genuine K^-. This is achieved by inverting its PID cut such that `probNNK>0.1`. The mass of the $\pi^+K^-\pi^+$ has to be within $\pm 20\,\mathrm{MeV}/c^2$ of the mass of D^0.

Moreover, the events where some tracks come from one B meson and the other from another one are removed. This condition is satisfied by requiring good alignment of the B^0 reconstructed momentum relative to the B^0 flight direction from the primary vertex with the cut acos(`B0_BPVDIRA`) < 0.02.

The $B \to D^{*-}D^0(X)$ and $B \to D^{*-}D^+(X)$ control samples comprise 10901 and 4059 events, respectively, from the collision data.

Figure 6.3 shows the q^2 distributions for the simulated (left) $B \to D^{*-}D^0(X)$ and (right) $B \to D^{*-}D^+(X)$ modes (top) before and (bottom) after the q^2 reweighting with `GBReweighter` [111].

Thus q^2-reweighted simulated samples are used to build templates for the signal yield fit described in Chapter 7.

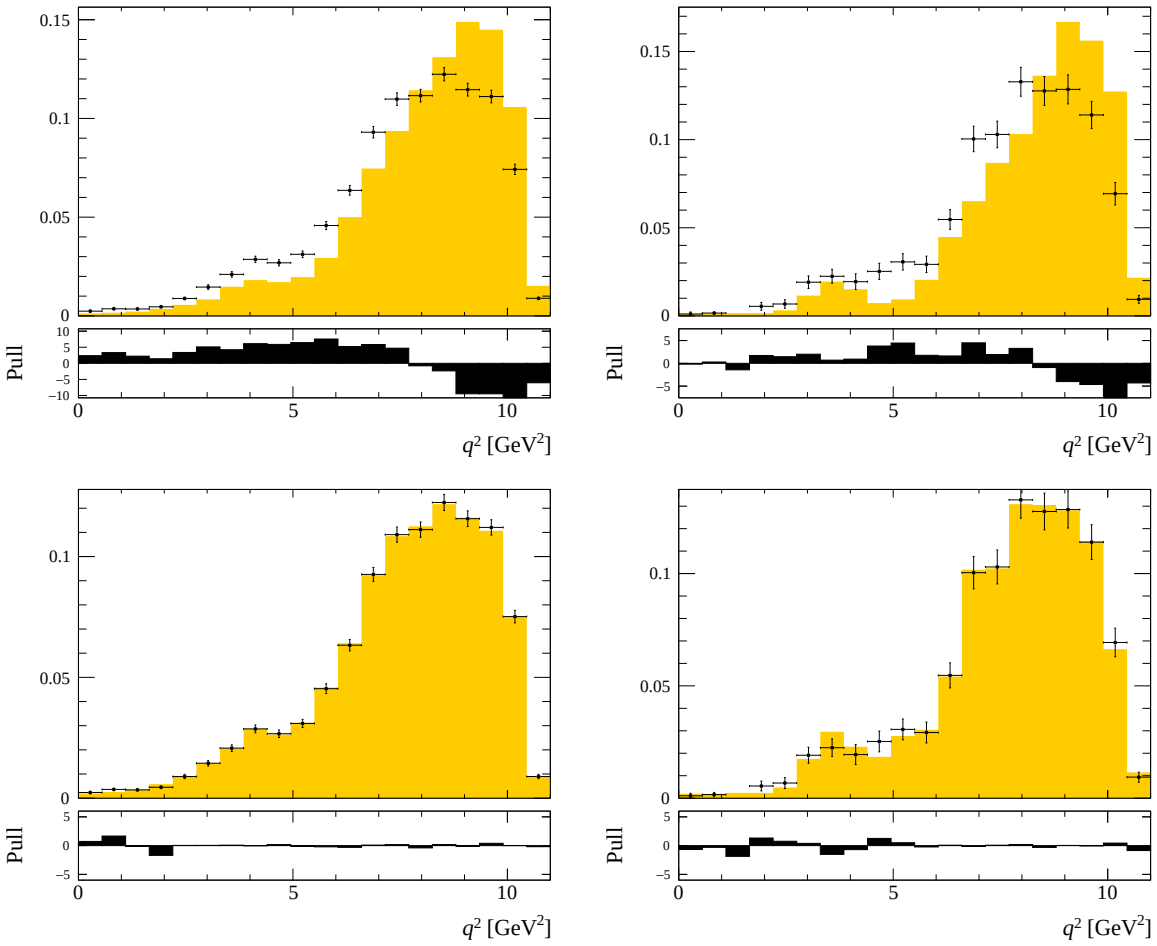


Figure 6.3 – Simulated (filled-in green bins) q^2 distributions for (left) $B \to D^{*-}D^0(X)$ and (right) $B \to D^{*-}D^+(X)$ modes (top) before and (bottom) after reweighting with `GBReweighter` based on the data control samples (data points shown).

Chapter 7

Signal and normalisation fits

The previous chapters discussed the selection (see Chapter 5) and control samples fits (see Chapter 6). The former provided the selected datasets (collision data and MC), while the latter, constraints and corrections for the double-charm backgrounds. In this chapter the signal and normalisation yields are extracted and the final $R(D^*)$ result is computed, following closely the strategy pioneered by Refs. [1, 2].

The fits to extract the signal and normalisation yields are described in Sections 7.1 and 7.2 respectively. The signal yield is currently blinded as the analysis has to undergo the LHCb review process after which the result will be unblinded. Finally, all the ingredients of the $R(D^*)$ ratio are combined and its value is computed in Section 7.3.

7.1 The signal yield fit

The signal yield is determined from a 3-dimensional maximum likelihood binned fit to q^2 (8 bins), lifetime of the τ^+-candidate (8 bins), t_τ, and the anti-D_s^+ BDT (6 bins).

There are 16 templates, 13 of them come from MC and three from data. The latter ones are combinatorial B, D^0 and D^* events. The templates are grouped into the 12 following categories, due to similar shapes:

- $B^0 \to D^{*-}\tau^+\nu_\tau$ – the signal; includes $\tau^+ \to 3\pi^-\overline{\nu}_\tau$ and $\tau^+ \to 3\pi^\pm\pi^0\overline{\nu}_\tau$
- $B^0 \to D^{**-}\tau^+\nu_\tau$ – excited D^{*-} states
- $B^0 \to D^{*-}D_s^{+(*)}$ – includes $B^0 \to D^{*-}D_s^+$, $B^0 \to D^{*-}D_s^+$, $B^0 \to D^{*-}D_{s0}^{*+}$ and $B^0 \to D^{*-}D_{s1}^{\prime+}$
- $B \to D^{**-}D_s^+X$
- $B_s^0 \to D^{*-}D_s^+X$
- $B \to D^{*-}D^+X$
- $B \to D^{*-}3\pi^\pm X$

- $B \to D^{*-} D^0 X$ SV, 'Same Vertex' where all 3 pions come from D^0
- $B \to D^{*-} D^0 X$ DV, 'Different Vertices' where at least 1 of the 3 pions does not come from the D^0 vertex, *e.g.* the slow pion from D^{*-} is reconstructed as coming from the D^0
- combinatorial B^0 whose template is made from the collision data with the $D^{*\pm}$ of the same sign as the $3\pi^{\pm}$ system (*i.e.* wrong sign data w.r.t. the signal)
- combinatorial D^0
- combinatorial D^{*-} but genuine D^0

The templates that come from MC are extracted with the MC-truth requirements.

7.1.1 The fit model

The model used to fit the data is the sum of 16 components, listed in the first column in Table 7.1. The shape of each component is modelled using histogram templates created from simulation samples and from the collision data samples in 3 cases. The parameters of the fit control the relative normalisation of each component in combinations listed in the second column of Table 7.1.

The meaning of each parameter is as follows:

- N_{sig} is the number of $B^0 \to D^{*-} \tau^+ \nu_\tau$ events, which is used as input when calculating $\mathcal{K}(D^{*-})$.
- $f_{\tau^+ \to 3\pi^{\pm} \overline{\nu}_\tau}$ is the fraction of $\tau^+ \to 3\pi^{\pm} \overline{\nu}_\tau$ decays relative to the sum of $\tau^+ \to 3\pi^{\pm} \overline{\nu}_\tau$ and $\tau^+ \to 3\pi^{\pm} \pi^0 \overline{\nu}_\tau$.
- $f_{D^{**} \tau \nu}$ is the amount of $B \to D^{**} \tau^+ \nu_\tau$ decays relative to $B^0 \to D^{*-} \tau^+ \nu_\tau$ decays.
- $N_{D^0}^{\text{same}}$ is the number of $B \to D^{*-} D^0 X$ events where all pions in the $3\pi^{\pm}$ system originate from the D^0 vertex.
- $f_{D^0}^{v_1 v_2}$ is the ratio of $B \to D^{*-} D^0 X$ decays where at least one pion originates from the D^0 vertex and the other pion(s) from a different vertex, normalised to $N_{D^0}^{\text{same}}$. This is the case when the soft pion from a D^{*-} decay is reconstructed as if it was produced at the 3π vertex.
- f_{D^+} is the ratio of $B \to D^{*-} D^+ X$ decays with respect to those containing a D_s^+ meson.
- N_{D_s} is the yield of events involving a D_s^+. The parameters $f_{D_s^+}$, $f_{D_{s0}^{*+}}$, $f_{D_{s1}^{\prime +}}$, $f_{D^{**} D_s X}$, $f_{B_s \to D^* D_s^+ X}$, obtained in Sec. 6.2, are used as constraints, after correcting for efficiency differences between the $B \to D^{*-} D_s^+ X$ and signal selections.
- $N_{B \to D^* 3\pi X}$ is the yield of $B \to D^{*-} 3\pi X$ events where the three pions come from the B vertex. This value is constrained by using the observed ratio between $B^0 \to D^{*-} 3\pi$ exclusive and $B \to D^{*-} 3\pi X$ inclusive decays, corrected for efficiency.

Table 7.1 – List of components in the signal yield extraction fit and how they are normalised.

Component	Normalisation
$B^0 \to D^{*-}\tau^+\nu_\tau$ ($\tau^+ \to 3\pi^\pm\overline{\nu}_\tau$)	$N_{\text{sig}} \times f_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau}$
$B^0 \to D^{*-}\tau^+\nu_\tau$ ($\tau^+ \to 3\pi^\pm\pi^0\overline{\nu}_\tau$)	$N_{\text{sig}} \times (1 - f_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau})$
$B \to D^{**}\tau^+\nu_\tau$	$N_{\text{sig}} \times f_{D^{**}\tau\nu}$
$B \to D^{*-}D^0X$ same vertex (SV)	$N^{\text{same}}_{D^0}$
$B \to D^{*-}D^0X$ different vertex (DV)	$N^{\text{same}}_{D^0} \times f^{v_1-v_2}_{D^0}$
$B \to D^{*-}D^+X$	$N_{D_s^+} \times f_{D^+}$
$B^0 \to D^{*-}D_s^+$	$N_{D_s^+} \times f_{D_s^+}/k$
$B^0 \to D^{*-}D_s^{*+}$	$N_{D_s^+} \times 1/k$
$B^0 \to D^{*-}D_{s0}^{*+}$	$N_{D_s^+} \times f_{D_{s0}^{*+}}/k$
$B^0 \to D^{*-}D_{s1}^{\prime+}$	$N_{D_s^+} \times f_{D_{s1}^{\prime+}}/k$
$B \to D^{**-}D_s^+X$	$N_{D_s^+} \times f_{D^{**}D_sX}/k$
$B_s^0 \to D^{*-}D_s^+X$	$N_{D_s^+} \times f_{B_s\to D^*D_s^+X}/k$
$B \to D^{*-}3\pi^\pm X$	$N_{B\to D^{*-}3\pi^\pm X}$
Combinatoric B	$N_{B_1-B_2}$
Combinatoric D^0	$N_{\text{fake } D^0}$
Combinatoric D^*	$N_{\text{fake } D^*}$

- $N_{B_1-B_2}$ is the yield of combinatorial background events where the D^{*-} and the 3π system come from different B decays. Its yield, scaling from MC, is found negligible, hence it is fixed to zero. Nonetheless, a more precise, data-driven estimation is being developed.
- $N_{\text{fake}D^0}$ is the combinatorial background yield with a fake D^0. Its value is fixed from a fit to $m(K^-\pi^+)$ and $m(D^{*-}) - m(K^-\pi^+)$
- $N_{\text{fake}D^*}$ is the combinatorial background yield with a fake D^{*-}. Its value is fixed from a fit to $m(K^-\pi^+)$ and $m(D^{*-}) - m(K^-\pi^+)$

The total probability density function, for N_{total} events total, is therefore:

$$
\begin{aligned}
\mathcal{P}_{\text{total}}(q^2, t_\tau, \text{BDT}) = 1/N_{\text{total}} \times \{ \\
N_{\text{sig}}[& f_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau}\mathcal{P}_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau} \\
& + (1 - f_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau})\mathcal{P}_{\tau^+\to\pi^+\pi^-\pi^+\pi^0\overline{\nu}_\tau} \\
& + f_{D^{**}\tau\nu}\mathcal{P}_{B\to D^{**}\tau^+\nu_\tau}] \\
+N_{D^0}^{\text{same}}[& \mathcal{P}_{B\to D^{*-}D^0X \text{ SV}} \\
& + f_{D^0}^{v_1-v_2}\mathcal{P}_{B\to D^{*-}D^0X \text{ DV}}] \\
+N_{D_s^+}/k \times [& \mathcal{P}_{B^0\to D^{*-}D_s^{*+}} \\
& + f_{D_s^+}\mathcal{P}_{B^0\to D^{*-}D_s^+} \\
& + f_{D_{s0}^{*+}}\mathcal{P}_{B^0\to D^{*-}D_{s0}^{*+}} \\
& + f_{D_{s1}^{\prime+}}\mathcal{P}_{B^0\to D^{*-}D_{s1}^{\prime+}} \\
& + f_{D^{**}D_sX}\mathcal{P}_{B\to D^{**-}D_s^+X} \\
& + f_{B_s\to D^*D_s^+X}\mathcal{P}_{B_s^0\to D^{*-}D_s^+X}] \\
+N_{D_s^+} & f_{D^+}\mathcal{P}_{B\to D^{*-}D^+X} \\
+N_{B\to D^{*-}3\pi^\pm X} & \mathcal{P}_{B\to D^{*-}3\pi^\pm X} \\
+N_{B_1-B_2} & \mathcal{P}_{\text{combinatoric } B} \\
+N_{\text{fake } D^0} & \mathcal{P}_{\text{combinatoric } D^0} \\
+N_{\text{fake } D^*} & \mathcal{P}_{\text{combinatoric } D^{*-}}\},
\end{aligned}
$$

where each $\mathcal{P}_i = \mathcal{P}_i(q^2, t_\tau, \text{BDT})$ is one of the 15 histogram templates and $k = 1 + f_{D_s^+} + f_{D_{s0}^{*+}} + f_{D_{s1}^{\prime+}} + f_{D^{**}D_sX} + f_{B_s\to D^*D_s^+X}$.

The 16 fit parameters can be grouped into three categories depending whether they are free, Gaussian-constrained or fixed in the fit:

- **free parameters (5)**: N_{sig}, $N_{D_s^+}$,$N_{B\to D^{*-}3\pi^\pm X}$, f_{D^+} and $f_{D^0}^{v_1-v_2}$
- **Gaussian-constrained parameters (5)**: $f_{D_s^+}$, $f_{D_{s0}^{*+}}$, $f_{D_{s1}^{\prime+}}$, $f_{D^{**}D_sX}$, $f_{B_s\to D^*D_s^+X}$
- **fixed parameters (6)**: $N_{B_1-B_2}$, $N_{\text{fake } D^0}$, $N_{\text{fake } D^*}$, $N_{D^0}^{\text{same}}$, $f_{D^{**}\tau\nu}$ and $f_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau}$

7.1.1.1 The Gaussian-constrained parameters

Most of the Gaussian-constrained parameters are constrained by Gaussian distributions of μ and σ corresponding to the mean and uncertainty as obtained in the $D^*D_s^+X$ fit (see Section 6.2). An exception is made for $N_{B\to D^{*-}3\pi^\pm X}$, whose central value is approximated by the relative yield of such events in the inclusive $b\bar{b}\to D^{*-}3\pi^\pm X$ MC sample, normalised to the total number of events in the fitted sample that equals 783 events. Since the relative yield of the 'prompt' events in the collision data sample may differ, a tentative standard deviation of 200 events is chosen in the Gaussian constraint. The normalisation of $N_{B\to D^{*-}3\pi^\pm X}$ is a work-in-progress in the analysis.

The yield of prompt events is estimated from MC with a data-driven correction. The prompt events can be from the exclusive $B^0 \to D^{*-}\ 3\pi^\pm$ or inclusive $B^0 \to D^{*-}\ 3\pi^\pm\ X$

modes. The former are straightforward to pinpoint as they form the peak in $m(D^{*-}3\pi^{\pm})$ around the mass of the B^0. The latter, are hard to distinguish from other backgrounds and it is them that remain after the full selection (due to the cut $m(D^{*-}3\pi^{\pm}) < 5100\,\mathrm{MeV}/c^2$). Selecting truth-matched prompt events from the inclusive $b\bar{b} \to D^{*-}\ 3\pi^{\pm}\ X$ enables to estimate the ratio of inclusive to exclusive prompt events after the full selection (except for the $m(D^{*-}3\pi^{\pm})$ cut) as $r_{\text{MC}} = 6.06$. Multiplying this result by the exclusive yield in the data sample, $N_{\text{prompt, peak}} = 182$ events, yields 673 expected prompt events in the fit region. However, r_{MC}, might not model well the data sample. To tackle this issue, the derived result needs to be corrected by the factor $c = r_{\text{data}}/r_{\text{MC}}$. To select a pure prompt sample from data, one may apply a negative requirement on the $3\pi^{\pm}$ vertex detachment, *e.g.* $[\text{vtx}_z(\tau^+) - \text{vtx}_z(B^0)]/\text{error} < -2$. Because the positive vertex detachment is already applied in the workflow, dedicated samples are made with the negative vertex detachment requirement, detachment BDT skipped and with all other signal selection cuts applied. This brings the correction factor of 1.72, resulting in the final estimate of the prompt yield as $N_{\text{prompt, fit}} = c \times r_{\text{MC}} \times N_{\text{prompt, peak}} \approx 1894$. An arbitrary uncertainty of ± 200 events is assigned as 1σ of the Gaussian constraint on the prompt yield in the fit.

7.1.1.2 The fixed parameters

The combinatoric B is estimated from the wrong-sign (WS) sample. First, full selection is applied to the WS sample, except for the combinatorial and isolation BDT's, and the $m(D^{*-}3\pi^{\pm})$. The BDT's skipped would reject the combinatorial B events. Then, $\text{acos}(\text{DIRA}(B^0, \text{PV})) > 0.03$ and reversed isolation requirement, isolation BDT < 0 are applied to obtain a pure combinatorial B sample. The ratio of the selected events in the fitted $m(D^{*-}3\pi^{\pm})$ region, *i.e.* $m(D^{*-}3\pi^{\pm}) < 5100\,\mathrm{MeV}/c^2$ to those in high $m(D^{*-}3\pi^{\pm})$ region, $m(D^{*-}3\pi^{\pm}) > 5100\,\mathrm{MeV}/c^2$ is measured. Finally, this ratio is multiplied by the yield of events in the fitted (right-sign) data sample in the high $m(D^{*-}3\pi^{\pm})$ region, requiring $\text{acos}(\text{DIRA}(B^0, \text{PV})) > 0.03$, thus obtaining an estimated yield of combinatorial B events in the fit region, $N_{B_1-B_2} = 237$.

The combinatoric D^0 and D^{*-} yields are obtained from a two-dimensional unbinned maximum-likelihood fit to the reconstructed mass of the D^0, $m(K^-\pi^+)$, and the difference of the reconstructed masses of the D^{*-} and D^0, $\Delta M \equiv m(D^{*-}) - m(K^-\pi^+)$. Their distributions are shown in Figure 7.1.

The combinatorial D^0 events do not peak in either of the fit variables. To ensure no overlap of the two categories, only such combinatorial D^{*-} events are chosen which correspond to the genuine D^0, *i.e.* where the $K^-\pi^+$ system comes from the D^0, while the pion from D^{*-} and D^0 do not come from the same D^{*-}. These events peak in $m(K^-\pi^+)$, but not in ΔM, separating the two categories.

The signal, defined as genuine D^0 and D^{*-} candidates, is modelled in $m(K^-\pi^+)$ with a sum of the Crystal Ball and Gaussian, whereas in ΔM, with a sum of two Crystal Balls (with left- and right-hand tails) and Gaussian. The combinatorial D^0 is modelled with an exponential in $m(K^-\pi^+)$ and a dedicated RooFit function, `RooDstD0Bkg`, in ΔM. The combinatorial D^{*-} (but genuine D^0) is modelled by the signal PDF in $m(K^-\pi^+)$

and like the combinatorial D^0 in ΔM. In addition, in the fitted sample there are misidentified D^{*-} events, whose shape is obtained from MC samples by requiring MC-truth criteria. They are modelled similarly to the signal.

Figure 7.2 shows the fit projections on the two fitted variables and the pull distributions, where the agreement, mostly within 3σ and occasionally within around 4σ, is observed. The constructed templates are used in the signal fit, normalised to the obtained yields of fake D^0, $N_{\text{fake } D^0} = 466$ and fake D^{*-}, $N_{\text{fake } D^{*-}} = 547$.

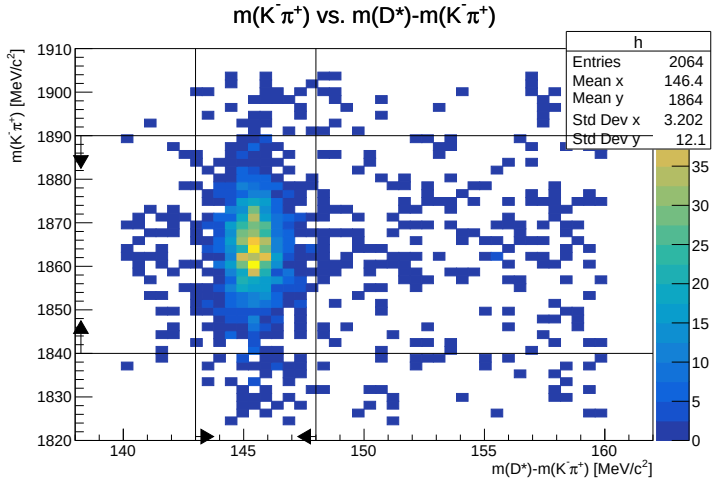


Figure 7.1 – $m(K^-\pi^+)$ as a function of $m(D^{*-}) - m(K^-\pi^+)$ for the data sample after the full selection except the cuts on these variables, which are marked by arrows.

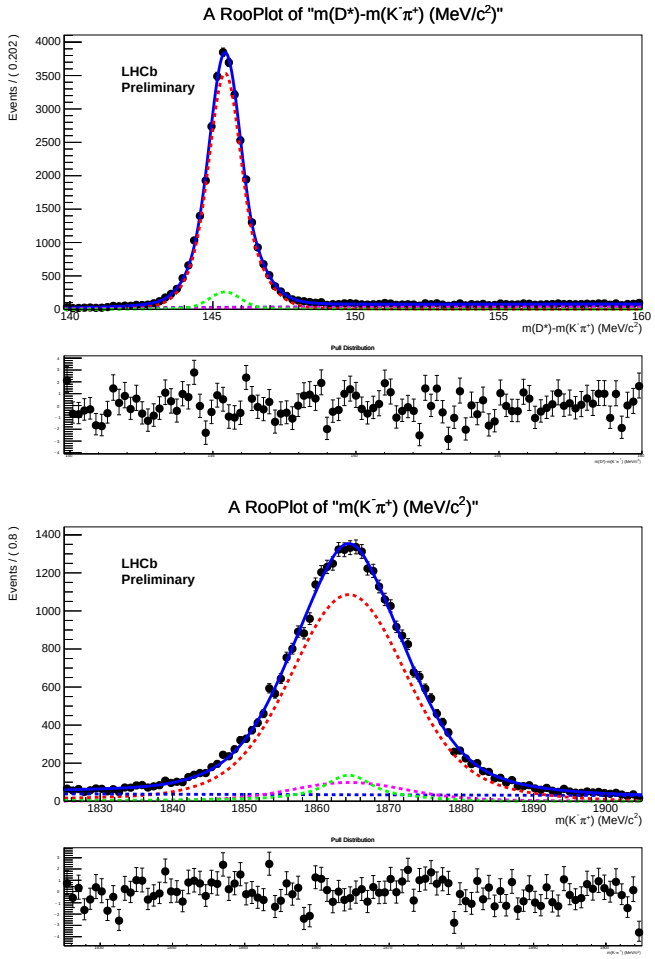


Figure 7.2 – Projections of the two-dimensional fit to the $m(K^-\pi^+)$ (left) and $\delta M \equiv m(D^{*-}) - m(K^-\pi^+)$ (right) for the collision data sample after the full selection. Signal (red), combinatorial D^0 (blue), combinatorial D^{*-} (magenta), mis-identified D^{*-} (green). See text for details.

Number of D^0 events decaying into three pions, $N^{\text{same}}_{D^0}$, is obtained by scaling the truth-matched yield from simulation by the ratio of yields of $D^0 \to K^-\pi^+\pi^-\pi^+$ decays in data and MC respectively as shown in Figure 7.3. The yields are extracted from fitting a Gaussian for signal and an exponential for the combinatorial backgrounds. Thus obtained value equals $N^{\text{same}}_{D^0} = 967 \pm 31$.

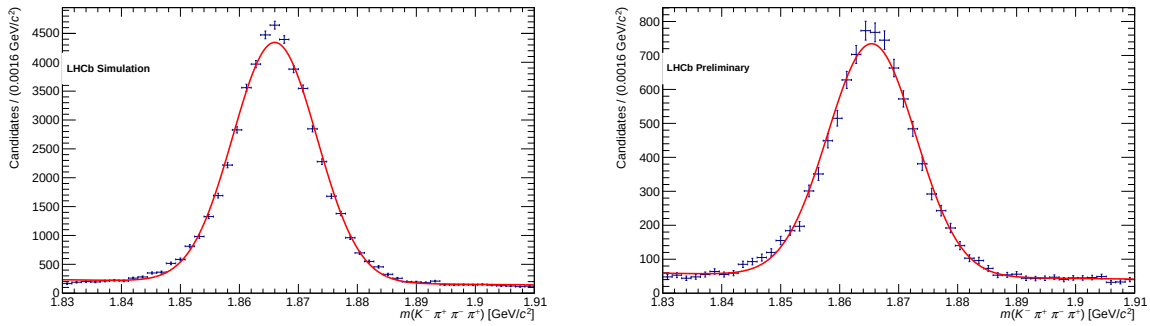


Figure 7.3 – Fit to $m(K^-\pi^+\pi^-\pi^+)$ in MC (left) and data (right) around the mass of D^0. The D^0 yields are used to scale the MC sample in order to deduce the $N_{D^0}^{\text{same}}$ parameter.

The fraction of $B^0 \to D^{**}\tau\nu$ events relative to the signal ones is temporarily fixed as $f_{D^{**}\tau\nu} = 11\%$, as a midpoint of the theoretically estimated range $[5.5, 16.5]\%$.

The relative abundance of the $\tau^+ \to 3\pi^\pm\overline{\nu}_\tau$ to the sum of $\tau^+ \to 3\pi^\pm\overline{\nu}_\tau$ and $\tau^+ \to 3\pi^\pm\pi^0\overline{\nu}_\tau$ is computed as: $f_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau} = \frac{\varepsilon_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau}\mathcal{B}(\tau^+\to 3\pi^\pm\overline{\nu}_\tau)}{\varepsilon_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau}\mathcal{B}(\tau^+\to 3\pi^\pm\overline{\nu}_\tau)+\varepsilon_{\tau^+\to 3\pi^\pm\pi^0\overline{\nu}_\tau}\mathcal{B}(\tau^+\to 3\pi^\pm\pi^0\overline{\nu}_\tau)} = 0.7742 \pm 0.0035$, with the efficiencies of the two signal modes, $\varepsilon_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau}$ and $\varepsilon_{\tau^+\to 3\pi^\pm\pi^0\overline{\nu}_\tau}$, are reported in Section 5.11.

7.1.1.3 The blinding scheme

The signal yield shown, N_{sig}, is blinded at this stage of analysis in order not to bias the analysts by the preliminary results. To verify the fit quality and to study systematics, we chose to blind it by multiplying the signal yield, N_{sig}, by a random number, x, from the Gaussian distribution of $\mu = 1.0$ and $\sigma = 0.3$ constrained to the range $[0.1, 1.9]$. To preserve the correct normalisation of the components parametrised by N_{sig}, what is actually passed to the fit is N_{sig}/x, i.e. the actual signal yield, while the blinded N_{sig} is reported by the fitting program [1]. Consequently, both N_{sig} and its uncertainty are multiplied by the same factor, hence the relative uncertainty, $\sigma_{N_{\text{sig}}}/N_{\text{sig}}$, remains unblinded and can be monitored.

At this point, even though N_{sig} is blinded, one might deduce it from the known dataset size, after subtracting all the background yields. Therefore, the difference between the blinded and actual yields of the three N_{sig}-dependent components, $(N_{\text{sig}} - N_{\text{sig}}/x)(1 + f_{D^{**}\tau\nu})$, is split evenly among the backgrounds, skipping those whose yields are fixed before the fit is performed.

To summarise, the fitting procedure uses unchanged probability density function and only the reported values of floating parameters are altered.

[1] The fit adjusts only the blinded N_{sig}, since x is a constant. Because only the floating parameters are reported during the fit, x remains undisclosed, rendering the actual signal yield N_{sig}/x blinded.

7.1.2 Fit results and projections

The results of a preliminary fit are reported in Table 7.2 and projected in Figure 7.5. The fit quality expressed in χ^2 per number of degrees of freedom is $\chi^2/374 = 1.20$. The agreement of the fit model and data may be inspected also in the flattened distribution shown in Figure 7.6, where the content of each 3-dimensional bin is compared against the fitted model.

Figure 7.4 shows the correlation coefficients between the fitted parameters. The N_{D_s} parameters exhibits correlations above 50% with $f_{B_s\to D^* D_s X}$, $f_{D^{**} D_s X}$, f_{D^+} and $f_{D^0}^{v_1-v_2}$. This may point to a difficulty in separating the relevant categories from one another in the fit.

The relative statistical uncertainty of the signal yield, $N_{\rm sig}$, equals $\frac{\sigma_{N_{\rm sig}}}{N_{\rm sig}} = 5.91\%$. To compare it with the Run1 result, the parameters f_{D^+} and $f_{D^0}^{v_1-v_2}$ need to be fixed to their central values obtained from the nominal fit (as it was the case in Run1, where the uncertainties of these two parameters were incorporated in the systematic uncertainty). Consequently, the resulting relative statistical uncertainty drops to 5.42%.

This value is lower from the Run1 value of 6.64% [1, 2]. In Run2, with, approximately factor 1.5 more data [2] and improved signal efficiency one would expect a value below 6%, as it is indeed the case.

[2] Taking into account approximately twice the $b\bar{b}$ cross-section in Run2 than Run1, relative luminosities, $\int \mathcal{L}(\text{Run2})/\int \mathcal{L}(\text{Run1}) = 2\,\text{fb}^{-1}/3.2\,\text{fb}^{-1}$, and improved Run2 trigger efficiency by around 15%, one would expect roughly factor $2 \times \frac{2}{3.2} \times 1.15 = 1.5$ more signal events in Run2 (2015-16).

Table 7.2 – The signal yield fit to the 2015-2016 data sample. The signal yield, N_{sig}, and other floating parameters are blinded (see text).

Parameter	Fit result
free	
N_{D_s}	18447 ± 348
N_{sig}	1997 ± 118
f_{D^+}	0.13 ± 0.01
$f_{D^0}^{v_1-v_2}$	3.5 ± 0.3
constrained	
$N_{B\to D^{*-}3\pi^{\pm}X}$	2323 ± 175
$f_{B_s\to D^* D_s X}$	0.06 ± 0.02
$f_{D_{s1}'^+}$	0.39 ± 0.03
f_{D_s}	0.59 ± 0.02
$f_{D_{s0}^{*+}}$	0.09 ± 0.03
$f_{D^{**}D_s X}$	0.39 ± 0.03
fixed	
$N_{B_1-B_2}$	226
$N_{D^0}^{\text{same}}$	971
$N_{\text{fake}D^0}$	458
$N_{\text{fake}D^*}$	562
$f_{D^{**}\tau\nu}$	0.11
$f_{\tau^+\to 3\pi^{\pm}\overline{\nu}_\tau}$	0.77

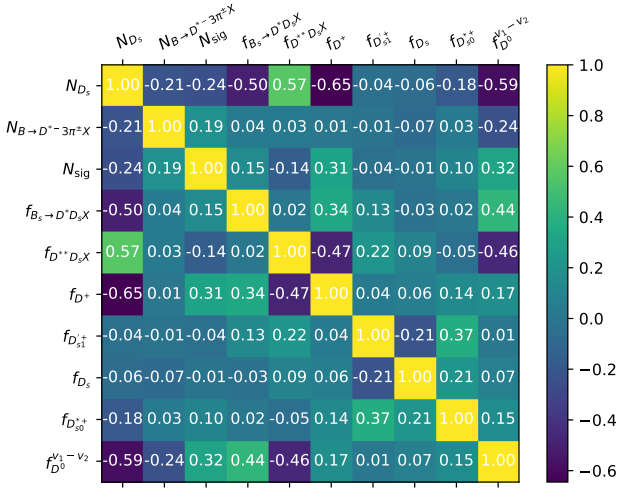


Figure 7.4 – Correlation matrix of the signal yield fit parameters for the 2015-2016 dataset. See Table E.2 for a more precise format.

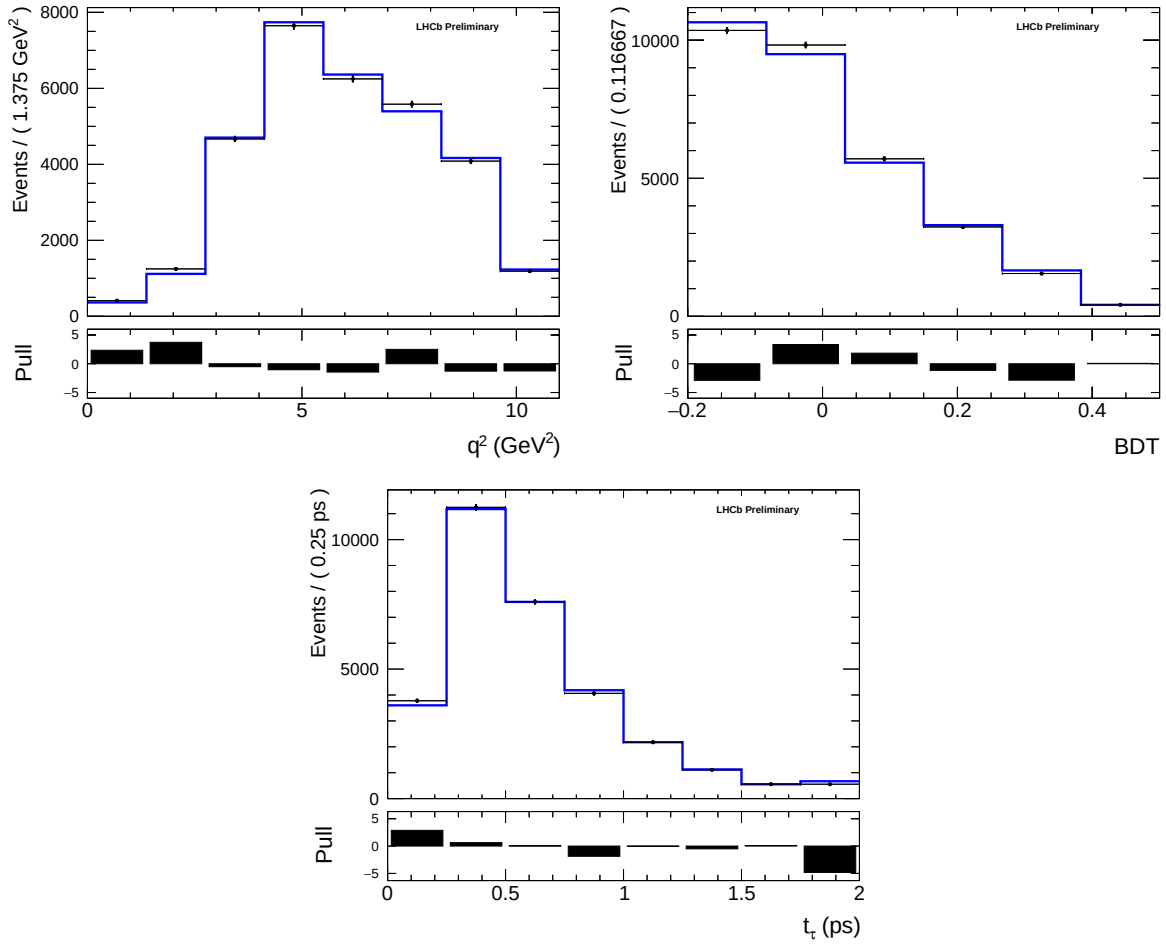


Figure 7.5 – The signal fit projections on the fit variables to the 2015-2016 data sample. The components are not drawn to prevent unblinding. The total PDF is unblinded.

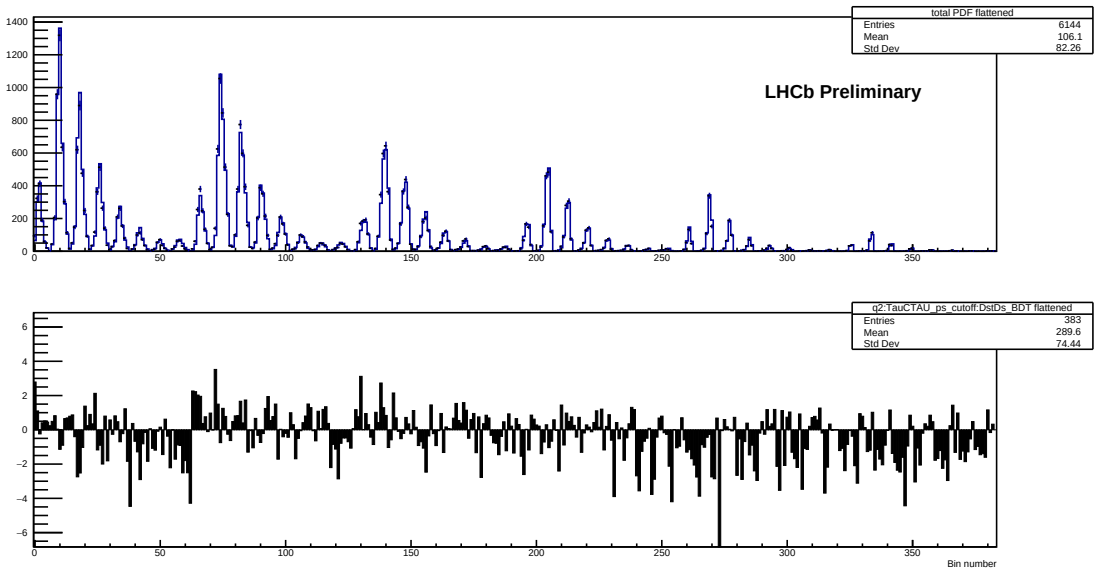


Figure 7.6 – The signal fit flattened projection across its 256 bins for the 2015-2016 data sample (top) and its pull distribution (bottom). The $x-$axis represents an arbitrary bin index.

The subsequent Table 7.3 and plots, Figures 7.7 7.8 and 7.9, show the similar distributions for the fit with fixed f_{D^+} and $f_{D^0}^{v_1-v_2}$. The fit quality expressed in χ^2 per number of degrees of freedom is $\chi^2/376 = 1.20$. Interestingly, the previously seen correlations disappear and the only correlation above 50% is between N_{D_s} and $N_{B\to D^{*-}3\pi^{\pm}X}$ and equals -56%.

Table 7.3 – The signal yield fit with fixed f_{D^+} and $f_{D^0}^{v_1-v_2}$ to the 2015-2016 data sample. The signal yield, N_{sig}, and other floating parameters are blinded (see text).

Parameter	Fit result
free	
N_{sig}	1997 ± 108
N_{D_s}	18446 ± 205
constrained	
$N_{B\to D^{*-}3\pi^{\pm}X}$	2324 ± 170
$f_{B_s\to D^* D_s X}$	0.06 ± 0.02
$f_{D_{s1}'^+}$	0.39 ± 0.03
f_{D_s}	0.59 ± 0.02
$f_{D_{s0}^{*+}}$	0.09 ± 0.03
$f_{D^{**}D_s X}$	0.39 ± 0.02
fixed	
$N_{B_1-B_2}$	226
$N_{D^0}^{\text{same}}$	971
$N_{\text{fake}D^0}$	458
$N_{\text{fake}D^*}$	562
$f_{D^{**}\tau\nu}$	0.11
$f_{\tau^+\to 3\pi^{\pm}\overline{\nu}_\tau}$	0.77
f_{D^+}	0.13 ± 0.01
$f_{D^0}^{v_1-v_2}$	3.5 ± 0.3

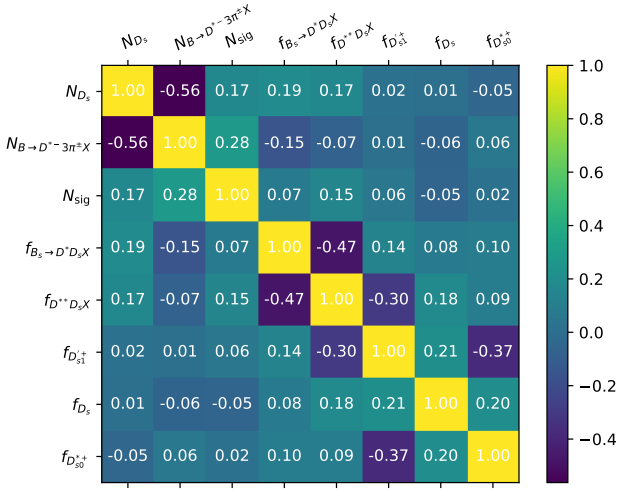


Figure 7.7 – Correlation matrix of the signal yield fit parameters with fixed f_{D^+} and $f_{D^0}^{v_1-v_2}$ for the 2015-2016 dataset. See Table E.2 for a more precise format.

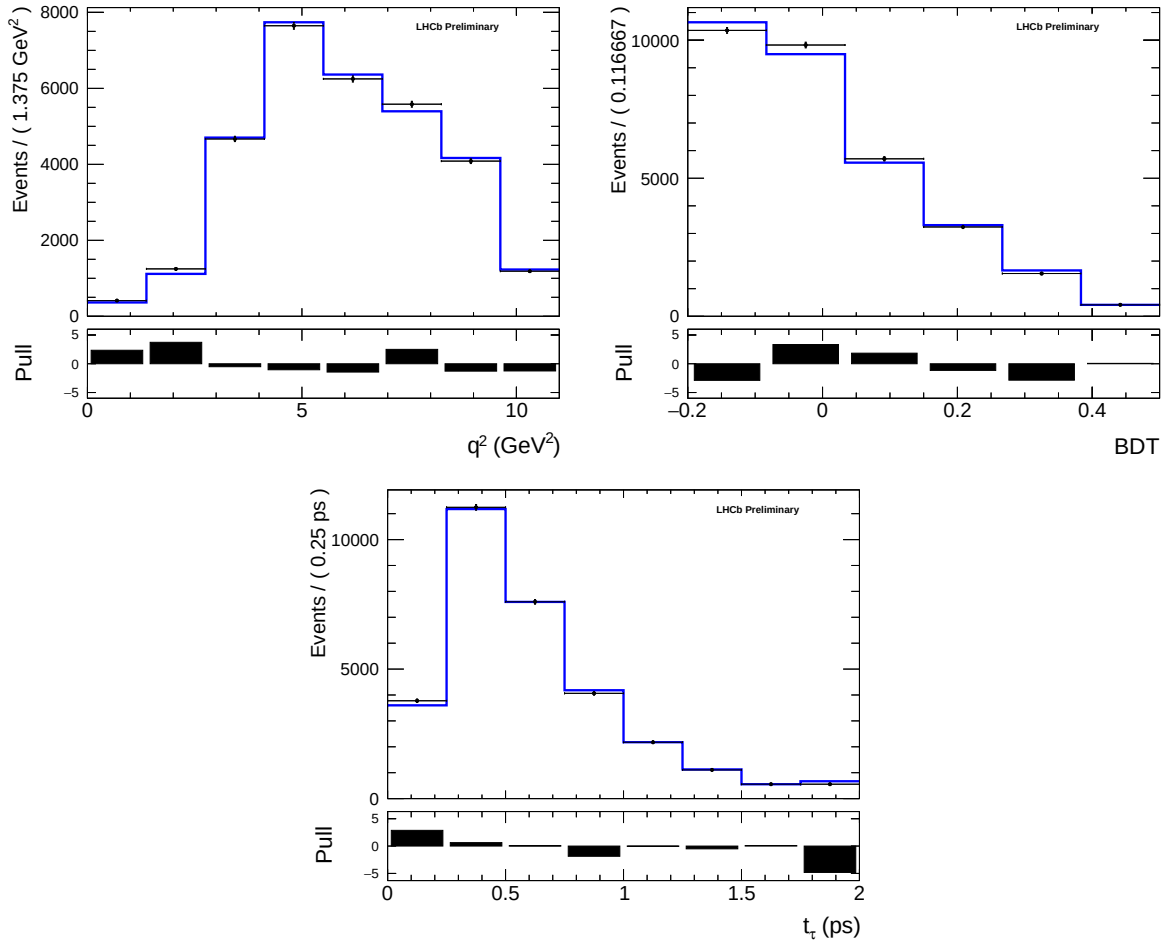


Figure 7.8 – The signal fit projections on the fit variables with fixed f_{D^+} and $f_{D^0}^{v_1-v_2}$ to the 2015-2016 data sample. The components are not drawn to prevent unblinding. The total PDF is unblinded.

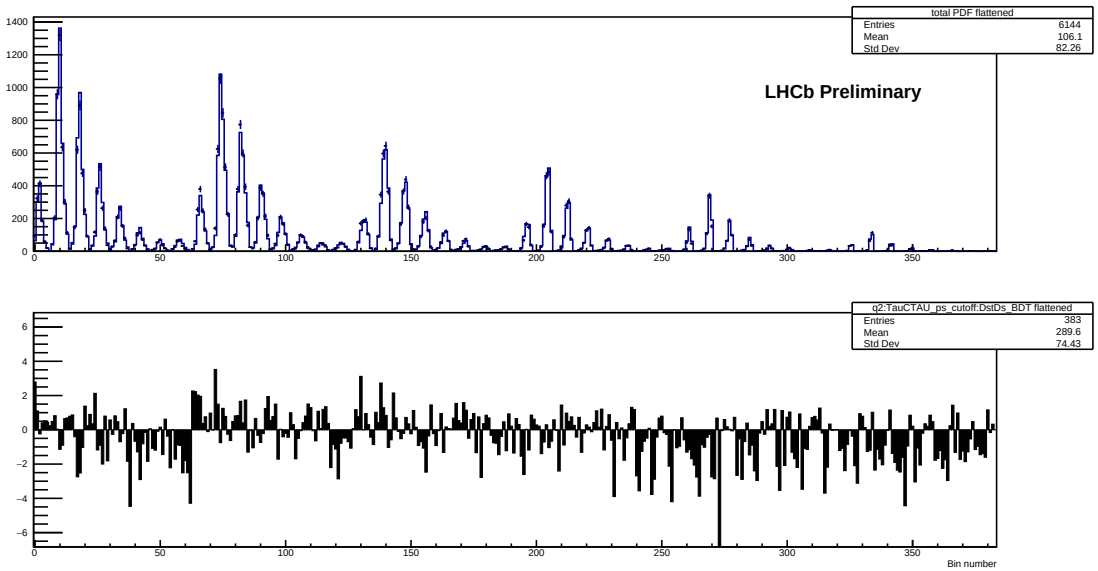


Figure 7.9 – The signal fit flattened projection across its 256 bins for the 2015-2016 data sample (top) and its pull distribution (bottom). The parameters f_{D^+} and $f_{D^0}^{v_1-v_2}$ are fixed. The $x-$axis represents an arbitrary bin index.

7.1.3 The Monte Carlo pseudo-experiment study

The fit is validated using simulated pseudo-experiments, referred to as toys. After the nominal fit is performed to the data sample of N_{tot} events, $N = 1000$ pseudo-experiments are made. In each the Poisson(N_{tot}) events are generated from the nominal fit model and the fit is performed, yielding $N_{\text{sig toy}}$ and other parameters.

Figure 7.10 shows the pull distribution, defined as $\text{pull}(N_{\text{sig}}) = \frac{N_{\text{sig toy}} - N_{\text{sig}}}{\sigma_{N_{\text{sig}}}}$.

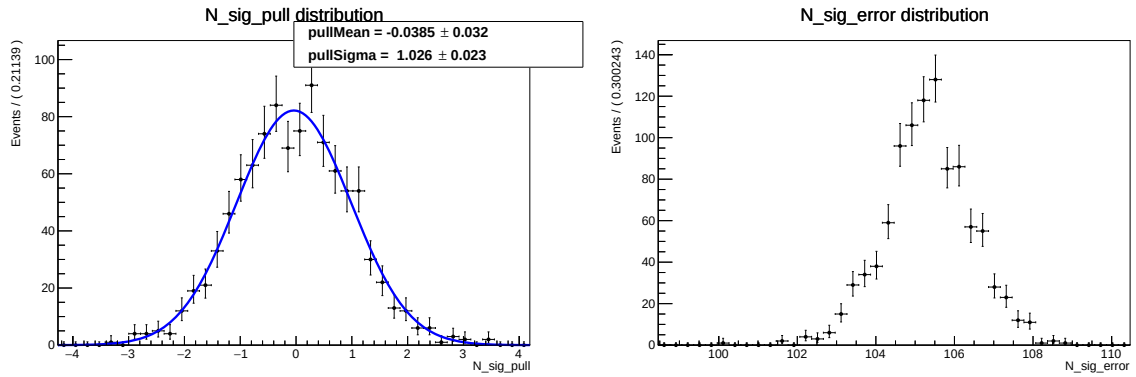


Figure 7.10 – Pull distribution, $\frac{N_{\text{sig toy}} - N_{\text{sig}}}{\sigma_{N_{\text{sig}}}}$ (right), from a toy study with 1000 generated events (left) and the uncertainty, $\sigma_{N_{\text{sig}}}$, distribution of N_{sig}.

The pull mean and standard deviation are compatible with zero and unity, respectively. Therefore, there is no evidence of bias in the signal yield nor in its uncertainty, as obtained from the fit.

Distributions of the other fit parameters are shown in Appendix F, where all parameters exhibit pull distributions of mean and standard deviation compatible with zero and unity, respectively.

Table 7.4 confirms the correlation coefficients of the nominal signal fit, whereby only N_{D_s} with $N_{B\to D^{*-}3\pi^{\pm}X}$exhibit absolute correlation above 50%, specifically -57%.

Table 7.4 – Correlation coefficients of the signal fit variables in the 2015-2016 collision MC pseudo-experiment fit.

	N_{D_s}	$N_{B\to D^{*-}3\pi^{\pm}X}$	N_{sig}	$f_{B_s\to D^* D_s X}$	$f_{D^{**} D_s X}$	$f_{D_{s1}'^+}$	f_{D_s}	$f_{D_{s0}^{*+}}$
N_{D_s}	1.00	−0.57	−0.15	0.20	0.17	0.02	0.02	−0.04
$N_{B\to D^{*-}3\pi^{\pm}X}$	−0.57	1.00	−0.30	−0.17	−0.07	0.02	−0.05	0.05
N_{sig}	−0.15	−0.30	1.00	−0.07	−0.15	−0.06	0.05	−0.02
$f_{B_s\to D^* D_s X}$	0.20	−0.17	−0.07	1.00	−0.49	0.14	0.07	0.09
$f_{D^{**} D_s X}$	0.17	−0.07	−0.15	−0.49	1.00	−0.30	0.18	0.10
$f_{D_{s1}'^+}$	0.02	0.02	−0.06	0.14	−0.30	1.00	0.23	−0.37
f_{D_s}	0.02	−0.05	0.05	0.07	0.18	0.23	1.00	0.22
$f_{D_{s0}^{*+}}$	−0.04	0.05	−0.02	0.09	0.10	−0.37	0.22	1.00

7.1.4 The fit to an MC sample

To verify the fitting procedure a similar fit is performed to a simulated inclusive $b\bar{b} \to D^{*-}3\pi^{\pm}X$ sample. This fit is still very preliminary.

The $b\bar{b} \to D^{*-}3\pi^{\pm}X$ is split into two approximately equal in size datasets, based on the LHCb magnet polarity [3]. One half is used to extract the templates, most of which come from the $b\bar{b} \to D^{*-}3\pi^{\pm}X$ MC. The other half is used as a pseudo-data sample, to independently evaluate the fit quality. The fitted MC sample is factor 2.45 larger than the collision data, whose fit is described in the previous section.

The fixed parameters are extracted from the MC-truth values in the fitted sample. In particular, $f_{D^{**}\tau\nu}$ is found to be around 6% in the simulation, while this parameter is fixed to 11% in the data sample fit.

The fit results are reported in Table 7.5, compared against the truth-matched parameters, and projected in Figure 7.12. The fit quality expressed in χ^2 per number of degrees of freedom is $\chi^2/374 = 1.28$. The agreement of the fit model and data may be analysed also in the flattened distribution shown in Figure 7.13, where the content of each 3-dimensional bin is compared against the fitted model.

The signal yield, $N_{\rm sig}$, is found to be compatible within 2 standard deviations (σ) with the MC-truth value. The relative signal yield uncertainty is $\frac{\sigma_{N_{\rm sig}}}{N_{\rm sig}} = 3.6\%$, significantly smaller than in the fit to the collision data, despite a worse fit quality as expressed in $\chi^2/{\rm ndf}$. However, $N_{D_s^+}$, N_{D^+}, $f_{D_{s1}'^+}$, and $f_{D^{**}D_sX}$ are, respectively, 18σ, 17σ, 21σ and 6σ different with respect to the MC-truth values. These large discrepancies might not be correct, since they arise from truth-matching criteria used to select the templates, some of which come from dedicated MC samples. In case of the inclusive MC sample, more stringent requirements need to be used, which is a subject of ongoing improvements.

Figure 7.11 shows the correlation coefficients between the fitted parameters. The highest correlation is between $N_{D_s^+}$ and f_{D^+} and equals -66%. The correlations above 50% are found between the same parameters as in the data fit and pseudo-experiment MC study, confirming the effect.

[3] At LHCb approximately half of the Run2 *pp* collisions took place with either magnet down or up polarity. The impact of the polarity on the measured observables is negligible for this exercise.

Table 7.5 – The unblinded signal yield fit to the 2015-2016 MC. The parameters f_{D^+} and $f_{D^0}^{v_1-v_2}$ have been fixed to their values from a previous fit to MC.

Parameter	Fit result	MC-truth	rel. difference [σ]
	free		
N_{D_s}	21647 ± 332	15505	18
N_{sig}	4720 ± 172	4085	4
f_{D^+}	0.30 ± 0.01	0.47	17
$f_{D^0}^{v_1-v_2}$	2.3 ± 0.2	1.8	3
	constrained		
$N_{B\to D^{*-}3\pi^\pm X}$	1512 ± 176	2319	5
$f_{B_s\to D^* D_s X}$	0.09 ± 0.02	0.21	6
$f_{D_{s1}'^+}$	0.53 ± 0.02	0.10	21
f_{D_s}	0.49 ± 0.02	0.49	0.09
$f_{D_{s0}^{*+}}$	0.07 ± 0.03	0.08	0.4
$f_{D^{**}D_s X}$	0.33 ± 0.03	0.15	6
	fixed		
$N_{B_1-B_2}$	278		
$N_{D^0}^{\text{same}}$	1617		
$N_{\text{fake}D^0}$	30		
$N_{\text{fake}D^*}$	64		
$f_{D^{**}\tau\nu}$	0.060		
$f_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau}$	0.78		

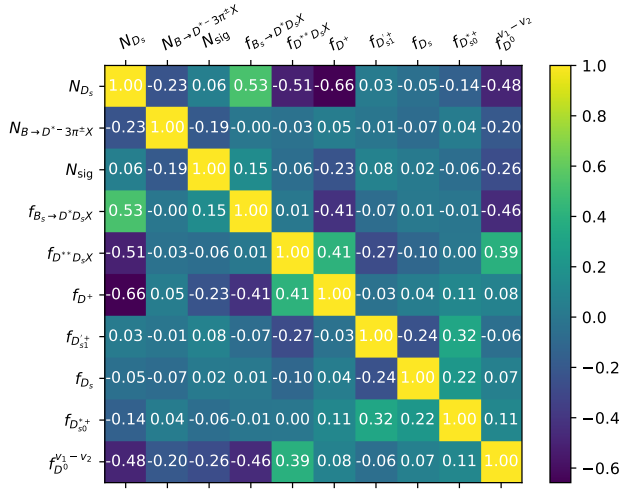


Figure 7.11 – Correlation matrix of the signal yield fit parameters for the 2015-2016 MC. See Table E.3 for a more precise format.

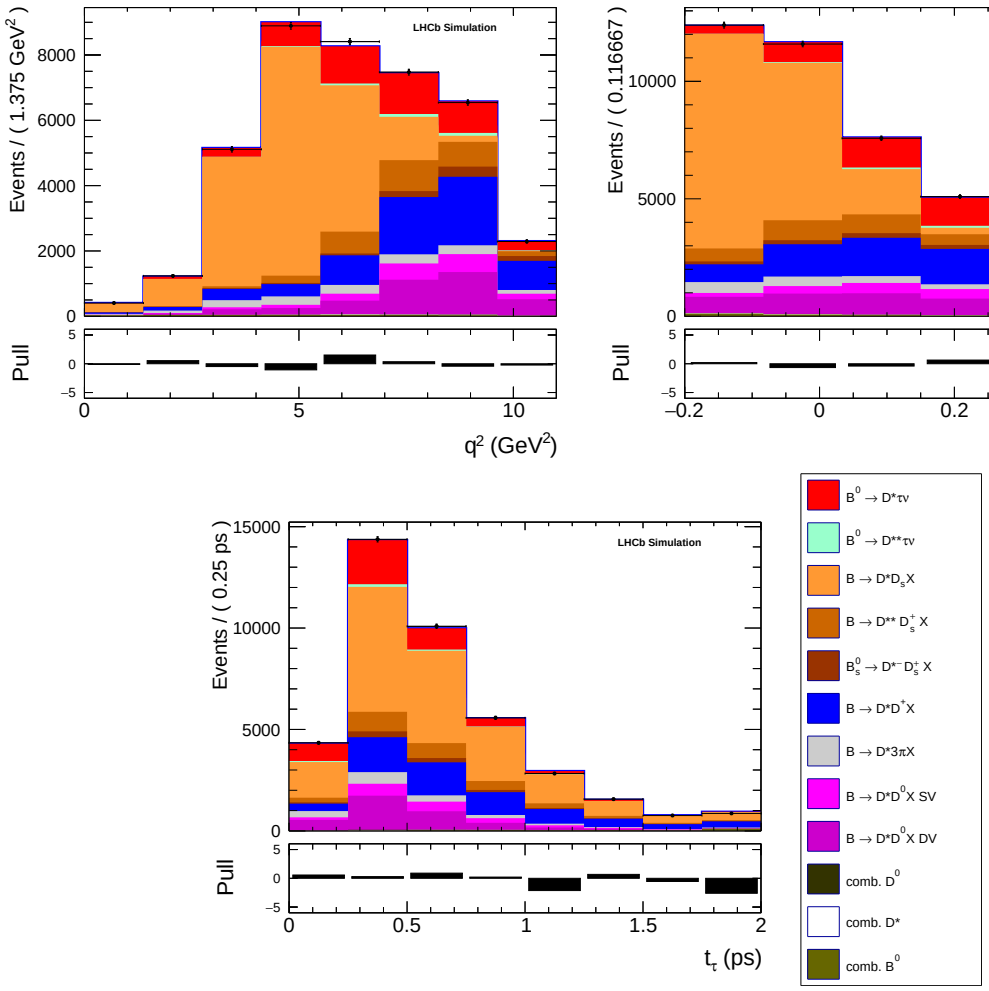


Figure 7.12 – The signal fit projections on the fit variables to the inclusive $b\bar{b} \to D^{*-}3\pi^{\pm}X$ MC sample.

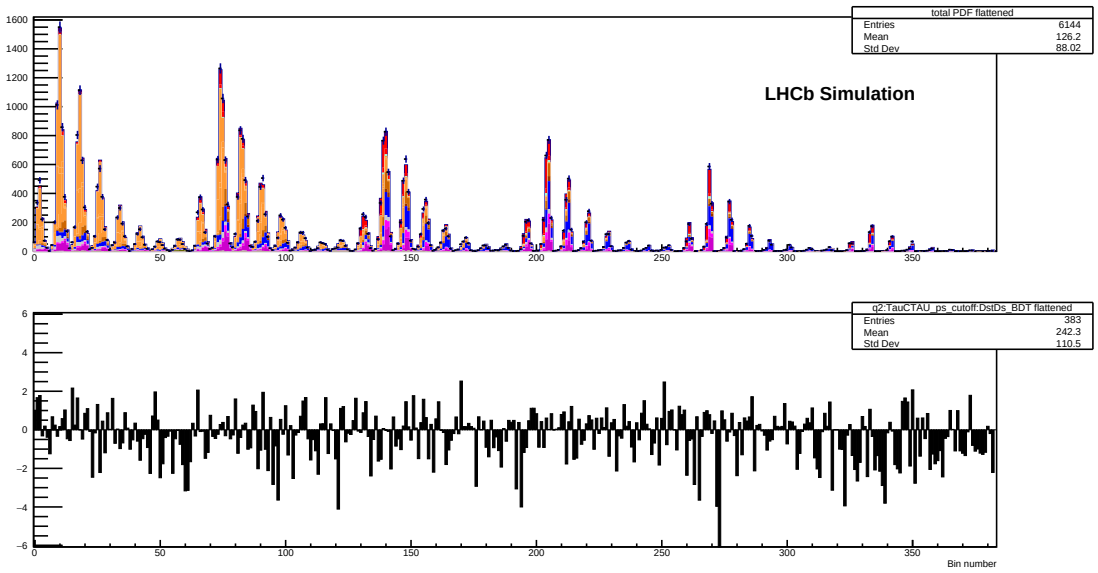


Figure 7.13 – The signal fit flattened projection across its 256 bins for the MC sample (top) and its pull distribution (bottom). The $x-$axis represents an arbitrary bin index.

7.2 The normalisation yield fit

The normalisation data sample is extracted as explained in Section 5.10. Its yield is obtained from an unbinned maximum likelihood fit to the mass of the $D^{*-}\pi^+\pi^-\pi^+$ system as shown in Figure 7.14 (left). Its signal is modelled as a sum of the Gaussian and the Crystal Ball of the same mean values, whereas the combinatorial background is modelled with an exponential. The $D^{*-}\pi^+\pi^-\pi^+$ yield obtained in the fit equals 29999 ± 201.

As shown in Figure 7.14, some events come from the $D_s^+ \to 3\pi^\pm$ decay.

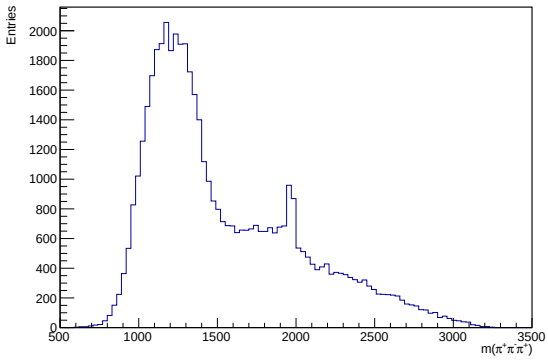


Figure 7.14 – Distribution of $m(\pi^+\pi^-\pi^+)$ in the fitted normalisation data sample.

The result is corrected by subtracting the $D_s^+ \to \pi^+\pi^-\pi^+$ decays in the mass range $1.8 < m(3\pi^\pm) < 2.1\,\text{GeV}/c^2$ fitted with double Gaussian as signal and an exponential to model the combinatorial pions, as shown in Figure 7.15 (right). The $D_s^+ \to \pi^+\pi^-\pi^+$ yield obtained in the fit equals 426 ± 34.

Finally, the normalisation yield is taken as 29572 ± 167.

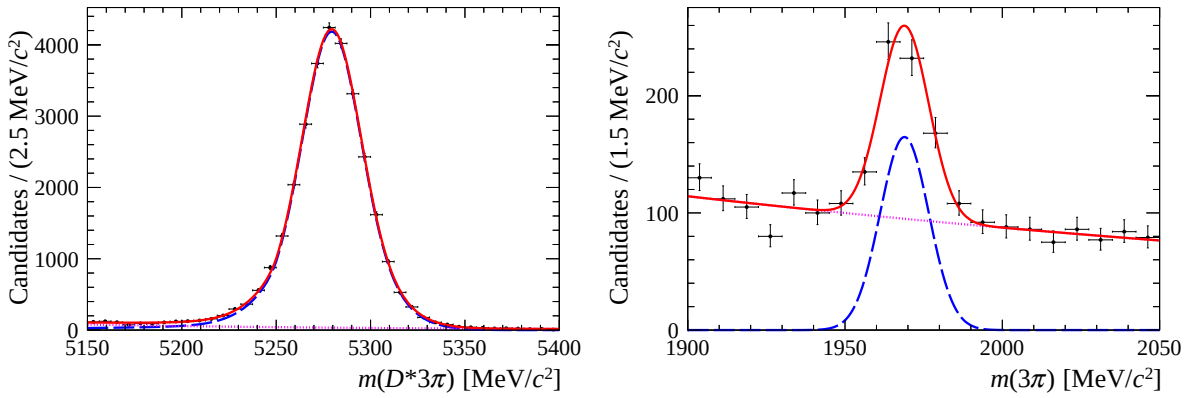


Figure 7.15 – $m(D^{*-}\pi^+\pi^-\pi^+)$ fit (left) and $m(\pi^+\pi^-\pi^+)$ fit (right) to extract the normalisation yield.

7.3 $R(D^*)$ determination

The external branching fractions used for $\mathcal{K}(D^*)$ and $R(D^*)$ are reported in Table 7.6. They depend on the $R(D^*)$ central value and introduce 0.09 absolute uncertainty to it.

Using the obtained signal and normalisation yields, the $\mathcal{K}(D^*)$ is:

$$\mathcal{K}(D^*) = \frac{N_{\text{sig}}}{N_{\text{norm}}} \frac{\varepsilon_{\text{norm}}}{\varepsilon_{\text{sig}}} \frac{1}{\mathcal{B}(\tau^+ \to 3\pi\overline{\nu}_\tau) + \mathcal{B}(\tau^+ \to 3\pi\pi^0\overline{\nu}_\tau)} = 1.389 \pm 0.079. \tag{7.1}$$

Finally, $R(D^*)$ is obtained as:

$$R(D^*) = \mathcal{K}(D^{*-}) \frac{\mathcal{B}(B^0 \to D^{*-}\pi^+\pi^-\pi^+)}{\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)} = 0.198 \pm 0.011\,(\text{stat}) \pm 0.009(\text{ext}), \tag{7.2}$$

where the first uncertainty is statistical and the second one stems from the uncertainty of the external branching fractions.

Table 7.6 – The external branching fractions used for the $\mathcal{K}(D^{*-})$ and $R(D^*)$ computation.

Decay	$\mathcal{B}$	Source
$\tau^- \to \pi^+\pi^-\pi^+\nu_\tau$ (excluding $K^0_S \to \pi^+\pi^-$)	$(9.02 \pm 0.05)\%$	PDG [14]
$\tau^- \to \pi^+\pi^-\pi^+\pi^0\nu_\tau$ (excluding $K^0_S \to \pi^+\pi^-$)	$(4.49 \pm 0.05)\%$	PDG [14]
$B^0 \to D^{*-}\ell\nu_\tau$, where $\ell = e, \mu$	$(5.06 \pm 0.02\,(\text{stat}) \pm 0.12\,(\text{syst}))\%$	HFLAV [30]
$B^0 \to D^{*-}\pi^+\pi^-\pi^+$	$(7.21 \pm 0.29) \times 10^{-3}$	PDG [14]

The Run1 result as published in Ref. [1, 2] is $R(D^*) = 0.291 \pm 0.019\,(\text{stat}) \pm 0.026\,(\text{syst}) \pm 0.013$(ext). However, HFLAV has since reported an updated average of the $B^0 \to D^{*-}\ell\nu_\tau$ branching fraction, which shifts the $R(D^*)$ measurement down to $R(D^*) = 0.281 \pm 0.019\,(\text{stat}) \pm 0.026\,(\text{syst}) \pm 0.013$(ext).

The following chapter discusses the systematic uncertainties to be considered in addition to the statistical and external ones shown above.

Chapter 8

Systematic uncertainties

In this chapter the systematic uncertainties of the $R(D^*)$ result are discussed. This part of the analysis is currently work-in-progress. Therefore, the already measured uncertainties are quoted, whereas a work plan for the remaining ones is laid out. Some of the individual systematics are subject to future updates and they are sensitive to the signal fit being improved. Nonetheless, the methods introduced should remain applicable.

Sections 8.1 and 8.2 describe the systematic uncertainties due to the signal and background modelling, respectively. Section 8.3 focuses on those that stem from the selection. Section 8.4 describes the uncertainty due to the limited size of the templates used in the signal fit.

Finally, all systematic uncertainties are combined in Section 8.5.

8.1 Signal model uncertainties

The fraction of tau decays to 3 pions is derived, as discussed in Section 7.1.1.2, as $f_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau} = \frac{\varepsilon_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau}\mathcal{B}(\tau^+\to 3\pi^\pm\overline{\nu}_\tau)}{\varepsilon_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau}\mathcal{B}(\tau^+\to 3\pi^\pm\overline{\nu}_\tau)+\varepsilon_{\tau^+\to 3\pi^\pm\pi^0\overline{\nu}_\tau}\mathcal{B}(\tau^+\to 3\pi^\pm\pi^0\overline{\nu}_\tau)} = 0.7742 \pm 0.0035$. In the signal fit, the central value is taken as a fixed parameter. An alternative fit is performed with this ratio fixed to the central value ± 1 standard deviation, resulting in $\sigma(f_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau}) = ^{+0.3\%}_{-0.27\%}$ relative uncertainty.

Another source of uncertainties comes from the knowledge of the form factors in $B^0 \to D^{*-}\tau^+\nu_\tau$. Pseudo-experiments should be made varying the form factor parameters $R_0(1)$, $R_1(1)$, $R_2(1)$ and ρ^2.

Other τ decays may contribute to the signal, especially those with three or five charged tracks in the final state, *e.g.* $K^+\pi^-\pi^+$, $K^+K^-\pi^+$, $\pi^+\pi^-\pi^+\pi^0\pi^0$. They all have very small branching fractions relative to the signal modes. We expect the systematic assigned in the Run1 analysis, $\sigma_{\text{other }\tau\text{ dec.}} = 1\%$, is still applicable in Run2.

The feeddown from $B \to D^{**}\tau\nu_\tau$ is fixed in the final fit as $f_{D^{**}\tau\nu} = 11\%$ as a midpoint of the theoretically predicted range 5.5%-16.5%. The upper limit is known from preliminary experimental results from the ongoing $R(D^{**})$ analysis [112] as 13.3%. Two alternative fits are performed for the fixed $B \to D^{**}\tau\nu_\tau$ fractions to 5.5% and 13.3%. The relative differences with respect to the nominal value of 11% are taken as a systematic of $\sigma_{f_{D^{**}\tau\nu}} = ^{+5.95\%}_{-2.18\%}$.

8.2 Background model uncertainties

The background model uncertainties stem from modelling of: the D_s^+ decay, double-charm backgrounds and the combinatorial backgrounds.

The D_s^+ decay model uncertainties can be estimated by varying the weights of the D_s^+ contributions according to the Gaussian distribution taking into account their uncertainties. (Only central values are used in the template weights before the signal fit is performed).

8.3 Uncertainties due to the selection

The particle Identity (PID) variables ProbNNpi and ProbNNk are used to select the pions from D^* and τ^+ and to reject K^- misidentified as π^- in the $\tau^+ \to \pi^+\pi^-\pi^+$ system. Their efficiencies for the given cut values are factors in the signal and normalisation efficiencies. In the end, their ratio impacts the $\mathcal{K}(D^*)$ result.

The PID efficiencies are estimated from a data-driven method using the LHCb package PIDCalib [96]. It uses a *calibration* sample of pions to measure their PID efficiencies at the cut values required. The efficiency is parametrised by the phase-space region of a given pion track. Hence efficiency calibration histograms are made. Next, the MC samples of signal or normalisation, with all selection applied except the PID itself, are used. The efficiency of each simulated pion is assigned, based on the phase-space bin it populates, to the corresponding value from the calibration histogram. The total efficiency for an event is computed as the product of the efficiencies for the tracks considered.

This approach contributes to the systematic uncertainties in at least two ways:

- The granularity of the binning scheme
- Assumption that the PID efficiency is fully parametrisable by the binning variables (*e.g.* track quality is ignored, but might impact the efficiency)

8.4 Template samples size

The dominant systematic is due to the limited size of the templates used in the signal fit. The signal fit templates, as discussed in the previous Chapter 7, are built mostly from the MC samples. Therefore, their size, after the complete selection, impacts the signal yield in the final fit, N_{sig}. The relative uncertainty on N_{sig} directly translates to the one of $R(D^*)$.

A usual procedure to estimate this uncertainty relies on bootstrapping the templates and performing the signal fit to every thus obtained templates set. The bootstrapping in most analyses comprises resampling with replacement individual events. In this project, however, for all simulated templates full ReDecay blocks ideally would be resampled, as discussed already in Section 4.2. Nonetheless, due to negligible bin-to-bin correlations in ReDecay, event-based bootstrapping is expected to be a very good approximation of the systematic uncertainty due to the templates size, and is used in the following.

The templates bootstrapping procedure can be summarised as:

1. Generate the nominal templates from MC or data (for combinatorial backgrounds)
2. Resample with replacement events from the templates, resulting in new ones of size Poisson(n), where n is the original template size
3. Fit the signal yield N_{sig}
4. Repeat the steps 2-3 N times and obtain the mean signal yield $\overline{N_{\text{sig}}}$ and its deviation $\sigma_{\text{bootstrap}}$
5. Infer the relative uncertainty due to the templates size as $\sigma_{\text{templ. stat.}} = \frac{\sigma_{\text{bootstrap}}}{\overline{N_{\text{sig}}}}$

Thanks to the multiplicative blinding scheme, both N_{sig} and its statistical uncertainty σ are multiplied by the same factor, which vanishes in the relative uncertainty ratio σ/N_{sig}.

Figure 8.1 shows the resulting distribution of N_{sig}. The resulting uncertainty due to the limited size of the templates is, therefore, $\sigma_{\text{templ. stat.}} = 40.65/2047 = 1.99\%$. This is a significant decrease with respect to the Run1 $R(D^*)$ hadronic measurement, where such uncertainty was equal 4.1%.

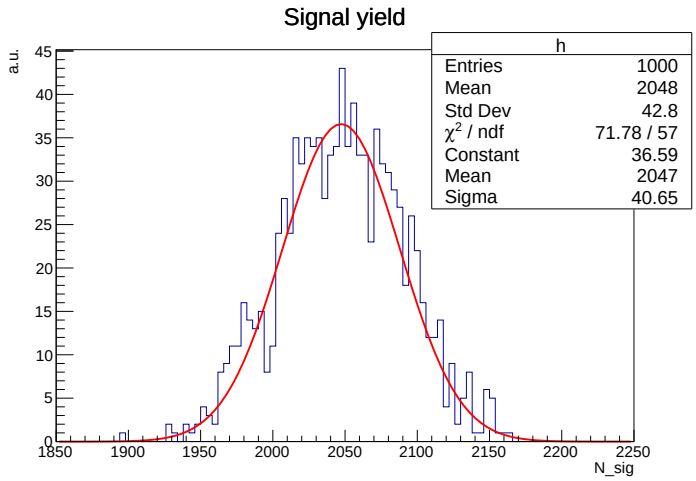


Figure 8.1 – Distribution of number of signal events, N_{sig}, for 1000 fits performed to bootstrapped templates.

8.5 Summary of systematic uncertainties

To summarise, a break-down of individual systematic uncertainties is reported in Table 8.1, where the current total systematic uncertainty, computed by adding the individual ones in quadrature, is given. This total uncertainty is a preliminary lower boundary and is expected to increase in the future updates of the analysis. Currently the largest systematic is due to the knowledge of the feeddown from $D^{**}\tau\nu$ to the signal mode.

Table 8.1 – Preliminary summary of relative systematic uncertainties to the $R(D^*)$ ratio, expressed in [%]. See text for definitions of the contributions.

Quantity	systematic uncertainty [%]
$\sigma_{\text{templ. stat.}}$	2.0
fixing f_{D^+}	1.70
fixing $f_{D^0}^{v_1-v_2}$	1.86
$\sigma(f_{D^{**}\tau\nu})$	$^{+5.95}_{-2.18}$
$\sigma_{\text{other } \tau \text{ dec.}}$	≈ 1
$\sigma(f_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau})$	$^{+0.30}_{-0.27}$
Total	$^{+6.84\%}_{-4.02\%}$

Conclusions and prospects

The presented work describes the measurement of the $R(D^*)$ ratio using the 2015-16 LHCb proton-proton collision dataset. The τ^+ is reconstructed in the $\tau^+ \to \pi^+\pi^-\pi^-$ (π^0) $\overline{\nu}_\tau$ modes.

The techniques used in this analysis improved the ones used in the previous, 2011-2012, $R(D^*)$ analysis at LHCb [1, 2]. The most important differences are: using the ReDecay algorithm for fast simulation; implementing an MVA-based charged track isolation and adding an anti-combinatorial background MVA.

Despite this project being constrained to the 2015-2016 datasets, a seamless continuation with the 2017-2018 datasets (*i.e.* the remainder of the Run2) was ensured. This was achieved by good software practises and following the reproducible research guidelines such as using a version control system and automating the analysis workflow.

The analysis is in a very advanced state, with a preliminary signal fit implemented and the $R(D^*)$ ratio measured. Almost a complete documentation of the analysis is reported in this manuscript. Nonetheless, the signal fit needs to be improved and several systematic uncertainties have to be estimated. For this reason, the final result is blinded by a multiplicative factor. Therefore, it cannot be yet compared with the other $R(D^*)$ measurements and their world average. However, the relative uncertainties are not blinded. The statistical uncertainty on $R(D^*)$ is tentatively estimated as 5.6%.

The author contributed mostly to: validating the new fast simulation algorithm "ReDecay"; implementing the selection, especially by tackling the charged track isolation with an MVA and computing the efficiencies across all the steps; studying the particle identity uncertainties; computing a systematic uncertainty due to the limited template size; helping organise the aforementioned analysis workflow.

The analysis described in this thesis will undergo an LHCb review in the upcoming months and will be sent for a publication afterwards. The measurement contributes to the cutting edge tests of the Lepton Flavour Universality. The published result shall update the current $R(D^*)$ world average.

The planned efforts in resolving the Lepton Flavour Universality puzzle shall expand in three ways: more decay channels considered; new observables measured; more collision data analysed.

Examples of $R(X)$ ratios in other decay modes are, in $b \to c\ell\nu_\ell$: $R(D^+)$, $R(D^0)$, $R(D_s^{-(*)})$, $R(\Lambda_c^{(*)})$; and in $b \to u\tau\nu$: $R(\Lambda_b^0 \to p\tau\nu)$, $R(B \to pp\tau\nu)$.

Apart from the ratios themselves, their spectra as a function of q^2 and angular observables, such as the τ^+ polarisation, already measured by Belle [36, 37], will help constrain the spin-structure of the New Physics models [84, 85].

Last but not least, new datasets are going to be collected and analysed over the next few years and decades. The remainder of the LHCb Run2 will bring further 3.9 fb^{-1} collision data, followed by a major detector upgrade, making it possible for LHCb to record 4 fb^{-1} each year for a decade, starting in 2022. LHCb is expecting to improve the precision of LFU observables by a factor of 3-4 by 2025 and by an order of magnitude with a further 300 fb^{-1} envisaged after the Phase-II Upgrade to be started around 2030. The Belle II experiment should bring an unprecedented data sample of 50 ab^{-1}. The large statistics of the collected data will enable to substantially decrease the systematic uncertainties, by more stringent signal selection.

Apart from these experimental aspects, new lattice QCD computations of form factors should help precisely estimate the Standard Model expected values of the LFU observables.

The upcoming years should reveal *whether* the New Physics contributions in the flavour changing charged/neutral currents processes are genuine. In case of a positive answer, we shall also discover *which* are these New Physics enhancements.

Appendix A

PDG review

This is a summary of essential branching fractions, lifetimes, and masses, used throughout the thesis. All the values come from the Particle Data Group Review [14].

Table A.1 – Selected branching fractions of the particles in the $B^0 \to D^{*-}\tau^+\nu_\tau$ signal decay chain. Highlighted are the modes chosen in this analysis allowing for good vertices reconstruction.

Particle	Decay	$\mathcal{B}$[%]
$D^*(2010)^-$	$\boldsymbol{D^*(2010)^- \to \overline{D}^0\pi^-}$	67.7 ± 0.5
	$D^*(2010)^- \to D^-\pi^0$	30.7 ± 0.5
	$D^*(2010)^- \to D^-\gamma$	1.6 ± 0.4
$\overline{D}^0$	$\boldsymbol{\overline{D}^0 \to K^+\pi^-}$	3.950 ± 0.031
τ^-	$\tau^- \to \mu^-\overline{\nu}_\mu\nu_\tau$	17.39 ± 0.04
	$\tau^- \to e^-\overline{\nu}_e\nu_\tau$	17.82 ± 0.04
	$\tau^- \to \pi^-\pi^0\nu_\tau$	25.49 ± 0.09
	$\tau^- \to \pi^-\nu_\tau$	10.82 ± 0.05
	$\boldsymbol{\tau^- \to \pi^+\pi^-\pi^+\nu_\tau}$ (excluding $\boldsymbol{K^0_S \to \pi^+\pi^-}$)	9.02 ± 0.05
	$\boldsymbol{\tau^- \to \pi^+\pi^-\pi^+\pi^0\nu_\tau}$ (excluding $\boldsymbol{K^0_S \to \pi^+\pi^-}$)	4.49 ± 0.05

Table A.2 – Selected branching fractions of the particles relevant to the D_s^+ decay chain.

Particle	Decay	$\mathcal{B}$[%]
D_s^+	$\eta\ X$	$(29.9 \pm 2.8)\%$
	$\eta\pi^+$	$(1.68 \pm 0.10)\%$
	$\eta\rho^+$	$(8.9 \pm 0.8)\%$
	$\eta\pi^+\pi^0$	$(9.5 \pm 0.5)\%$
D_s^+	$\eta'\ X$	$(10.3 \pm 1.4)\%$
	$\eta'\pi^+$	$(3.94 \pm 0.25)\%$
	$\eta'\rho^+$	$(5.8 \pm 1.5)\%$
	$\eta'\pi^+\pi^0$	$(5.6 \pm 0.8)\%$
η'	$\rho^0\ \gamma$ (incl. non-resonant $\pi^+\pi^-$)	$(29.5 \pm 0.4)\%$
η'	$\pi^+\pi^-\ \eta$	$(42.5 \pm 0.5)\%$
η	$\pi^+\pi^-\pi^0$	$(22.92 \pm 0.28)\%$
η	$\pi^+\pi^-\ \gamma$	$(4.22 \pm 0.08)\%$
ρ^0	$\pi^+\pi^-$	100%
ρ^+	$\pi^+\pi^0$	100%

Table A.3 – Mean lifetimes of the selected particles relevant to this analysis.

Particle	Mean lifetime
B^0	$(1.519 \pm 0.004) \times 10^{-12}\,\mathrm{s}$
D_s^+	$(5.05 \pm 0.04) \times 10^{-13}\,\mathrm{s}$
D^+	$(1.040 \pm 0.007) \times 10^{-12}\,\mathrm{s}$
D^0	$(4.101 \pm 0.015) \times 10^{-13}\,\mathrm{s}$
τ^+	$(2.903 \pm 0.005) \times 10^{-13}\,\mathrm{s}$

Table A.4 – Invariant masses of the selected particles relevant to this analysis.

Particle	Invariant mass
B^0	$(5279.65 \pm 0.12)\,\mathrm{MeV}/c^2$
D^{*-}	$(2010.26 \pm 0.05)\,\mathrm{MeV}/c^2$
D_s^+	$(1968.34 \pm 0.07)\,\mathrm{MeV}/c^2$
D^+	$(1869.65 \pm 0.05)\,\mathrm{MeV}/c^2$
D^0	$(1864.83 \pm 0.05)\,\mathrm{MeV}/c^2$
τ^+	$(1776.86 \pm 0.12)\,\mathrm{MeV}/c^2$

Appendix B

Fast simulation

Adapted from D. Müller's thesis (CERN-THESIS-2017-257).
Bootstrapping is the valid strategy for statistical uncertainty estimation and is adapted for ReDecay as follows:

- Start with a sample of size n
- Make a pseudo-sample of n' random entries from the original sample
  - n' is drawn randomly from a Poisson distribution with mean n
- Non-independence of ReDecay events requires sampling whole 'blocks'
- Make many pseudo-samples and bin in histograms
- Take the mean n_i^{bs} and standard deviation σ_i^{bs} of each bin i across all bootstrapped histograms j to form the bootstrapped distribution with the following mean and standard deviation per bin, and the bin-to-bin correlations:

$$n_i^{\text{bs}} = \frac{1}{N}\sum_j^N n_i^j; \qquad \sigma_i^{\text{bs}} = \sqrt{\frac{1}{N}\sum_j^N \left(n_i^j - n_i^{\text{bs}}\right)^2} \tag{B.1}$$

$$\text{corr}_{k,l}^{\text{bs}} = \frac{1}{\sigma_k^{\text{bs}}\sigma_l^{\text{bs}}}\frac{1}{N}\sum_j^N \left(n_k^j - n_k^{\text{bs}}\right)\left(n_l^j - n_l^{\text{bs}}\right). \tag{B.2}$$

ReDecay validation distributions are shown in the following.

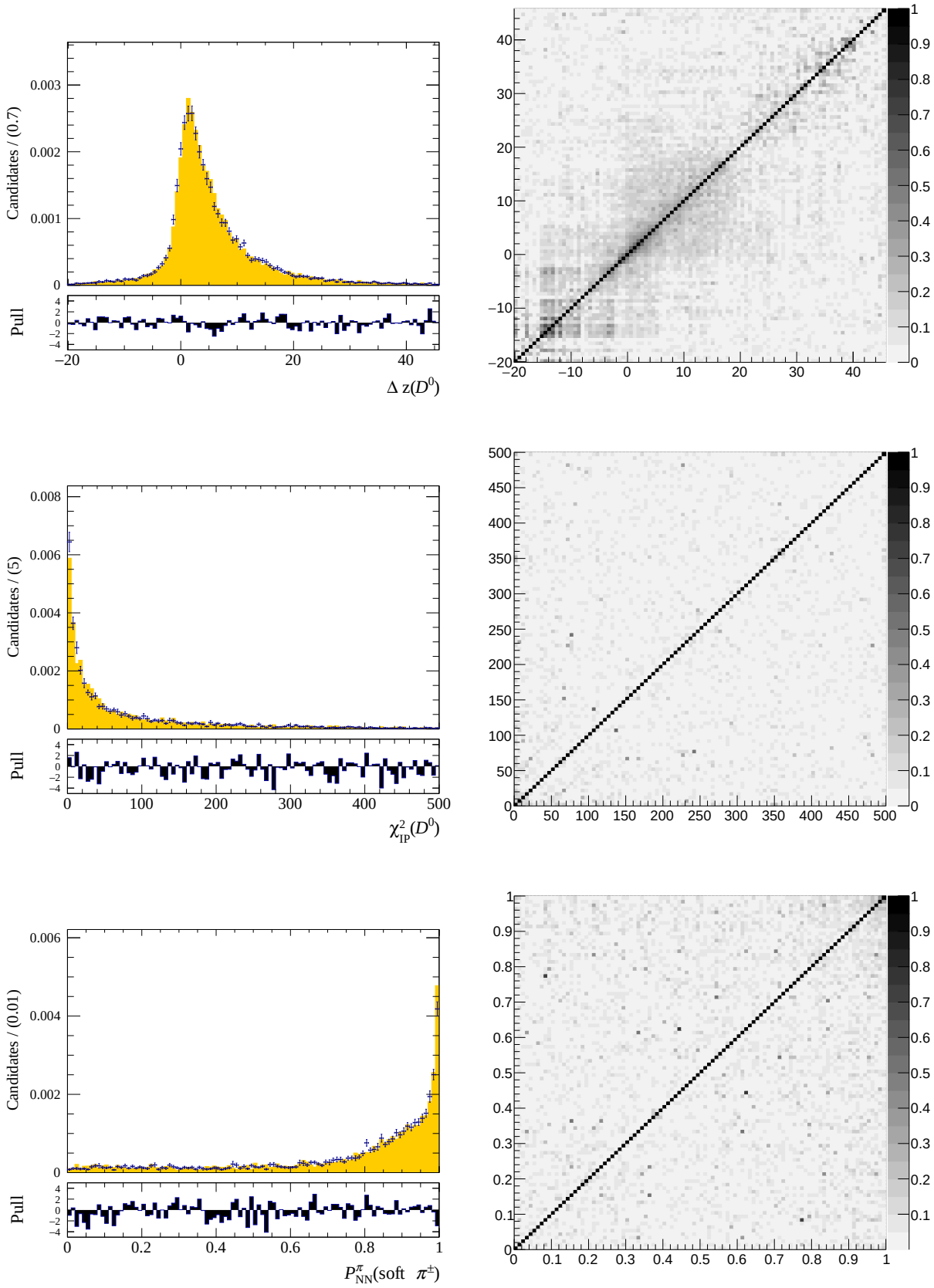


Figure B.1 – Left: distributions for the 28k $B^0 \to D^{*-}\tau^+\nu_\tau$ ($\tau^+ \to 3\pi^{\pm}\overline{\nu}_\tau$) mode events after stripping; full simulation is in yellow and ReDecay in bars with uncertainty. Right: Bin-to-bin correlation of $\eta(B^0)$.

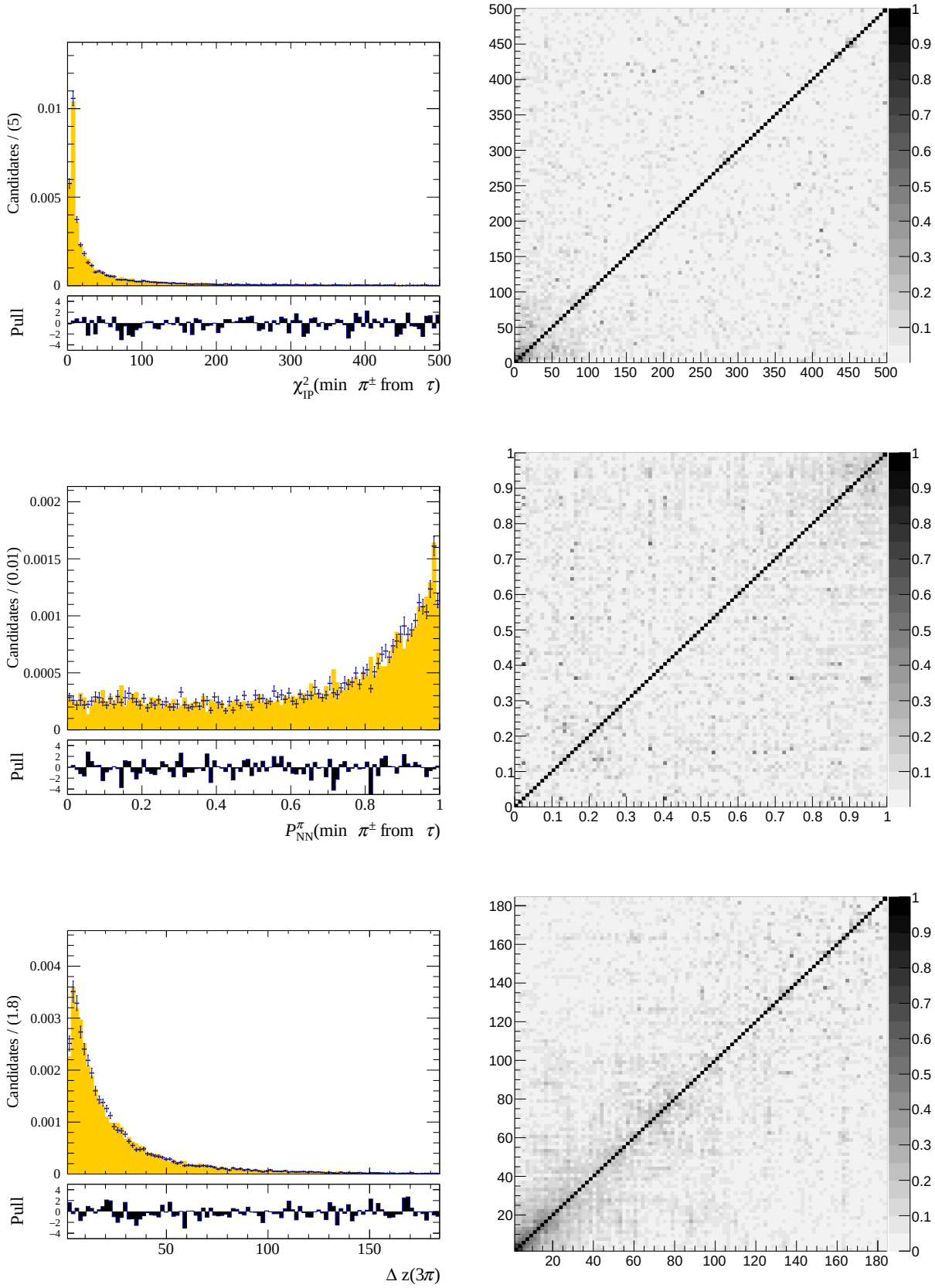


Figure B.2 – Left: distributions for the 28k $B^0 \to D^{*-}\tau^+\nu_\tau$ ($\tau^+ \to 3\pi^\pm\overline{\nu}_\tau$) mode events after stripping; full simulation is in yellow and ReDecay in bars with uncertainty. Right: Bin-to-bin correlation of $\eta(B^0)$.

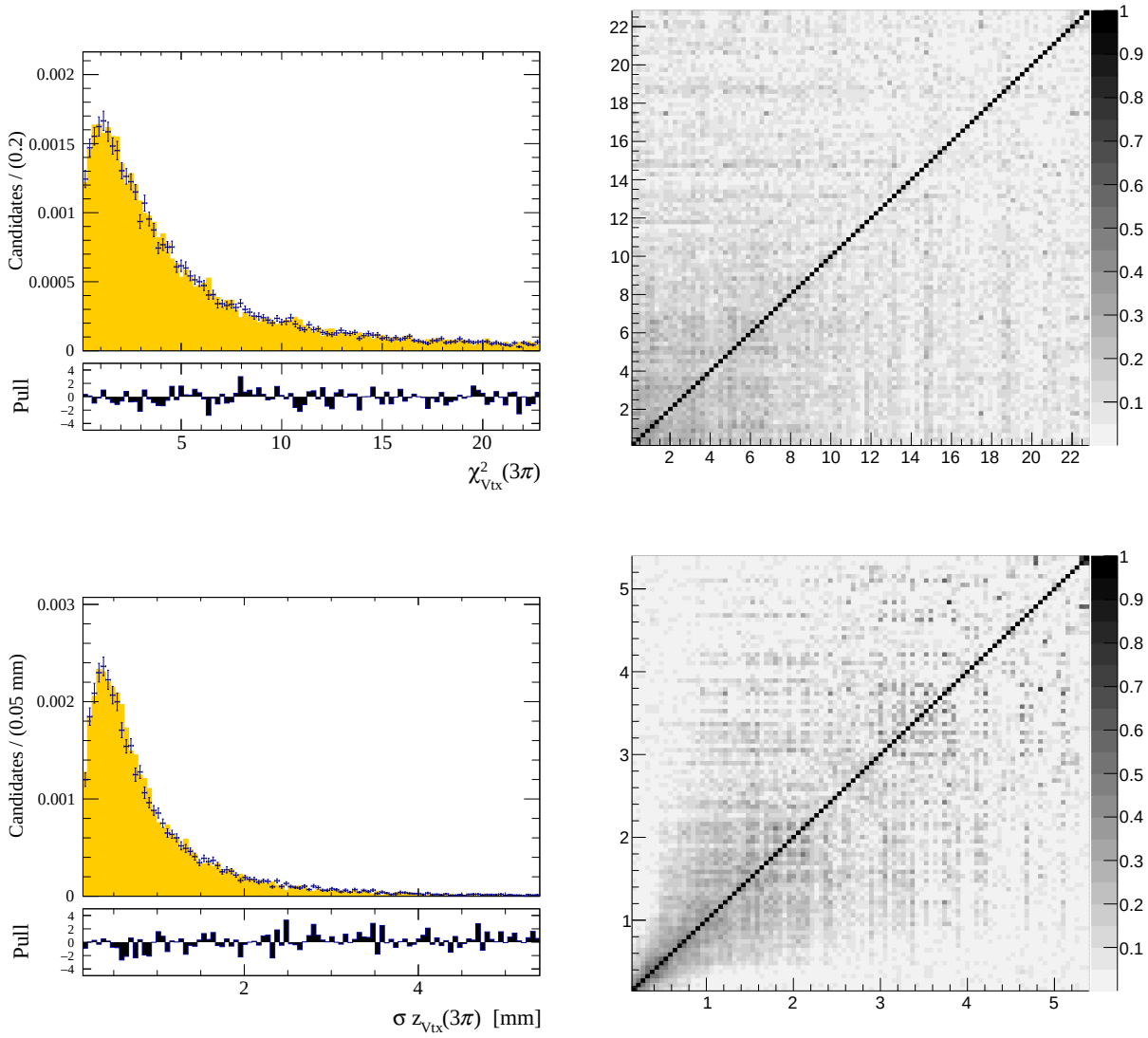


Figure B.3 – Left: distributions for the 28k $B^0 \to D^{*-}\tau^+\nu_\tau$ ($\tau^+ \to 3\pi^\pm\overline{\nu}_\tau$) mode events after stripping; full simulation is in yellow and ReDecay in bars with uncertainty. Right: Bin-to-bin correlation of $\eta(B^0)$.

Appendix C

Selection efficiencies

Selection efficiencies are broken down in more detail in Table C.1.

Table C.1 – 'Offline' selection absolute and cumulative efficiencies (in %) as a function of the selection stage for the simulated $b\bar{b} \to D^{*-}3\pi^{\pm}X$, $B^0 \to D^{*-}\tau^+\nu_\tau$ and $B^0 \to D^{*-}3\pi^{\pm}$ modes. The cells group cuts into the selection stages, where each cell assumes all the cuts from the previous stage have been applied.

Cut	Absolute efficiencies				Cumulative efficiencies			
	$b\bar{b} \to D^{*-}3\pi^{\pm}X$	$B^0 \to D^{*-}\tau^+\nu_\tau$ $3\pi\overline{\nu}_\tau$	$B^0 \to D^{*-}\tau^+\nu_\tau$ $3\pi\pi^0\overline{\nu}_\tau$	$B^0 \to D^{*-}3\pi^{\pm}$	$b\bar{b} \to D^{*-}3\pi^{\pm}X$	$B^0 \to D^{*-}\tau^+\nu_\tau$ $3\pi\overline{\nu}_\tau$	$B^0 \to D^{*-}\tau^+\nu_\tau$ $3\pi\pi^0\overline{\nu}_\tau$	$B^0 \to D^{*-}3\pi^{\pm}$
Initial selection								
`L0`	91.60 ± 0.01	89.51 ± 0.04	86.60 ± 0.07	89.08 ± 0.06	91.60 ± 0.01	89.51 ± 0.04	86.60 ± 0.07	89.08 ± 0.06
`Hlt1`	91.12 ± 0.01	89.76 ± 0.04	87.32 ± 0.07	90.92 ± 0.06	88.35 ± 0.01	87.14 ± 0.05	83.88 ± 0.08	88.02 ± 0.06
`Hlt2`	79.50 ± 0.02	79.90 ± 0.05	77.31 ± 0.09	90.33 ± 0.06	73.77 ± 0.02	73.25 ± 0.06	69.02 ± 0.10	85.10 ± 0.07
$PV(\overline{D}^0) = PV(\tau^+)$	69.88 ± 0.02	69.76 ± 0.06	65.73 ± 0.10	79.94 ± 0.08	69.88 ± 0.02	69.76 ± 0.06	65.73 ± 0.10	79.94 ± 0.08
`totCandidates` $= 1$	42.22 ± 0.02	60.89 ± 0.07	52.22 ± 0.11	71.97 ± 0.09	39.97 ± 0.02	58.06 ± 0.07	49.87 ± 0.11	67.75 ± 0.09
$[\mathrm{vtx}_z(\tau^+) - \mathrm{vtx}_z(PV)]/\mathrm{error} > 10$	70.68 ± 0.02	71.66 ± 0.06	66.59 ± 0.10	78.60 ± 0.08	38.24 ± 0.02	57.01 ± 0.07	48.29 ± 0.11	62.64 ± 0.10
`nSPDHits` < 450	72.63 ± 0.02	72.24 ± 0.06	67.78 ± 0.10	83.97 ± 0.07	37.76 ± 0.02	56.37 ± 0.07	47.56 ± 0.11	61.99 ± 0.10
Signal selection								
$m(D^{*-}) - m(\overline{D}^0) \in [143, 148]\,\mathrm{MeV}$	92.04 ± 0.025	94.631 ± 0.044	93.98 ± 0.079	92.04 ± 0.025	94.631 ± 0.044	93.98 ± 0.079		
$m(K^-\pi^+) \in [1840, 1890]\,\mathrm{MeV}/c^2$	96.804 ± 0.016	97.363 ± 0.031	97.387 ± 0.053	89.467 ± 0.028	92.276 ± 0.052	91.701 ± 0.091		
$m(3\pi) < 1825\,\mathrm{MeV}/c^2$	69.204 ± 0.042	98.242 ± 0.026	98.772 ± 0.037	61.975 ± 0.044	90.734 ± 0.057	90.682 ± 0.096		
$m(B^0) < 5350\,\mathrm{MeV}/c^2$	71.093 ± 0.041	99.29 ± 0.016	99.031 ± 0.032	52.578 ± 0.045	90.463 ± 0.057	90.267 ± 0.098		
$q^2 \in [0, 12]\,\mathrm{GeV}^2/c^4$	78.073 ± 0.038	97.52 ± 0.03	97.224 ± 0.054	49.047 ± 0.045	88.736 ± 0.062	88.53 ± 0.11		
combinatorial BDTD > 0	65.427 ± 0.043	80.367 ± 0.077	76.71 ± 0.14	34.554 ± 0.043	74.723 ± 0.085	71.89 ± 0.15		
$[\mathrm{vtx}_z(\tau^+) - \mathrm{vtx}_z(B^0)]/\mathrm{error} > 2$	77.028 ± 0.038	99.8126 ± 0.0084	99.775 ± 0.016	34.554 ± 0.043	74.723 ± 0.085	71.89 ± 0.15		
isolation BDT > 0	55.935 ± 0.045	87.849 ± 0.064	83.86 ± 0.12	17.31 ± 0.034	67.415 ± 0.091	62.41 ± 0.16		
anti D_s^+ BDT > -0.2	71.018 ± 0.041	98.296 ± 0.025	86.1 ± 0.11	13.925 ± 0.031	67.123 ± 0.092	54.87 ± 0.16		
PID	-	76.23 ± 0.20	78.86 ± 0.16	-	-	-	-	-
Normalisation selection								
$[\mathrm{vtx}_z(\overline{D}^0) - \mathrm{vtx}_z(\tau^+)]/\mathrm{error} > 4$	-	-	-	94.30 ± 0.08	-	-	-	94.30 ± 0.08
$m(D^*3\pi^{\pm}) \in [5150, 5400]\,\mathrm{MeV}$	-	-	-	97.87 ± 0.05	-	-	-	93.32 ± 0.09
$m(D^{*-}) - m(\overline{D}^0) \in [143, 148]\,\mathrm{MeV}$	-	-	-	94.97 ± 0.08	-	-	-	89.04 ± 0.11
combinatorial BDTD > 0	-	-	-	81.37 ± 0.13	-	-	-	74.19 ± 0.15
isolation BDT > 0	-	-	-	88.33 ± 0.11	-	-	-	66.94 ± 0.16
PID	-	-	-	73.96 ± 0.24	-	-	-	-

Appendix D

MC/data comparison

MC/data comparisons of the input variables to the MVA variables from the Chapter 5 are shown here.

D.1 Combinatorial BDT

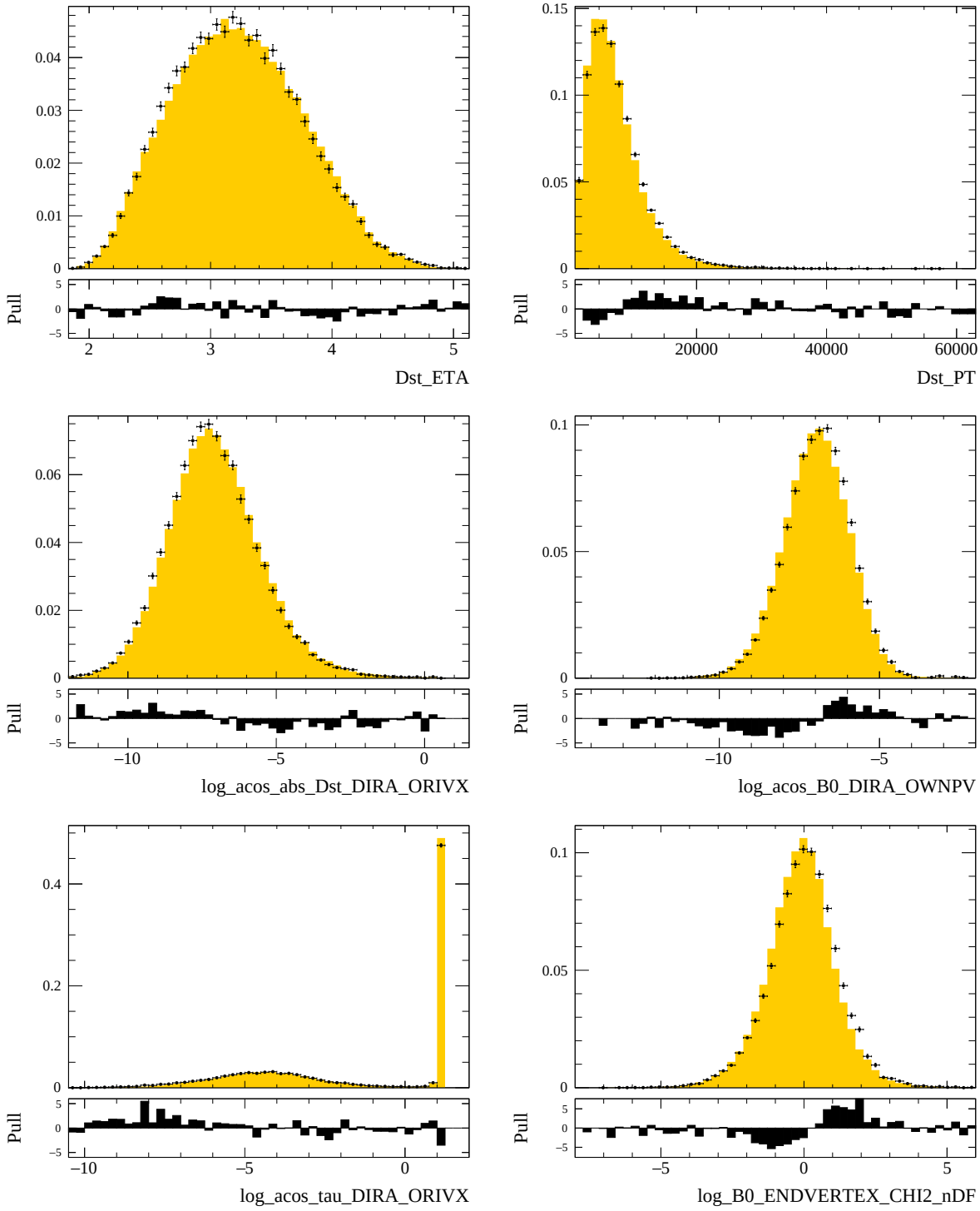


Figure D.1 – Data (black points) / MC (coloured bins) comparison for the `combinatorialBDTD` input variables.

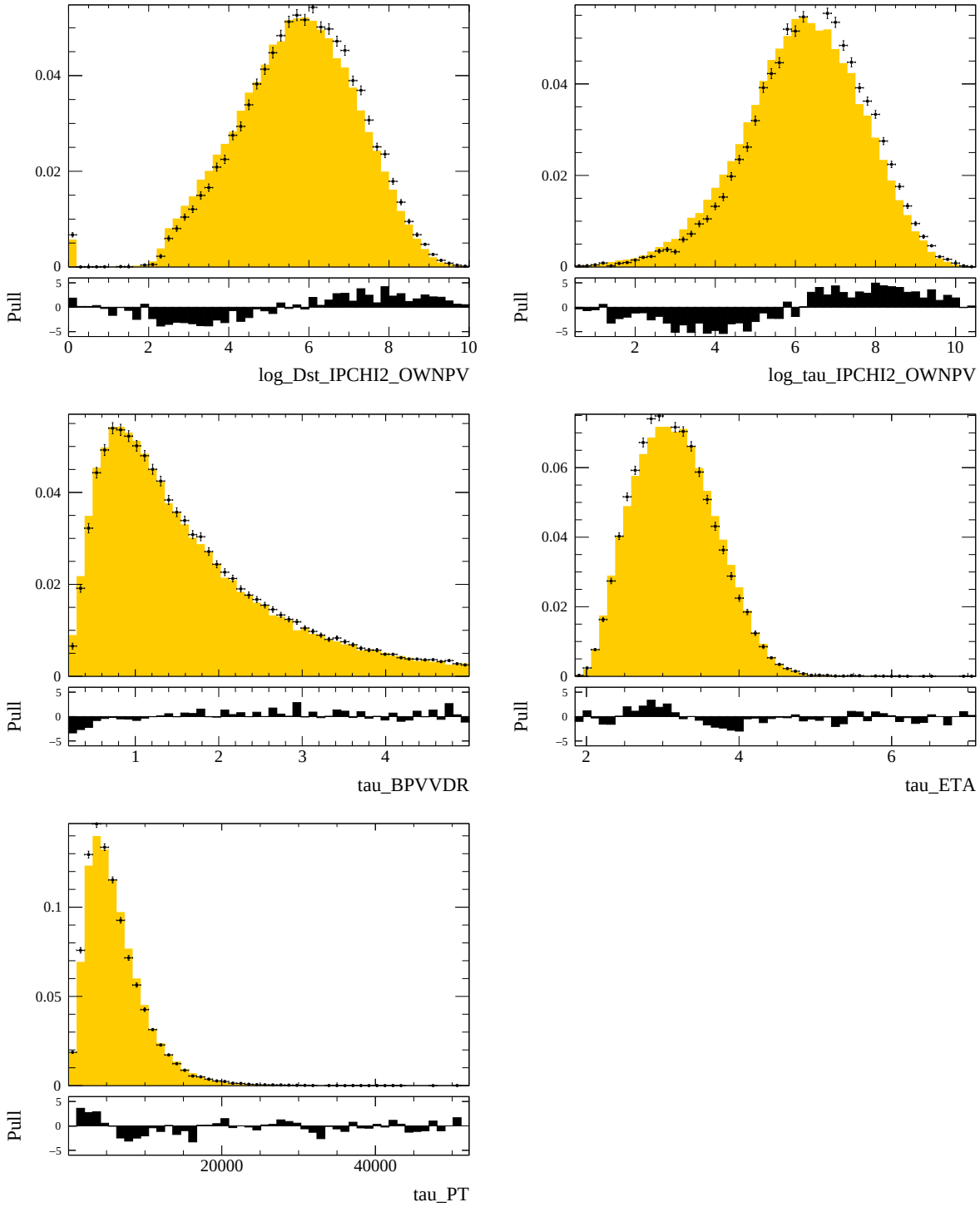


Figure D.2 – Data (black points) / MC (coloured bins) comparison for the `combinatorialBDTD` input variables.

D.2 Isolation BDT

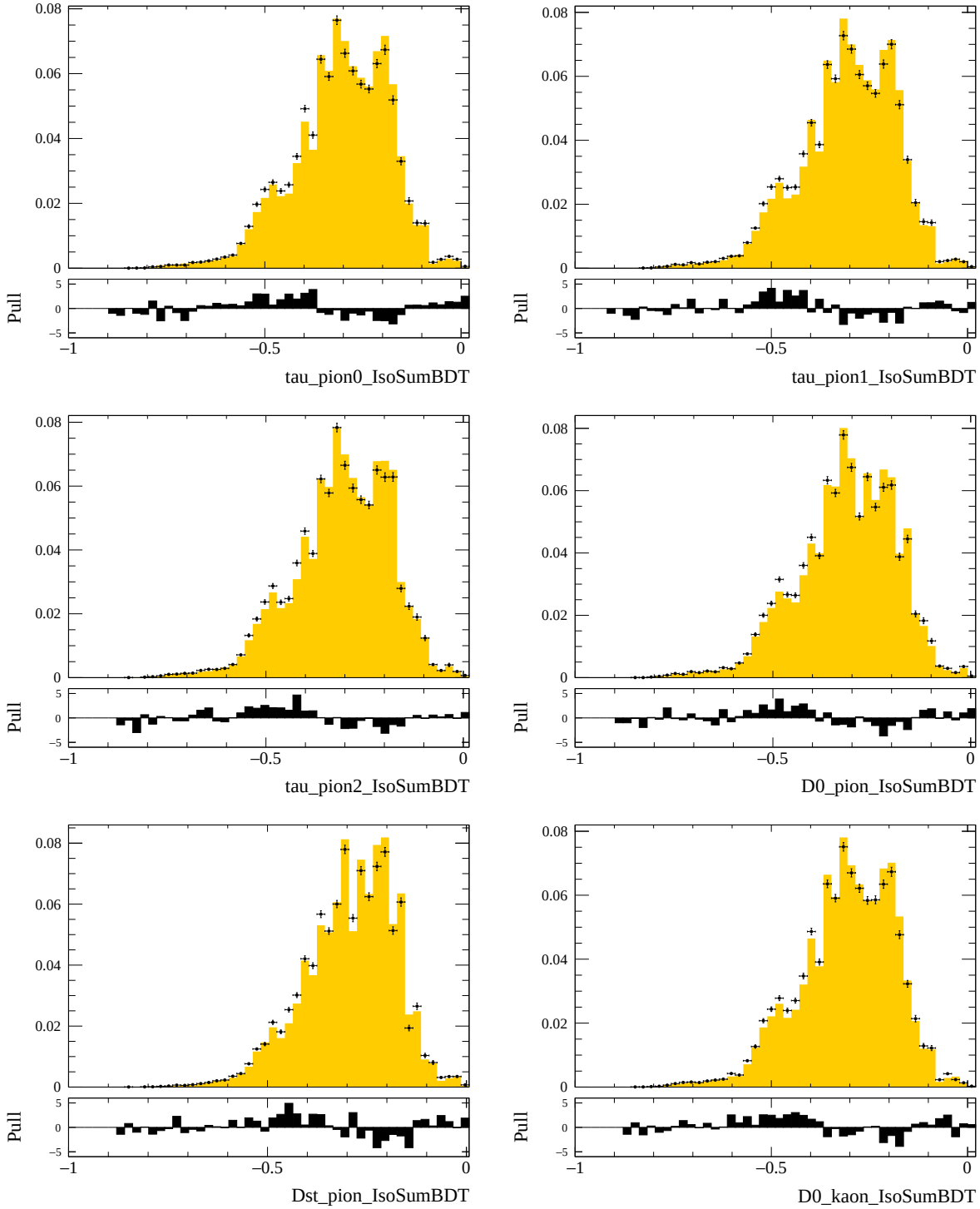


Figure D.3 – Data (black points) / MC (coloured bins) comparison for the charged isolation input variables.

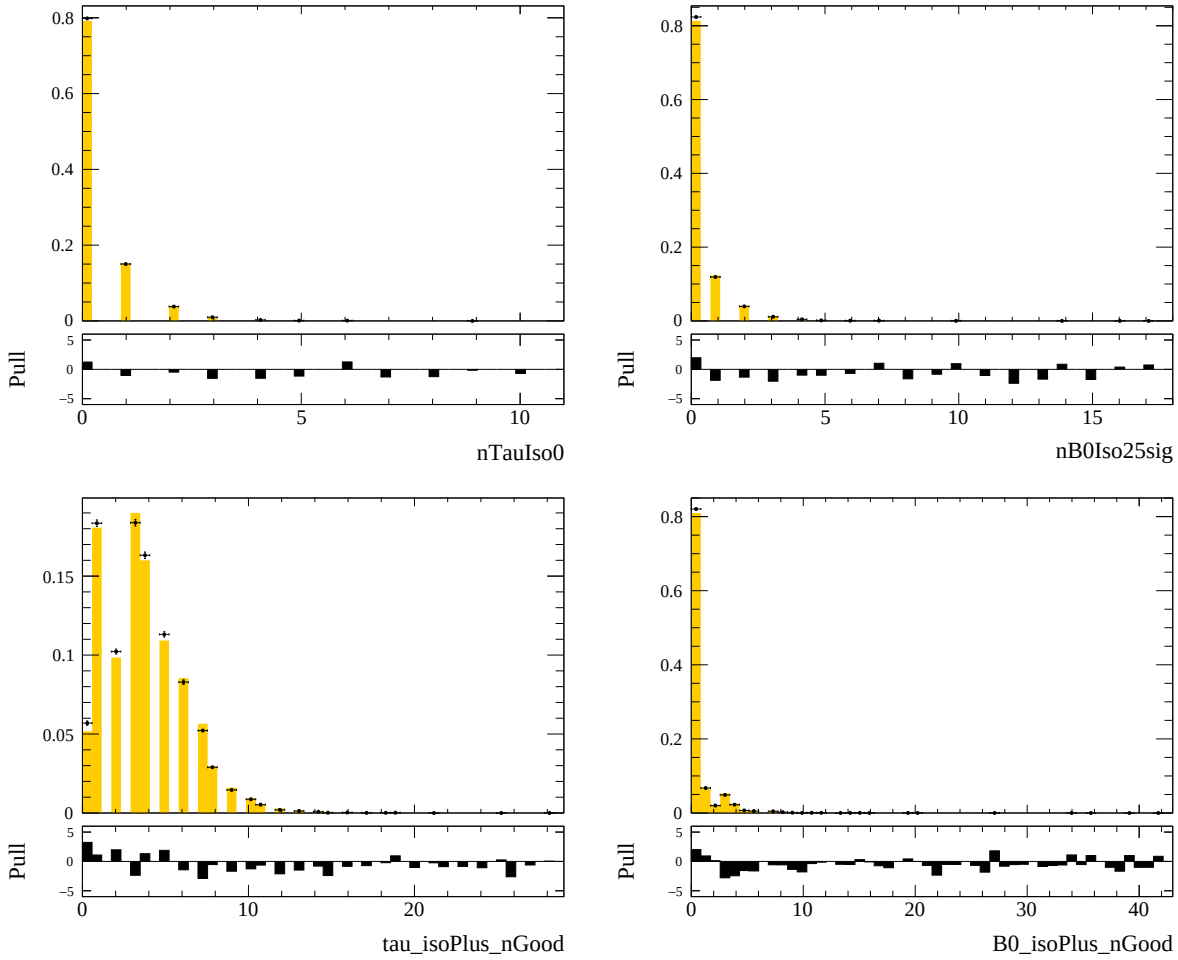


Figure D.4 – Data (black points) / MC (coloured bins) comparison for the charged isolation input variables.

D.3 Detachment BDT

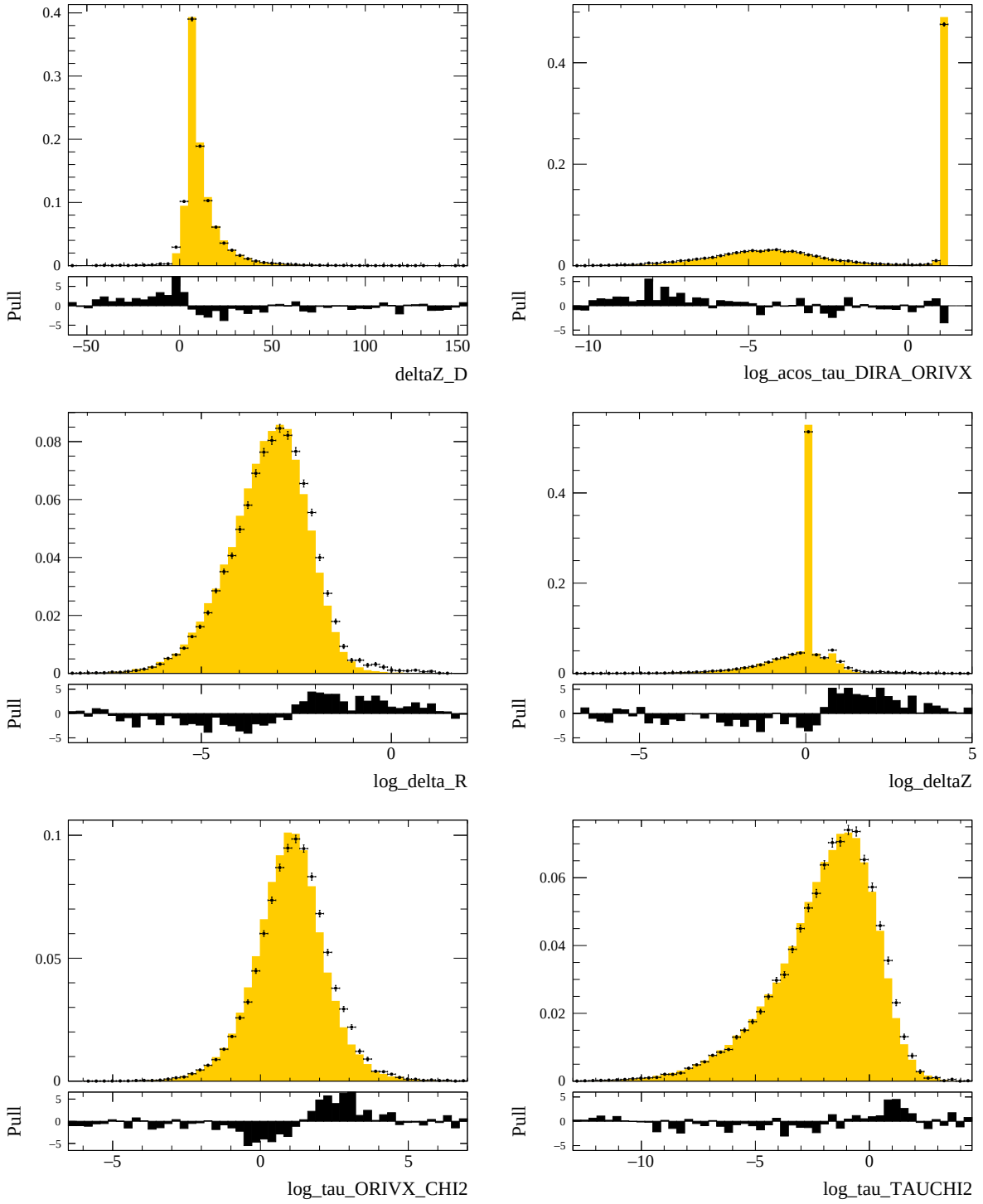


Figure D.5 – Data (black points) / MC (coloured bins) comparison for the `detachment_BDTG` input variables.

D.4 Anti-D_s^+ BDT

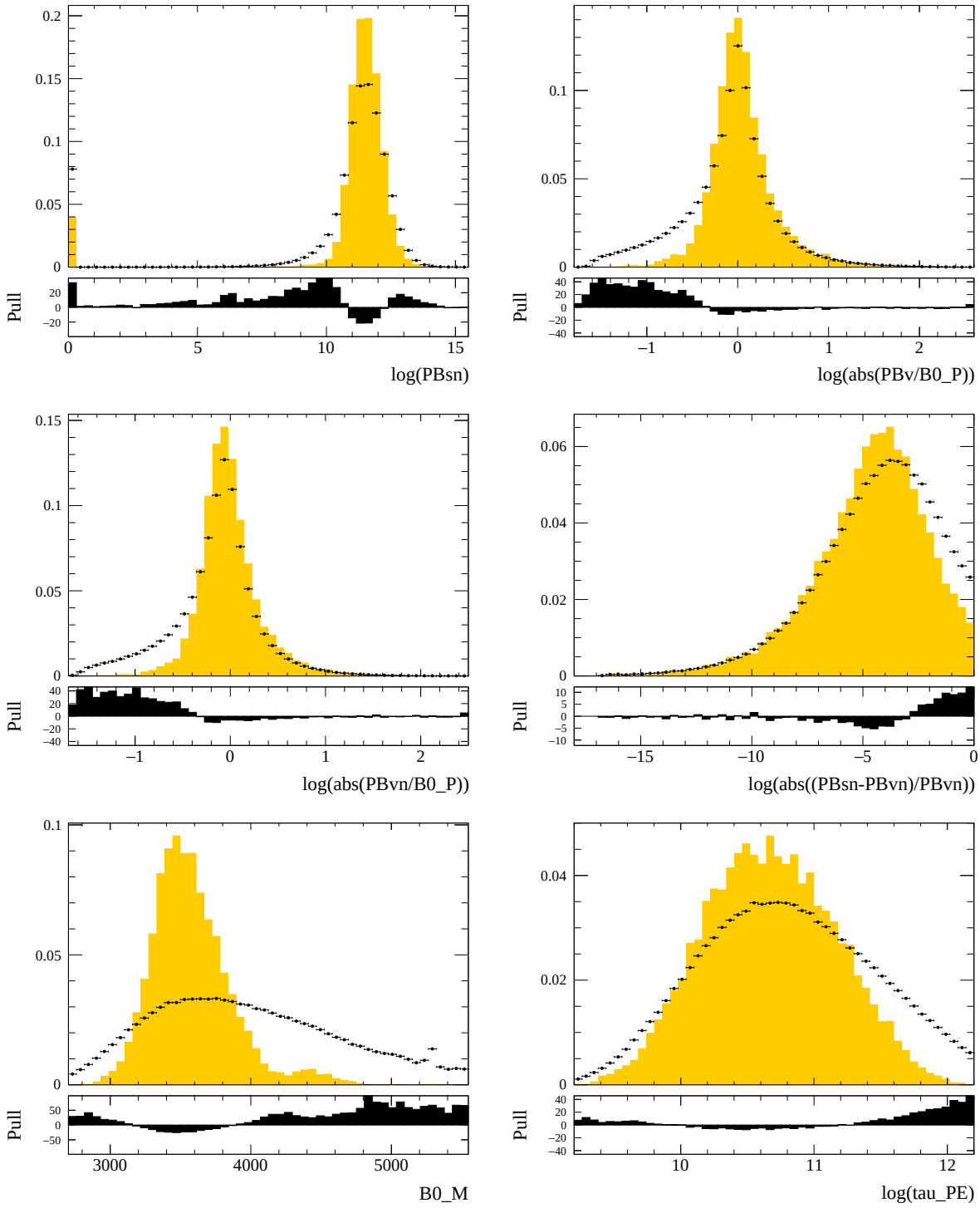


Figure D.6 – Data (black points) / MC (coloured bins) comparison of the anti-D_s^+ input variables for the $B \to D^{*-} D^0\ X$ control mode. The MC sample is extracted from the inclusive $b\bar{b} \to D^{*-} 3\pi^{\pm}\ X$ with the $B \to D^{*-} D^0\ X$ control mode selection.

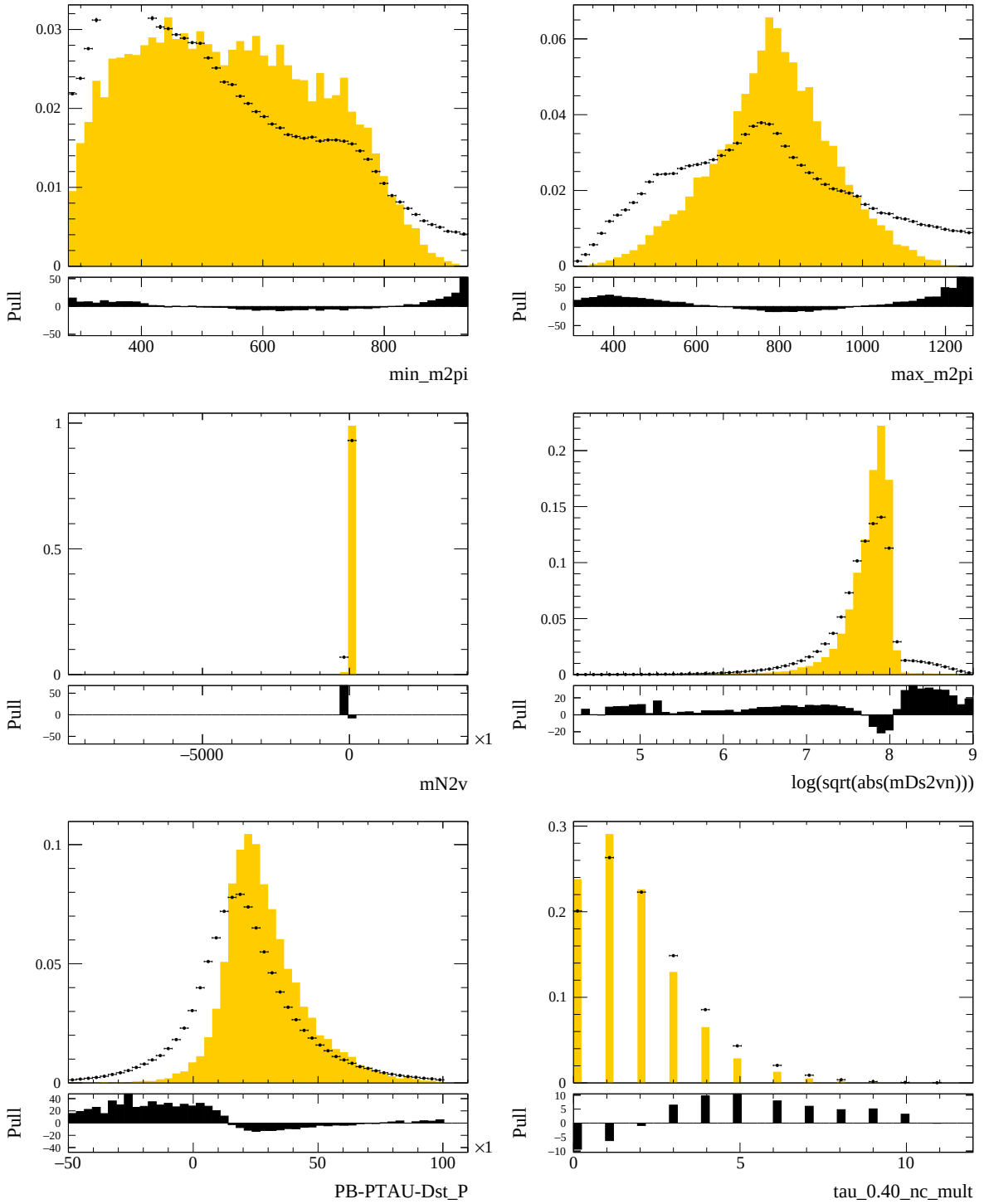


Figure D.7 – Data (black points) / MC (coloured bins) comparison of the anti-D_s^+ input variables for the $B \to D^{*-} D^0\ X$ control mode. The MC sample is extracted from the inclusive $b\bar{b} \to D^{*-} 3\pi^{\pm}\ X$ with the $B \to D^{*-} D^0\ X$ control mode selection.

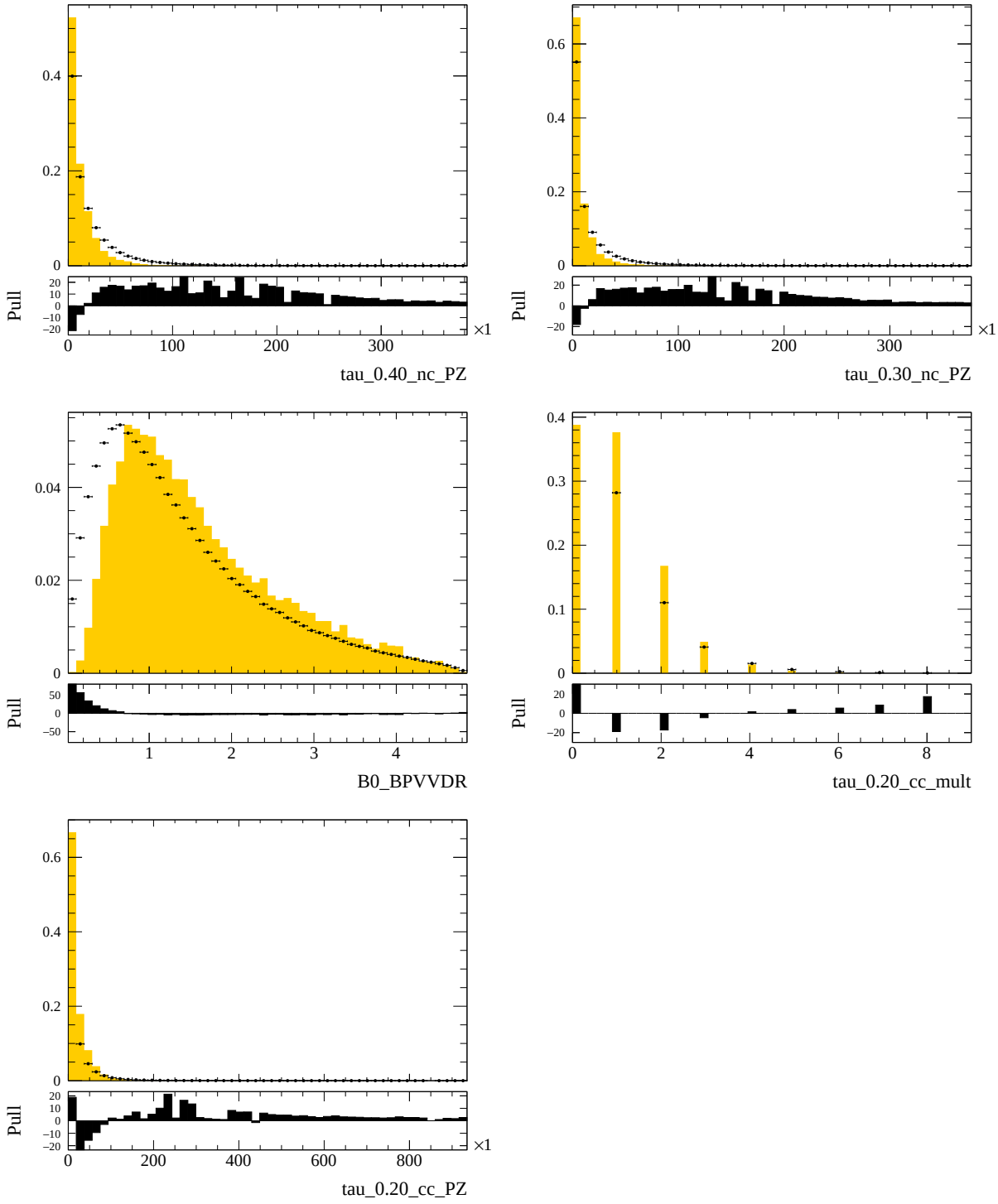


Figure D.8 – Data (black points) / MC (coloured bins) comparison of the anti-D_s^+ input variables for the $B \to D^{*-} D^0\ X$ control mode. The MC sample is extracted from the inclusive $b\bar{b} \to D^{*-} 3\pi^{\pm}\ X$ with the $B \to D^{*-} D^0\ X$ control mode selection.

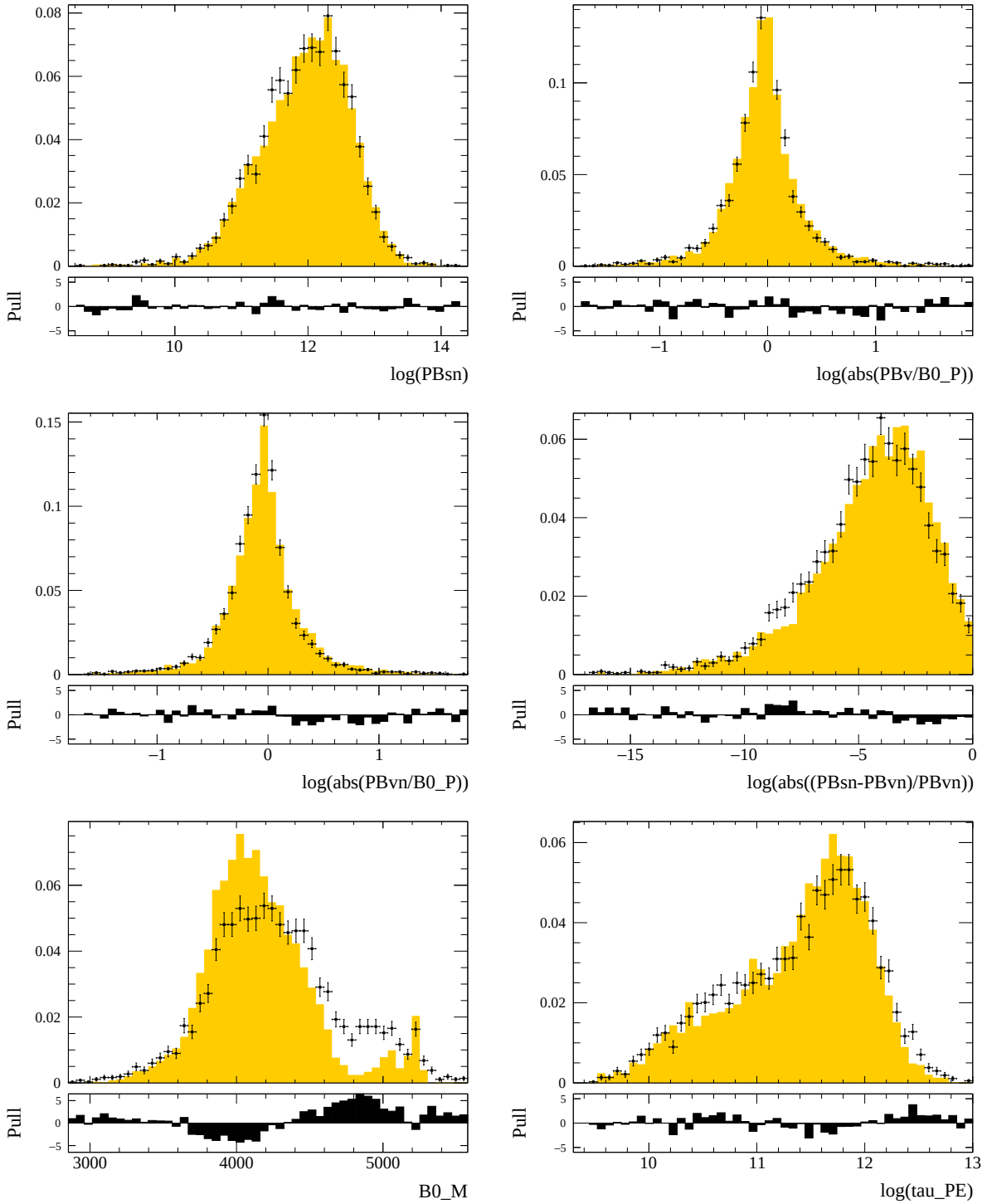


Figure D.9 – Data (black points) / MC (coloured bins) comparison of the anti-D_s^+ input variables for the $B \to D^{*-} D^+ X$ control mode. The MC sample is extracted from the inclusive $b\bar{b} \to D^{*-} 3\pi^{\pm} X$ with the $B \to D^{*-} D^+ X$ control mode selection.

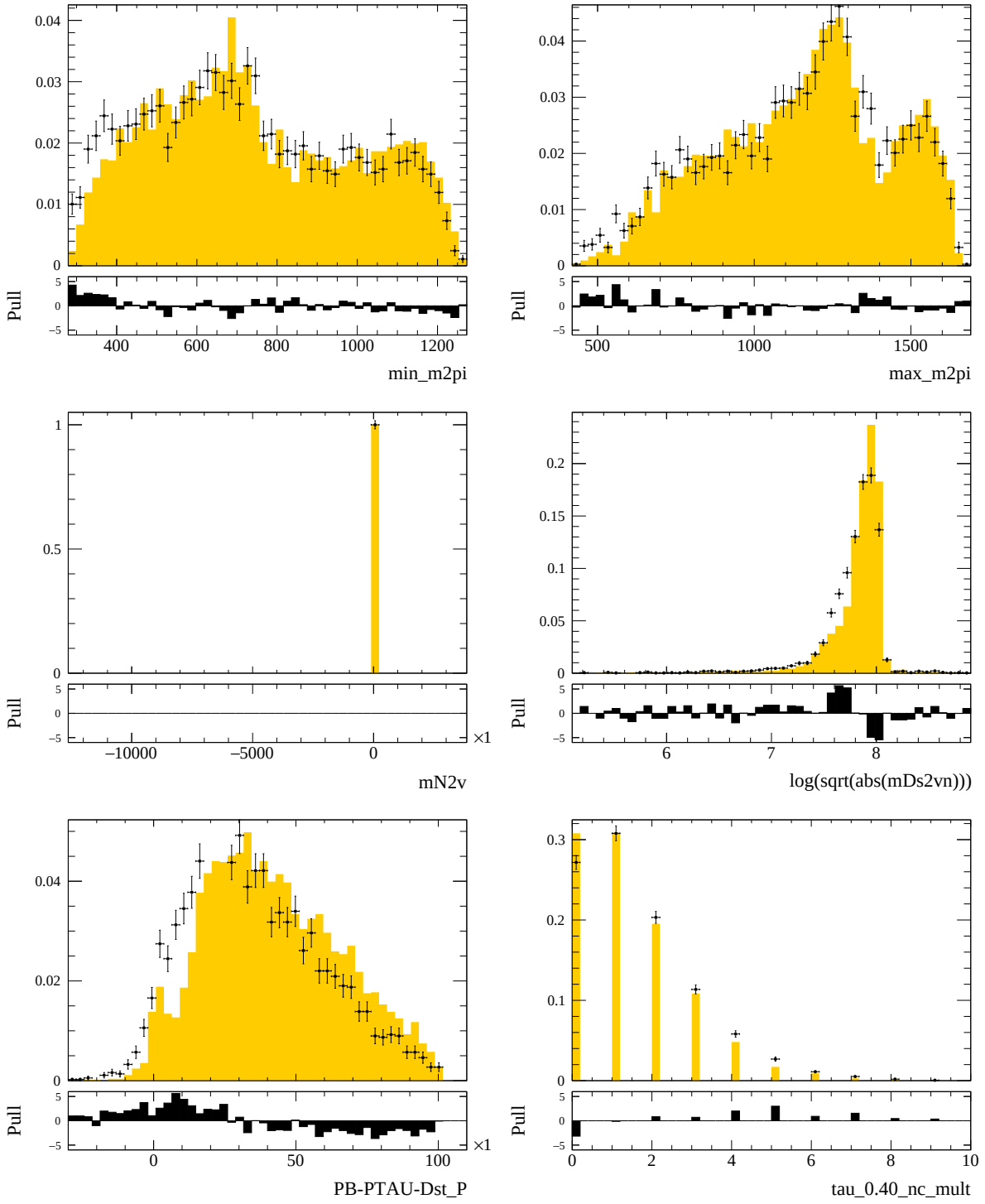


Figure D.10 – Data (black points) / MC (coloured bins) comparison of the anti-D_s^+ input variables for the $B \to D^{*-} D^+ X$ control mode. The MC sample is extracted from the inclusive $b\bar{b} \to D^{*-}\ 3\pi^{\pm}\ X$ with the $B \to D^{*-} D^+ X$ control mode selection.

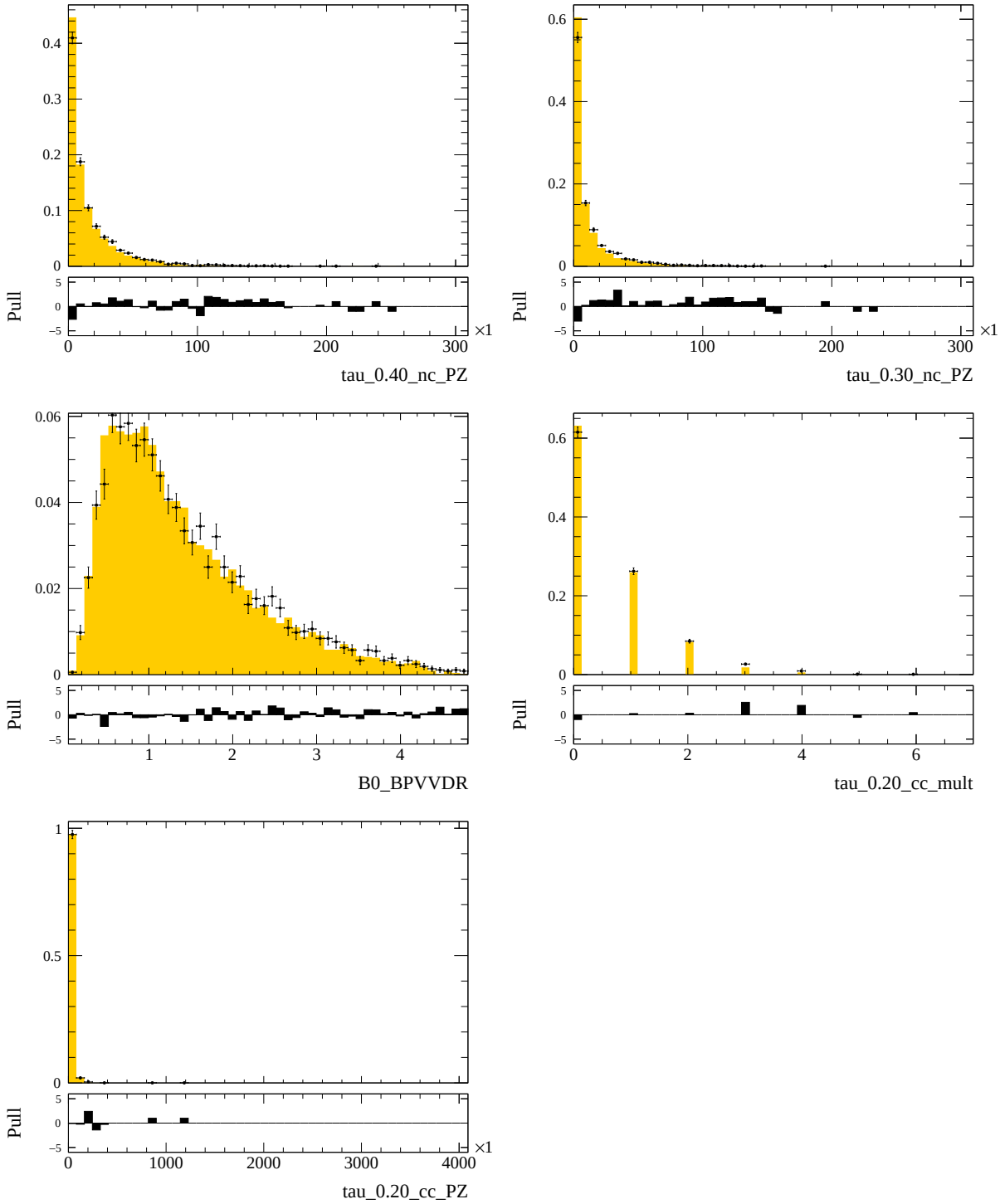


Figure D.11 – Data (black points) / MC (coloured bins) comparison of the anti-D_s^+ input variables for the $B \to D^{*-} D^+ X$ control mode. The MC sample is extracted from the inclusive $b\bar{b} \to D^{*-}\ 3\pi^{\pm}\ X$ with the $B \to D^{*-} D^+ X$ control mode selection.

Appendix E

Correlation matrices of the signal yield fits

Table E.1 – Correlation coefficients of the signal fit variables in the 2015-2016 collision data fit.

	N_{D_s}	$N_{B\to D^{*-}3\pi^{\pm}X}$	N_{sig}	$f_{B_s\to D^*D_sX}$	$f_{D^{**}D_sX}$	f_{D^+}	$f_{D_{s1}'^+}$	f_{D_s}	$f_{D_{s0}^{*+}}$	$f_{D^0}^{v_1-v_2}$
N_{D_s}	1.000	−0.209	−0.243	−0.499	0.572	−0.645	−0.037	−0.059	−0.184	−0.587
$N_{B\to D^{*-}3\pi^{\pm}X}$	−0.209	1.000	0.189	0.036	0.034	0.007	−0.015	−0.070	0.028	−0.237
N_{sig}	−0.243	0.189	1.000	0.151	−0.145	0.311	−0.037	−0.013	0.098	0.317
$f_{B_s\to D^*D_sX}$	−0.499	0.036	0.151	1.000	0.016	0.335	0.130	−0.028	0.016	0.436
$f_{D^{**}D_sX}$	0.572	0.034	−0.145	0.016	1.000	−0.475	0.216	0.089	−0.046	−0.457
f_{D^+}	−0.645	0.007	0.311	0.335	−0.475	1.000	0.037	0.058	0.141	0.166
$f_{D_{s1}'^+}$	−0.037	−0.015	−0.037	0.130	0.216	0.037	1.000	−0.211	0.371	0.011
f_{D_s}	−0.059	−0.070	−0.013	−0.028	0.089	0.058	−0.211	1.000	0.213	0.070
$f_{D_{s0}^{*+}}$	−0.184	0.028	0.098	0.016	−0.046	0.141	0.371	0.213	1.000	0.154
$f_{D^0}^{v_1-v_2}$	−0.587	−0.237	0.317	0.436	−0.457	0.166	0.011	0.070	0.154	1.000

Table E.2 – Correlation coefficients of the signal fit variables with fixed f_{D^+} and $f_{D^0}^{v_1-v_2}$ in the 2015-2016 collision data fit.

	N_{D_s}	$N_{B\to D^{*-}3\pi^{\pm}X}$	N_{sig}	$f_{B_s\to D^*D_sX}$	$f_{D^{**}D_sX}$	$f_{D_{s1}'^+}$	f_{D_s}	$f_{D_{s0}^{*+}}$
N_{D_s}	1.000	−0.563	0.165	0.189	0.169	0.018	0.014	−0.049
$N_{B\to D^{*-}3\pi^{\pm}X}$	−0.563	1.000	0.284	−0.152	−0.072	0.015	−0.057	0.062
N_{sig}	0.165	0.284	1.000	0.073	0.147	0.055	−0.052	0.020
$f_{B_s\to D^*D_sX}$	0.189	−0.152	0.073	1.000	−0.475	0.135	0.083	0.098
$f_{D^{**}D_sX}$	0.169	−0.072	0.147	−0.475	1.000	−0.297	0.177	0.092
$f_{D_{s1}'^+}$	0.018	0.015	0.055	0.135	−0.297	1.000	0.214	−0.372
f_{D_s}	0.014	−0.057	−0.052	0.083	0.177	0.214	1.000	0.201
$f_{D_{s0}^{*+}}$	−0.049	0.062	0.020	0.098	0.092	−0.372	0.201	1.000

Table E.3 – Correlation coefficients of the signal fit variables in the 2015-2016 MC fit.

	N_{D_s}	$N_{B\to D^*-3\pi^\pm X}$	N_{sig}	$f_{B_s\to D^*D_sX}$	$f_{D^{**}D_sX}$	f_{D^+}	$f_{D'^+_{s1}}$	f_{D_s}	$f_{D^{*+}_{s0}}$	$f_{D^0}^{v_1-v_2}$
N_{D_s}	1.000	−0.228	0.055	0.526	−0.508	−0.658	0.031	−0.049	−0.141	−0.485
$N_{B\to D^*-3\pi^\pm X}$	−0.228	1.000	−0.189	−0.002	−0.030	0.046	−0.010	−0.075	0.040	−0.197
N_{sig}	0.055	−0.189	1.000	0.149	−0.060	−0.226	0.076	0.021	−0.061	−0.263
$f_{B_s\to D^*D_sX}$	0.526	−0.002	0.149	1.000	0.009	−0.415	−0.065	0.011	−0.008	−0.457
$f_{D^{**}D_sX}$	−0.508	−0.030	−0.060	0.009	1.000	0.414	−0.274	−0.104	0.002	0.393
f_{D^+}	−0.658	0.046	−0.226	−0.415	0.414	1.000	−0.031	0.042	0.112	0.078
$f_{D'^+_{s1}}$	0.031	−0.010	0.076	−0.065	−0.274	−0.031	1.000	−0.239	0.323	−0.055
f_{D_s}	−0.049	−0.075	0.021	0.011	−0.104	0.042	−0.239	1.000	0.220	0.066
$f_{D^{*+}_{s0}}$	−0.141	0.040	−0.061	−0.008	0.002	0.112	0.323	0.220	1.000	0.110
$f_{D^0}^{v_1-v_2}$	−0.485	−0.197	−0.263	−0.457	0.393	0.078	−0.055	0.066	0.110	1.000

Appendix F

The Monte Carlo pseudo-experiment studies

In this section all distributions of the pseudo-experiment Monte Carlo studies for the signal fit parameters (*cf.* Section 7.1.3) are shown.

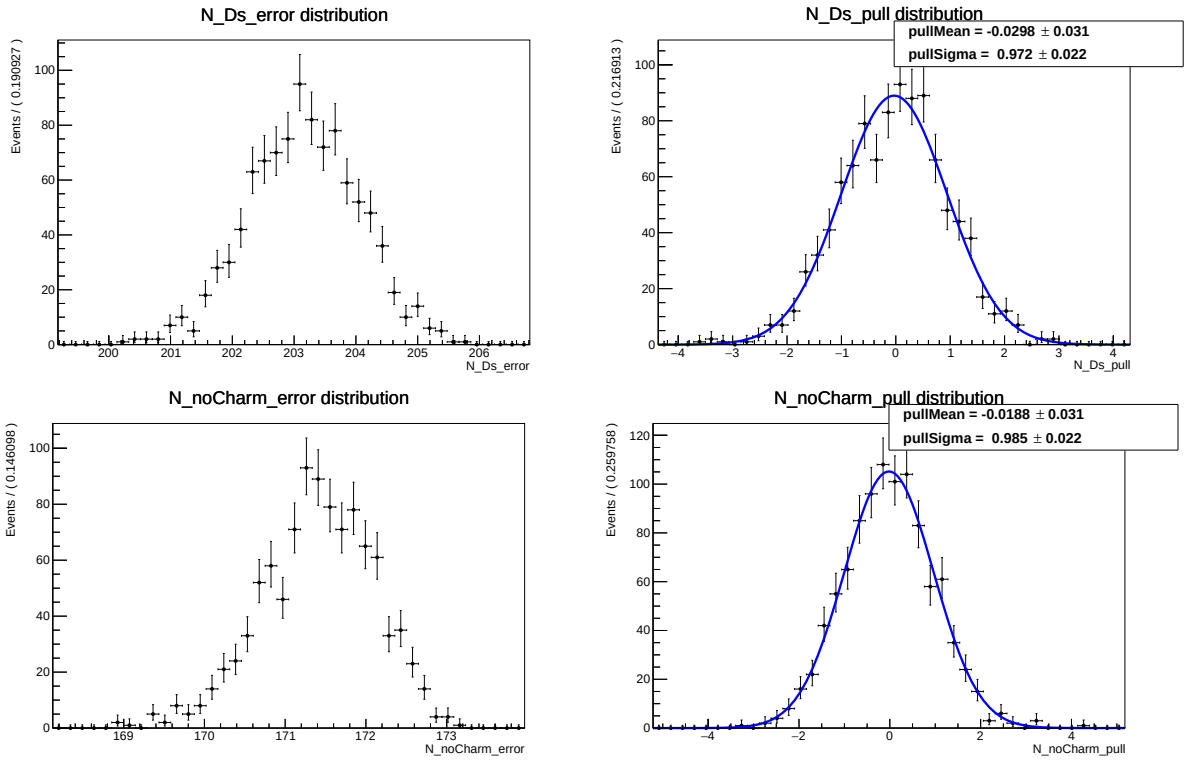


Figure F.1 – Distribution of the relative difference of the fitted and nominal yields (left) and of the pull from a toy study with 1000 generated events.

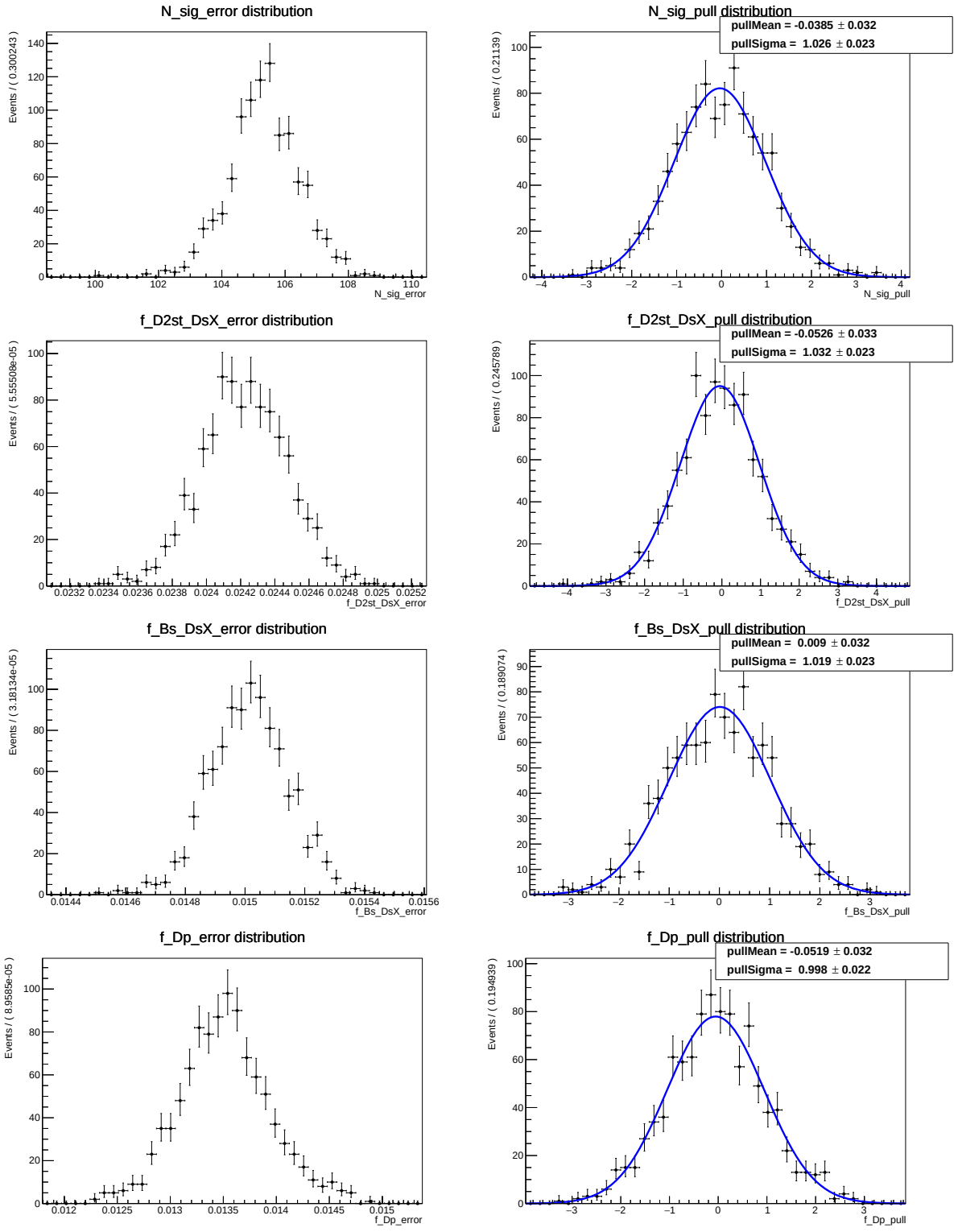


Figure F.2 – Distribution of the relative difference of the fitted and nominal yields (left) and of the pull (right) from a toy study with 1000 generated events.

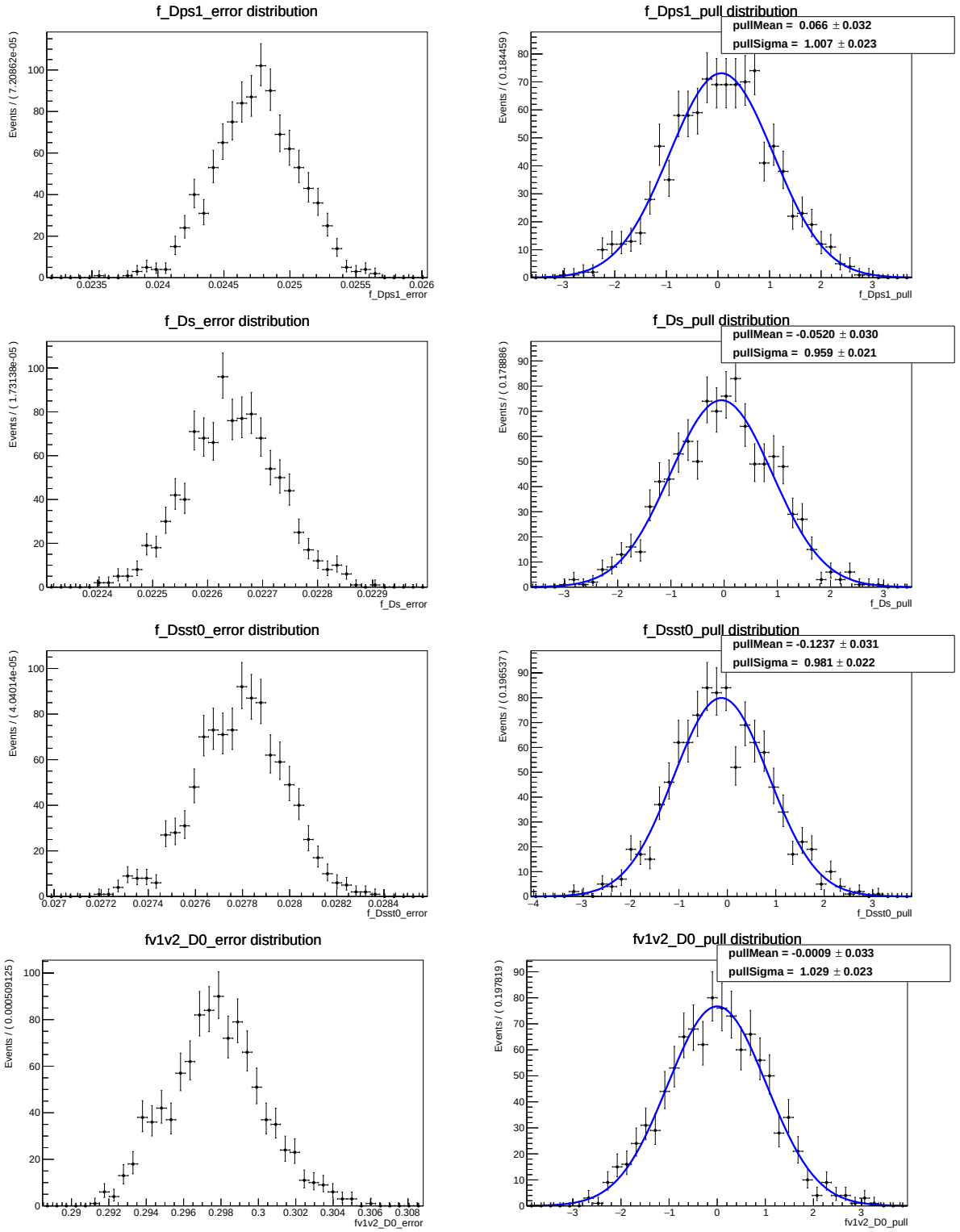


Figure F.3 – Distribution of the relative difference of the fitted and nominal yields (left) and of the pull (right) from a toy study with 1000 generated events.

Bibliography

[1] R. Aaij et al. Measurement of the ratio of the $\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)$ and $\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)$ branching fractions using three-prong τ-lepton decays. *Phys. Rev. Lett.*, 120 (2018), p. 171802. DOI: 10.1103/PhysRevLett.120.171802. arXiv: 1708.08856 [hep-ex]

[2] R. Aaij et al. Test of lepton flavor universality by the measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays. *Phys. Rev.*, D97 (2018), p. 072013. DOI: 10.1103/PhysRevD.97.072013. arXiv: 1711.02505 [hep-ex]

[3] G. Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716.1 (Sept. 2012), pp. 1–29. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.020. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.020

[4] S. Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716.1 (Sept. 2012), pp. 30–61. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.021. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.021

[5] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc. C*, 680519 (1968), pp. 367–377. DOI: 10.1142/9789812795915_0034

[6] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19 (21 Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264

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

[8] Bruce T. Cleveland et al. Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector. *The Astrophysical Journal*, 496.1 (Mar. 1998), pp. 505–526. DOI: 10.1086/305343. URL: https://doi.org/10.1086%2F305343

[9] Y. Fukuda et al. Evidence for Oscillation of Atmospheric Neutrinos. *Physical Review Letters*, 81.8 (Aug. 1998), pp. 1562–1567. ISSN: 1079-7114. DOI: 10.1103/physrevlett.81.1562. URL: http://dx.doi.org/10.1103/PhysRevLett.81.1562

[10] M.V. Diwan et al. Long-Baseline Neutrino Experiments. *Annual Review of Nuclear and Particle Science*, 66.1 (Oct. 2016), pp. 47–71. ISSN: 1545-4134. DOI: 10.1146/

annurev-nucl-102014-021939. URL: http://dx.doi.org/10.1146/annurev-nucl-102014-021939

[11] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. Remarks on the Unified Model of Elementary Particles. *Progress of Theoretical Physics*, 28.5 (Nov. 1962), pp. 870–880. ISSN: 0033-068X. DOI: 10.1143/PTP.28.870. eprint: https://academic.oup.com/ptp/article-pdf/28/5/870/5258750/28-5-870.pdf. URL: https://doi.org/10.1143/PTP.28.870

[12] Mark Thomson. *Modern particle physics*. New York: Cambridge University Press, 2013. ISBN: 9781107034266. URL: http://www-spires.fnal.gov/spires/find/books/www?cl=QC793.2.T46::2013

[13] Makoto Kobayashi and Toshihide Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X. DOI: 10.1143/PTP.49.652. eprint: https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf. URL: https://doi.org/10.1143/PTP.49.652

[14] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev. D*, 98 (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001

[15] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak Interactions with Lepton-Hadron Symmetry. *Phys. Rev. D*, 2 (7 Oct. 1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285. URL: https://link.aps.org/doi/10.1103/PhysRevD.2.1285

[16] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321

[17] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508

[18] Peter W. Higgs. Spontaneous Symmetry Breakdown without Massless Bosons. *Phys. Rev.*, 145 (4 May 1966), pp. 1156–1163. DOI: 10.1103/PhysRev.145.1156. URL: https://link.aps.org/doi/10.1103/PhysRev.145.1156

[19] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett.*, 13 (20 Nov. 1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585

[20] S.-K. Choi et al. Observation of a Resonancelike Structure in the $\pi^{+-}\psi'$ Mass Distribution in Exclusive $B \to K\pi^{+-}\psi'$ Decays. *Phys. Rev. Lett.*, 100 (14 Apr. 2008), p. 142001. DOI: 10.1103/PhysRevLett.100.142001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.142001

[21] R. Aaij et al. Observation of the Resonant Character of the Z(4430) State. *Physical Review Letters*, 112.22 (June 2014). ISSN: 1079-7114. DOI: 10.1103/physrevlett.112.222002. URL: http://dx.doi.org/10.1103/PhysRevLett.112.222002

[22] 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 (7 Aug. 2015), p. 072001. DOI: 10.1103/PhysRevLett.115.072001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.115.072001

[23] M. Fael and C. Greub. Next-to-leading order prediction for the decay $\mu \to e(e^+e^-)\nu\overline{\nu}$. *Journal of High Energy Physics*, 2017.1 (Jan. 2017). ISSN: 1029-8479. DOI: 10.1007/jhep01(2017)084. URL: http://dx.doi.org/10.1007/JHEP01(2017)084

[24] A. Pich. Tau physics: theoretical perspective. *Nuclear Physics B - Proceedings Supplements*, 98.1-3 (Apr. 2001), pp. 385–396. ISSN: 0920-5632. DOI: 10.1016/s0920-5632(01)01254-3. URL: http://dx.doi.org/10.1016/S0920-5632(01)01254-3

[25] Svjetlana Fajfer, Jernej F. Kamenik, and Ivan Nišandžić. $B \to D^*\tau\overline{\nu}_\tau$ sensitivity to new physics. *Physical Review D*, 85.9 (May 2012). ISSN: 1550-2368. DOI: 10.1103/physrevd.85.094025. URL: http://dx.doi.org/10.1103/PhysRevD.85.094025

[26] Florian U. Bernlochner et al. Combined analysis of semileptonic B decays to D and D*: R(D(*)), |Vcb|, and new physics. *Physical Review D*, 95.11 (June 2017). ISSN: 2470-0029. DOI: 10.1103/physrevd.95.115008. URL: http://dx.doi.org/10.1103/PhysRevD.95.115008

[27] Dante Bigi, Paolo Gambino, and Stefan Schacht. $R(D)$, V_{cb}, and the Heavy Quark Symmetry relations between form factors. *Journal of High Energy Physics*, 2017.11 (Nov. 2017). ISSN: 1029-8479. DOI: 10.1007/jhep11(2017)061. URL: http://dx.doi.org/10.1007/JHEP11(2017)061

[28] Sneha Jaiswal, Soumitra Nandi, and Sunando Kumar Patra. Extraction of $|V_{cb}|$ from $B \to D^{(*)}\ell\nu_\ell$ and the Standard Model predictions of $R(D^{(*)})$. *Journal of High Energy Physics*, 2017.12 (Dec. 2017). ISSN: 1029-8479. DOI: 10.1007/jhep12(2017)060. URL: http://dx.doi.org/10.1007/JHEP12(2017)060

[29] The Belle Collaboration et al. *Precise determination of the CKM matrix element $|V_{cb}|$ with $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*. 2017. arXiv: 1702.01521 [hep-ex]

[30] Heavy Flavour Averaging Group (HFLAV). *Updates of Semileptonic Results for Spring 2019*. 2019. URL: https://hflav-eos.web.cern.ch/hflav-eos/semi/spring19/html/RDsDsstar/RDRDs.html (visited on 12/29/2019)

[31] J. P. Lees et al. Evidence for an Excess of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ Decays. *Phys. Rev. Lett.*, 109 (10 Sept. 2012), p. 101802. DOI: 10.1103/PhysRevLett.109.101802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.109.101802

[32] J. P. Lees et al. Measurement of an excess ofB$^-$→D(*) − $^-$ decays and implications for charged Higgs bosons. *Physical Review D*, 88.7 (Oct. 2013). ISSN: 1550-2368. DOI: 10.1103/physrevd.88.072012. URL: http://dx.doi.org/10.1103/PhysRevD.88.072012

[33] M. Huschle et al. Measurement of the branching ratio ofB$^-$→D(*) − $^-$ relative toB$^-$→D(*) − $^-$ decays with hadronic tagging at Belle. *Physical Review D*, 92.7 (Oct. 2015). ISSN: 1550-2368. DOI: 10.1103/physrevd.92.072014. URL: http://dx.doi.org/10.1103/PhysRevD.92.072014

[34] R. Aaij et al. Measurement of the Ratio of Branching Fractions $\mathcal{B}(\overline{B}^0 \to D^{*+}\tau^-\overline{\nu}_\tau)/\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)$. *Phys. Rev. Lett.*, 115 (11 Sept. 2015), p. 111803. DOI: 10.1103/PhysRevLett.115.111803. URL: https://link.aps.org/doi/10.1103/PhysRevLett.115.111803

[35] 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 (7 Oct. 2016), p. 072007. DOI: 10.1103/PhysRevD.94.072007. URL: https://link.aps.org/doi/10.1103/PhysRevD.94.072007

[36] S. Hirose et al. Measurement of the τ Lepton Polarization and $R(D^*)$ in the Decay $\overline{B} \to D^*\tau^-\overline{\nu}_\tau$. *Phys. Rev. Lett.*, 118 (21 May 2017), p. 211801. DOI: 10.1103/PhysRevLett.118.211801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.118.211801

[37] S. Hirose et al. Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\overline{B} \to D^*\tau^-\overline{\nu}_\tau$ with one-prong hadronic τ decays at Belle. *Phys. Rev. D*, 97 (1 Jan. 2018), p. 012004. DOI: 10.1103/PhysRevD.97.012004. URL: https://link.aps.org/doi/10.1103/PhysRevD.97.012004

[38] R. Aaij et al. Measurement of the Ratio of the B0→D*− + and B0→D*− + Branching Fractions Using Three-Prong -Lepton Decays. *Physical Review Letters*, 120.17 (Apr. 2018). ISSN: 1079-7114. DOI: 10.1103/physrevlett.120.171802. URL: http://dx.doi.org/10.1103/PhysRevLett.120.171802

[39] R. Aaij et al. Test of lepton flavor universality by the measurement of the B0→D*− + branching fraction using three-prong decays. *Physical Review D*, 97.7 (Apr. 2018). ISSN: 2470-0029. DOI: 10.1103/physrevd.97.072013. URL: http://dx.doi.org/10.1103/PhysRevD.97.072013

[40] A. Abdesselam et al. Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method (Apr. 2019). arXiv: 1904.08794 [hep-ex]

[41] Marzia Bordone et al. Heavy-Quark expansion for $\bar{B}_s \to D_s^{(*)}$ form factors and unitarity bounds beyond the $SU(3)_F$ limit. *Eur. Phys. J. C*, 80.4 (2020), p. 347. DOI: 10.1140/epjc/s10052-020-7850-9. arXiv: 1912.09335 [hep-ph]

[42] Marat Freytsis, Zoltan Ligeti, and Joshua T. Ruderman. Flavor models for B−→D(*) −. *Physical Review D*, 92.5 (Sept. 2015). ISSN: 1550-2368. DOI: 10.1103/physrevd.92.054018. URL: http://dx.doi.org/10.1103/PhysRevD.92.054018

[43] R. Aaij et al. Measurement of the Ratio of Branching Fractions B(Bc+→J/ +)/B(Bc+→J/ +). *Physical Review Letters*, 120.12 (Mar. 2018). ISSN: 1079-7114. DOI: 10.1103/physrevlett.120.121801. URL: http://dx.doi.org/10.1103/PhysRevLett.120.121801

[44] Judd Harrison, Christine T.H. Davies, and Andrew Lytle. $R(J/\psi)$ and $B_c^- \to J/\psi\ell^-\overline{\nu}_\ell$ Lepton Flavor Universality Violating Observables from Lattice QCD (July 2020). arXiv: 2007.06956 [hep-lat]

[45] R. Aaij et al. Test of Lepton Universality UsingB+→K+ + −Decays. *Physical Review Letters*, 113.15 (Oct. 2014). ISSN: 1079-7114. DOI: 10.1103/physrevlett.113.151601. URL: http://dx.doi.org/10.1103/PhysRevLett.113.151601

[46] R. Aaij et al. Search for Lepton-Universality Violation in B+→K+ + − Decays. *Physical Review Letters*, 122.19 (May 2019). ISSN: 1079-7114. DOI: 10.1103/physrevlett.122.191801. URL: http://dx.doi.org/10.1103/PhysRevLett.122.191801

[47] J. P. Lees et al. Measurement of branching fractions and rate asymmetries in the rare decaysB→K(*) + −. *Physical Review D*, 86.3 (Aug. 2012). ISSN: 1550-2368. DOI: 10.1103/physrevd.86.032012. URL: http://dx.doi.org/10.1103/PhysRevD.86.032012

[48] J.-T. Wei et al. Measurement of the Differential Branching Fraction and Forward-Backward Asymmetry forB→K(*)l+l−. *Physical Review Letters*, 103.17 (Oct. 2009). ISSN: 1079-7114. DOI: 10.1103/physrevlett.103.171801. URL: http://dx.doi.org/10.1103/PhysRevLett.103.171801

[49] R. Aaij et al. Test of lepton universality with B0 → K*0 + − decays. *Journal of High Energy Physics*, 2017.8 (Aug. 2017). ISSN: 1029-8479. DOI: 10.1007/jhep08(2017)055. URL: http://dx.doi.org/10.1007/JHEP08(2017)055

[50] R. Aaij et al. Test of lepton universality with $\Lambda_b^0 \to pK^-\ell^+\ell^-$ decays. *Journal of High Energy Physics*, 2020.5 (May 2020). ISSN: 1029-8479. DOI: 10.1007/jhep05(2020)040. URL: http://dx.doi.org/10.1007/JHEP05(2020)040

[51] LHCb collaboration et al. *Measurement of CP-averaged observables in the $B^0 \to K^{*0+-}$ decay*. 2020. arXiv: 2003.04831 [hep-ex]

[52] A. Abdesselam et al. Test of lepton flavor universality in $B \to K\ell^+\ell^-$ decays (Aug. 2019). arXiv: 1908.01848 [hep-ex]

[53] A. Abdesselam et al. Test of lepton flavor universality in $B \to K^*\ell^+\ell^-$ decays at Belle (Apr. 2019). arXiv: 1904.02440 [hep-ex]

[54] Marcel Algueró et al. Emerging patterns of New Physics with and without Lepton Flavour Universal contributions. *The European Physical Journal C*, 79.8 (Aug. 2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7216-3. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7216-3

[55] Gerhard Buchalla, Andrzej J. Buras, and Markus E. Lautenbacher. Weak decays beyond leading logarithms. *Reviews of Modern Physics*, 68.4 (Oct. 1996), pp. 1125–1244. ISSN: 1539-0756. DOI: 10.1103/revmodphys.68.1125. URL: http://dx.doi.org/10.1103/RevModPhys.68.1125

[56] Andrzej J. Buras. *Weak Hamiltonian, CP Violation and Rare Decays*. 1998. arXiv: hep-ph/9806471 [hep-ph]

[57] Simone Bifani et al. Review of lepton universality tests in B decays. *Journal of Physics G: Nuclear and Particle Physics*, 46.2 (Dec. 2018), p. 023001. ISSN: 1361-6471. DOI: 10.1088/1361-6471/aaf5de. URL: http://dx.doi.org/10.1088/1361-6471/aaf5de

[58] Irinel Caprini, Laurent Lellouch, and Matthias Neubert. Dispersive bounds on the shape of $\overline{B} \to D^{(*)}\ell\bar{\nu}$ form factors. *Nucl. Phys.*, B530 (1998), pp. 153–181. DOI: 10.1016/S0550-3213(98)00350-2. arXiv: hep-ph/9712417 [hep-ph]

[59] C. Glenn Boyd, Benjamin Grinstein, and Richard F. Lebed. Constraints on form factors for exclusive semileptonic heavy to light meson decays. *Phys. Rev. Lett.*, 74 (1995), pp. 4603–4606. DOI: 10.1103/PhysRevLett.74.4603. arXiv: hep-ph/9412324 [hep-ph]

[60] C. Glenn Boyd, Benjamin Grinstein, and Richard F. Lebed. Model-independent determinations of $\overline{B} \to D\ell\bar{\nu}, D^*\ell\bar{\nu}$ form factors. *Nucl. Phys.*, B461 (1996), pp. 493–511. DOI: 10.1016/0550-3213(95)00653-2. arXiv: hep-ph/9508211 [hep-ph]

[61] C. Glenn Boyd, Benjamin Grinstein, and Richard F. Lebed. Precision corrections to dispersive bounds on form factors. *Phys. Rev.*, D56 (1997), pp. 6895–6911. DOI: 10.1103/PhysRevD.56.6895. arXiv: hep-ph/9705252 [hep-ph]

[62] R. Aaij et al. Measurement of $|V_{cb}|$ with $B_s^0 \to D_s^{(*)-}\mu^+\nu_\mu$ decays. *Phys. Rev. D*, 101 (7 Apr. 2020), p. 072004. DOI: 10.1103/PhysRevD.101.072004. URL: https://link.aps.org/doi/10.1103/PhysRevD.101.072004

[63] S. Aoki et al. FLAG Review 2019. *The European Physical Journal C*, 80.2 (Feb. 2020). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7354-7. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7354-7

[64] Jon A. Bailey et al. Update of $|V_{cb}|$ from the $\overline{B} \to D^* \ell \overline{\nu}$ form factor at zero recoil with three-flavor lattice QCD. *Phys. Rev. D*, 89 (11 June 2014), p. 114504. DOI: 10.1103/PhysRevD.89.114504. URL: https://link.aps.org/doi/10.1103/PhysRevD.89.114504

[65] Judd Harrison, Christine T. H. Davies, and Matthew Wingate. Lattice QCD calculation of the $B_{(s)} \to D^*_{(s)} \ell \nu$ form factors at zero recoil and implications for $|V_{cb}|$. *Phys. Rev. D*, 97 (5 Mar. 2018), p. 054502. DOI: 10.1103/PhysRevD.97.054502. URL: https://link.aps.org/doi/10.1103/PhysRevD.97.054502

[66] T. Kaneko et al. *$B \to D^{(*)} \ell \nu$ form factors from lattice QCD with relativistic heavy quarks*. 2019. arXiv: 1912.11770 [hep-lat]

[67] A. Vaquero et al. *$B \to D^* \ell \nu$ at non-zero recoil*. 2019. arXiv: 1906.01019 [hep-lat]

[68] Andreas Crivellin, Dario Müller, and Toshihiko Ota. Simultaneous explanation of R(D (∗)) and b→s + −: the last scalar leptoquarks standing. *Journal of High Energy Physics*, 2017.9 (Sept. 2017). ISSN: 1029-8479. DOI: 10.1007/jhep09(2017)040. URL: http://dx.doi.org/10.1007/JHEP09(2017)040

[69] Andreas Crivellin, Julian Heeck, and Peter Stoffer. Perturbed Lepton-Specific Two-Higgs-Doublet Model Facing Experimental Hints for Physics beyond the Standard Model. *Physical Review Letters*, 116.8 (Feb. 2016). ISSN: 1079-7114. DOI: 10.1103/physrevlett.116.081801. URL: http://dx.doi.org/10.1103/PhysRevLett.116.081801

[70] Admir Greljo, Gino Isidori, and David Marzocca. *On the breaking of Lepton Flavor Universality in B decays*. 2015. arXiv: 1506.01705 [hep-ph]

[71] S. Fajfer and N. Košnik. Vector leptoquark resolution of R and RD( ) puzzles. *Physics Letters B*, 755 (Apr. 2016), pp. 270–274. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2016.02.018. URL: http://dx.doi.org/10.1016/j.physletb.2016.02.018

[72] Damir Bećirević et al. Leptoquark model to explain the B-physics anomalies, R_K and R_D. *Physical Review D*, 94 (Aug. 2016). DOI: 10.1103/PhysRevD.94.115021

[73] M. Aaboud et al. Searches for scalar leptoquarks and differential cross-section measurements in dilepton–dijet events in proton–proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV with the ATLAS experiment. *The European Physical Journal C*, 79.9 (Sept. 2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7181-x. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7181-x

[74] A. M. Sirunyan et al. Search for pair production of first-generation scalar leptoquarks at $\sqrt{s} = 13$ TeV. *Phys. Rev. D*, 99 (5 Mar. 2019), p. 052002. DOI: 10.1103/PhysRevD.99.052002. URL: https://link.aps.org/doi/10.1103/PhysRevD.99.052002

[75] A. M. Sirunyan et al. Search for pair production of second-generation leptoquarks at $\sqrt{s} = 13$ TeV. *Phys. Rev. D*, 99 (3 Feb. 2019), p. 032014. DOI: 10.1103/PhysRevD.99.032014. URL: https://link.aps.org/doi/10.1103/PhysRevD.99.032014

[76] M. Aaboud et al. Searches for third-generation scalar leptoquarks in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector. *Journal of High Energy Physics*, 2019.6 (June 2019). ISSN: 1029-8479. DOI: 10.1007/jhep06(2019)144. URL: http://dx.doi.org/10.1007/JHEP06(2019)144

[77] A. M. Sirunyan et al. Search for heavy neutrinos and third-generation leptoquarks in hadronic states of two leptons and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Journal of High Energy Physics*, 2019.3 (Mar. 2019). ISSN: 1029-8479. DOI: 10.1007/jhep03(2019)170. URL: http://dx.doi.org/10.1007/JHEP03(2019)170

[78] A. M. Sirunyan et al. Search for third-generation scalar leptoquarks decaying to a top quark and a τ lepton at $\sqrt{s} = 13$ TeV. *The European Physical Journal C*, 78.9 (Sept. 2018). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-018-6143-z. URL: http://dx.doi.org/10.1140/epjc/s10052-018-6143-z

[79] A. M. Sirunyan et al. Search for a singly produced third-generation scalar leptoquark decaying to a lepton and a bottom quark in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Journal of High Energy Physics*, 2018.7 (July 2018). ISSN: 1029-8479. DOI: 10.1007/jhep07(2018)115. URL: http://dx.doi.org/10.1007/JHEP07(2018)115

[80] A. M. Sirunyan et al. Constraints on models of scalar and vector leptoquarks decaying to a quark and a neutrino at $\sqrt{s} = 13$ TeV. *Phys. Rev. D*, 98 (3 Aug. 2018), p. 032005. DOI: 10.1103/PhysRevD.98.032005. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.032005

[81] A. M. Sirunyan et al. Search for Leptoquarks Coupled to Third-Generation Quarks in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. Lett.*, 121 (24 Dec. 2018), p. 241802. DOI: 10.1103/PhysRevLett.121.241802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.241802

[82] A. Angelescu et al. Closing the window on single leptoquark solutions to the B-physics anomalies. *Journal of High Energy Physics*, 2018.10 (Oct. 2018). ISSN: 1029-8479. DOI: 10.1007/jhep10(2018)183. URL: http://dx.doi.org/10.1007/JHEP10(2018)183

[83] Damir Be čirevi ć et al. Scalar leptoquarks from grand unified theories to accommodate the B-physics anomalies. *Phys. Rev. D*, 98 (5 Sept. 2018), p. 055003. DOI: 10.1103/PhysRevD.98.055003. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.055003

[84] Damir Becirevic et al. *Lepton Flavor Universality tests through angular observables of $\overline{B} \to D^{(*)}\ell\overline{\nu}$ decay modes*. 2019. arXiv: 1907.02257 [hep-ph]

[85] Donal Hill et al. Model-independent method for measuring the angular coefficients of $B^0 \to D^{*-}\tau^+\overline{\nu}_\tau$ decays. *Journal of High Energy Physics*, 2019.11 (Nov. 2019). ISSN: 1029-8479. DOI: 10.1007/jhep11(2019)133. URL: http://dx.doi.org/10.1007/JHEP11(2019)133

[86] Roel Aaij et al. Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions. *Phys. Rev. Lett.*, 118.5 (2017). [Erratum: Phys. Rev. Lett.119,no.16,169901(2017)], p. 052002. DOI: 10.1103/PhysRevLett.119.169901,10.1103/PhysRevLett.118.052002. arXiv: 1612.05140 [hep-ex]

[87] Lyndon Evans and Philip Bryant. LHC Machine. *Journal of Instrumentation*, 3.08 (Aug. 2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001. URL: https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08001

[88] Adam Morris. *Measurements of charmless B_s^0 meson decays at LHCb*. Presented 29 Sep 2017. Aug. 2017. URL: https://cds.cern.ch/record/2293045

[89] A Augusto Alves et al. The LHCb Detector at the LHC. *JINST*, 3.LHCb-DP-2008-001. CERN-LHCb-DP-2008-001 (2008). Also published by CERN Geneva in 2010, S08005. DOI: 10.1088/1748-0221/3/08/S08005. URL: https://cds.cern.ch/record/1129809

[90] Michel De Cian et al. *Fast neural-net based fake track rejection in the LHCb reconstruction*. Tech. rep. LHCb-PUB-2017-011. CERN-LHCb-PUB-2017-011. Geneva: CERN, Mar. 2017. URL: https://cds.cern.ch/record/2255039

[91] Roel Aaij et al. Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC. Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC. *JINST*, 14.arXiv:1812.10790. 04 (Dec. 2018). 46 pages, 35 figures, 1 table. All figures and tables are available at https://cern.ch/lhcbproject/Publications/LHCbProjectPublic/LHCb-DP-2019-001.html, P04013. 43 p. DOI: 10.1088/1748-0221/14/04/P04013. URL: https://cds.cern.ch/record/2652801

[92] LHCb detector performance. *International Journal of Modern Physics A*, 30.07 (Mar. 2015), p. 1530022. ISSN: 1793-656X. DOI: 10.1142/s0217751x15300227. URL: http://dx.doi.org/10.1142/S0217751X15300227

[93] Roel Aaij et al. Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2. *EPJ Tech. Instrum.*, 6.1 (2019), p. 1. DOI: 10.1140/epjti/s40485-019-0050-z. arXiv: 1803.00824 [hep-ex]

[94] R. Brun and F. Rademakers. ROOT: An object oriented data analysis framework. *Nucl. Instrum. Meth. A*, 389 (1997). Ed. by M. Werlen and D. Perret-Gallix, pp. 81–86. DOI: 10.1016/S0168-9002(97)00048-X

[95] Muriel Pivk and Francois R. Le Diberder. SPlot: A Statistical tool to unfold data distributions. *Nucl. Instrum. Meth. A*, 555 (2005), pp. 356–369. DOI: 10.1016/j.nima.2005.08.106. arXiv: physics/0402083

[96] Lucio Anderlini et al. *The PIDCalib package*. Tech. rep. LHCb-PUB-2016-021. CERN-LHCb-PUB-2016-021. Geneva: CERN, July 2016. URL: https://cds.cern.ch/record/2202412

[97] Roel Aaij et al. Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC. Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC. *JINST*, 14.arXiv:1812.10790. 04 (Dec. 2018). 46 pages, 35 figures, 1 table. All figures and tables are available at https://cern.ch/lhcbproject/Publications/LHCbProjectPublic/LHCb-DP-2019-001.html, P04013. 43 p. DOI: 10.1088/1748-0221/14/04/P04013. URL: https://cds.cern.ch/record/2652801

[98] *The Gaudi Framework*. https://gaudi.web.cern.ch/gaudi/

[99] M Clemencic et al. The LHCb Simulation Application, Gauss: Design, Evolution and Experience. *Journal of Physics: Conference Series*, 331.3 (Dec. 2011),

p. 032023. DOI: 10.1088/1742-6596/331/3/032023. URL: https://doi.org/10.1088%2F1742-6596%2F331%2F3%2F032023

[100] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178.11 (2008), pp. 852–867. ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2008.01.036. URL: http://www.sciencedirect.com/science/article/pii/S0010465508000441

[101] D.J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth. A*, 462 (2001). Ed. by S. Erhan, P. Schlein, and Y. Rozen, pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4

[102] S. Agostinelli et al. GEANT4–a simulation toolkit. *Nucl. Instrum. Meth. A*, 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8

[103] Yasmine Sara Amhis et al. Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018 (2019). updated results and plots available at https://hflav.web.cern.ch/. arXiv: 1909.12524 [hep-ex]

[104] R. Aaij et al. Study ofB0→D*− + − +andB0→D*−K+ − +decays. *Physical Review D*, 87.9 (May 2013). ISSN: 1550-2368. DOI: 10.1103/physrevd.87.092001. URL: http://dx.doi.org/10.1103/PhysRevD.87.092001

[105] J. P. Lees et al. Measurement of the B0→D*− + − + branching fraction. *Physical Review D*, 94.9 (Nov. 2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.94.091101. URL: http://dx.doi.org/10.1103/PhysRevD.94.091101

[106] G. Majumder et al. Observation of B0→D*-(5)+, B+→D*-(4)++ and B+→D¯*0 (5)+. *Physical Review D*, 70.11 (Dec. 2004). ISSN: 1550-2368. DOI: 10.1103/physrevd.70.111103. URL: http://dx.doi.org/10.1103/PhysRevD.70.111103

[107] R. Aaij et al. Test of lepton flavor universality by the measurement of the B0→D*− + branching fraction using three-prong decays. *Physical Review D*, 97.7 (Apr. 2018). ISSN: 2470-0029. DOI: 10.1103/physrevd.97.072013. URL: http://dx.doi.org/10.1103/PhysRevD.97.072013

[108] Dominik Müller et al. ReDecay: A novel approach to speed up the simulation at LHCb. *Eur. Phys. J.*, C78 (2018), p. 1009. DOI: 10.1140/epjc/s10052-018-6469-6. arXiv: 1810.10362 [hep-ex]

[109] Stanislaw Jadach, Johann H. Kuhn, and Zbigniew Was. TAUOLA: A Library of Monte Carlo programs to simulate decays of polarized tau leptons. *Comput. Phys. Commun.*, 64 (1990), pp. 275–299. DOI: 10.1016/0010-4655(91)90038-M

[110] D. Müller et al. ReDecay: a novel approach to speed up the simulation at LHCb. *The European Physical Journal C*, 78.12 (Dec. 2018). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-018-6469-6. URL: http://dx.doi.org/10.1140/epjc/s10052-018-6469-6

[111] Alex Rogozhnikov. Reweighting with Boosted Decision Trees. *Journal of Physics: Conference Series*, 762 (Oct. 2016), p. 012036. ISSN: 1742-6596. DOI: 10.1088/1742-6596/762/1/012036. URL: http://dx.doi.org/10.1088/1742-6596/762/1/012036

[112] Guy Wormser Davide Fazzini. *Measurement of $D^{**}\tau\nu$ production from B^0 and B^+ decays*. (LHCb internal note). Aug. 2020