

Measurements of production cross-sections and mixing of charm mesons at LHCb

A thesis submitted to The University of Manchester
for the degree of
Doctor of Philosophy
in the Faculty of Science and Engineering

Dominik Müller



The University of Manchester

Particle Physics Group
The School of Physics and Astronomy
2017



Abstract	5
Declaration	7
Copyright	9
Acknowledgements	13
Preface	15
1 Introduction	17
2 The giants' shoulders	19
2.1 The Standard Model of particle physics	20
2.2 Charm meson production	24
2.3 Mixing of neutral mesons	30
2.4 The decay amplitude – An isobar model	43
3 Experimental Setup	55
3.1 The Large Hadron Collider	55
3.2 The LHCb experiment	57
3.3 The LHCb detector	60
4 Fast event simulation in LHCb	71
4.1 Monte Carlo event generation – A brief overview	71
4.2 Monte Carlo samples at LHCb	74
4.3 The need for speed	77
4.4 ReDecay – An LHCb fast simulation solution	81
5 Charm production at 5 and 13 TeV	95
5.1 Analysis overview	96
5.2 The datasets	98
5.3 The event selection	102
5.4 Fitting – Extracting the prompt yield	107
5.5 Reconstruction and selection efficiencies	122
5.6 Systematic uncertainties	135
5.7 Results	148
5.8 Summary and impact	159
6 Mixing in charm four-body decays	161
6.1 Selection of the dataset	162
6.2 The mixing fit	175
6.3 The sensitivity study	187
6.4 Summary – impact on world average of x and y	194
7 A glimpse into the future	199
A References	201

Word count: 46000

Name of the university	The University of Manchester
Candidate name	Dominik Müller
Degree title	Doctor of Philosophy
Thesis title	Measurements of production cross-sections and mixing of charm mesons at LHCb
Submission date	25 th of September 2017
Resubmission date	22 nd of November 2017

The start of the second running period of the Large Hadron Collider in 2015, delivering proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, marks a new era in the study of particles containing charm quarks with samples of unprecedented size being recorded by the LHCb experiment. This thesis first presents cross-section measurements for D^0 , D^+ , D_s^+ , and D^{*+} at $\sqrt{s} = 13$ TeV, supplemented by a measurement of the same cross-sections at $\sqrt{s} = 5$ TeV, leading to very precise, double-differential cross-section ratios that are hugely beneficial for constraining theoretical uncertainties in quantum chromodynamics. This is followed by a study into the potential for extracting the parameters describing the oscillation of neutral D^0 mesons in $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ decays, exploiting the very large data samples recorded at 13 TeV. These studies show that this decay can be used to heavily constrain the mixing parameters, leading to a substantial improvement of the uncertainties on them when included in the world average. However, these large datasets create computational challenges due to the sheer amount of recorded data, which is especially true for the LHCb simulation effort. Therefore, this thesis also discusses work undertaken to increase the speed at which simulated samples can be created by one order of magnitude.

Declaration

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

Date and signature

1. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.
2. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.
3. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.
4. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property University IP Policy¹, in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations² and in The University’s policy on Presentation of Theses.

¹ see: <http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420>

² see: <http://www.library.manchester.ac.uk/about/regulations/>

*Für Mama und Papa.
Unendlichen Dank für alles.*

Acknowledgements

Over the past three years, countless people have done so much for me and their support really made this thesis into what is. Therefore, my deepest gratitudes to all of you.

To my supervisor Marco Gersabeck: thank you for all your explanations, discussions and your time. While you always made sure that we roughly stay on course, you always encouraged me to do things my way, even if that involved taking a slight detour (for the record: making plots pretty is not a detour!). And a big thank you for letting me invest a lot of time into the LHCb fast simulation effort, which worked out pretty well in the end. I can definitely say that you did an amazing job!

Another big thank you goes to Gloria Corti for her invaluable support and sharing her immense knowledge of the LHCb simulation framework, and for putting her trust into me.

To Chris and Alex if anyone deserves thanks, it's you two for braving the cross-section measurements with me. I am proud of what we achieved as a team, and despite the long hours, especially in the final months, it was always a joyful experience and our truly collaborative approach has probably spoilt me for the rest of my career. And a big thank you for making me learn python – turns out you were absolutely correct.

To Christoph, thank you for your help concerning the implementation of the mixing fit and for helping me hack in everything I needed. I am filled with gratitude that you had to do most of the work on Goofit, not me!

For financial support, I would like to take the opportunity and thank The University of Manchester and STFC.

I also want to thank Andy Buckley and Yvonne Peters for a pleasant viva.

Dear Fabian, thank you so invaluablely much for our long journey together, you have taught me so many skills over the years and I would not be here without you! Another thank you goes to all my office mates over the years which have been an important part of creating a pleasant working environment. I also want to apologise for my sometimes rather vocal displays of disappointment at the computer when it and I had disagreements.

Dearest Nicola, during my PhD, nothing compares to the joy of having met you and enjoying the privilege of your company. Your support means more to me than you can ever imagine and I am longing for our future life together.

Finally the biggest thank you goes to my family: your immeasurable support during almost two decades of pursuing education is far beyond anything one could ever wish for. Your support, especially during those times I have felt insecure, sets an example to which I strive to live up to for the rest of my life and I cannot thank you enough.

Und noch einen Nachsatz den jeder in meiner Familie versteht, und weil er gleichzeitig der Wichtigste ist: Ihr seid einfach nur wunderbar. Eure Unterstützung in den letzten zwei Jahrzehnten meiner Ausbildung hat mich zu dem Menschen gemacht der ich bin und ohne euch wäre ich niemals hier um diese Zeilen zu schreiben. Ich werde mein Bestes tun um dem Beispiel zu folgen das ihr gesetzt habt und ich kann euch nicht genug danken.

All the work in this thesis was performed as a member of the LHCb collaboration, consisting of around 800 researchers and engineers, many of whom have implicitly contributed to this thesis and it would not have been possible without all their contributions. The author mostly worked on the analysis of data after they were collected and processed by the global LHCb software chain. Additionally, the author contributed to the LHCb simulation software framework.

This work on the simulation framework is detailed in Chapter 4, where the presented ReDecay approach has been developed, implemented and extensively tested by the author.

Chapter 5 presents cross-section measurements of different charm-mesons at $\sqrt{s} = 13$ and 5 TeV. The 13 TeV measurement was a collaboration of one Master's student, Christopher Burr, two PhD students, Alex Pearce and the author, and four seasoned researchers, Sajan Easo, Marco Gersabeck, Vladimir Gligorov, and Patrick Spradlin. Almost all of the work was carried out by the three students in a truly collaborative manner. While different areas of focus were assigned, every analyst contributed significantly to all parts of the analysis. The author's main responsibilities were the evaluation of the efficiencies and systematic uncertainties, the development of the fitting approach and the correct treatment of statistical and systematic uncertainties when combining the different aspects of the measurement into the final cross-section result. The follow-up measurement at 5 TeV, albeit based on the 13 TeV one, received substantial updates and is the sole work of the author. All results are published [1, 2].

Chapter 6 presents sensitivity studies for the measurement of neutral charm meson mixing parameters, and is a collaboration between two PhD students, Christoph Hasse and the author. The author was responsible for all parts of the study except the implementation of the amplitude model in the fitting framework.

1. Introduction

The ongoing advancement of humankind is driven by the irrepressible aspiration to understand how things function and what they are made of. Elementary particle physics is at the forefront of this endeavour as it strives to describe the smallest constituents of matter. It is the key to our endless curiosity of understanding the most fundamental questions: when and – more importantly – *how* did the universe begin? Although the ultimate answer is yet unknown, the Standard Model of elementary particle physics is the current theoretical framework describing the elementary particles and their interactions. It is the culmination of work done over the past decades while constantly being improved and extended. Nonetheless, the core concepts of the model withstand all efforts trying to falsify them using experiments of ever increasing magnitude. Huge laboratories operate particle accelerators reaching higher and higher energies to test the predictions of the Standard Model and to advance our knowledge of elementary particles and their interactions. One of the apparent short-comings of the Standard Model is its lacking explanation for the observable matter-antimatter difference in the universe. While it does allow processes that behave differently under a charge-parity transformation (CP), these effects are too small to account for the visible difference in the universe itself. The violation of CP symmetry has been discovered in the decays in the strange and beauty quark sector. However, all of these observations are in agreement with the predictions of the Standard Model. This sets the stage for the charm quark for which no CP violation has yet been observed and Standard Model predictions are tiny. Thus any kind of substantial CP violation in charm physics is a sign of physics beyond the Standard Model and might be a hint towards the missing piece of the puzzle to explain the visible matter-antimatter difference.

In Summer 2015, the Large Hadron Collider (LHC) resumed operations after two years of technical overhauls and updates. Commonly referred to as Run 2, it delivers proton-proton collisions at $\sqrt{s} = 13$ TeV, almost twice the energy as before and the highest energy ever reached in human-made collisions of particles. This heralds a new age in charm physics where the LHCb experiment located at the LHC will record samples of charm decays of unheard-of sizes, allowing the ability to probe new physics in the charm sector with unprecedented precision. The main part of this thesis is constituted by a measurement of charm meson production cross-sections using Run 2 data at $\sqrt{s} = 13$ TeV as well as 5 TeV. Intended to confirm the expected large sample sizes that will be recorded in the future due to the increased cross-sections at 13 TeV, those measurements also constitute valuable input to QCD calculations where they can, especially in their combination, constrain the large theoretical uncertainties.

Additionally, this thesis explores the potential of performing a time- and model-dependent amplitude analysis of $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ decays. This decay can proceed via two different paths and their interference: the direct decay, which is suppressed, and the oscillation of a D^0 to an $\bar{D}^0$ followed by the favoured decay, $D^0 \rightarrow \bar{D}^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$. The study aims to highlight the tremendous increase in precision on the parameters governing this mixing (and subsequently indirect CP violation) achievable due to the large amount of data available in Run 2 of the LHC. However, these large samples of charm decays do not come without a cost: their analysis is increasingly strenuous on the available computing resources, which are not increasing in proportion with the data. An area especially affected by this is the simulation of events, which are crucial to understand the detector response and to extract the physics knowledge from the data. Therefore, this thesis also presents developments to improve the speed of the simulation, thus enabling the exploitation of the large datasets of charm decays to their fullest potential.

This thesis is structured as follows: Chapter 2 discusses the background of the presented work. This includes theoretical discussions as well as summaries of the current experimental status. The experimental setup, namely the LHC and the LHCb detector, are the topic of Chapter 3. The simulation of events in general and within LHCb in particular are discussed in Chapter 4, focusing on the author's work on increasing the speed of the event simulation by implementing a novel approach. Following this, the main body of work, namely the measurements of charm meson production cross-sections at $\sqrt{s} = 5 \text{ TeV}$ and 13 TeV , is presented in Chapter 5 before studies on the measurement of mixing parameters are explored in Chapter 6. Ultimately, this thesis concludes with a brief summary and an outlook for the future of charm mixing measurements at LHCb in Chapter 7.

Natural units $\hbar = c = e = 1$ are used throughout this document and all given values for particle properties are taken from Ref. [3] unless otherwise stated. Furthermore, charge conjugation is implied wherever certain particles or decay chains are discussed unless explicitly revoked.

2. The giants' shoulders

2.1	The Standard Model of particle physics	20
2.1.1	The elementary particles	20
2.1.2	The electroweak interaction	21
2.1.3	The strong interaction	22
2.1.4	The CKM matrix	23
2.2	Charm meson production	24
2.2.1	Parton density functions constraints	25
2.2.2	Atmospheric neutrino flux	27
2.2.3	Heavy ion physics	28
2.2.4	Theoretical predictions	29
2.3	Mixing of neutral mesons	30
2.3.1	Formalism	30
2.3.2	The decay $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$	35
2.3.3	Model-independent measurements in $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$. .	37
2.3.4	Model-dependent measurements in $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$. .	38
2.3.5	The breakdown of the approximations	41
2.3.6	The current world averages	42
2.4	The decay amplitude – An isobar model	43
2.4.1	Form factor $\mathcal{F}$	44
2.4.2	Spin factor $\mathcal{S}$	45
2.4.3	Lineshapes $\mathcal{L}$	51
2.4.4	Shortcomings and future developments	53

This section intends to give an overview of the context in which the measurements constituting the thesis are performed. Typically, this context refers to the Standard Model of particle physics which forms the underlying theoretical description that the measurements rely upon and thus is discussed briefly in Section 2.1. However, this connection is rather loose in flavour physics, which tends to be driven by phenomenological approaches and effective theories. Therefore, this section aims to give a narrative as to why the measurements are performed by highlighting the need of cross-section measurements with wide-ranging applications in Section 2.2. This is followed by an overview of the current theoretical and experimental approach to measurements of parameters describing mixing in the neutral charm system in Section 2.3 before a discussion on the construction of amplitude models needed for these measurements is given in Section 2.4.

2.1 The Standard Model of particle physics

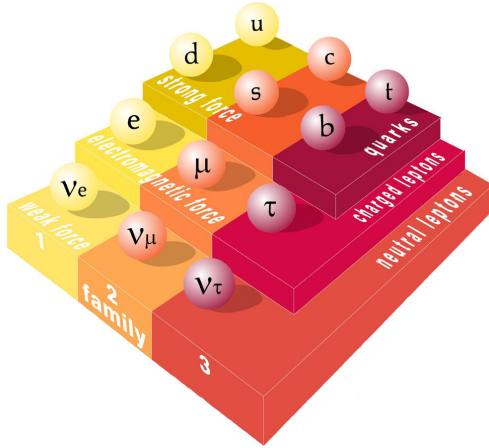


Fig. 2.1: Overview of all known elementary fermions and their affiliation to one of the three families.

The current Standard Model of elementary particle physics serves as a theoretical framework to understand the behaviour and properties of the microcosm. It is based on quantum field theories, which describe the interactions between fermions – half integer spin particles – mediated by bosons with integer spins. The Standard Model currently incorporates three fundamental forces: the electromagnetic, weak and strong force. The fourth fundamental force – gravity – is not included.

2.1.1 The elementary particles

Elementary particles are the basic building blocks of matter and have no known substructure. The Standard Model is currently comprised of twelve different spin- $1/2$ particles – the so-called fermions. One distinguishes between two distinct kinds of elementary fermions: quarks and leptons. An overview of all known elementary fermions is given in Figure 2.1. The set of six known quarks is further divided into two groups: up- and down-type quarks, while the known set of leptons is separated into electrically charged and neutral particles, known as the neutrinos. All particles within one of these four groups share the same quantum numbers. As visualised in Figure 2.1, a second ordering scheme is implied which places the known elementary fermions in one of three generations. All observed stable matter is solely built by the members of the first generation – the electron, the electron neutrino, the up- and the down-quark. The particles in the other two families are unstable and can only be observed for a brief moment after their creation in highly energetic processes, e.g. at a particle collider. Furthermore, due to their significantly higher masses, the charm, bottom and top quark are also referred to as the heavy quarks.

The aforementioned separation into distinct types of fermions is made due to their different quantum numbers. These quantum numbers determine the coupling strength of the fermions to the different forces and hence also which forces can act upon them.

Electric charge

The defining property that separates the four different types of particles is the electric charge Q . Charged leptons carry $Q = -1$ whereas their counterpart – the neutrinos

– do not have an electric charge and hence will not participate in electromagnetic processes. The up-type quarks carry an electric charge of $Q = +2/3$ whereas the down-type quarks carry $Q = -1/3$.

Isospin

The isospin, typically described by its third component T_3 , is the charge of the weak interaction. Due to the vector minus axial vector ($V - A$) structure of the charged current interaction, it only couples to left-handed particles. The handedness refers to the chirality of the particle, which is a quantity invariant under Lorentz boosts and obtained by using the chirality operator γ^5 on a spinor representing the particle. Therefore, the left-handed leptons and quarks are ordered in isospin doublets ($T_3 = \pm 1/2$) while the right-handed particles are present in isospin singlets ($T_3 = 0$). Neutrinos solely couple to $V - A$ structures and are assumed to be massless in the Standard Model. Conclusively, only left-handed neutrinos participate in the interactions described by the Standard Model. However, the observation of neutrino oscillation indicates that neutrinos are indeed massive [4, 5]. This implies the existence of physics beyond the Standard Model.

Colour charge

The colour charge is the charge of the strong interaction and only carried by quarks. It can take three different values which are called, in analogy to visual colours, red, blue and green. Therefore, leptons do not participate in the strong interaction.

For every particle there exists an antiparticle of equal mass but inverted quantum numbers. The interactions are included in the Standard Model by requiring the Lagrangian of a free particle to be invariant under local gauge transformations. They are described in the following.

2.1.2 The electroweak interaction

The electromagnetic interaction is included in the Lagrangian by requiring it to be invariant under the local gauge transformation $U(1)_Q$. To achieve the gauge invariance, one has to modify the derivative and introduce a local gauge field A_μ corresponding to the photon, the mediating boson of the electromagnetic interaction. The same can be done to include the weak interaction by using $SU(2)_L$ as the symmetry group. This introduces three additional fields $W_\mu^{1,2,3}$ of which W_μ^1 and W_μ^2 are linearly combined to form the W^+ and W^- bosons. However, no correspondence to any known particle has been found for the third component W_μ^3 .

A solution to this problem was first developed by Weinberg, Salam and Glashow [6–8], who proposed to unify the descriptions of electromagnetic and weak interactions into the electroweak interaction described by a common gauge symmetry group:

$$U(1)_Y \times SU(2)_L$$

where the hypercharge $Y = 2Q - T_3$ is the generator of the $U(1)_Y$ interaction, which introduces a field B_μ to the Lagrangian to ensure gauge invariance. Mixing W_μ^3 and B_μ using a unitary matrix characterised by the Weinberg-angle θ_W , leads to two gauge fields:

$$\begin{aligned} A_\mu &= W_\mu^3 \sin(\theta_W) + B_\mu \cos(\theta_W) \text{ and} \\ Z_\mu &= W_\mu^3 \cos(\theta_W) - B_\mu \sin(\theta_W) . \end{aligned}$$

A_μ corresponds to the photon and only couples to electrically charged particles while the new gauge field Z_μ corresponds to the Z^0 boson, which is the mediating particle of the weak neutral-current interaction. The Higgs mechanism [9, 10], which was experimentally confirmed in 2012 [11, 12], spontaneously breaks the symmetry of the electroweak Lagrangian, which generates mass terms for the W and Z boson field, while none is obtained for the A field and thus keeping the photon massless.

2.1.3 The strong interaction

The strong interaction is described by quantum chromodynamics and is given by the gauge symmetry group $SU(3)_C$. It describes the interactions of particles carrying a colour charge. The eight generators of $SU(3)_C$ result in eight different gluons. Those are the mediating particles in the strong interaction. Furthermore, they can couple to each other due to carrying colour charges themselves. The so-called *confinement* requires quarks to be bound in colourless hadrons because states with a net colour charge are not invariant under $SU(3)_C$ transformations. Therefore, quarks cannot be observed as free particles in nature. These colourless hadrons are either mesons $|q\bar{q}\rangle$ combining two quarks carrying a colour and corresponding anticolour charge, or baryons $|qqq\rangle$ combining all three colours or anticolours. Furthermore, exotic states, which could be explained as four-quark and five-quark bound states, have been observed [13–15]. The strong coupling, which is bound by gluons, increases with increasing distance between two quarks which are bound by gluons in contrast to the weak coupling that decreases with increasing distance. If two quarks are separated, there is a point where it is energetically more favourable to produce an additional quark-antiquark pair instead of increasing the distance any further. This results in *hadronisation* – a quark in the final state of any process spawns quark-antiquark pairs until the energy is too low and bound states between these quarks are able

to form. These bound states are the aforementioned hadrons – the actual particles which are observable in experiments.

2.1.4 The CKM matrix

Experimental results indicate that the quark states participating in the charged current interaction given by the $SU(2)_L$ are not identical to the mass eigenstates (which correspond to the states coupling to the electromagnetic and strong force). This is described by the unitary CKM matrix [16, 17], which connects the weak down-type eigenstates $|d'\rangle$ to the flavour eigenstates $|d\rangle$:

$$\begin{pmatrix} |d'\rangle \\ |s'\rangle \\ |b'\rangle \end{pmatrix} = V_{\text{CKM}} \begin{pmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{pmatrix}. \quad (2.1)$$

This has important consequences: the charged-current interactions are unique as they can couple quarks of different generations by accessing the off-diagonal elements in the CKM matrix. The square of the absolute value for each entry of the CKM matrix gives the probability for finding a specific flavour in a given weak eigenstate. As the magnitudes of the off-diagonal elements in the matrix are significantly smaller than those on the diagonal, processes involving off-diagonal transitions are called Cabibbo-suppressed. Furthermore, a 3×3 unitary matrix allows for a complex phase, which introduces CP violation to the weak decays. This CP violation was discovered in B^0 decays [18, 19]. The CKM matrix has four free parameters and is commonly expressed by three rotation angles, θ_{13} , θ_{12} and θ_{23} as well as a complex phase δ_{13} , leading to a combination of three rotation matrices:

$$V_{\text{CKM}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{13}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{13}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (2.2)$$

where s_{ij} and c_{ij} are the sine and cosine of the rotation angle ϕ_{ij} . A common parameterisation of the CKM matrix is also given by the Wolfenstein parameterisation [20], which is centred around the parameter

$$\lambda = s_{12}. \quad (2.3)$$

The additional three parameters A , ρ and η are defined as

$$s_{23} = A\lambda^2, \text{ and} \quad (2.4)$$

$$s_{13}e^{-i\delta_{13}} = A\lambda^3(\rho - i\eta). \quad (2.5)$$

The benefit of choosing this parameterisation is the magnitude of the different parameters. While λ is approximately 0.22, the remaining parameters are all $\mathcal{O}(1)$, which allows the expansion of the CKM matrix in orders of λ to quickly judge the magnitudes of different quark transitions and which of these transitions are most sensitive to the effect of the CP violating phase. The expansion to λ^3 in the real parts and to λ^5 in the imaginary parts of the CKM matrix yields:

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta + i\eta\frac{1}{2}\lambda^2) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 - i\eta A^2\lambda^4 & A\lambda^2(1 + i\eta\lambda^2) \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \quad (2.6)$$

Therefore, up to this expansion, CP violation enters through the complex terms in V_{td} , V_{cs} , V_{ub} and V_{cb} . For charm, the CKM matrix is effectively reduced to a 2×2 matrix, significantly suppressing CP violation for charm mesons in the SM, and thus any substantial amount of CP violation in the charm sector is a clear sign of new physics.

2.2 Charm meson production

Chapter 5 presents measurements of charm meson production cross-sections in proton-proton collisions at centre-of-mass energies of $\sqrt{s} = 5$ TeV and 13 TeV. These cross-section measurements have previously been performed in different kinematic regions for various centre-of-mass energies in the TeV range. Measurements by the CDF experiment at the Tevatron cover the central rapidity region $|y| < 1$ and transverse momenta, p_T , between 5.5 GeV/c and 20 GeV/c at $\sqrt{s} = 1.96$ TeV in $p\bar{p}$ collisions [21]. At the LHC, charm cross-sections in pp collisions have been measured in the $|y| < 0.5$ region for $p_T > 1$ GeV/c at $\sqrt{s} = 2.76$ TeV and for $p_T > 0$ GeV/c at $\sqrt{s} = 7$ TeV by the ALICE experiment [22–25], and for pseudorapidity $|\eta| < 2.1$ in the p_T region $3.5 < p_T < 100$ GeV/c at $\sqrt{s} = 7$ TeV by the ATLAS experiment [26]. The LHCb experiment has recorded the world's largest dataset of charm hadrons to date and this has led to numerous high-precision measurements of their production and decay properties. LHCb measured the cross-sections in the forward region $2.0 < y < 4.5$ for $0 < p_T < 8$ GeV/c at $\sqrt{s} = 7$ TeV [27].

The LHCb experiment's unique acceptance, which is discussed in Section 3.2, allows for these cross-sections to probe relatively large rapidities with values reaching up to 4.5, where the rapidity of the meson is defined as:

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}. \quad (2.7)$$

While the existence of charm mesons is hardly a surprise and the abundance in pp collisions is well known, precise measurements, especially at different centre-of-mass

energies in the forward region, are an important ingredient or constraint for other applications. These applications cover a wide range of scenarios, for example, they are used to constrain theoretical uncertainties in QCD calculations which, in turn, help to reduce experimental uncertainties in many areas, for example high-energy neutrino physics. The measurements in pp collisions are also used as reference inputs for analyses of heavy ion collisions. This section discusses the different applications and describes the theoretical predictions which are compared to data and is discussed in Chapter 5.

2.2.1 Parton density functions constraints

The proton is a composite particle comprised of two up and one down quark tightly bound by the strong force. These partons form the so-called valence quarks. They are accompanied by the so-called sea quarks, which can be momentarily created as fluctuations in the strong field. The high energy collisions at the Large Hadron Collider probe this substructure of the protons and the actual interactions happening between the different partons of the colliding protons. The total cross-section for a given process, e.g. $pp \rightarrow c\bar{c}$, is then given by a combination of all allowed parton combinations and can be written as:

$$d\sigma(pp \rightarrow c\bar{c}) = \sum_{i,j} \frac{1}{x_i x_j s} |M_{ij \rightarrow c\bar{c}}|^2 f_i(x_i, Q^2) f_j(x_j, Q^2) dx_i dx_j d\Phi_n, \quad (2.8)$$

where the sum is taken over all allowed permutations of partons, $M_{ij \rightarrow c\bar{c}}$ is the matrix element for the specific process, and Φ_n is a point in the phase space of the process. The momentum fractions x_i and x_j give the fraction of the proton's momentum carried by parton i and j respectively and f_i and f_j are the respective parton density functions (PDF), which give the probabilities for the different partons as functions of the momentum fraction x . Additionally, the PDFs depend on the energy scale Q^2 at which they are being probed. As the dynamics inside a proton are driven by low energetic, soft QCD effects, perturbation theory cannot be applied and thus no calculations can be performed to get the value of f for a given x and Q^2 . Therefore, the values need to be extracted from measurements. However, this poses an immediate problem: different experiments cover different regions in Q^2 . Figure 2.2a illustrates the different regions, which do not necessarily overlap and thus it is not immediately clear how to apply already known PDFs to new experiments, e.g. how the PDFs derived from HERA data can be used in LHC analyses which cover a much higher Q^2 range. The Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) differential equations [29–31] allow to compute the evolution of the PDF in Q^2 from any known reference value at Q_0^2 , thus allowing the transfer of a measured PDF to different Q^2 regions. Nonetheless, the dependence on x still needs to be extracted from measurements.

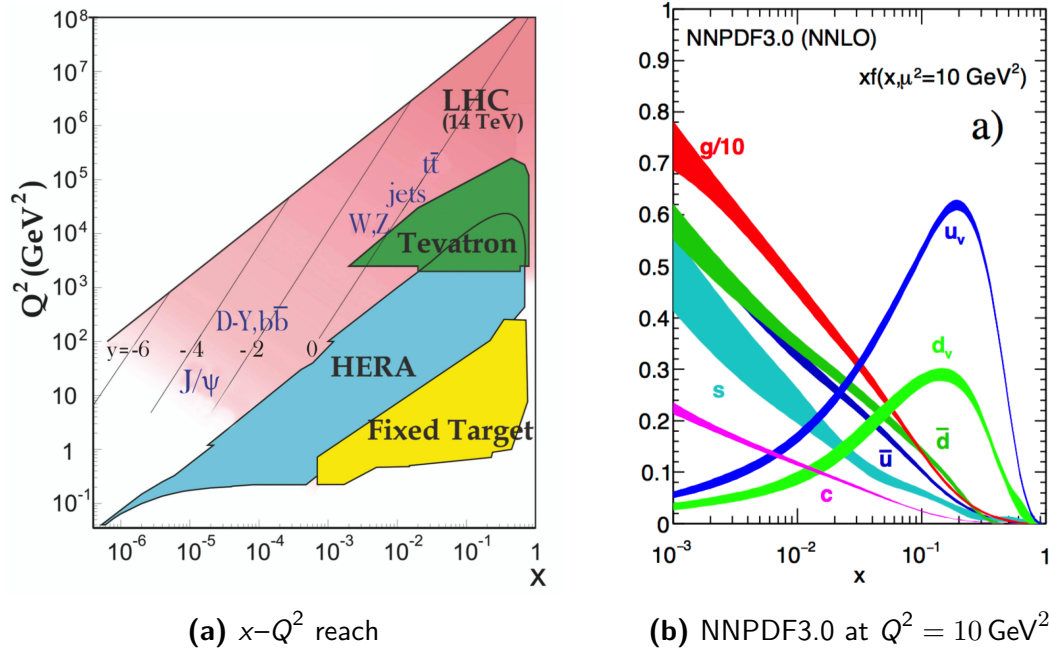


Fig. 2.2: (a) Region probed by various experiments, depending on the process and rapidity where charm meson production occupies the Q^2 region starting at $(2m(D))^2 \approx 16 \text{ GeV}^2$. (b) NNPDF3.0 set evaluated at roughly this momentum transfer, obtained from a global analysis described in Ref. [28]. Both figures are taken from Ref. [3].

When extracting the PDF for the different partons from experimental data, some choices need to be made that are to some extent arbitrary due to the phenomenological treatment of soft QCD effects. These choices include the decision on which experimental data to consider, the level of expansion in the coupling constant when computing the perturbative matrix element and the treatment of soft QCD effects. For simplicity, Equation (2.8) ignored the additional contributions due to the fact that an experiment does not observe a $c\bar{c}$ final state but measures only the resulting hadrons after hadronisation. This requires the inclusion of fragmentation models and their uncertainties. Due to these different choices that need to be made, several different theory groups published different sets of PDFs. For example, the theoretical predictions which are compared to the measured cross-sections in this thesis use the NNPDF3.0 [28] (also shown in Figure 2.2) and CT10 [32] sets.

Figure 2.2 reveals that LHCb measurements, which are able to cover rapidity values up to 4.5, are potentially very useful to extract the PDF for low values of x down to 10^{-6} , a region that no other experiment can probe. In this region, processes such as $pp \rightarrow c\bar{c}$ are completely dominated by the gluon contribution and measurements of charm meson production can constrain the very large uncertainties in the gluon PDF. Using the measured production cross-sections of D^0 and D^+ at $\sqrt{s} = 7 \text{ TeV}$ by LHCb [27], Gauld *et al.* [33] managed to significantly reduce the uncertainty of the gluon PDF in the NNPDF3.0 set with the results shown in Figure 2.3. This is expected to further improve drastically with the production measurements presented

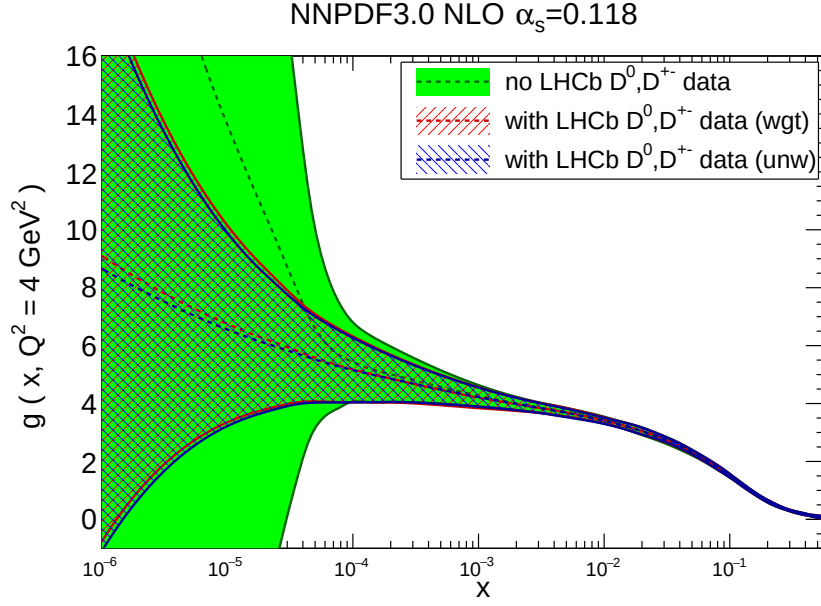


Fig. 2.3: Impact of the LHCb charm meson production data at $\sqrt{s} = 7$ TeV on the precision of the gluon PDF at low x . Taken from Ref. [33]

in this thesis and has implications beyond particle physics at the LHC. In fact, Gauld *et al.* have updated their calculations with the cross-sections presented in this thesis and the results are discussed in Section 5.8. One application benefiting from an improved gluon PDF is presented in the following.

2.2.2 Atmospheric neutrino flux

The previously described improvements are very beneficial for experiments measuring neutrinos of astrophysical origin. Using the IceCube neutrino observatory [34], the IceCube collaboration studies neutrinos with energies up to several dozen PeV. Unlike the measurements presented in this thesis, where backgrounds in the measurement can usually be understood and separated by combining all the products of the studied decay and looking at variables like the invariant mass, such an approach is not possible for neutrino observatories. Therefore, it is crucial to understand the rate of expected background events. One such source of background is referred to as atmospheric background. While this background is still comprised of neutrinos, their origins are collisions of highly energetic cosmic rays (mostly protons) with the planet's atmosphere. In these collisions, abundantly produced pions and kaons can decay to neutrinos, which could in principle be detected by the IceCube detector. However, due to their very high energies from the initial boost of the cosmic ray and their long life-times, pions and kaons elastically scatter and lose energy while traversing the atmosphere, which alters and softens the energy spectrum, effectively suppressing their contribution. This is not true for the produced charm mesons, which have lifetimes so short that the energy spectrum of the resulting neutrino is almost unchanged. Figure 2.4 shows the distribution of deposited energy in the IceCube detector

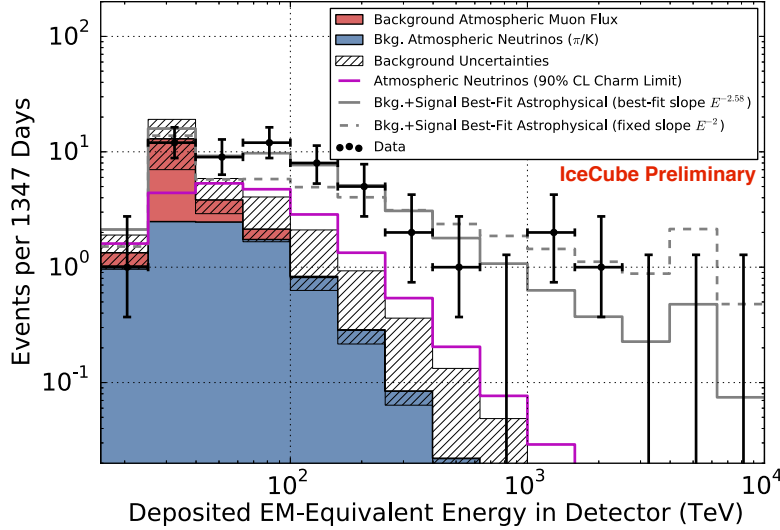


Fig. 2.4: Distribution of deposited energy in the IceCube detector compared to the different expected backgrounds. As can be seen, the contribution from charm mesons becomes the dominant background at high energies. Taken from Ref. [35].

compared to different sources of background. As can be seen, the contributions from charm mesons become the dominant background towards higher energies. However, the histogram only gives an upper limit due to the large uncertainties on those predictions. Comparing the cosmic-ray-with-atmosphere interactions, which are essentially fixed target experiments, to symmetric proton-proton collisions, production measurements performed by LHCb at 5, 7 and 13 TeV correspond to cosmic ray energies of 12.5, 24.5 and 84.5 PeV. Therefore, measurements performed by LHCb probe the same part of the gluon PDF and thus the improvements of these PDFs from LHCb data help to constrain the atmospheric neutrino background [33].

2.2.3 Heavy ion physics

In the study of collisions between heavy ions or heavy ions with protons, the difference in measured cross-sections when compared to the equivalent proton-proton collisions is of great interest to understand the effects the nuclear structure has on the cross-sections themselves. As an example, the nuclear modification factor in proton-lead collisions (which are an alternative operation mode of the Large Hadron Collider) is given by:

$$R_{pPb} \equiv \frac{1}{A} \cdot \frac{\sigma_{pPb}(y^*, p_T, \sqrt{s}) / (dy^* dp_T)}{\sigma_{pp}(y^*, p_T, \sqrt{s}) / (dy^* dp_T)}, \quad (2.9)$$

with $A = 208$ being the atomic mass number of lead and y^* being the rapidity of the proton in the frame of the lead nucleus. For these quantities, reference measurements need to be performed at the equivalent pp centre-of-mass energy. Figure 2.5 shows preliminary results relying on extrapolated pp cross-sections for

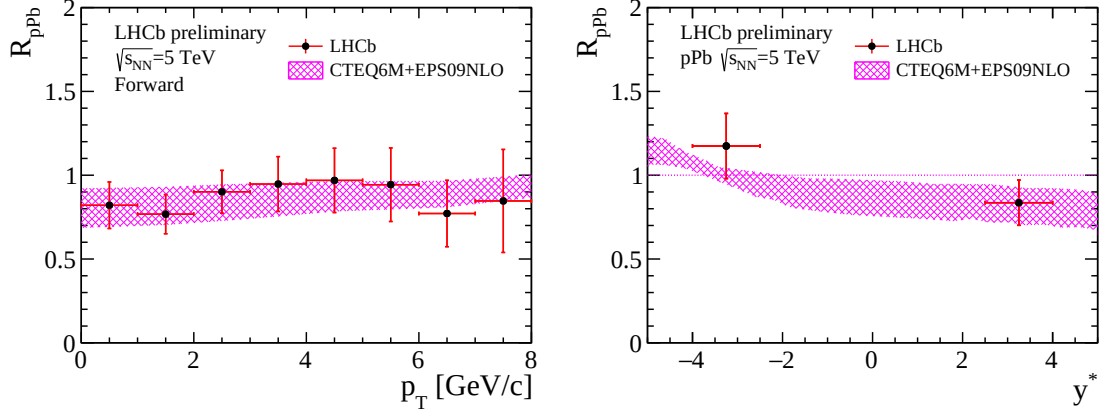


Fig. 2.5: Measurements of the nuclear modification factor R_{pPb} by LHCb [36] using data recorded at 5 TeV. The presented results are preliminary as the pp cross-sections at 5 TeV have been extrapolated from the measured values at 7 TeV and 13 TeV and an update is in preparation utilising the results presented in this thesis.

pPb collisions recorded at a nuclear centre-of-mass energy of 5 TeV. These results will be updated with the results presented in this thesis.

2.2.4 Theoretical predictions

Theoretical calculations for charm meson production cross-sections in pp collisions at 5 TeV and 13 TeV, detailed in Refs. [33] (POWHEG+NNPDF3.0L), [37] (FONLL) and [38] (GMVFNS), have been provided by the authors. All sets of calculations are performed at next-to-leading order (NLO precision), and each includes estimates of theoretical uncertainties due to the renormalisation and factorisation scales. The theoretical uncertainties provided with the POWHEG+NNPDF3.0L predictions also include contributions due to uncertainties in the effective charm quark mass and the parton distribution functions.

The FONLL predictions are provided in the form of D^0 , D^+ , and D^{*+} production cross-sections for pp collisions at 5 TeV and 13 TeV for different (p_T, y) bins. Predicted ratios of these cross-sections, $R_{13/5}$, and with respect to 7 TeV, $R_{13/7}$, are also supplied. The calculations use the NNPDF3.0 NLO [28] parton densities. These FONLL calculations of the meson differential production cross-sections assume $f(c \rightarrow D) = 1$ and are multiplied by the transition probabilities measured at e^+e^- colliders for comparison to the current measurements. No dedicated FONLL cross-section calculation for D_s^+ production is available.

The POWHEG+NNPDF3.0L predictions provide D^0 , D^+ , D_s^+ (only 5 TeV) and D^{*+} integrated production cross-sections for pp collisions in the same p_T and y binning used for the measurements. The ratios $R_{13/7}$ and $R_{13/5}$ are also predicted for D^0 , D^+ and D_s^+ . They are obtained with POWHEG [39] matched to Pythia 8 [40] parton showers and an improved version of the NNPDF3.0 NLO parton distribution function

set designated NNPDF3.0+LHCb [33]. To produce this improved set, the authors of Ref. [33] weight the NNPDF3.0 NLO set in order to match FONLL calculations to the LHCb charm cross-section measurements at 7 TeV [27]. This results in a significant reduction in the uncertainties for the gluon distribution function at small momentum fractions x . Predictions for the integrated cross-section are provided for each meson.

The GMVFNS calculations include theoretical predictions for the absolute cross-sections for all mesons studied in this analysis. Results are provided for $p_T > 3 \text{ GeV}/c$. Additionally, central values for $R_{13/5}$ are provided. Here the CT10 [32] set of parton distributions is used. The GMVFNS theoretical framework includes the convolution with fragmentation functions describing the transition $c \rightarrow H_c$ that are normalised to the respective total transition probabilities [41, 42]. The fragmentation functions are taken from a fit to production measurements at e^+e^- colliders, where no attempt is made to separate direct production and feed-down from higher resonances.

2.3 Mixing of neutral mesons

Chapter 6 presents studies on the potential for measuring the parameters describing the oscillation of $D^0 \leftrightarrow \bar{D}^0$ using four-body decays. Therefore, this section discusses the mixing formalism for neutral mesons and presents results for a model-independent extraction of the parameters describing the oscillation from $D^0 \rightarrow K^+\pi^-\pi^+\pi^-$ decays. The same decay is also used in the studies presented in this thesis, albeit following a model-dependent approach whose intricacies are also highlighted.

2.3.1 Formalism

In neutral meson systems, the weak force gives rise to oscillations between a meson and its antiparticle, as the weak eigenstates of the quarks differ from their mass eigenstates as previously discussed. The following discussion focuses on the oscillations in the neutral D^0 meson system, which occur via box diagrams as depicted in Figure 2.6. These transitions occur via down-type quarks on internal lines of the diagram resulting in at least two suppressed quark transitions accessing off-diagonal elements of the CKM matrix. Inspecting the leading orders of the accessed CKM matrix elements for the different down-type quarks, one finds $V_{cd}V_{ud}^* \approx -\lambda$, $V_{cs}V_{us}^* \approx \lambda$, and $V_{cb}V_{ub}^* \propto \lambda^5$. Therefore, contributions from the b quark on the internal lines are strongly suppressed and negligible even if the moderate enhancement due to the b quark mass is considered. This enhancement of heavy quarks is found in calculations of the loop [43] and is caused by the very heavy W boson also occurring in the loop. As the masses of s and u quarks are similar and the CKM factor only differs by the sign, these two contributions almost cancel each other, thus leading to a GIM

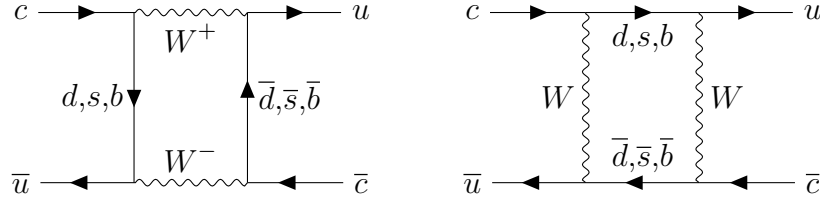


Fig. 2.6: Box diagrams for the oscillation of $D^0 \leftrightarrow \bar{D}^0$ via intermediate down-type quarks on internal lines.

suppression [44] of the oscillation, making it experimentally challenging to measure. This is unlike the neutral B system where similar transitions occur via up-type quarks on the internal lines. While the CKM factors for all three up-type quarks are of similar order in λ , the contribution from the t quark is substantially increased due to its very high mass compared to all other quarks [43]. This results in experimentally well established measurements of B meson oscillation [3]. Nonetheless, mixing of D^0 mesons has been observed in $D^0 \rightarrow K^+ \pi^-$ decays [45] as well as in combinations of other measurements.

The neutral charm meson system can be described as a two quantum state system governed by Schrödinger's equation for an effective non-hermitian Hamiltonian $\mathcal{H}$:

$$i \frac{d}{dt} |\psi\rangle = \mathcal{H} |\psi\rangle = \left(\mathbf{M} + \frac{i}{2} \mathbf{\Gamma} \right) |\psi\rangle, \quad (2.10)$$

with the hermitian matrices $\mathbf{M}$ and $\mathbf{\Gamma}$:

$$\mathbf{M} = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^* & M_{11} \end{pmatrix} \quad \text{and} \quad \mathbf{\Gamma} = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma_{11} \end{pmatrix}, \quad (2.11)$$

where CPT invariance is assumed, resulting in the elements on the diagonal being identical. In general, the off-diagonal elements of the matrices $\mathbf{M}$ and $\mathbf{\Gamma}$ are not zero and thus the two states $|D^0\rangle$ and $|\bar{D}^0\rangle$ are not identical to the physical eigenstates $|D_1\rangle$ and $|D_2\rangle$. Their eigenvalues for the mass $m_{1,2}$, decay width $\Gamma_{1,2}$ and the resulting eigenvalues for $\mathcal{H}$ are:

$$\lambda_{1,2} = m_{1,2} + \frac{i}{2} \Gamma_{1,2}, \quad (2.12)$$

where

$$m_{1,2} = M_{11} \pm |M_{12}| \quad \text{and} \quad \Gamma_{1,2} = \Gamma_{11} \pm |\Gamma_{12}|. \quad (2.13)$$

These eigenstates of $\mathcal{H}$ can be expressed as linear combinations of the flavour eigenstates $|D^0\rangle$ and $|\bar{D}^0\rangle$:

$$|D_{1,2}\rangle = p |D^0\rangle \pm q |\bar{D}^0\rangle, \quad (2.14)$$

where $p, q \in \mathbb{C}$ and the normalisation of $\langle D_{1,2} | D_{1,2} \rangle$ implies $|p|^2 + |q|^2 = 1$. Solving $(\mathcal{H} - \lambda_{1,2} \mathbf{I}) = 0$ yields an additional constraint relating them to the off-diagonal elements in the Hamiltonian:

$$\frac{p}{q} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} = \left| \frac{p}{q} \right| e^{i\phi} . \quad (2.15)$$

Using the physical eigenstates, the evolution of the state $|D_{1,2}(t=0)\rangle$ in time is then given by:

$$|D_{1,2}(t)\rangle = \hat{U}_{1,2}(t) |D_{1,2}(0)\rangle , \quad (2.16)$$

where $\hat{U}_{1,2}(t)$ are the time evolution operators for the physical eigenstates given by:

$$\hat{U}_{1,2}(t) = e^{-i\lambda_{1,2}t} = e^{-im_{1,2}t} e^{-\frac{i}{2}\Gamma_{1,2}t} . \quad (2.17)$$

It is convenient to introduce here the average mass and decay width of the physical eigenstates,

$$m \equiv \frac{m_1 + m_2}{2} \quad \text{and} \quad \Gamma \equiv \frac{\Gamma_1 + \Gamma_2}{2} , \quad (2.18)$$

and to define the differences of the respective eigenvalues as dimensionless quantities, which are the main observables when studying neutral meson mixing:

$$x \equiv \frac{m_1 - m_2}{\Gamma} \quad \text{and} \quad y \equiv \frac{\Gamma_1 - \Gamma_2}{2\Gamma} . \quad (2.19)$$

With these definitions, the time evolution operator can be transformed into a form more useful later on:

$$\hat{U}_{1,2}(t) = e^{-imt} e^{-\frac{1}{2}\Gamma t} e^{\mp \frac{1}{2}(y+ix)\Gamma t} . \quad (2.20)$$

When studying time dependent decay rates one in fact studies the time evolution of the flavour eigenstates, which is obtained by expressing the flavour eigenstates as combinations of mass eigenstates (by inverting Equation (2.14)):

$$|D^0(t)\rangle = \frac{1}{2p} \left(\hat{U}_1(t) |D_1(0)\rangle + U_2(t) |D_2(0)\rangle \right), \quad (2.21)$$

$$|\bar{D}^0(t)\rangle = \frac{1}{2q} \left(\hat{U}_1(t) |D_1(0)\rangle - U_2(t) |D_2(0)\rangle \right). \quad (2.22)$$

Expressing the mass eigenstates at $t = 0$ again as flavour eigenstates yields the final equations for the time evolution of the flavour states:

$$|D^0(t)\rangle = \frac{\hat{U}_1(t) + \hat{U}_2(t)}{2} |D^0\rangle + \frac{q}{p} \times \frac{\hat{U}_1(t) - \hat{U}_2(t)}{2} |\bar{D}^0\rangle, \quad (2.23)$$

$$|\bar{D}^0(t)\rangle = \frac{\hat{U}_1(t) + \hat{U}_2(t)}{2} |\bar{D}^0\rangle + \frac{q}{p} \times \frac{\hat{U}_1(t) - \hat{U}_2(t)}{2} |D^0\rangle. \quad (2.24)$$

In a measurement, this time dependence is studied by measuring the time-dependent decay rate to a certain final state $\langle f|$. The time dependent decay amplitudes are then given as:

$$A_f(t) \equiv \langle f|\mathcal{H}|D^0(t)\rangle = \frac{\hat{U}_1(t) + \hat{U}_2(t)}{2} A_f + \frac{q}{p} \times \frac{\hat{U}_1(t) - \hat{U}_2(t)}{2} \bar{A}_f, \quad (2.25)$$

$$\bar{A}_f(t) \equiv \langle f|\mathcal{H}|\bar{D}^0(t)\rangle = \frac{\hat{U}_1(t) + \hat{U}_2(t)}{2} \bar{A}_f + \frac{q}{p} \times \frac{\hat{U}_1(t) - \hat{U}_2(t)}{2} A_f, \quad (2.26)$$

where $A_f \equiv \langle f|\mathcal{H}|D^0\rangle$ and $\bar{A}_f \equiv \langle f|\mathcal{H}|\bar{D}^0\rangle$. Depending on the final state, these decay amplitudes are usually a function of the position in phase-space Φ_n . For readability, this dependence is not explicitly noted here. Additionally, a second set of equations exists for the CP conjugated decay that can be computed analogously to the calculations presented here. The measured decay rate as a function of the decay time is proportional to the squared amplitude $|A_f|^2$. Computing this can be simplified by reordering the right-hand side of Equation (2.25):

$$A_f(t) = \frac{\hat{U}_1(t)}{2} \left(A_f + \frac{q}{p} \bar{A}_f \right) + \frac{\hat{U}_2(t)}{2} \left(A_f - \frac{q}{p} \bar{A}_f \right). \quad (2.27)$$

With this, the amplitude squared is computed as:

$$\begin{aligned} 4|A_f|^2 &= 2A_f^* \times 2A_f \\ &= \left| \hat{U}_1(t) \right|^2 \left(A_f + \frac{q}{p} \bar{A}_f \right) \left(A_f^* + \left[\frac{q}{p} \bar{A}_f \right]^* \right) \\ &\quad + \left| \hat{U}_2(t) \right|^2 \left(A_f - \frac{q}{p} \bar{A}_f \right) \left(A_f^* - \left[\frac{q}{p} \bar{A}_f \right]^* \right) \\ &\quad + \hat{U}_1(t) \hat{U}_2^*(t) \left(A_f + \frac{q}{p} \bar{A}_f \right) \left(A_f^* - \left[\frac{q}{p} \bar{A}_f \right]^* \right) \\ &\quad + \hat{U}_1^*(t) \hat{U}_2(t) \left(A_f^* + \left[\frac{q}{p} \bar{A}_f \right]^* \right) \left(A_f - \frac{q}{p} \bar{A}_f \right). \end{aligned} \quad (2.28) \quad (2.29)$$

Given a set of identities for the combinations of the time evolution operators,

$$\left| \hat{U}_1(t) \right|^2 = e^{-\Gamma t} e^{-y\Gamma t}, \quad (2.30)$$

$$\left| \hat{U}_2(t) \right|^2 = e^{-\Gamma t} e^{+y\Gamma t}, \quad (2.31)$$

$$\hat{U}_1(t) \hat{U}_2^*(t) = e^{-\Gamma t} e^{-ix\Gamma t}, \quad (2.32)$$

$$\hat{U}_1^*(t) \hat{U}_2(t) = e^{-\Gamma t} e^{+ix\Gamma t}, \quad (2.33)$$

and some general identities for any two complex numbers a and b , namely:

$$2i\Im(ab^*) = (ab^* - ba^*) = -2i\Im(a^*b) , \quad (2.34)$$

$$2\Re(ab^*) = (ab^* + ba^*) = 2\Re(ba^*) . \quad (2.35)$$

This can be used to simplify Equation (2.29):

$$\begin{aligned} 4e^{\Gamma t}|A_f|^2(t) = & e^{-y\Gamma t} \left\{ |A_f|^2 + \left| \frac{q}{p} \bar{A}_f \right|^2 + 2\Re \left(A_f \left[\frac{q}{p} \bar{A}_f \right]^* \right) \right\} \\ & + e^{+y\Gamma t} \left\{ |A_f|^2 + \left| \frac{q}{p} \bar{A}_f \right|^2 - 2\Re \left(A_f \left[\frac{q}{p} \bar{A}_f \right]^* \right) \right\} \\ & + e^{-ix\Gamma t} \left\{ |A_f|^2 - \left| \frac{q}{p} \bar{A}_f \right|^2 - 2i\Im \left(A_f \left[\frac{q}{p} \bar{A}_f \right]^* \right) \right\} \\ & + e^{+ix\Gamma t} \left\{ |A_f|^2 - \left| \frac{q}{p} \bar{A}_f \right|^2 + 2i\Im \left(A_f \left[\frac{q}{p} \bar{A}_f \right]^* \right) \right\}, \end{aligned} \quad (2.36)$$

and finally, the exponential functions can be accumulated into trigonometric and hyperbolic functions, revealing the oscillating components in the time dependent decay rate:

$$\begin{aligned} |A_f|^2(t, \Phi_n) = & \frac{1}{2} e^{-\Gamma t} \left[\left\{ |A_f(\Phi_n)|^2 + \left| \frac{q}{p} \bar{A}_f(\Phi_n) \right|^2 \right\} \cosh(y\Gamma t) \right. \\ & + \left\{ |A_f(\Phi_n)|^2 - \left| \frac{q}{p} \bar{A}_f(\Phi_n) \right|^2 \right\} \cos(x\Gamma t) \\ & - 2\Re \left\{ \left[\frac{q}{p} \right]^* A_f(\Phi_n) \bar{A}_f^*(\Phi_n) \right\} \sinh(y\Gamma t) \\ & \left. - 2\Im \left\{ \left[\frac{q}{p} \right]^* A_f(\Phi_n) \bar{A}_f^*(\Phi_n) \right\} \sin(x\Gamma t) \right]. \end{aligned} \quad (2.37)$$

The corresponding equation for the CP conjugated decay can analogously be obtained:

$$\begin{aligned} |\bar{A}_{\bar{f}}|^2(t, \Phi_n) = & \frac{1}{2} e^{-\Gamma t} \left[\left\{ |\bar{A}_{\bar{f}}(\Phi_n)|^2 + \left| \frac{p}{q} A_{\bar{f}}(\Phi_n) \right|^2 \right\} \cosh(y\Gamma t) \right. \\ & + \left\{ |\bar{A}_{\bar{f}}(\Phi_n)|^2 - \left| \frac{p}{q} A_{\bar{f}}(\Phi_n) \right|^2 \right\} \cos(x\Gamma t) \\ & - 2\Re \left\{ \left[\frac{p}{q} \right]^* \bar{A}_{\bar{f}}(\Phi_n) A_{\bar{f}}^*(\Phi_n) \right\} \sinh(y\Gamma t) \\ & \left. - 2\Im \left\{ \left[\frac{p}{q} \right]^* \bar{A}_{\bar{f}}(\Phi_n) A_{\bar{f}}^*(\Phi_n) \right\} \sin(x\Gamma t) \right]. \end{aligned} \quad (2.38)$$

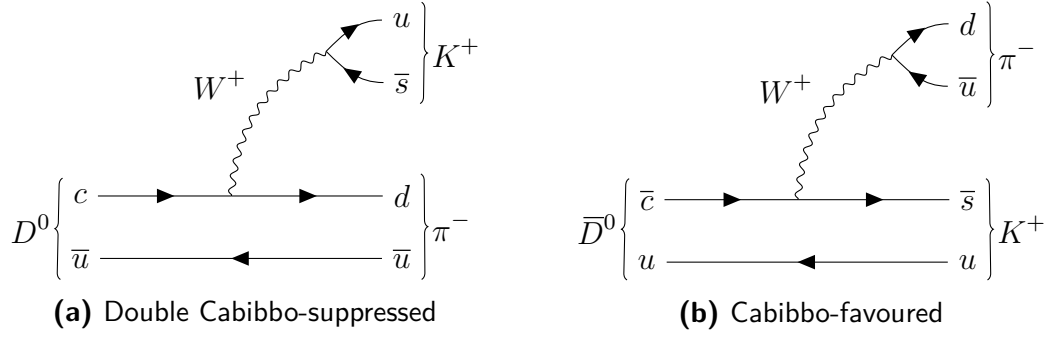


Fig. 2.7: Feynman diagrams for the double Cabibbo-suppressed decay of a D^0 (a) and the Cabibbo-favoured decay of a $\bar{D}^0$ (b) into the same final state, $K^+\pi^-\pi^+\pi^-$. Only the weak components of the decays are depicted, QCD radiation creates an additional $u\bar{u}$ and $d\bar{d}$ pair to form the remaining two pions.

In the absence of direct CP violation in the decay amplitudes themselves, implying $A_f = \bar{A}_{\bar{f}}$ and $\bar{A}_f = A_{\bar{f}}$, indirect CP violation can be introduced for $t > 0$ through the mixing and interference terms in Equations (2.37) and (2.38) if $q/p \neq 1$.

2.3.2 The decay $D^0 \rightarrow K^+\pi^-\pi^+\pi^-$

A possible candidate to study the mixing of D^0 and $\bar{D}^0$ mesons is the decay $D^0 \rightarrow K^+\pi^-\pi^+\pi^-$. This decay can either occur directly as illustrated in the diagram in Figure 2.7a or, as discussed in the previous section, by the transition into a $\bar{D}^0$ first and then following the decay depicted in Figure 2.7b to the same final state. The two amplitudes involved have a striking difference: while the decay amplitude A_{DCS} for the $D^0 \rightarrow K^+\pi^-\pi^+\pi^-$ decay is strongly suppressed due to two unfavoured quark transitions, the amplitude $\bar{A}_{\text{CF}}$ for $\bar{D}^0 \rightarrow K^+\pi^-\pi^+\pi^-$ proceeds only via favoured transitions. Therefore, $|\bar{A}_{\text{CF}}|^2 \gg |A_{\text{DCS}}|^2$ holds for the phase-space integrated amplitudes which has important consequences when studying mixing in this decay.

As previously discussed, the charm mixing diagrams in Figure 2.6 are strongly suppressed, which is also reflected by the current world average on the mixing parameters x and y , which have values well below 1%. This results in the arguments of the trigonometric and hyperbolic functions, $x\Gamma t$ and $y\Gamma t$ being $\ll 1$ for any experimentally reasonable t . Therefore, the effect of mixing when measuring the rate of Cabibbo favoured $D^0 \rightarrow K^-\pi^+\pi^-\pi^+$ decays is completely negligible and the time dependent decay amplitude simply becomes

$$|A_{\text{RS}}(t, \Phi_n)|^2 = e^{-\Gamma t} |\bar{A}_{\text{CF}}(\Phi_n)|^2, \quad (2.39)$$

where the subscript RS stands for the right-sign decay to differentiate the measured process more clearly from the pure decay amplitude. While this seems arbitrary for RS, the corresponding wrong-sign (WS) decay $D^0 \rightarrow K^+\pi^-\pi^+\pi^-$ is not simply

obtained by multiplying the A_{DCS} amplitude with an exponential decay time term. In this decay, the double Cabibbo-suppressed final state $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ receives significant contributions from mixing, where small mixing parameters and the much larger Cabibbo-favoured amplitude balance. This leads to contributions similar to the direct decay to the suppressed final state on the order of a few decay times. This follows from Equation (2.37), where $A_f = A_{\text{DCS}}$ and $\bar{A}_f = \bar{A}_{\text{CF}}$. Expanding this equation to second order in x and y leads to an approximated relation where the different contributions to the time-dependent amplitude are separated:

$$\begin{aligned}
 |A_{WS}(t, \Phi_n)|^2 = e^{-\Gamma t} & \left[|A_{\text{DCS}}(\Phi_n)|^2 \right. \\
 & + |\bar{A}_{\text{CF}}(\Phi_n)|^2 \left| \frac{q}{p} \right|^2 \frac{y^2 + x^2}{4} (\Gamma t)^2 \\
 & + |A_{\text{DCS}}(\Phi_n)|^2 \frac{y^2 - x^2}{4} (\Gamma t)^2 \\
 & - |A_{\text{DCS}}(\Phi_n)| |\bar{A}_{\text{CF}}(\Phi_n)| \left| \frac{q}{p} \right| \Gamma t \\
 & \left. \times \{y \cos(\delta(\Phi_n) - \phi) - x \sin(\delta(\Phi_n) - \phi)\} \right], \quad (2.40)
 \end{aligned}$$

where the argument of the trigonometric functions is defined as the argument of the interference term:

$$\delta(\Phi_n) - \phi = -\arg \left(\left[\frac{q}{p} \right]^* A_f(\Phi_n) \bar{A}_f^*(\Phi_n) \right). \quad (2.41)$$

Here, ϕ is the phase of q/p and $\delta(\Phi_n)$ is the phase difference between A_{DCS} and $\bar{A}_{\text{CF}}$, which can vary across the phase-space. Typically, ϕ is referred to as the weak phase while δ is called the strong phase difference. Equation (2.40) reveals the different contributions: a term constant in t gives the direct decay amplitude and is the only remaining term for $t = 0$. With $t > 0$, a term quadratic in t gives the contribution from oscillating into a $\bar{D}^0$ followed by the favoured decay and a linear contribution gives the interference between these two processes and strongly depends on the phase difference between the two decay amplitudes. In the following, two different approaches are discussed to extract information on the mixing parameters from $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ decays. The main difference is the treatment of the dependence on the position in phase-space Φ_n : in a model-independent approach, the mixing equation is reformulated using phase-space averaged quantities, which are then extracted alongside the mixing parameters themselves. In contrast, in a model-dependent approach, the full amplitude models are constructed. For a four-body decay, this results in five-dimensional complex models. As will be demonstrated in the following, such an approach allows direct access to the parameters of interest

while it introduces additional complexity and sources for systematic uncertainties due to the choice of the model.

2.3.3 Model-independent measurements in $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$

In a model-independent measurement, all phase-space dependent quantities in Equation (2.40) are replaced by phase-space averaged quantities. Furthermore, contributions proportional to $|A_{\text{DCS}}|^2 x^2$ and $|A_{\text{DCS}}|^2 y^2$ are ignored as they are suppressed both by the small mixing parameters and the suppressed amplitude. Performing the average (assuming no CP violation and thus $p/q = 1$) one obtains:

$$|A_{WS}(t)|^2 = e^{-\Gamma t} \left[|A_{\text{DCS}}|^2 + |\bar{A}_{\text{CF}}|^2 \times \frac{y^2 + x^2}{4} (\Gamma t)^2 - |A_{\text{DCS}}| |\bar{A}_{\text{CF}}| \times \{y \langle \cos(\delta) \rangle - x \langle \sin(\delta) \rangle\} \times \Gamma t \right], \quad (2.42)$$

where $\langle \cos(\delta) \rangle$ and $\langle \sin(\delta) \rangle$ are the phase space averaged sine and cosine values for the strong phase difference. Typically, these averages are replaced by defining an average phase-space difference $\bar{\delta}$ defined such that

$$R_c e^{i\bar{\delta}} = \langle \cos(\delta) \rangle + i \langle \sin(\delta) \rangle \quad (2.43)$$

holds with R_c being a coherence factor, which encapsulates the motion of the phase across the phase-space, resulting in:

$$|A_{WS}(t)|^2 = e^{-\Gamma t} \left[|A_{\text{DCS}}|^2 + |\bar{A}_{\text{CF}}|^2 \times \frac{y^2 + x^2}{4} (\Gamma t)^2 - |A_{\text{DCS}}| |\bar{A}_{\text{CF}}| \times R_c \times \{y \cos(\bar{\delta}) - x \sin(\bar{\delta})\} \times \Gamma t \right]. \quad (2.44)$$

Following this approach, a model-independent measurement has been performed by LHCb which is described in Ref. [46]. In this measurement, the ratio of $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ to $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ decays is extracted in bins of decay time, which is modelled using

$$R(t) = \frac{|A_{WS}(t)|^2}{|A_{RS}(t)|^2} = r^2 - r \times R y' \times \Gamma t + \frac{x^2 + y^2}{4} (\Gamma t)^2 \quad (2.45)$$

with the substitutions

$$r = \frac{|A_{\text{DCS}}|}{|\bar{A}_{\text{CF}}|} \text{ and} \quad (2.46)$$

$$y' = y \cos(\bar{\delta}) - x \sin(\bar{\delta}). \quad (2.47)$$

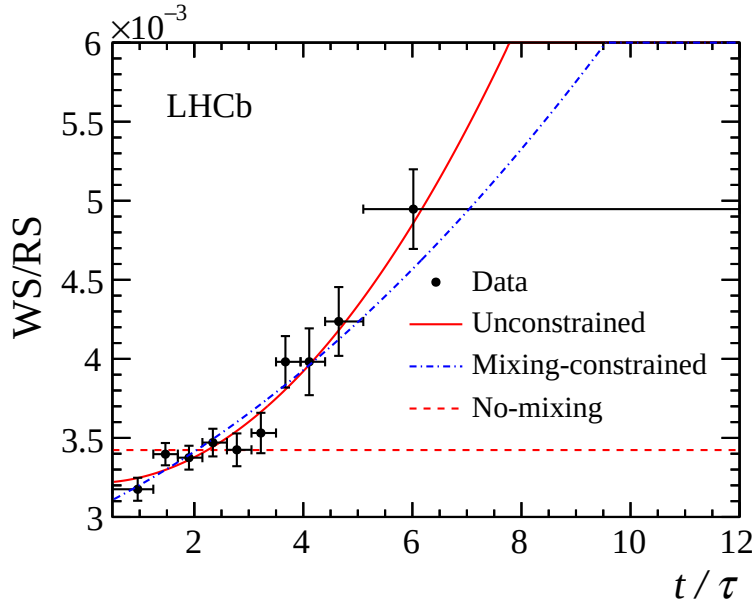


Fig. 2.8: Fit of $R(t)$, see Equation (2.45), to data measured by LHCb. Three different fits are shown: an unconstrained fit with the floating parameters as described in the text, a fit where the mixing parameters x and y are constrained to $(0.371 \pm 0.158) \times 10^{-2}$ and $(0.656 \pm 0.080) \times 10^{-2}$ [47] as well as a fit without mixing and thus fitting $R(t) = r^2$. The plot is taken from Ref. [46].

The new observable, y' , is the result of rotating y in the x - y plane by an unknown angle $\bar{\delta}$. Furthermore, y' cannot be extracted from a fit of $R(t)$ as only the combined product $R_c y'$ can be extracted alongside values for r and $x^2 + y^2$. Figure 2.8 shows the result of a fit to the LHCb data that reveals a clear dependence of the ratio $R(t)$ on the decay time, and good agreement with the obtained curves for a fit constrained to the world average values for x and y . A fit of a constant, implying no mixing, does not agree with the data at a significance of 8.1σ . This constitutes both the first observation of D^0 meson oscillation in $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ decays and, more importantly, the first single-measurement observation of such a process in a channel other than $D^0 \rightarrow K^+ \pi^-$.

2.3.4 Model-dependent measurements in $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$

In the previous section, the model-independent approach showed that an additional term is introduced due to taking the average across the phase-space: the coherence factor R_c . Due to this factor, parameters for the mixing (or their rotated versions x' and y') cannot be determined from the terms linear in decay time as an ambiguity exists between R_c and y' . The behaviour of $R(t)$ is dominated by the linear contributions which, in turn, would contribute most to constrain the uncertainty on the extracted mixing parameters. Following a model-dependent approach, which will be discussed in the following, this can be partly remedied.

In a model-dependent approach, instead of extracting parameters representing phase-space averaged quantities, amplitude models for both Cabibbo-favoured $\bar{A}_{\text{CF}}(\Phi_n)$ and doubly Cabibbo-suppressed $A_{\text{DCS}}(\Phi_n)$ amplitudes are used. In the case of the decays discussed here, these models are five-dimensional complex descriptions that can be constructed in various ways highlighted in the following section, while here one can simply assume their existence. Having these models one has access to the phase-difference and its variation across the phase-space, and can thus deduce the strong phase-space difference at every point. However, one caveat born due to practical reasons exists: as the amplitude for the favoured decay is usually constrained by a sample of favoured decays using Equation (2.39), its overall phase is arbitrary as it vanishes when squaring the amplitude. Nonetheless, this lack of knowledge can be incorporated into the phase of the interference term where Equation (2.41) is now modified:

$$-\arg\left(\left[\frac{q}{p}\right]^* A_f(\Phi_n) \bar{A}_f^*(\Phi_n)\right) = \delta'(\Phi_n) + \delta_0 - \phi, \quad (2.48)$$

where $\delta'(\Phi_n)$ describes the phase motion across the phase-space and is known from the amplitude models while δ_0 is an unknown offset identical for all phase-space points. Going back to Equation (2.40) and substituting this modified phase (for simplicity here, again no CP violation is assumed) yields:

$$\begin{aligned} |A_{WS}(t, \Phi_n)|^2 = e^{-\Gamma t} & \left[|A_{\text{DCS}}(\Phi_n)|^2 \right. \\ & + |\bar{A}_{\text{CF}}(\Phi_n)|^2 \times \frac{y^2 + x^2}{4} (\Gamma t)^2 \\ & + |A_{\text{DCS}}(\Phi_n)|^2 \times \frac{y^2 - x^2}{4} (\Gamma t)^2 \\ & - |A_{\text{DCS}}(\Phi_n)| |\bar{A}_{\text{CF}}(\Phi_n)| \Gamma t \\ & \left. \times \{y \cos(\delta'(\Phi_n) + \delta_0) - x \sin(\delta'(\Phi_n) + \delta_0)\} \right]. \quad (2.49) \end{aligned}$$

By expanding the sine and cosine, this becomes

$$\begin{aligned} |A_{WS}(t, \Phi_n)|^2 = e^{-\Gamma t} & \left[|A_{\text{DCS}}(\Phi_n)|^2 \right. \\ & + |\bar{A}_{\text{CF}}(\Phi_n)|^2 \times \frac{y^2 + x^2}{4} (\Gamma t)^2 \\ & + |A_{\text{DCS}}(\Phi_n)|^2 \times \frac{y^2 - x^2}{4} (\Gamma t)^2 \\ & - |A_{\text{DCS}}(\Phi_n)| |\bar{A}_{\text{CF}}(\Phi_n)| \Gamma t \\ & \times \left(\{y \cos(\delta_0) - x \sin(\delta_0)\} \cos(\delta'(\Phi_n)) - \right. \\ & \left. \left. - \{x \cos(\delta_0) + y \sin(\delta_0)\} \sin(\delta'(\Phi_n)) \right) \right]. \quad (2.50) \end{aligned}$$

This shows that the constant global strong phase difference δ_0 acts like a rotation of x and y in the linear term of the equation. The obvious substitution is given by;

$$y''' \equiv y \cos(\delta_0) - x \sin(\delta_0) \text{ and} \quad (2.51)$$

$$x''' \equiv x \cos(\delta_0) + y \sin(\delta_0) , \quad (2.52)$$

where three primes are deliberately chosen as these variables differ due to the unknown phase from other measurements listed in Ref. [47]. In order to fully absorb the unknown phase into a redefinition of the mixing parameters, the term proportional to $y^2 - x^2$ needs to be negligible as it is, unlike $x^2 + y^2$, not invariant under a rotation of x and y . In the model-independent approach, it was argued that the corresponding term can be neglected because the phase-space-averaged doubly-Cabibbo-suppressed amplitude is very small compared to the favoured one. However, in the model-dependent approach it is not a priori clear that this holds true for all regions of phase-space as it is imaginable to find an area in phase-space for which the two amplitudes are much closer. This needs to be checked with the actual models being used in an analysis and depends on the achievable experimental precision as well. Nonetheless, for the sake of the argument here, the term is neglected and the final relation is obtained:

$$\begin{aligned} |A_{WS}(t, \Phi_n)|^2 = e^{-\Gamma t} & \left[|A_{DCS}(\Phi_n)|^2 \right. \\ & + |\bar{A}_{CF}(\Phi_n)|^2 \times \frac{y'''^2 + x'''^2}{4} (\Gamma t)^2 \\ & - |A_{DCS}(\Phi_n)| |\bar{A}_{CF}(\Phi_n)| \Gamma t \\ & \left. \times \{y''' \cos(\delta'(\Phi_n)) - x''' \sin(\delta'(\Phi_n))\} \right] . \end{aligned} \quad (2.53)$$

No coherence factor was introduced and the rotated x and y parameters can be directly extracted in a fit. While the unknown rotation leads, in fact, only to an estimate of $x^2 + y^2$, which can also be obtained in a model-independent approach, it is now constrained in the linear part of the relation, which will result in substantially stronger constraints on the uncertainty of the extracted parameter.

Turning the argument around, in the limit of the approximation discussed, the unknown global phase difference cannot be constrained by the fit itself as any arbitrarily added additional phase can always be absorbed in a redefinition of x and y thus leading to an ambiguity, which would prevent a convergence of any likelihood fit. Finally, it is worthwhile to consider what would happen if the approximations are not appropriate anymore and neglected cubic contributions of the expansion of Equation (2.37) become relevant. In this case, δ_0 cannot be absorbed into a redefinition of x and y and thus the ambiguity vanishes and a likelihood fit can

extract values for x , y and δ_0 at the same time. In the following, estimates are provided for how many WS events are required to be sensitive to the omitted terms in the expansion of $|A_{WS}(t, \Phi_n)|^2$.

2.3.5 The breakdown of the approximations

To obtain an estimate for the number of WS events necessary and to illustrate at which statistics the different terms become relevant, a simplified calculation is performed. In this calculation, $|\bar{A}_{CF}(\Phi_n)| = 1$, $|A_{DCS}(\Phi_n)| = 1/\sqrt{300}$ and $\delta'(\Phi_n) = \pi/2$ are assumed and the values for x''' and y''' are fixed to the world average values for x and y as presented in the next section. These values are chosen to approximate the actual difference in magnitude between the two amplitudes while the phase-space dependence is removed to simplify the following calculations. Distributions for different hypotheses are then constructed, where the null hypothesis is obtained by normalising Equation (2.53), and alternative hypotheses are constructed by adding the terms omitted in the expansion, namely the suppressed term,

$$|A_{DCS}(\Phi_n)|^2 \times \frac{y'''^2 - x'''^2}{4} (\Gamma t)^2, \quad (2.54)$$

and the term given by the expansion to the third order in Γt ,

$$|A_{DCS}(\Phi_n)| |\bar{A}_{CF}(\Phi_n)| \times \left\{ y'''^3 \cos(\delta'(\Phi_n)) + x'''^3 \sin(\delta'(\Phi_n)) \right\} \frac{(\Gamma t)^3}{6}. \quad (2.55)$$

The linear approximation, where the term proportional to $|\bar{A}_{CF}(\Phi_n)|^2 (\Gamma t)^2$ is removed, is also studied for comparison. To study how many events are necessary to statistically separate the different hypotheses and which region in Γt is most sensitive, the hypotheses are integrated in Γt from a starting value $(\Gamma t)_0$ to infinity, yielding the expected fraction of the total number of events N under the null hypothesis (f_0) and one of the alternatives (f') for which $\Gamma t > (\Gamma t)_0$. For each value of $(\Gamma t)_0$, Poisson statistics is used to define the minimal N at which the expectation value under that alternative hypothesis, Nf' , does not lie within the central 68% interval of a Poisson distribution with a mean of Nf_0 . Figure 2.9 depicts the obtained N as a function of $(\Gamma t)_0$ for the different alternative hypotheses. Generally, a similar trend is seen for all three different hypotheses, with a minimum around five life times and an exponential increase towards higher life times. This exponential increase is due to the overall $\exp(-\Gamma t)$ factor that overwhelms the increase in relative difference between the null and alternative hypotheses at larger life times. Concerning the actual numbers required, a drastically different picture is seen for different terms explored. The term proportional to $|\bar{A}_{CF}(\Phi_n)| (\Gamma t)^2$, which is not omitted in Equation (2.53), becomes statistically significant at around 10^4 WS events. This is a number already seen in the model-independent measurement discussed in Section 2.3.3 and thus

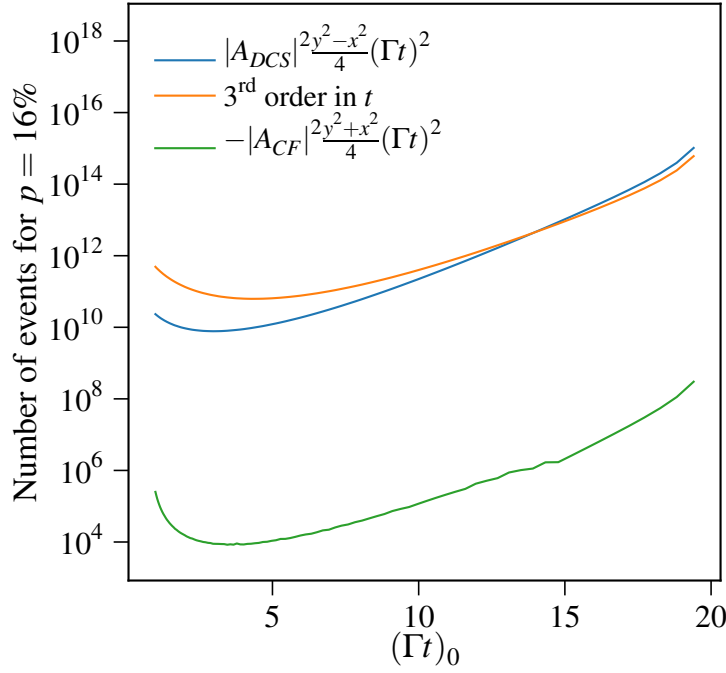


Fig. 2.9: Illustration of the number of events necessary to statistically separate the used mixing approximation (given in Equation (2.53)) from versions where higher order terms are included. For reference, a comparison to the linear expansion, obtained by omitting the term proportional to $|\bar{A}_{CF}(\phi_n)|^2(\Gamma t)^2$, is also given. A minimal value for $(\Gamma t)_0$ is chosen as 0.001 as the problem diverges when $(\Gamma t)_0$ approaches zero.

its inclusion is not only justified but necessary. In contrast, the next two terms in the expansion become statistically significant at around $10^{10} \sim 10^{11}$ events. This is several magnitudes beyond the currently available data and thus any higher order terms are completely negligible.

2.3.6 The current world averages

The Heavy Flavor Averaging Group (HFLAV) [47] calculates world averages for a variety of properties in flavour physics each combining as many experimental results as possible. The current world averages [47] for the mixing parameters in the neutral charm meson system are:

$$x = (0.46 \pm 0.15)\% \text{ and} \quad (2.56)$$

$$y = (0.62 \pm 0.08)\% . \quad (2.57)$$

These averages have been obtained in a global fit of 31 individual measurements under the assumption of no CP violation. A visualisation of the two-dimensional contour on these averages is given in Figure 2.10. Unlike for y , the current world average values for x are not significantly different from zero.

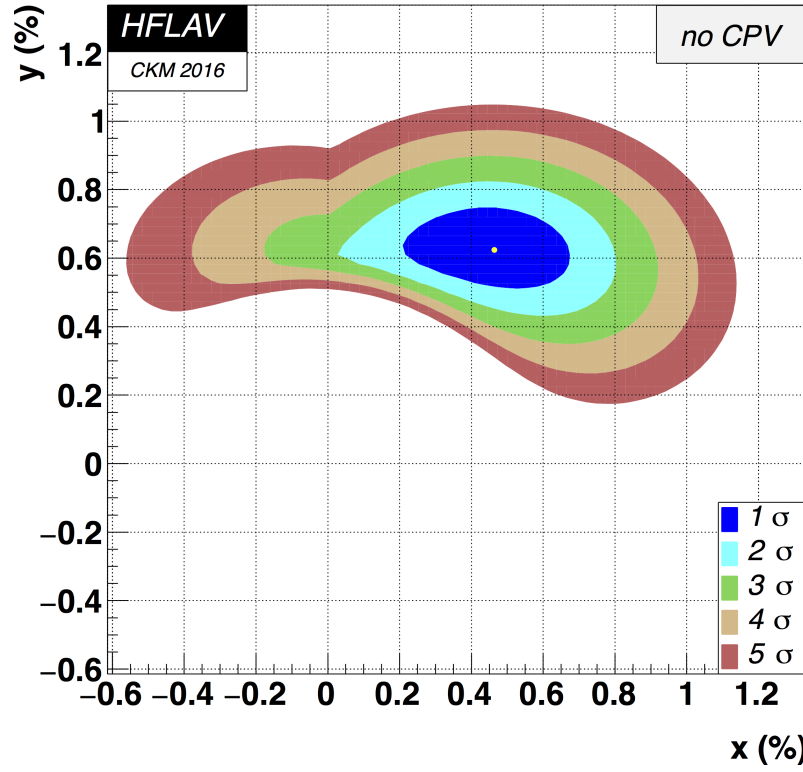


Fig. 2.10: Visualisation of the two-dimensional contour of the world averages on the D^0 mixing parameters x and y . Figure taken from Ref. [47]. 31 individual measurements of charm mixing parameters are used for this average.

2.4 The decay amplitude – An isobar model

In the previous section on the backdrop of model-dependent mixing measurements, the general concept of an amplitude model was introduced as a prerequisite for those measurements. This section will discuss the widely used isobar approach [48] to construct amplitude models for $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ and $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ decays, focusing on the specific treatment of these two decays where necessary instead of giving a general introduction to amplitude modelling. While experimentally successfully practiced, the isobar model suffers from theoretical insufficiencies as well as experimental shortcomings whose discussion concludes this section.

An amplitude model $A(\Phi_n)$ gives a complex number as a function of the position in phase-space. Unlike two-body decays, which have a trivial decay structure, three- and four-body decays can have different, interfering resonances to the decay and the total amplitude is given as

$$A(\Phi_n) = \sum_i c_i A_i(\Phi_n), \quad (2.58)$$

where the coefficients c_i are complex numbers typically extracted from fits to data and A_i is the individual resonance. In the isobar approach one assumes that the decay and all its resonances can be approximated as a series of independent two-body decays.

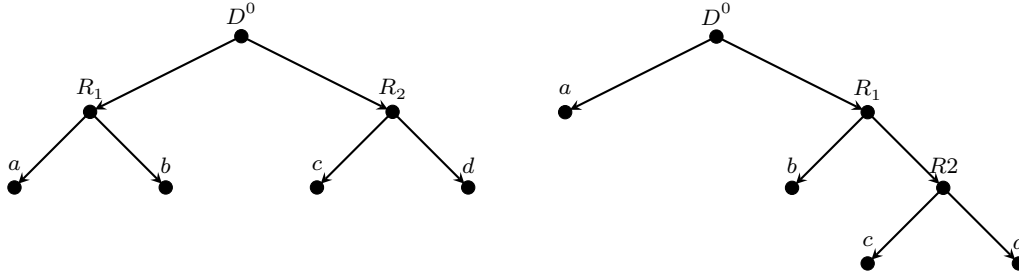


Fig. 2.11: Possible arrangement of quasi two-body decays in a four-body D^0 decay. Direct decay into two resonances (left) and cascading decay (right).

Two pathways are illustrated in Figure 2.11: the decay can occur via two independent resonances (e.g. $D^0 \rightarrow R_1(\rightarrow K^+\pi^-)R_2(\rightarrow \pi^+\pi^-)$) or in a cascade where the second resonance is the first one's decay product (e.g. $D^0 \rightarrow R_1(\rightarrow K^+R_2[\rightarrow \pi^+\pi^-])\pi^-$). To illustrate how the amplitudes A_i are constructed, the decay via $K^*(892)$ and $\rho(770)$ is used as an example. In this case, the resonant amplitude is constructed as:

$$A_i = \frac{1}{\sqrt{2}} [\mathcal{F}(K^+\pi_1^-\pi^+\pi_2^-) \cdot \mathcal{S}(K^+\pi_1^-\pi^+\pi_2^-) \cdot \mathcal{L}(K^+\pi_1^-) \cdot \mathcal{L}'(\pi^+\pi_2^-) \\ \mathcal{F}(K^+\pi_2^-\pi^+\pi_1^-) \cdot \mathcal{S}(K^+\pi_1^-\pi^+\pi_2^-) \cdot \mathcal{L}(K^+\pi_2^-) \cdot \mathcal{L}'(\pi^+\pi_1^-)] , \quad (2.59)$$

where the amplitude is symmetric under the exchange of the indistinguishable pions with the same electric charge. Equation (2.59) is comprised of three different terms, a form factor $\mathcal{F}$, a spin factor $\mathcal{S}$ and the lineshape $\mathcal{L}$ describing the propagator for the specific resonance. Lastly, due to the different form factors, different orbital angular momenta for the same decay are treated as individual amplitudes A_i .

2.4.1 Form factor $\mathcal{F}$

The form factor $\mathcal{F}$ for the D^0 decay is given by the Blatt-Weisskopf centrifugal barrier factors $B_L(q)$ [49, 50] that are given in Table 2.1. They are functions of the meson radius $d \approx 1 \text{ fm} \sim 5 \text{ GeV}^{-1}$ and the momentum transfer

$$q^2(m, m_1, m_2) = \frac{m}{4} - \frac{m_1 + m_2}{2} + \frac{(m_1 - m_2)^2}{4m}, \quad (2.60)$$

where m , m_1 and m_2 are the reconstructed masses of the particle and its decay products. The D^0 is a spin-less meson and thus the total angular momentum of the decay must always be $J = 0$. To allow for resonances R_1 and R_2 carrying spin, this implies an angular momentum different from zero to ensure conservation of the total angular momentum. However, experimental observations show that in decays where the available momentum is small, the decay products have difficulties generating sufficient angular momentum, resulting in a suppression of such contributions. The form factors $\mathcal{F}$ in Equation (2.59) only incorporate this suppression for the decay

Tab. 2.1: Blatt-Weisskopf barrier factors [49, 50]. The meson radius d and breakup momentum q are explained in the text.

L	$B_L(q)$	$B'_L(q, q_0)$
0	1	1
1	$\sqrt{\frac{2q^2 d^2}{1+q^2 d^2}}$	$\sqrt{\frac{q_0^2 d^2}{1+q^2 d^2}}$
2	$\sqrt{\frac{13q^4 d^4}{(q^2 d^2 - 3)^2 + 9q^2 d^2}}$	$\sqrt{\frac{(q_0^2 d^2 - 3)^2 + 9q_0^2 d^2}{(q^2 d^2 - 3)^2 + 9q^2 d^2}}$

of the D^0 , $D^0 \rightarrow R_1 R_2$ or $D^0 \rightarrow a R_1$, while the corresponding form factors for the decays of the intermediate resonances are part of the lineshapes discussed later.

Table 2.1 also gives a modified version of Ball-Weisskopf factors $B'_L(q, q_0)$, which are later used in the definition of certain lineshapes. The additional parameter is defined as $q_0^2 \equiv q^2(m_0, m_1, m_2)$ where m_0 is the pole mass of the decaying particle or resonance.

2.4.2 Spin factor $\mathcal{S}$

As stated before, the spin-less D^0 implies that the total angular momentum of the entire decay needs to be zero, coupling the intrinsic spin of the different possible resonances to the angular momentum between decay products. This is responsible for the observable distributions of angular variables between the final state particles. These constraints are included in the amplitude A_i by so-called spin factors $\mathcal{S}$ which are formulated in the covariant spin tensor formalism [51–55]. Due to the aforementioned suppression of high angular momentum configurations, this requires the derivation of spin-1 and spin-2 tensor representations with the corresponding projection operators in order to construct the spin factors. This section aims to give a summary of the necessary spin formalism and is based on Ref. [56].

Spin-1 particles

The first step requires the construction of an explicit basis for the spin states $|1m\rangle$ with $m \in [-1, 0, 1]$ which fulfills the usual normalisation constraints:

$$\langle 1m' | 1m \rangle = \delta_{mm'} \quad \text{and} \quad \sum_m |1m\rangle \langle 1m| = 1. \quad (2.61)$$

These are fulfilled by the three-dimensional polarisation vectors:

$$|11\rangle = -\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \quad |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{and} \quad |1-1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}. \quad (2.62)$$

These refer to the spin measured with the particle at rest. Expressing the states in the momentum space hence implies that the resulting four-momentum cannot have a time component:

$$e^\mu(m) = \{0, |1m\rangle\} . \quad (2.63)$$

This object now behaves just like a normal four-vector and applying a Lorentz boost $L(\vec{p})$ allows access to the covariant spin bases for states with momentum $\vec{p}$,

$$e^\mu(\vec{p}, m) = L_\nu^\mu(\vec{p}) e^\nu(m) . \quad (2.64)$$

Following on from this, some useful relations are

$$p^\mu e_\mu(\vec{p}, m) = 0 , \quad (2.65)$$

$$e_\mu^*(\vec{p}, m') e^\mu(\vec{p}, m) = -\delta_{mm'} , \quad (2.66)$$

$$P_{\mu\nu}^{(1)} = \sum_m e_\mu(\vec{p}, m) e_\nu^*(\vec{p}, m) = \bar{g}_{\mu\nu} , \text{ and} \quad (2.67)$$

$$\bar{g}_{\mu\nu} = -g_{\mu\nu} + \frac{p_\mu p_\nu}{M^2} , \quad (2.68)$$

where M is the mass of the particle with momentum $\vec{p}$ and $g_{\mu\nu}$ the Minkowski metric with negative spatial components. $P_{\mu\nu}^{(1)}$ is the spin-1 projection operator, which transforms any four-vector x^μ into x'_ν which is orthogonal to p^μ .

Spin-2 particles

Covariant representations for spin-2 states are constructed in a similar way as above. As the states $|2m\rangle$ with $m \in [-2, -1, 0, 1, 2]$ require a five-dimensional space, the association with known covariant objects is not as straightforward. Instead, the covariant representation of the spin state with momentum $\vec{p}$ is given by combining the two spin-1 states into a rank-2 tensor (a matrix in Minkowski space),

$$e_{\mu\nu}(\vec{p}, 2m) = \sum_{m_1 m_2} (1m_1 1m_2 | 2m) e_\mu(\vec{p}, m_1) e_\nu(\vec{p}, m_2) . \quad (2.69)$$

Here, the individual e_μ and e_ν are defined as above and the first terms in each element of the sum are the Clebsch-Gordan coefficients. As a rank-2 tensor has, in principle, sixteen degrees of freedom, additional constraints are required to reduce this to five degrees of freedom. These constraints are not unlike those for spin-1 and are given by the Rarita-Schwinger conditions for integral spin tensors:

$$p^\mu e_{\mu\nu}(\vec{p}, 2m) = 0 , \quad (2.70)$$

$$e_{\mu\nu}(\vec{p}, 2m) = e_{\nu\mu}(\vec{p}, 2m) \text{ and} \quad (2.71)$$

$$g^{\mu\nu} e_{\mu\nu}(\vec{p}, 2m) = 0 . \quad (2.72)$$

Equation (2.70) implies that the rank-2 tensors are orthogonal to the four-momentum, a consequence of the spin-1 states having no time component in the rest frame, while Equations (2.71) and (2.72) demand that the $e_{\mu\nu}$ is symmetric and traceless. Analog to the spin-1 case, a projection operator is defined as

$$P_{\alpha\beta\mu\nu}^{(2)} = \sum_m e_{\mu\nu}(\vec{p}, 2m) e_{\alpha\beta}^*(\vec{p}, 2m) \quad (2.73)$$

$$= \frac{1}{2}(\bar{g}_{\mu\alpha}\bar{g}_{\nu\beta} + \bar{g}_{\mu\beta}\bar{g}_{\nu\alpha}) - \frac{1}{3}\bar{g}_{\mu\nu}\bar{g}_{\alpha\beta} , \quad (2.74)$$

Orbital angular momentum tensor

After constructing the spin component of $\mathcal{S}$, a second contribution comes in due to the orbital angular momentum L between the two decay products in each decay. Assuming the two decay products have the four-momenta p_a and p_b , one defines their sum and difference as $p_{ab} \equiv (p_a + p_b)$ and $q_{ab} \equiv (p_a - p_b)$ and uses q_{ab} to form a rank- L tensor:

$$Q_{ab}^{(L)} = q_{ab}^{\nu_1} q_{ab}^{\nu_2} \dots q_{ab}^{\nu_L} . \quad (2.75)$$

Making use of the observation above that applying the projection operators for rank- L tensors yield a spin- L state where the implicitly enforced Rarita-Schwinger conditions reduce the dimensionality of the tensor to $2L + 1$, an orbital angular momentum state is constructed as:

$$L_{\mu_1\mu_2\dots\mu_L}^{(L)} = (-1)^L P_{\mu_1\mu_2\dots\mu_L\nu_1\nu_2\dots\nu_L}^{(L)}(p_{ab}) Q_{ab}^{(L)} . \quad (2.76)$$

As this relation is rather cumbersome, the relevant tensors for the remainder of this section are given by:

$$L^{(0)}(p_{ab}, q_{ab}) = 1 , \quad (2.77)$$

$$L_{\mu}^{(1)}(p_{ab}, q_{ab}) = -P_{\mu\nu}^{(1)}(p_{ab}) q_{ab}^{\nu} , \quad (2.78)$$

$$L_{\mu\nu}^{(2)}(p_{ab}, q_{ab}) = P_{\mu\nu\alpha\beta}^{(2)}(p_{ab}) q_{ab}^{\alpha} q_{ab}^{\beta} , \quad (2.79)$$

Construction of spin factors $\mathcal{S}$

The relations given in Equations (2.67), (2.74) and (2.77) to (2.79) can now be used to construct the spin factors entering in Equation (2.59). One factor is derived for each of the quasi two-body decays and the final factor is obtained as a product of these individual terms. Additionally, the sum is taken over all possible values for

the internal spins of the resonances appearing in the decay. Unlike the initial state particle (the D^0) and the final state particles (one $K^\pm$ and three $\pi^\pm$) that are all pseudoscalar particles with spin-0 and parity -1, the resonances R_1 and R_2 with spins S_1 and S_2 might not be. Looking at a general two-body decay, $D \rightarrow ab$, the D has a given initial spin and, as a one-particle system, no orbital angular momentum. However, the final state system can take any value for the spin constrained by

$$|S_a - S_b| \leq S_{ab} \leq S_a + S_b , \quad (2.80)$$

and thus introducing orbital angular momentum to preserve the total angular momentum given by the spin of the D . Additionally, possible values for the angular momentum L_{ab} and thus for the spin S_{ab} might arise depending on the specific process: electromagnetic and strong decays preserve the parity and thus add an additional constraint

$$\eta_D = \eta_a \eta_b (-1)^{L_{ab}} , \quad (2.81)$$

where $\eta_{D/a/b}$ is the intrinsic parity of the involved particles. Depending on the specific particles involved in the decay, this results in strong and electromagnetic decays being restricted to either even or odd values of L_{ab} . As stated above, the different L_{ab} configurations are treated as individual amplitudes contributing to the overall decay.

In the following, the spin factors $\mathcal{S}$ are constructed for the same example previously used, where the D^0 decays weakly into two resonances, $K^*(892)$ and $\rho(770)$, which then decay strongly to $K^+\pi^-$ and $\pi^+\pi^-$. Both resonances are vector particles with spin-1 and negative parity. The derivation of the factors follows a similar description as presented in Ref. [57]. As a first step, the factor for the $D^0 \rightarrow K^*(892)\rho(770)$ decay is computed. Following Equation (2.80) and keeping in mind that parity does not need to be conserved in weak decays, the overall spin of the $K^*(892)\rho(770)$ system can be in three different configurations with $S_{K^*\rho} \in [0, 1, 2]$. This results in three possible configurations for the orbital angular momentum: $L_{K^*\rho} = 0$ (the $\mathcal{S}$ -wave), $L_{K^*\rho} = 1$ (the $\mathcal{P}$ -wave) and $L_{K^*\rho} = 2$ (the $\mathcal{D}$ -wave). For the $\mathcal{S}$ -wave, the transition probability is obtained as

$$\langle K^*\rho, S=0, L=0 | M_S | D^0 \rangle = I(D^0) e_\mu^*(K^*) e^{\mu*}(\rho) , \quad (2.82)$$

where M_S is a general transition operator encapsulating the spin component of the total matrix element governing the decay, $I(D^0) = 1$ is the representation of the D^0 spin state and only given here for completeness but omitted in the following, and the remaining terms give the spin-1 states for decay products. For readability, the dependence on the momentum $\vec{p}$ and minor spin number m is not explicitly stated.

Furthermore, the complex conjugate indicates that the respective particles are the product of the decay. The $\mathcal{P}$ -wave takes a similar form where an $L = 1$ state, $L_{(1)\mu}$, is contracted with a spin-1 state $A_{(1)}^\mu$:

$$\langle K^* \rho, 1, 1 | M_S | D^0 \rangle = L_{(1)\mu}(D^0) A_{(1)}^\mu . \quad (2.83)$$

The spin-1 state, $A_{(1)}^\mu$, is constructed by coupling the two individual spin-1 states and using the projection operator to project out the component perpendicular to the momentum vector of the D^0 and hence into the same subspace as the orbital angular momentum state. The state is then given by

$$A_{(1)}^\mu = P_{(1)}^{\mu\nu}(D^0) e^{*\alpha}(K^*) e^{*\beta}(\rho) \varepsilon_{\nu\alpha\beta\lambda} p^\lambda (D^0) , \quad (2.84)$$

where $\varepsilon_{\nu\alpha\beta\lambda}$ is the totally antisymmetric Levi-Cevita symbol. The addition of the rank-3 tensor $\varepsilon_{\nu\alpha\beta\lambda} p^\lambda$ is necessary to obtain a rank-1 representation consistent with a spin-1 state [51]. Substituting this back into the expression yields

$$\langle K^* \rho, 1, 1 | M_S | D^0 \rangle = L_{(1)}^\nu(D^0) e^{*\alpha}(K^*) e^{*\beta}(\rho) \varepsilon_{\nu\alpha\beta\lambda} p^\lambda (D^0) , \quad (2.85)$$

with the simplification $L_{(1)\mu}(D^0) P_{(1)}^{\mu\nu}(D^0) = L_{(1)}^\nu(D^0)$ as the angular momentum state is already, by construction, perpendicular to the direction of the momentum vector. Lastly, the $\mathcal{D}$ -wave component is constructed in a similar way, leading to

$$\langle K^* \rho, 2, 2 | M_S | D^0 \rangle = L_{(2)\mu\nu}(D^0) e^{*\mu}(K^*) e^{*\nu}(\rho) , \quad (2.86)$$

where the projection operator that projects the combination of spin-1 states is omitted for the same reason as above.

The next step involves the construction of the terms for the subsequent decays of the resonances to the pseudoscalars. In the case of this example, these decays must proceed as $\mathcal{P}$ -wave leading to:

$$\langle K^+ \pi^-, 1, 0 | M_S | K^* \rangle = e_\xi(K^*) L_{(1)}^\xi(K^*) \text{ and} \quad (2.87)$$

$$\langle \pi^+ \pi^-, 1, 0 | M_S | \rho \rangle = e_\zeta(\rho) L_{(1)}^\zeta(\rho) . \quad (2.88)$$

The final spin factor is then obtained as a product of the individual terms derived above and is written as:

$$\mathcal{S}_{K^* \rho}^S = \sum_{m_{K^*}, m_\rho} e_\mu^*(K^*) e^{\mu*}(\rho) e_\xi(K^*) L_{(1)}^\xi(K^*) e_\zeta(\rho) L_{(1)}^\zeta(\rho) . \quad (2.89)$$

Tab. 2.2: Spin factors $\mathcal{S}$ necessary for the construction of a decay model. All considered quasi two-body combinations of the decay $D^0 \rightarrow P_1 P_2 P_3 P_4$ are considered where the particle names P , V , A and T correspond to a pseudoscalar, vector, axialvector and tensor particle, respectively. Unless otherwise indicated, the decays proceed as S -wave.

#	Decay Topology	Spin Factor
1	$D^0 \rightarrow V_1(P_1 P_2) V_2(P_3 P_4)$	$L_{(1)}^\mu(V_1) L_{(1)\mu}(V_2)$
2	$D^0 \rightarrow [V_1(P_1 P_2) V_2(P_3 P_4)]^{\mathcal{P}\text{-wave}}$	$L_{(1)}^\alpha(D^0) \epsilon_{\alpha\beta\gamma\delta} p_{D^0}^\delta L_{(1)}^\beta(V_1) L_{(1)}^\gamma(V_2)$
3	$D^0 \rightarrow [V_1(P_1 P_2) V_2(P_3 P_4)]^{\mathcal{D}\text{-wave}}$	$L_{(2)\mu\nu}(D^0) L_{(1)}^\mu(V_1) L_{(1)}^\mu(V_2)$
4	$D^0 \rightarrow V(P_1 P_2) S(P_3 P_4)$	$L_{(1)}^\mu(D^0) L_{(1)\mu}(V)$
5	$D^0 \rightarrow P_1 P(P_2 V(P_3 P_4))$	$L_{(1)}^\mu(P) L_{(1)\mu}(V)$
6	$D^0 \rightarrow P_1 P(P_2 S(P_3 P_4))$	1
7	$D^0 \rightarrow P_1 A(P_2 V(P_3 P_4))$	$L_{(1)}^\mu(D^0) P_{(1)\mu\nu}(A) L_{(1)}^\nu(V)$
8	$D^0 \rightarrow P_1 A[P_2 V(P_3 P_4)]^{\mathcal{D}\text{-wave}}$	$L_{(1)}^\mu(D^0) L_{(2)\mu\nu}(A) L_{(1)}^\nu(V)$
9	$D^0 \rightarrow P_1 A(P_2 S(P_3 P_4))$	$L_{(1)}^\mu(D^0) L_{(1)\mu}(A)$
10	$D^0 \rightarrow P_1 T(P_2 V(P_3 P_4))$	$L_{(2)}^{\mu\alpha}(D^0) \epsilon_{\alpha\beta\gamma\delta} p_{D^0}^\delta L_{(2)\mu}^\beta(T) P_{(1)}^{\gamma\rho}(T) L_{(1)\rho}(V)$

Rearranging the terms allows a significant simplification of the equation:

$$\mathcal{S}_{K^*\rho}^{\mathcal{S}} = \left[\sum_{m_{K^*}} e_\mu^*(K^*) e_\xi(K^*) \right] L_{(1)}^\xi(K^*) \left[\sum_{m_\rho} e^{\mu*}(\rho) e^\zeta(\rho) \right] L_{(1)\zeta}(\rho) . \quad (2.90)$$

As can be seen, the terms in the rectangular braces are the projections as defined in Equation (2.67) and can be omitted because, as before, the angular momentum states are already projected by construction. This leads to the simple relation:

$$\mathcal{S}_{K^*\rho}^{\mathcal{S}} = L_{(1)\mu}(K^*) L_{(1)}^\mu(\rho) . \quad (2.91)$$

Using the same simplifications, the corresponding factors for the remaining orbital angular momentum configurations become:

$$\mathcal{S}_{K^*\rho}^{\mathcal{P}} = L_{(1)}^\alpha(D^0) \epsilon_{\alpha\beta\gamma\delta} p_{D^0}^\delta L_{(1)}^\beta(\rho) L_{(1)}^\gamma(K^*) \text{ and} \quad (2.92)$$

$$\mathcal{S}_{K^*\rho}^{\mathcal{D}} = L_{(2)\mu\nu}(D^0) L_{(1)}^\mu(K^*) L_{(1)}^\mu(\rho) . \quad (2.93)$$

This concludes the construction of the three different spin factors for the decay via a $K^*(892)$ and $\rho(770)$, one for each orbital angular momentum configuration, resulting in three different amplitudes. Following the same procedure as above, spin factors are derived for all other decay paths, which are considered in the construction of an amplitude model and given in Table 2.2.

2.4.3 Lineshapes $\mathcal{L}$

The final term missing in Equation (2.59) are the lineshapes $\mathcal{L}$. They describe the magnitude and phase of the different intermediate resonances. Several different choices are available and an overview is given in the following.

Relativistic Breit-Wigner

The most commonly used lineshape is the relativistic Breit-Wigner propagator given by:

$$\mathcal{L}(s, q) = \frac{B'_L(q, q_0)}{m_0^2 - s - im_0\Gamma(s, q)}, \quad (2.94)$$

where s is the square of the measured mass of the resonance, m_0 is the pole mass of the resonance, and B'_L , q and q_0 are defined as in Section 2.4.1. For the resonances, a smaller meson radius of $d \approx 0.3 \text{ fm} \sim 1.5 \text{ GeV}^{-1}$ is used. Furthermore, the running decay width is :

$$\Gamma(s, q) = \frac{\Gamma_0 q m_0 B_L(q)^2}{q_0 \sqrt{s} B_L(q_0)^2}. \quad (2.95)$$

This shape is used as the default for most of the resonances considered and generally models isolated, narrow resonances well.

LASS – the $K\pi$ $\mathcal{S}$ -wave

The $K\pi$ $\mathcal{S}$ -wave (*e.g.* $K_0^*(1430) \rightarrow K^+\pi^-$) gives a contribution too broad to be correctly modelled by a Breit-Wigner lineshape due to shortcomings in the isobar approach that are detailed in Section 2.4.4. Therefore, a sum of a non-resonant and an inelastic Breit-Wigner term, tuned using data from a $K^-p \rightarrow K^- \pi^+ n$ scattering experiment [58], is used:

$$\mathcal{L}(s, q) = \frac{\sqrt{s}}{q(s) \cot(\delta_B) - iq(s)} + e^{2i\delta_B} \frac{m_0 \Gamma(s)}{s - m_0 - im_0 \Gamma(s)}, \quad (2.96)$$

where q is defined in Equation (2.60). The phase δ_B is mass dependent and parameterised as:

$$\cot(\delta_B) = \frac{1}{aq(s)} + \frac{rq(s)}{2}. \quad (2.97)$$

The scattering length a and effective range r are free parameters that are determined from fits to the scattering dataset [58]:

$$a = (2.07 \pm 0.10) \text{ GeV} , \quad (2.98)$$

$$r = (3.32 \pm 0.34) \text{ GeV} . \quad (2.99)$$

Bugg – the $\pi\pi$ $\mathcal{S}$ -wave

Similar to the LASS parameterisation for the $\mathcal{S}$ -wave $K\pi$ decays, the Bugg parameterisation [59] is used for $\pi\pi$ $\mathcal{S}$ -wave (*e.g.* $f_0(600) \rightarrow \pi^- \pi^+$). It is based on a Breit-Wigner lineshape but takes into account contributions to the decay from multiple channels, namely $\pi\pi$, KK , $\eta\eta$, and 4π , resulting in a description given by,

$$\mathcal{L}(s, q) = \frac{m_0 \Gamma_1(s)}{m_0^2 - s - g_1 \frac{s-s_A}{m_0^2 - s_A} z(s) - i m_0 \Gamma_{tot}(s)} , \quad (2.100)$$

with a total width described as the sum over the individual channels:

$$\Gamma_{tot}(s) = \sum_{i=0}^4 \Gamma_i(s) , \quad (2.101)$$

where the different Γ_i are given, in the same order as above, by:

$$m_0 \Gamma_1(s) = g_1^2 \frac{s - s_A}{m_0^2 - s_A} \rho_1(s) , \quad (2.102)$$

$$m_0 \Gamma_2(s) = 0.6 g_1^2(s) (s/m_0^2) e^{-\alpha|s-4m_K^2|} \rho_2(s) , \quad (2.103)$$

$$m_0 \Gamma_3(s) = 0.2 g_1^2(s) (s/m_0^2) e^{-\alpha|s-4m_\eta^2|} \rho_3(s) , \quad (2.104)$$

$$m_0 \Gamma_4(s) = m_r g_{4\pi} \rho_{4\pi}(s) / \rho_{4\pi}(m_r^2) . \quad (2.105)$$

The remaining substitutions are:

$$g_1^2 = m_0 (b_1 + b_2 s) e^{-(s-m_0^2)/A} , \quad (2.106)$$

$$z(s) = j_1(s) - j_1(m_0^2) , \quad (2.107)$$

$$j_1 = \frac{1}{\pi} \left[2 + \rho_1 \log \left(\frac{1 - \rho_1}{1 + \rho_1} \right) \right] , \quad (2.108)$$

$$\rho_{4\pi} = 1.0 / (1 + e^{7.082 - 2.845s}) , \quad (2.109)$$

$$\rho_i(s) = \sqrt{1 - \frac{4m_i^2}{s}} , \quad i = \pi, K, \eta . \quad (2.110)$$

This leaves the free parameters m_0 , $g_{4\pi}$, s_A , α , b_1 , b_2 and A , which have been determined using BES III data [59] and are listed in Table 2.3.

Tab. 2.3: Free parameters of the Bugg parametrization as given in [59].

Parameter	m_0 [GeV]	$g_{4\pi}$ [GeV]	s_A [GeV]	α [GeV]	b_1 [GeV ²]	b_2 [GeV ²]	A [GeV ²]
Value	0.953	0.011	$0.41m_\pi^2$	1.3	1.302	0.340	2.426

2.4.4 Shortcomings and future developments

While the isobar model is almost exclusively used in analyses that construct amplitude models to describe multi-body decays, it suffers from several experimental and theoretical drawbacks that need to be addressed. Concerning the experimental side, isobar models typically have many parameters that need to be determined from fits to data, leading to many floating parameters, usually resulting in instabilities in the fits. Furthermore, there tend to be ambiguities when it comes to the question about which resonances to include as different combinations often lead to the same fit quality.

The fundamental theoretical assumption of the isobar model is the independence of the individual two-body decays, which do not interact with each other. However, Figure 2.12 shows possible extensions where the final state particles are coupled again. Ignoring these additional contributions can lead to rather large inaccuracies of the extracted phase, which can have important consequences in the search for CP violation as it can affect the phase in Equation (2.15) [60]. This effect is enhanced by the use of Breit-Wigner propagators, which are only a good approximation if the resonances have narrow widths and are isolated such that their tails do not overlap. If there are multiple amplitudes close in phase-space or amplitudes that are very wide such that significant interference occurs, a region where the rescattering is likely, the obtained isobar amplitudes violate the unitarity of the overall scattering matrix [60, 61]. Using the isobar approach in such cases typically leads to fit fractions F_i , which are defined as

$$F_i = \frac{\int |c_i A_i(\Phi_n)| d\Phi_n}{\int |\sum c_i A_i(\Phi_n)| d\Phi_n}, \quad (2.111)$$

to sum up to something far away from one. This indicates large contributions due to interference and the violation of the unitarity. As discussed before, data-driven alternatives exist, which can be used instead of Breit-Wigner lineshapes to address wide resonances such as the $\pi^+\pi^-$ $\mathcal{S}$ -wave. Lastly, due to treating the decay as independent two-body resonances it is not a priori clear which resonances to include, especially in the case of very large datasets. Usually, many different amplitude models lead to identical fit qualities (but potentially different behaviour in the phases) and a

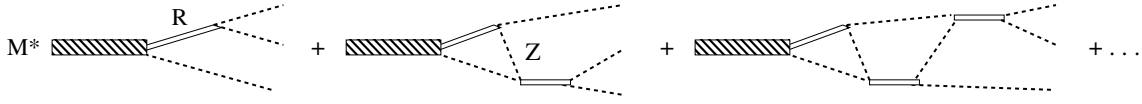


Fig. 2.12: Illustration of rescattering effects in a three-body decay of a meson M^* . In the isobar approach, only the leftmost contribution is considered. Illustration taken from Ref. [60].

large number of free parameters. Several alternatives to the classic isobar approach have been developed [62, 63] that respect the unitarity and incorporate rescattering effects into the amplitudes. In comparisons to data a good agreement of these new models is observed, which often have fewer free parameters than comparable isobar models. However, theoretical work to build such models for charm decays has been focused on $D^0 \rightarrow K_s^0 \pi^+ \pi^-$ decays and transferring the results to four-body decays is not trivial. To the best of the author's knowledge, none of these new models is currently applied to $D^0 \rightarrow K^\pm \pi^\mp \pi^\pm \pi^\mp$ decays at the time of writing this thesis and hence they are not discussed further.

3. Experimental Setup

3.1	The Large Hadron Collider	55
3.2	The LHCb experiment	57
3.2.1	The design requirements	58
3.3	The LHCb detector	60
3.3.1	Tracking	62
3.3.2	Particle identification	64
3.3.3	The calorimeters	66
3.3.4	The muon system	67
3.3.5	The trigger	67
3.3.6	The LHCb upgrade in a nutshell	69

The measurements presented in this thesis use data recorded by the LHCb detector at the Large Hadron Collider and hence this chapter describes both the accelerator and the detector. It goes without saying that giving a full account of all the technical details for both machines extends well beyond the scope of this thesis. Therefore, only a concise description focusing on the parts relevant for the presented measurements is given. Furthermore, the simulation, which is also part of the experimental conditions, is described in more detail in Chapter 4 to acknowledge its central role in my research.

3.1 The Large Hadron Collider

The LHC [65] is located at the CERN research centre close to Geneva under the Franco-Swiss border. With a circumference of 27 km, it is the largest machine and the most powerful particle accelerator in terms of centre-of-mass energy between the accelerated particles ever built by humankind. It was originally designed to accelerate protons to an energy of 7 TeV resulting in collisions at a centre-of-mass energy of 14 TeV but was first operated in 2010 and 2011 at 3.5 TeV beam energy followed by a data taking period at a collision energy of 8 TeV until early 2013. After an upgrade phase, the LHC restarted its operation in 2015 with a new data-taking period (commonly referred to as Run 2) at a centre-of-mass energy of 13 TeV. The LHC is also able to accelerate lead ions in order to study the physics in proton-lead and lead-lead collisions; however this mode is not addressed in this analysis and thus not discussed in further detail. Nonetheless, in preparation for a short proton-lead data taking period at the end of 2015, three days of pp collisions at a centre-of-mass energy of 5 TeV were recorded, providing a new data-point to study the evolution of production cross-sections with the centre-of-mass energy.

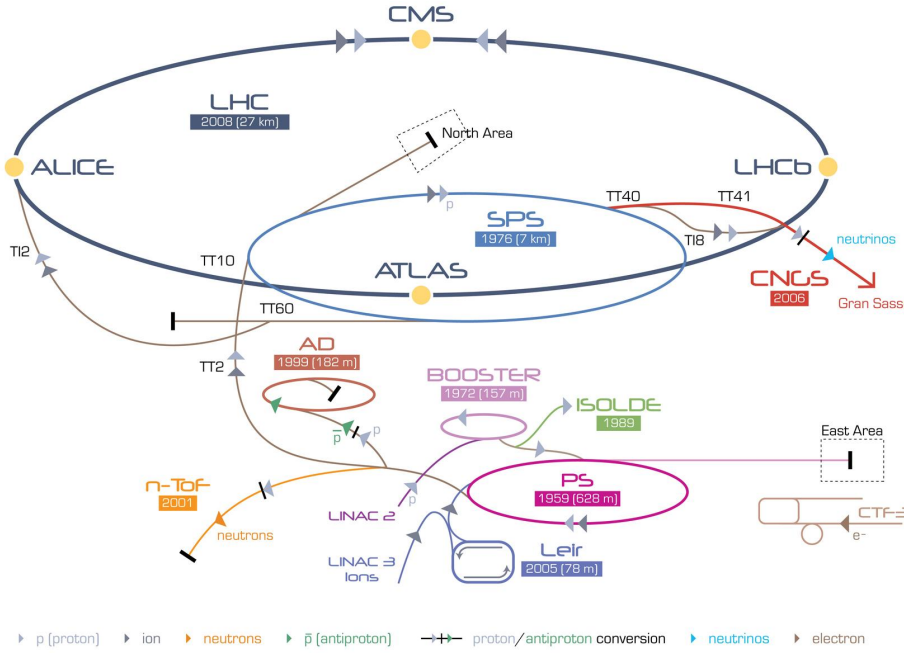


Fig. 3.1: Overview of the CERN accelerator complex [64].

The protons are extracted from a hydrogen source after stripping off the electrons and are then bunched together. Before those protons are injected into the LHC, they pass through a chain of accelerators [66], which is depicted in Figure 3.1. The first accelerator, Linac 2, accelerates the protons to an energy of 50 MeV. These protons are then injected into the Proton Synchrotron Booster that raises the energy to 1.4 GeV, followed by the Proton Synchrotron, which increases it up to 25 GeV. Upon injection into the main LHC ring, the protons have an energy of 450 GeV after passing the Super Proton Synchrotron. In the main ring, they are further accelerated until they reach their nominal energy. The entire acceleration process takes about 20 minutes; however once the protons reach their nominal energy, they can be stored in the main ring for several hours while the experiments located at the LHC take data. The actual acceleration of the protons is done by using electromagnetic fields within a high-vacuum beam pipe. Since time-constant fields cannot reach gradients that are large enough, radio-frequency cavities creating a time-oscillating field are used. Therefore, particles are present in bunches of $\sim 10^{11}$ protons whose revolution frequency of 11.246 kHz is synchronised with the frequency of the electric field such that they pass the cavities during an acceleration phase. Because both beams consist of positively charged particles, they are kept in separate beam pipes which only intersect at the interaction points where the different experiments are located. Up to 2800 bunches are present in the accelerator, which corresponds to the minimal possible spacing between successive bunches, achieving a bunch crossing rate at the interaction points of about 30 MHz. During Run 1 and the first bit of the 13 TeV data-taking period (the one used for the cross-section measurement presented in this thesis), bunch crossings were separated by 50 ns. In the remaining part of Run 2,

including the 5 TeV data, this separation was reduced to 25 ns, which introduces an effect called spillover where the electronic signals created in the detector, as a result of a pp collision, can spill into the previous event due to pulse shapes in the electronics that are longer than 25 ns.

Superconducting dipole magnets with fields up to 8 T are used to keep the protons in a circular orbit. In addition, quadrupole magnets are employed to focus the beam in the plane transverse to the beam direction. Focusing the beam at the interaction point increases the probability for proton-proton interactions to occur. Additionally, sextupole magnets are used to correct the dispersion effects due to different bending and focusing of particles that have slightly different momenta [67].

Besides the centre-of-mass energy, the luminosity is another important quantity that characterises the collider. It can be calculated from the beam parameters of the collider:

$$\mathcal{L} = \frac{1}{4\pi} \cdot \frac{N_1 N_2 f N_b}{\sigma_x \sigma_y} \quad \left[\frac{1}{\text{cm}^2 \text{s}} \right],$$

where N_1 and N_2 are the number of particles in the colliding bunches, f is the revolution frequency of 11.246 kHz and N_b is the number of bunches. σ_x and σ_y describe the beam profile in the transverse plane under the assumption of a Gaussian shape. This instantaneous luminosity relates the rate of events $\dot{N}$ to the cross-section σ of a certain process:

$$\dot{N} = \mathcal{L} \sigma.$$

Besides the LHCb experiment, which focuses on flavour physics in the forward region, the LHC provides particle collisions for three additional major experiments. The ATLAS [68] and CMS [69] experiments use multi-purpose detectors to investigate the physics in the central region with an almost 4π coverage around the interaction point. Lastly, the ALICE [70] experiment studies the collisions of heavy ions, in which it searches especially for quark-gluon plasma.

3.2 The LHCb experiment

The LHCb detector [71] is located at interaction point 8 of the LHC. Unlike the two largest experiments, ATLAS and CMS, which focus on a broad physics programme, especially the direct search for physics beyond the standard model, the LHCb experiment was designed to focus on just one specialised area of particle physics in which it outperforms all other experiments located at the LHC. The goal is to study the area of flavour physics, which, broadly speaking, is the study of the production

and transitions of beauty and charm quarks by measuring the properties and decays of the hadrons that contain them. LHCb is the first dedicated experiment for this area operating at a hadron collider. While previous experiments have been situated at e^+e^- colliders and benefited from their clean experimental conditions, the production of heavy flavour hadrons in pp collisions has substantially higher cross-sections and hence a very large dataset can be recorded. This large dataset, which is expected to increase manifold in the future, has already been used to measure decays that had not previously been observed, and to achieve the most precise measurements of heavy flavour properties to date. Just to name a few examples, LHCb was first to observe mixing of neutral charm mesons ($D^0 \leftrightarrow \bar{D}^0$) [45] in a single measurement as well as bound states formed by more than three quarks [14, 15]. LHCb also contributed important milestones to the study of matter-antimatter differences by performing the first measurement of such differences in baryons [72]. Furthermore, many measurements have been performed to test the precise SM predictions for rare decays of beauty and charm mesons [73, 74], some of which show consistent tensions with those predictions when comparing branching fractions with different leptons in the final state [75, 76]. While flavour physics in pp collisions is the core of LHCb's physics program, it extends beyond this by analysing the physics in heavy ion or minimum bias collisions, contributing results in a unique kinematical region. In the following, the unique design requirements for the LHCb detector, enabling measurements such as those mentioned above, are going to be discussed before the actual detector implementing these requirements is presented.

3.2.1 The design requirements

Section 2.2 already discussed the parton distribution functions for the different partons in a proton and showed the large dominance of the gluon component for small momentum fractions. As a result, only a small fraction of the centre-of-mass energy of pp collisions is required to produce the $q\bar{q}$ pairs of interest – $c\bar{c}$ and $b\bar{b}$ – via the fusion of two gluons as depicted in Figure 3.2a. However, due to the steeply rising gluon PDF towards low x , it is much more likely for this collision to occur when one of the participating gluons has a tiny momentum fraction, resulting in a highly boosted system. This is illustrated in Figure 3.2b, where the quarks are highly collimated along the beam pipe in the forward or backward region. The polar angle θ with respect to the direction of beam is commonly expressed as the pseudorapidity

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right), \quad (3.1)$$

where the region of maximal $q\bar{q}$ production corresponds to $2 < |\eta| < 5$. A detector focusing on flavour physics needs to instrument this region to maximise the amount of data recorded.

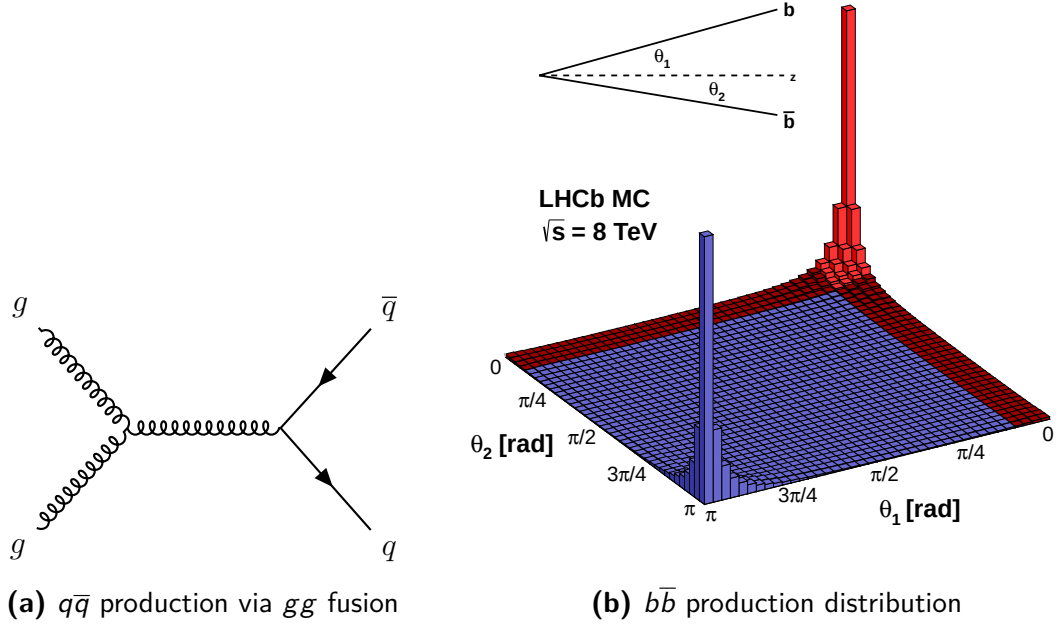


Fig. 3.2: Diagram for quark pair production via gluon-gluon fusion (a), and a simulation of the angular distribution of $b\bar{b}$ production at the LHC at $\sqrt{s} = 8$ TeV (b). The area in red marks the acceptance of the LHCb detector.

The large boost of the $q\bar{q}$ system produced leads to a substantial time dilation experienced by the produced particles with respect to the lab frame. Typical lifetimes for charm and beauty hadrons are $\mathcal{O}(0.1 - 1\text{ps})$ [77], resulting in several millimetres of flight distance in the lab frame before they decay. A sufficiently precise detector can identify that their decay products form a separate vertex that is not compatible with the primary vertex of the pp interaction. This allows a measurement of the flight distance, which is crucial in many heavy flavour measurements: the displacement from the primary vertex is a very effective tool in reducing backgrounds in the measurement, and with knowledge of the hadron's four-momentum, the decay-time can be calculated, which is an important observable in the analysis presented in Chapter 6. Furthermore, precise tracking and vertexing enables the measurement of a track's impact parameter (IP), defined as the minimum distance of the track to a primary vertex as illustrated in Figure 3.3. Large impact parameters indicate that a particle did not originate in the primary vertex, allowing to distinguish particles produced promptly in the primary vertex from those produced in the decay of heavy, long-lived particles.

The detector must be able to reconstruct the four-momenta of all decay products to infer the charm or beauty hadron's four-momentum. The momentum component of charged particles is estimated by measuring their bending in a known magnetic field, using the excellent tracking system already required for locating vertices with additional tracking systems before and after the magnet. Estimating the missing

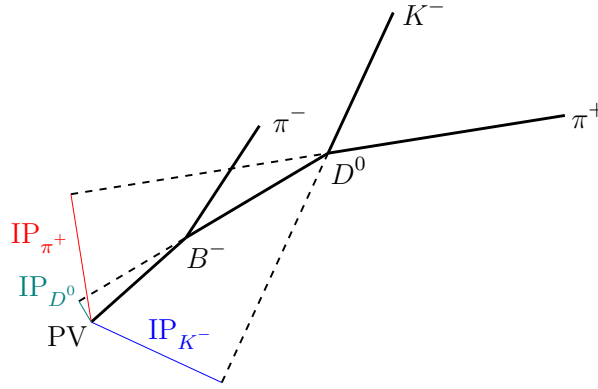


Fig. 3.3: Definition of the impact parameter variable, which gives the closest approach of a track with respect to a vertex, typically the primary vertex.

component – mass or energy – is more difficult. In principle, calorimeter systems can be used, which can associate an energy to a track, however; the measured energy in a calorimeter tends to have a large uncertainty, too large for the physics goals of LHCb. Instead, different hypotheses on the identity of a measured track are formed and the corresponding known masses are assigned. In order to assign probabilities to the different hypotheses on a track by track basis, a dedicated particle identification (PID) component is necessary. Such a system also helps to avoid the mis-identification of decays, especially in cases where the decay of interest has a smaller branching fraction than similar decays which might be mis-identified. An example is shown in Figure 3.4a, where the invariant $\pi^+\pi^-$ mass in a measurement of $B^0 \rightarrow \pi^+\pi^-$ is shown. The sample is dominated by the mis-identified $K^+\pi^-$ decays where the K^+ is incorrectly assigned the pion hypothesis, resulting in a slightly smaller invariant mass. This picture changes entirely in Figure 3.4b once PID requirements are imposed, which completely suppresses the contamination due to mis-identification. This illustrates how crucial a good PID performance is for the measurements performed in heavy flavour physics.

In the following, the design of the LHCb detector is described, detailing how the important requirements identified above are implemented by various detector subsystems.

3.3 The LHCb detector

In order to describe positions and locations, a right-handed Cartesian coordinate system is defined with the z -axis pointing along the beam axis, the x -axis being directed away from the centre of the main accelerator ring and the y -axis pointing upwards. The azimuthal angle ϕ is defined as the angle in the xy plane with respect to the positive x -axis and rotating clock-wise around the positive z -axis, and the polar angle θ is measured against the positive z -axis.

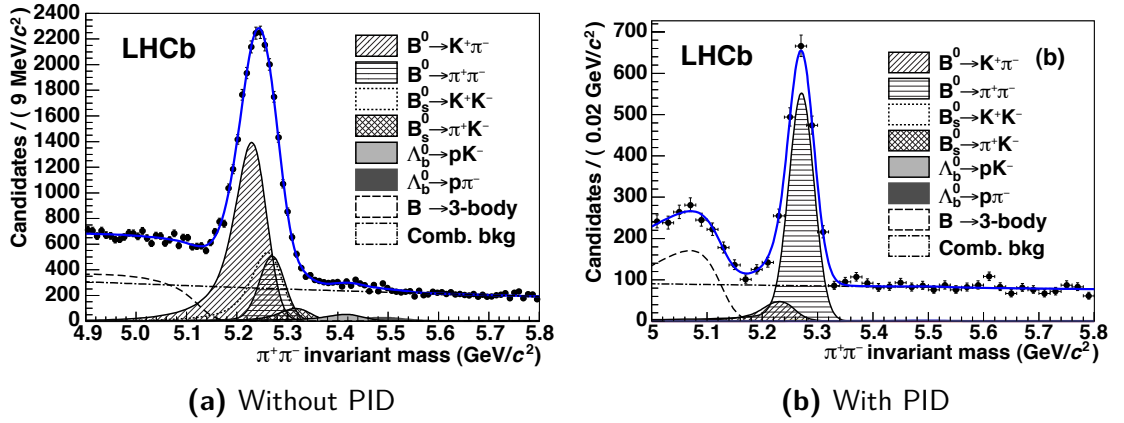


Fig. 3.4: Illustration of the power of a PID system in a measurement of $B^0 \rightarrow \pi^+\pi^-$ [78]. Invariant $\pi^+\pi^-$ distribution without PID requirements where the dominant contributions are real $B^0 \rightarrow K^+\pi^-$ decays with a mis-identified K^+ (a). The same distribution with PID requirements, dominated by real $B^0 \rightarrow \pi^+\pi^-$ (b).

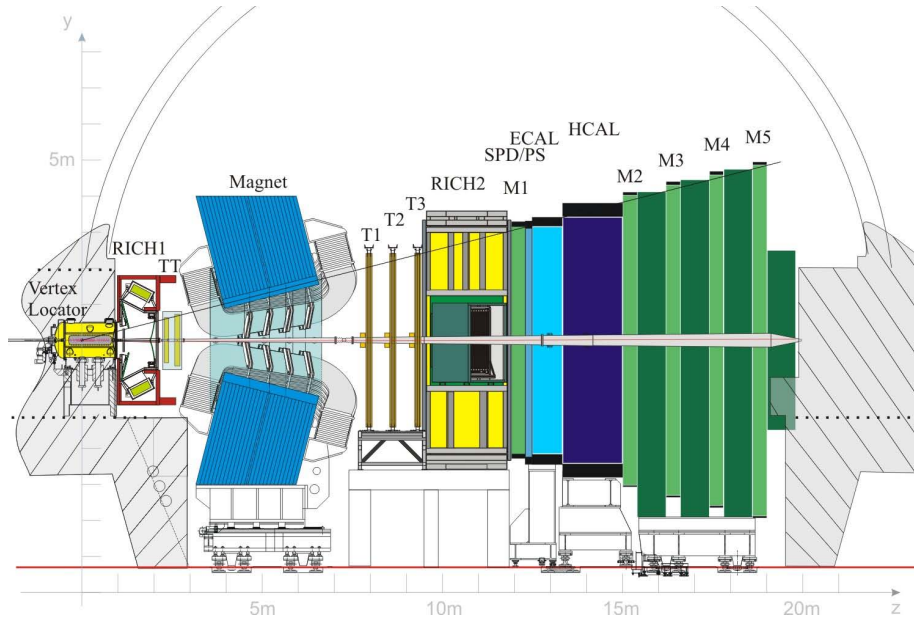


Fig. 3.5: Schematic overview of the LHCb detector showing the different subsystems traversed by particles from left to right. In this depiction, the x-axis points into the page. The thin black line indicates the acceptance in the polar angle θ .

The LHCb detector is a forward spectrometer covering a pseudorapidity range of $2 < \eta < 5$, ideally setup to study the produced $q\bar{q}$ pairs as visualised in Figure 3.2b, where the LHCb acceptance is highlighted in red. Figure 3.5 gives a schematic overview of the detector. The vertex locator (VELO), surrounding the interaction region, is responsible for the excellent vertex resolution. Following the VELO, produced particles traverse the first of two ring-imaging Cherenkov (RICH) detectors used for particle identification, and the first set of tracking stations. A 4T dipole magnet then bends charged particles in the xz plane, and the amount of bending is measured by a second set of tracking stations after the magnet, allowing to infer the momentum of the traversing charged particles. Immediately downstream is the second RICH detector, followed by the calorimeters, which can detect neutral particles and measure energies directly. Finally, a muon identification system forms the last part of the detector.

Recording and processing all pp collisions is far beyond the capabilities of the available computing resources. Furthermore, the production cross-sections for the particles typically studied by LHCb are much smaller than the total inelastic pp cross-section and thus most interactions do not contain anything interesting. Therefore, a trigger system is employed to reduce the rate of data that needs to be analysed.

The remainder of this chapter describes the different sub-detectors in greater detail and gives an overview of the trigger system and its planned evolution in the future. All measurements presented in this thesis use only Run 2 data and hence only the Run 2 configuration of the detector is discussed. However, at the time of writing this thesis, no publicly available performance studies for Run 2 exist. Therefore, Run 1 performance studies are used to quantify the detector performance, which is expected to be very similar in both LHC running periods.

3.3.1 Tracking

The LHCb tracking system is responsible for identifying tracks of charged particles traversing the detector and determining their momentum. When charged particles traverse the detector material they are able to ionise it, creating electron-ion pairs or electron-hole pairs in case of a semiconductor materials. Tracking detectors exploit this by applying an electric field to their active material where a measured current then indicates the presence of separated charge carriers and thus a traversing particle, and this information is saved as a hit. These hits are then used to reconstruct tracks. The tracking system is comprised of three different subsystems: the VELO, the Tracker Turicensis (TT) and three tracking stations (T1–3).

The VELO [79] is the first sub-detector traversed by particles produced in the interaction of its centre, and arguably the most important one: its extremely high

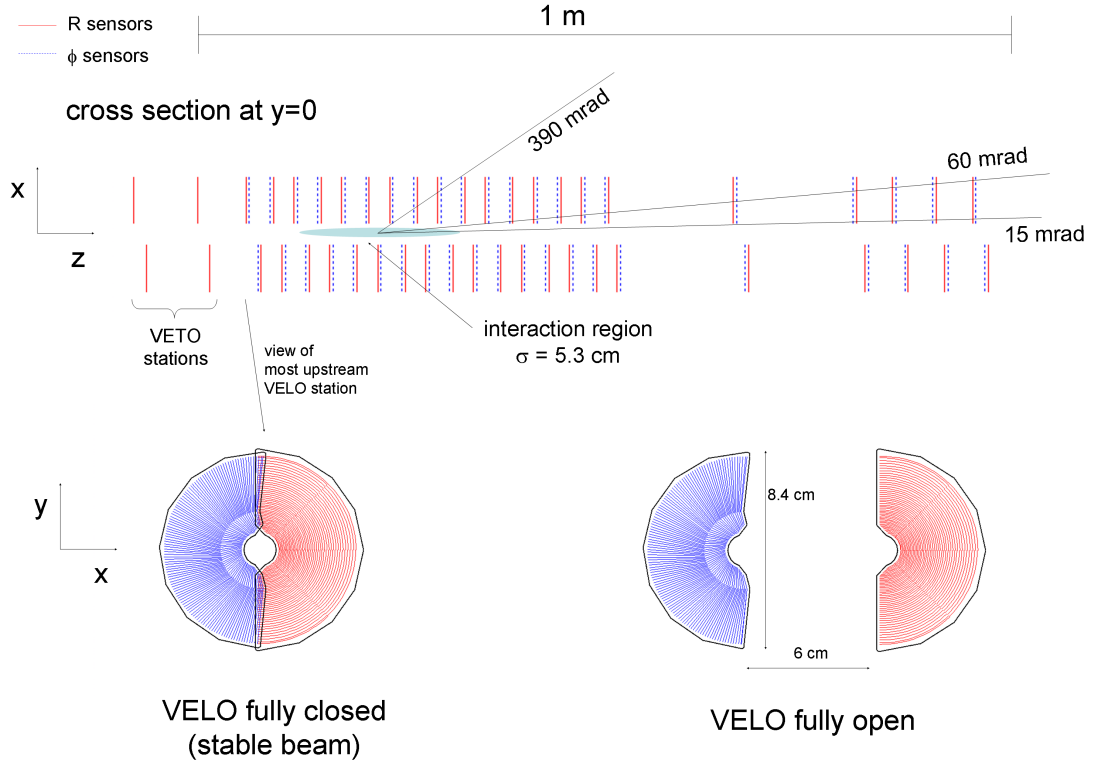


Fig. 3.6: Different schematic cross-sections of the VELO sub-detector showing the positions of the individual modules and the open and closed position. Figure taken from Ref. [71].

precision allows it to measure tracks accurately, allowing it to resolve different vertices in close proximity. The VELO is comprised of two separate, movable halves on either side of the beam pipe in the xz plane. The two halves close by moving in the horizontal direction once the beams circulating in the beam pipe are declared stable, resulting in a distance of only 8 mm between the beam and the active parts of the detector, closer than any other experiment at the LHC. A schematic view of the detector is given in Figure 3.6, showing the two halves of the detector and the open and closed position. The halves contain 21 silicon modules each, with the modules being comprised of two silicon strip sensors, segmented in r and ϕ direction, respectively. By collecting the produced electron-hole pairs in the individual strips and combining the information of neighbouring r and ϕ sensors, the spatial position of a hit created by a traversing charged particle is determined.

The LHCb tracking system contains two additional silicon strip detectors: the Tracker Turicensis (TT) and the Inner Tracker (IT), which comprises the inner part of the TT stations surrounding the beam pipe resulting in four modules in total. Each of these modules consists of four layers of silicon strips. To identify the three-dimensional point where the charged particle passed through, the silicon strips in different layers cannot be aligned. Therefore, the strips in the first and last layer are oriented vertically while the ones in the second and third layer are slightly rotated. Due to

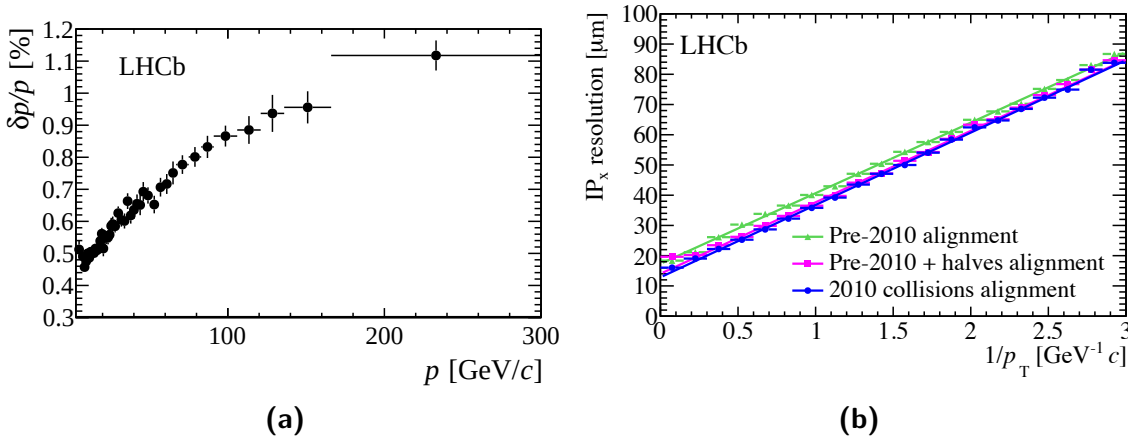


Fig. 3.7: Tracking system performance in Run 1. Momentum resolution as a function of particle momentum p (a). Impact parameter resolution as a function of the inverse transverse momentum of the particle $1/p_T$ (b). Figures are taken from Ref. [78].

costs, only the inner region of the T stations, which has a large particle occupancy as a result of its close proximity to the beam pipe, uses this silicon strip design. The Outer Tracker (OT), covering the remainder of the LHCb acceptance, uses straw tubes filled with a mixture of Argon, Oxygen and CO_2 . Electron-ion pairs produced in the gas are collected by an anode wire running along the centre of the tube and the tube walls. Additional timing information allows it to pinpoint the radial distance of the particle interaction in the tube and combining the information provided by the different layers of straw tubes again allows the calculation of a three-dimensional position where the charged particle traversed the detector.

The performance of the tracking system in Run 1 is illustrated in Figure 3.7: the momentum resolution reaches 0.5% for low momentum tracks and increases to 0.8% for $p = 100 \text{ GeV}/c$. The resolution on the impact parameter reaches $80 \mu\text{m}$ for particles with a low transverse momentum and reduces to $15 \mu\text{m}$ for those with a high transverse momentum. Using a tag-and-probe technique and hits made in the muon identification system, the average track reconstruction efficiency has been measured to be above 95% [80].

3.3.2 Particle identification

The previously described tracking system allows precise measurements of the momenta of a charged particle. To fully reconstruct its four-momentum, a measurement of its mass or energy is necessary. The velocity of a particle can be measured by exploiting Cherenkov radiation which in turn allows the computation of the mass of the particle using the relation between momentum and velocity. Cherenkov radiation occurs once a particle traverses a medium faster than the speed of light in this medium, and the

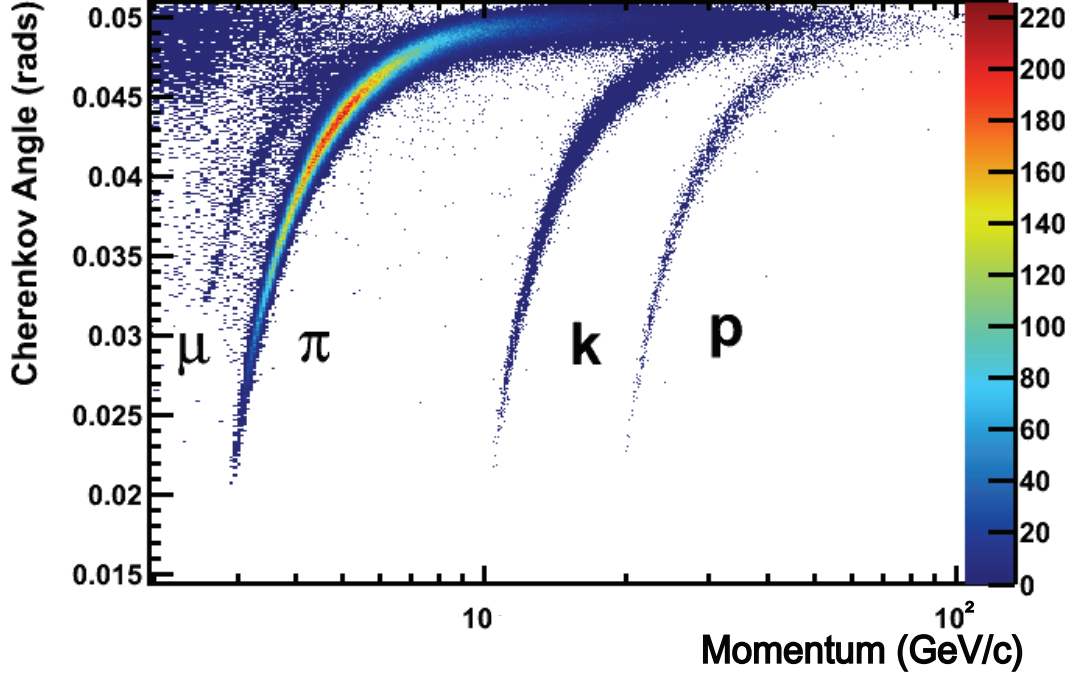


Fig. 3.8: Cherenkov angle θ_c measured with C_4F_{10} as a radiator. Contribution from muons, charged pions and kaons as well as protons are distinctly visible in the spectrum. Due to their small mass, no clear contribution from electrons can be identified. Figure taken from Ref. [81].

angle θ_c under which the radiation is emitted depends on the velocity:

$$\cos(\theta_c) = \frac{1}{n\beta}, \quad (3.2)$$

where $\beta = v/c$ and n is the known refractive index of the material. Therefore, a measurement of the angle θ_c allows an estimation of the velocity and subsequently the mass of the traversing particle.

In order to measure the angle θ_c , LHCb has two ring imaging Cherenkov sub-detectors, RICH1 and RICH2. RICH1 is located between the VELO and the TT before the magnet and is filled with C_4F_{10} as a radiator while RICH2, located behind the T stations further downstream, uses CF_4 mixed with a small amount of CO_2 to reduce scintillation light, which would otherwise impact the resolution. RICH1 covers the entire LHCb acceptance and is designed to efficiently measure the angle for particles in the momentum range of 3 to 60 GeV/c. RICH2 is intended to increase the momentum range to 100 GeV/c and covers a smaller, central area of the LHCb acceptance, which is sufficient because most high momentum particles cover the high η region. The optical photons produced by Cherenkov radiation are reflected outside of the LHCb acceptance by mirrors and collected by photo multiplier tubes (PMT).

Figure 3.8 shows the two dimensional distribution of particle momenta and corresponding angles for C_4F_{10} , with the contributions from different particle species

being clearly distinguishable. As the momentum is known from the tracking system, the compatibility of different particle species hypotheses can be tested for a given angle. It should be pointed out that the absence of Cherenkov light is a powerful discriminatory feature as well, *e.g.* a particle with 10 GeV/ c showing Cherenkov light cannot be a proton. In practice, each particle is initially assumed to be a pion and a likelihood is computed by comparing the number of observed optical photons in each PMT with the expected number for pions. The hypothesis is then changed, the expected number of optical photons is updated and the change in likelihood is computed. These variables are referred to as DLL (for delta-log-likelihood) and they exist for each possible particle hypothesis. As they all express the change with respect to the pion hypothesis, the value of DLL for the pion hypothesis is zero by construction.

3.3.3 The calorimeters

The sub-detectors discussed above allow for a precise measurement of a particle's four-momentum; however, two caveats apply: no measurements of neutral particles are possible and reconstructing tracks takes too long to be feasible at an event rate of 40 MHz. Both caveats are addressed by the calorimeters, which can measure the neutral particles and their information can be used to apply a meaningful trigger to reduce the event rate to 1 MHz at which the whole detector can be read out. The LHCb calorimeter system consists of an electromagnetic (ECAL) and hadronic calorimeter (HCAL). The HCAL is a sampling calorimeter consisting of alternating layers of iron acting as the absorber and scintillating material. Incoming hadrons deposit large amounts of energy in the absorber plates and create a particle shower. These shower particles then excite the scintillating material, which emits light measured by photo multiplier tubes. The same design principle is common to all LHCb calorimeters. In order to obtain positional information, the HCAL is segmented into square cells, 130 mm wide in the centre and up to twice that size at the edges to account for the lower particle multiplicities. The relative resolution on the energy of the hadrons in Run 1 has been measured to be $(69 \pm 5)\%/\sqrt{E/\text{GeV}}$ plus a constant $(8 \pm 2)\%$ [82]. The ECAL is positioned upstream of the HCAL and uses lead as the absorber material. It is designed to measure the energy of photons and to improve the PID information for electrons (Figure 3.8 illustrates the poor performance of the RICH detectors for electrons as no clear contribution can be identified). Furthermore, energy depositions can be used to trigger on electrons and photons. The ECAL is segmented into cells similar to the HCAL, albeit with a finer granularity to account for electromagnetic showers being smaller in the transverse plane. The relative energy resolution in the ECAL during Run 1 has been measured to be $(9 \pm 0.5)\%/\sqrt{E/\text{GeV}}$ plus a constant 0.8% [82]. Electron and photon clusters can be differentiated by checking for the existence of a matching track pointing to

the cluster in the ECAL. However, this requires that the tracking has already been performed for the specific events, something which is too slow for the lowest level of the LHCb trigger system. Therefore, two additional layers are added in front of the ECAL: the scintillating pad (SPD) and preshower (PS) detectors. Both are separated by a 15 mm thick layer of lead. An incoming electron already produces a signal in the SPD detector before showering in the lead layer while a photon will pass through the SPD unseen and start its electromagnetic shower in the lead layer. Both particles will then leave a signal in the PS detector and thus electrons are distinguished from photons by the presence of a signal in the SPD.

3.3.4 The muon system

Muons are able to traverse large amounts of material and thus are not stopped in the calorimeters described above. Therefore, a dedicated muon detection system, comprised of five individual stations (M1–5), is installed in LHCb. All stations consist of multi-wire proportional chambers with the exception of M1, which has a thin layer of gas electric multipliers in its centre. M1 is located upstream of the calorimeters while M2–5 are located downstream, constituting the final sub-detector. Additional 80 cm iron absorbers are placed between M2 and M5 to filter out particles other than muons that have punched through the calorimeters. The muon system fulfils several important roles (at least for analyses involving muons in the final state, a pristine group that none of the measurements presented in this thesis belongs to): providing additional PID information for muons, improving their track precision by adding additional hits to the fit and, most importantly, allowing for triggering on muons. A rudimentary p_T reconstruction is performed, allowing the system to trigger on both their existence and momentum.

3.3.5 The trigger

LHCb employs, like most high energy physics experiments, a trigger system to reduce the rate of pp collisions that require a full read out of the detector while selecting events that are more likely to contain interesting physics. The LHCb trigger system is comprised of three different stages: the hardware level-zero (L0) trigger and two high-level triggers (HLT1 and HLT2) that are implemented in software. All three implement multiple trigger lines. The levels have to handle successively lower rates and can thus employ more refined selection criteria as each level only processes events accepted by the previous level.

The L0 trigger accepts events at the full bunch crossing frequency of 40 MHz and its output rate is limited to 1 MHz by the processing capabilities of HLT1. At 40 MHz, only the calorimeter and muon system information can be read out and used in

the L0 decision. The muon stations allow the inference of the number of muons in the event and a rudimentary transverse momentum estimate from the muon track orientation allows the selection of muon decays, resulting in the single and dimuon trigger lines. The measurement presented in Chapter 6 relies heavily on a line utilising the information provided by the calorimeters by searching for the largest E_T in all 2×2 cells in the calorimeter, which is then required to pass a threshold of typically 3.5 to 4 GeV. Lastly, L0 can trigger randomly on bunch crossings, which does not bias the selected events and avoids a complicated estimation of the efficiencies of the hadron line while sacrificing a large amount of data. Nonetheless, such a trigger is beneficial when measuring production cross-sections as presented in this thesis.

If any of the L0 lines fired, the entire detector is read out and HLT1 performs a simplified reconstruction of the event. As a first step, all tracks in the VELO are reconstructed and primary vertices are identified. In a second step, tracks are identified that have been produced in the decay of a heavy particle instead of directly in the pp collision. To this end, all VELO tracks are removed that either have no significant displacement to any of the primary vertices or do not match a muon track reconstructed in the muon system. The remaining tracks are then matched to tracks formed in the T stations, enabling a measurement of the track's momentum from the bending in the magnet. From this information, multiple HLT1 lines are defined again for single and dimuon events as well as hadrons with a total output rate of 80 kHz.

Once an HLT1 line fired, the entire event is pushed to a buffer storage from which HLT2 takes events to process. However, before doing so, at the beginning of each LHC fill, the detector alignment is performed in real-time [83], allowing HLT2 to use the final alignment constants in its reconstruction. Once the alignment is completed, HLT2 starts processing the buffered events asynchronous to the LHC, allowing for an ideal use of the available computing resources. HLT2 performs a full event reconstruction, reaching the same precision as an off-line reconstruction of the data, and selecting a multitude of different lines reducing the effective output rate to 12.5 kHz. From here on, two pathways exist depending on the line that fired: for certain lines, especially those selecting inclusive decays, the entire event is written to disk and analysed later. For many exclusive lines however (this includes all lines used in this thesis), only the candidates reconstructed in the trigger that fired the specific lines are saved to disk. This path is called Turbostream [84], and is a new development for Run 2. It reduces the disk space requirements as well as enabling a faster availability of the recorded data for analysts.

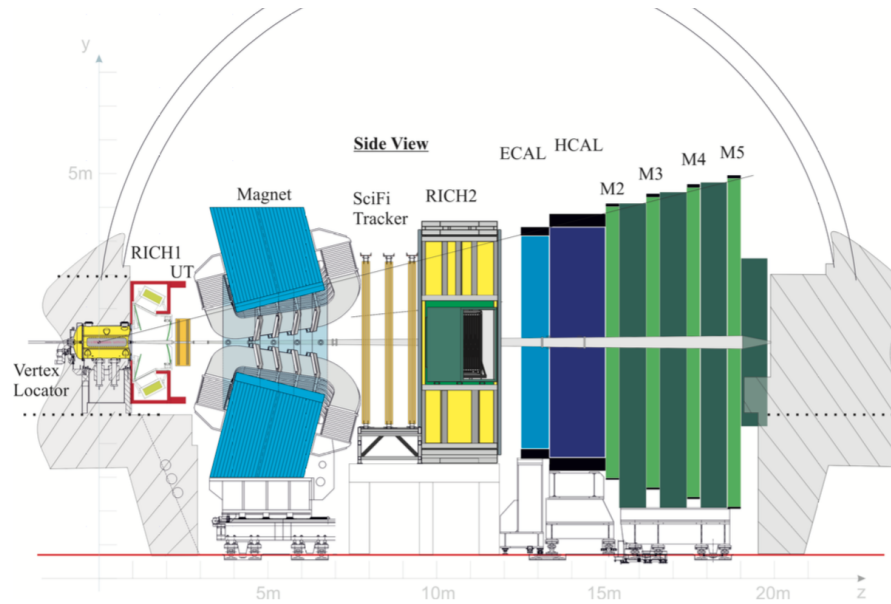


Fig. 3.9: Schematic overview of the LHCb detector after the upgrade. Figure taken from Ref. [85]. The current detector is shown in Figure 3.5.

3.3.6 The LHCb upgrade in a nutshell

The LHCb detector will be upgraded during the Long Shutdown 2 (LS2) in 2019 and 2020 and is planned to collect 50 fb^{-1} of integrated luminosity in the 10 years following LS2 during Run 3 and Run 4 of the LHC. To achieve this, the detector will run at a higher instantaneous luminosity of $2 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ and with an increased number of collisions per bunch crossing. The largest bottleneck in the current detector is the limit of 1 MHz read-out rate which is reached by the purely hardware L0 trigger selecting only basic signatures. This causes the largest inefficiencies in the trigger chain, especially for purely hadronic decays. Therefore, the upgraded LHCb detector will not use a hardware trigger, and instead will read out the entire detector at the full 40 MHz and use a full software trigger [86]. This lifts the restrictions on the information available in the first level of the trigger, for example, vertexing, tracking and momentum information would become available, significantly improving the efficiencies and purity after the first level.

In the current detector, only the muon system and the calorimeters are able to be read out at the required rate and thus most of the front-end readout electronics as well as some of the sub-detectors will be replaced in the upgrade. Figure 3.9 depicts a schematic overview showing the detector with the upgrade installed. The VELO with its silicon strip design will be replaced by a pixel detector [87] with an improved vertex resolution to tackle an increased number of primary vertices per bunch crossing. The remaining components of the tracking system will also be replaced: the TT will be replaced by the Upstream Tracker (UT) which uses high

granularity silicon micro strips and has a larger acceptance compared to the TT while the T stations (the IT and OT) will be replaced by a scintillating fiber detector (SciFi) [85]. The RICH detectors will remain, in principle, in place; however, updates to the optical system are planned to handle the increase in occupancy. Additionally, new PMTs will be installed which support the increased read-out frequency [88]. Lastly, due to the discontinuation of the hardware trigger, no fast information on electron and photon discrimination is necessary and thus the SPD and PS will be removed as well as M1 due to the increased particle flux. The ECAL and HCAL as well as the remaining muon stations M2–5 are expected to last until Long Shutdown 3 at which time they might need to be revised.

The upgrade of the detector will allow the study of heavy flavour phenomena with unprecedented statistical precision. However, an upgrade of the actual detector is only part of the challenges that need to be overcome. The creation of simulated samples, which are crucial to understand the effects of the detector, is required to be much faster to be able to harness the full potential of the upgraded detector. Therefore, the next chapter discusses the simulation effort in LHCb and developments towards making them faster are presented.

4. Fast event simulation in LHCb

4.1	Monte Carlo event generation – A brief overview	71
4.2	Monte Carlo samples at LHCb	74
4.2.1	Creation of a typical sample in LHCb	74
4.2.2	Simulation speed and LHCb production resources	76
4.3	The need for speed	77
4.3.1	Overview of different approaches	78
4.3.2	Particle Gun – A simulation of the exclusive final state . .	79
4.4	ReDecay – An LHCb fast simulation solution	81
4.4.1	Idea and procedure	81
4.4.2	Variants and implementation	82
4.4.3	Validation and performance	83
4.4.4	Statistical treatment	85
4.4.5	Summary and impact	91

Extracting physics knowledge from recorded data requires the correction of both resolution effects and inefficiencies in the detector. With the increasing complexity of the experimental environments present at collider experiments, analytic descriptions of the detector response are hardly obtainable, and data-driven methods cannot be applied to all problems. Therefore, simulated samples are a crucial tool in many analyses. This chapter presents an overview of the general as well as LHCb-specific procedures to create such samples before discussing contributions I have made to the LHCb simulation effort. These contributions result in an increase in simulation speed of one order of magnitude.

4.1 Monte Carlo event generation – A brief overview

The creation of a sample of Monte Carlo events is separated into two steps: the generation, where, based on physics models, an event is generated, and the simulation, which propagates the products of the generation step through a simulation of the detector to emulate the response of the real detector. While many details can differ between various implementations, this section gives a brief summary of these steps in a general way.

The generation of Monte Carlo events is a complicated process with a variety of challenges, each with its set of solutions and thus it is only possible to scratch the very surface in this summary. A detailed review is provided by A. Buckley *et al.* in Ref. [89]. In the interest of staying concise, this summary only discusses the case

of a central hard scattering process (*e.g.* the production of a $c\bar{c}$ or $b\bar{b}$ pair) being generated. The entire process relies on the factorisation theorem in QCD, which allows splitting the whole process into multiple phases: the hard scattering process resulting in partons, the radiation of additional hard QCD partons and the soft QCD interactions combining the partons into hadrons. This coincides with the transition from very high to low energies and hence from the perturbative to the non-perturbative regimes of QCD. Figure 4.1 depicts the different steps in the event generation which will each be discussed in more detail in the following.

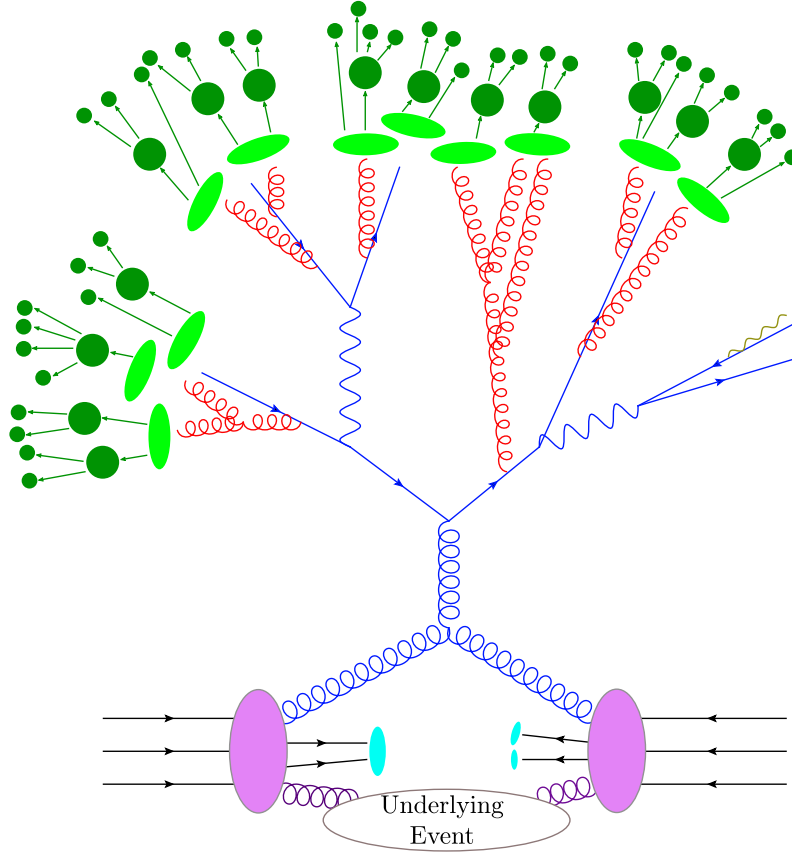


Fig. 4.1: Illustration of MC generation, depicting the colliding protons (purple), the hard scattering process (blue), QCD radiation (red), hadronisation (light green) and the decays of hadrons or leptons (dark green).

The generation of the hard process $pp \rightarrow n$ with a final state n is conceptually the same as calculating its cross-section:

$$\sigma(pp \rightarrow n) = \sum_{i,j} \int_0^1 dx_i \int_0^1 dx_j \int d\Phi_n f_i(x_i) f_j(x_j) \cdot \frac{1}{2\hat{s}} |M_{ij \rightarrow n}(\Phi_n)|^2, \quad (4.1)$$

where i and j are different partons within the protons, each carrying a fraction of the proton's momentum x_i and x_j respectively, and thus defining the initial state of the interaction. The parton density function f gives the probability to find a parton with this momentum fraction in the proton, see Section 2.2.1. As many different combinations of proton constituents may contribute to the process with

varying momentum fractions, the sum is taken over all possible combinations and then integrated over all possible momentum fractions. All final state properties are described by the phase-space element Φ_n , which includes the momentum components of the particles, resulting in a $3n - 4$ dimensional object as four degrees of freedom are constrained due to energy and momentum conservation. The variable $\hat{s}$ is the effective centre-of-mass energy of the parton collision $\hat{s} = x_i x_j s$ and $M_{ij \rightarrow n}$ is the matrix element for the hard process of interest. Monte Carlo integration is the only feasible method to solve the high dimensional integrals present in Equation (4.1):

$$\sigma(pp \rightarrow n) = \int d\vec{x} g(\vec{x}) \approx V \frac{1}{N} \sum_{i=0}^N g(\vec{x}_i) , \quad (4.2)$$

where g encapsulates the entire integrand in Equation (4.1) and $\vec{x}$ being its dependent variables spanning a total volume V . The sum is taken over N randomly generated values for $\vec{x}$. Using this integration technique to compute the integrated cross-section of a process is intrinsically linked to the generation of a sample of events following the distributions dictated by the physics encoded in the matrix element and parton density functions as such a sample is obtained by weighting the events $\vec{x}_i$ with a weight given by $g(\vec{x}_i)$. Furthermore, unweighted samples can be obtained by applying a hit and miss procedure to the weighted events.

In pp collisions, most final state partons are coloured particles and thus can radiate additional gluons. Due to the high energies of the partons from the hard scattering process, this process is still in the perturbative regime. A Markov chain is used to iteratively add more and more additional gluons to the event, creating parton showers in the event. This gradually reduces the energy per parton in the event until energies become so small that the involved QCD becomes non-perturbative. The same procedure is also applied to simulate interactions between other constituents of the proton that did not participate in the hard scattering process themselves.

At the low energy scale where the perturbative evolution of the partons is not possible anymore, and the partons form bound states, phenomenological models of the clustering take over. These are not directly based on QCD as no calculations in the non-perturbative regime can yet be performed. Instead of acting on individual partons, these models form colourless groups by combining several partons, which are then replaced by hadrons matching the parton content of the groups.

Most of the hadrons produced in the hadronisation step are unstable resonances which are decayed using a variety of different models, ranging from simple phase-space decays to decays via complicated, interfering amplitudes that are usually derived from measurements. An example for this is the $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ decay discussed in

this thesis. As the summary presented here is very general, the tools and procedure used in LHCb are discussed in detail in Section 4.2.1.

If the generated events are intended to be compared to measured data or to evaluate efficiencies, the effect of the detector on the stable particles needs to be modelled. In a simplified picture, this is done by moving the particle through the detector in many incremental steps. Typically, the step distance is dependent on the radiation length of the specific medium with additional steps being added at any transition between different materials. At each step, the energy and direction of the traversing particle are updated to reflect the energy deposition in the material between the current and previous step via ionisation. Additionally, at each step, the particle has a chance to undergo an inelastic interaction producing new particles which are then, in turn, also tracked through the remaining detector.

4.2 Monte Carlo samples at LHCb

Simulated samples in LHCb can be created using different generators with slightly different procedures. Providing a description for all of them exceeds the scope of this chapter. Nonetheless, a discussion of the most commonly used setup and its performance is presented in the following. This method is also the one used for the measurements presented in this thesis, discussed in Chapter 5. In general, simulated samples are only created for particular decays, either for the signal process of interest or physically motivated background. Combinatorial backgrounds are evaluated using data-driven methods. In the following, the term ‘signal particle’ refers to the head of the decay of interest. For example, if a sample for $D^0 \rightarrow K^- \pi^+$ should be generated, D^0 is the signal particle and the entire decay is referred to as signal decay.

4.2.1 Creation of a typical sample in LHCb

In LHCb, the typical event simulation starts with the generation of minimum bias¹ pp collisions using PYTHIA 8.1 [90, 91] with a particular LHCb configuration similar to the one described in Ref. [92]. PYTHIA performs all the steps detailed in Section 4.1 with the exception of the decays of hadronised particles and thus the unstable resonances are part of its output. Minimum bias events are preferred to the generation of specific hard scattering processes to achieve an unbiased mixture of the different hadrons after hadronisation. Pile-up, which refers to multiple pp interactions in one bunch crossing, is added by generating multiple such pp collisions within the same event. All hadronised particles that are heavier than the desired signal particle, and hence could be able to produce it, are then iteratively decayed. This process stops once all heavier particles have been decayed. If no signal particle

¹ An event where no particular hard scattering process is required.

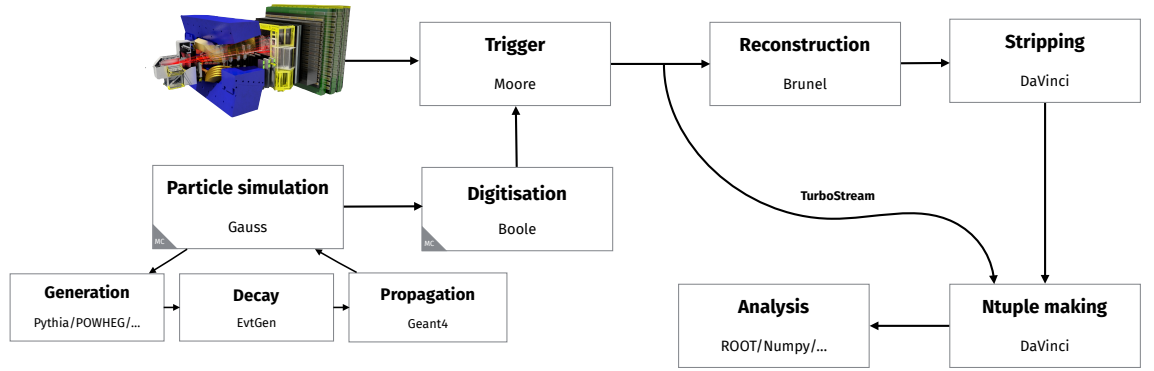


Fig. 4.2: The LHCb data flow through the various steps of the data processing in Run 2.

is identified among the undecayed particles, the process is repeated by generating a new minimum bias pp collisions. Once a matching particle is found, it is decayed to the desired final state. All remaining unstable particles then decay randomly to any of the available final states, respecting the known measured branching fractions. EVTGEN [93] performs the actual decays, either by generating simple phase-space distributions for the final state particles or by applying a particular amplitude model. Additional final-state QED radiation is then generated with PHOTOS [94]. At this stage, the event consists of only quasi-stable particles — mostly kaons, pions, electrons, muons, neutrinos, and photons — which have lifetimes long enough to traverse at least parts of the detector, and be measurable.

Propagating events through the simulation of the detector whose signal decay products have no chance of being reconstructed or selected is a waste of expensive computing resources. A situation like this can occur if decay products are outside of the geometrical acceptance of the detector and thus cannot be reconstructed, or in cases where the event selection will impose criteria that, even taking resolution effects on the reconstructed quantities into account, are impossible for the generated event to meet. These events are usually filtered out in this step of the sample creation to save computing resources. In order to classify the different simulated samples produced by LHCb, they are grouped by their event type, which combines the information on the signal particle and decay with the selection applied in the filtering. The implementation of the interaction of the generated particles with the detector, and its response, uses the GEANT4 toolkit [95] as described in Ref. [96].

All of these previous steps are controlled by GAUSS [96], the LHCb simulation framework which is built upon the GAUDI [97, 98] software framework and interfaces to the external implementations for the generation, the decay, and the detector propagation. For each event, the output of GAUSS consists of the information on the generated particles and the interaction with the detector, namely the spatial positions of interactions with the associated energy deposition.

The information about the detector interactions is then digitised and noise is added, resulting in discrete values associated with read-out channels like real data. This digitisation is implemented in BOOLE, and the LHCb software stack can process the resulting output in the same way as real data with the exception that the association between reconstructed and true particles is propagated and available to the user. The overall data flow for simulated events, compared to real data, is also depicted in Figure 4.2.

4.2.2 Simulation speed and LHCb production resources

Tab. 4.1: Average time per event spent in the different sub-detectors in simulation. Estimated from 10 minimum bias events.

Detector volume	Time / event [s]
RICH 2	31.2
Electronic cal.	21.3
RICH 1	14.2
Hadronic cal.	7.4
VELO	2.7
Muon stations	2.0
Tracking	1.5
Magnet	0.8
Other	9.6
Total	90.7

Among all the steps in the LHCb data flow that are required to obtain a simulated sample that can be used in an analysis, the propagation of the particles through the detector is, by far, the most time-consuming. Depending on the complexity of the event and whether spillover needs to be simulated, this step takes on the order of minutes while all other steps require only fractions of a second. Table 4.1 breaks down the time spent in the simulation of the LHCb detector. The timing has been obtained by monitoring the time spent in every GEANT4 tracking step when simulating ten minimum bias events on an Intel Xeon E3 at 2.6 GHz. The test was conducted in an uncontrolled environment, and thus the numbers shown

are only indicative. As shown, the RICH detectors and the calorimeters dominate the total time in the simulation. In the case of the former, a large number of optical photons are produced which need to be individually tracked, substantially slowing down the simulation. The same is true for the particles produced in the electromagnetic and hadronic showers in the calorimeters.

Figure 4.3a shows the total number of events simulated within the last year by the LHCb collaboration. During this year, the production system was running at full capacity, producing around three billion events in total. Regarding operating, upgrading and maintenance costs, creating around 1,000 events on the production system costs about 0.5 CHF. This significant financial investment is even more substantial as most of the simulated samples are only usable for one analysis and hence only contribute to one publication by the collaboration. To avoid the systematic uncertainties due to the size of the simulated sample becoming larger than the

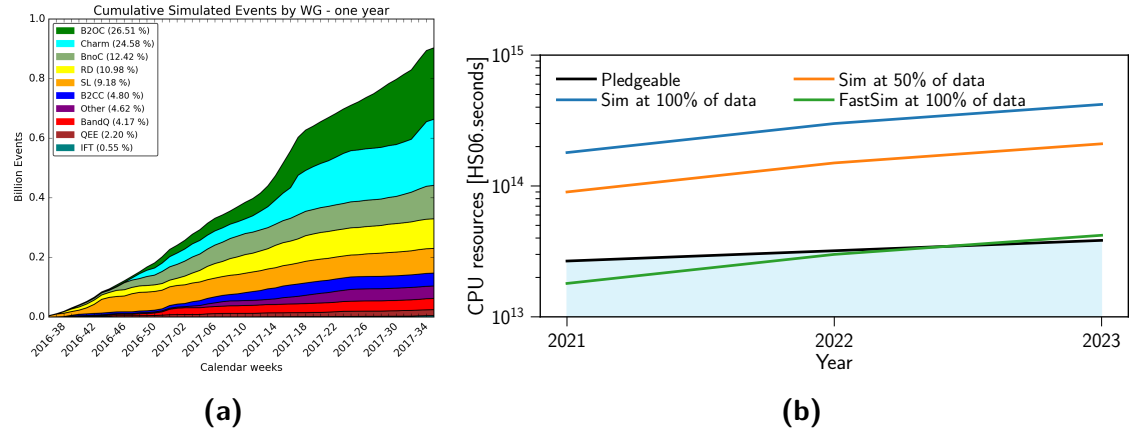


Fig. 4.3: Information on the LHCb production system: (a) number of cumulative events produced per LHCb working group, Individual contributions are stacked; (b) CPU resources required for Run 3 of the LHC for different scenarios compared to pledged.

statistical uncertainty and hence effectively cancelling the benefit of recording more and more data, the ratio of selected simulated to recorded signal events must remain constant. While the rate of pp collisions in the detector is utterly impossible to simulate, the relevant quantity for most measurements, the number of selected signal events, is achievable by applying a suitable filter before the detector simulation as described above. However, the production system is already running at full capacity, and the increase in available CPU resources is expected to be far smaller than the growth of data, especially considering the Run 3 upgrade of the LHCb detector, described in Section 3.3.6, after which the detector read-out will be 40 MHz instead of 1 MHz. This is illustrated in Figure 4.3b, which shows the required CPU resources for different ratios of simulated to real events and compares those with the pledged resources. Achieving the same number of events would require almost a factor of ten more resources than available. These predictions already include expected performance increases due to general software updates to GEANT4. Therefore, other means of increasing the speed need to be found, and different approaches are discussed in the following.

4.3 The need for speed

The need for faster simulation in LHCb is indisputable and their implementation is an absolute requirement to benefit from the increase in recorded data expected in the future. This section summarises the ongoing effort and describes an implementation that pre-dates the work in this thesis and whose shortcomings are the motivation for the this work.

4.3.1 Overview of different approaches

Simulated samples are used for different use-cases, *e.g.* analysis, analysis prototyping, systematic studies, or detector development. Many of these do not necessarily require the expensive full simulation and thus different approaches are investigated in parallel, tailored for different applications. Nonetheless, a common feature of all approaches is the need to reduce the time consumed by the simulation of the detector. Three groups of approaches can be identified and briefly summarised in the following.

Simulating no detector

As almost all the time is used by the detector simulation, the most straightforward approach is getting rid of the detector simulation altogether. The effect of the detector is mimicked by discarding generated particles to emulate efficiency effects and smearing the remaining ones to imitate imperfect resolutions. The amount of these corrections is typically based on parameterised descriptions obtained from data-driven detector performance studies or by comparing the differences of generated and reconstructed particles in the full simulation. One possible example providing such parameterisations for the LHCb detector is RapidSim [99]. Additionally, a Delphes [100] implementation is in development, which provides an improved parametric description of the LHCb detector and is fully integrated into the LHCb data flow. Using parameterised detector descriptions can increase the speed by several orders of magnitude. However, due to the lack of hits made in the detector, one sacrifices access to reconstruction-level quantities describing the quality of the tracks and vertices. Hence, these samples cannot be used to determine selection efficiencies as these quantities are often used during the event selection as they provide a powerful tool to suppress combinatorial backgrounds. Nonetheless, parametric descriptions can be used to study sources of systematic uncertainties or to test for contamination due to physical backgrounds, for example through partial reconstruction or mis-identified final state particles.

Simulating less detector

Instead of removing and replacing the entire detector simulation, the simulation can be customised to suit the particular needs by deactivating components of the detector that are not needed. Based on Table 4.1, ideal candidates are the RICH detectors and the calorimeters. It is possible to replace the simulated RICH detector response using purely data-driven techniques, which is detailed in the Section 5.5.2. Particles traversing the RICH detectors might scatter with the detector material and thus removing the detector and replacing it with air will lead to differences with respect to the full simulation beyond PID related quantities. However, it is possible to solely

deactivate the propagation of the optical photons but leave the detector itself in place. Tests have shown that this reduces the time spent in the RICH detectors to negligible levels while achieving perfect agreement with the full simulation. Simply removing the calorimeters, however, is impractical: their information is used to trigger on hadrons in the event and they significantly affect the number of particles reaching the muon stations, which in turn affects the muon based triggers. However, studies are ongoing to replace the full GEANT4 based calorimeter simulation by a parameterised approach where the positions of energy deposition in the calorimeter are sampled from previously created tables. This approach is expected to reduce the time required to simulate the calorimeters to negligible levels as well.

Simulating fewer particles

Instead of modifying the detector simulation to decrease the time spent in it, one can reduce the number of particles that need to be propagated through the detector. The total time required to simulate an event is roughly proportional to the number of particles that need to be tracked. Per event, several hundred particles need to be tracked through the detector, not including the secondary particles produced in the interaction between particles and the material. However, most LHCb measurements only explicitly look at the signal candidate in the simulation by selecting the reconstructed particles matched to the true particles stemming from the decay of interest and typical signal decays rarely have more than ten final state particles. Consequently, a substantial fraction of the total time per event is spent on particles that are never explicitly looked at. The most extreme approach trying to rectify this imbalance is a simulation of the final state particles of the signal only. This method is presented and its shortcomings are discussed in the following.

4.3.2 Particle Gun – A simulation of the exclusive final state

The simulation speed can be significantly improved by only simulating the decay of interest without any underlying event. This approach has been implemented and used by LHCb and is referred to as Particle Gun. In the following, the method is described using $D^{*+} \rightarrow D^0 \pi^+$ with $D^0 \rightarrow K^- \pi^+$ as an example.

Instead of generating a pp collision, which subsequently produces the D^{*+} meson, an origin vertex for the D^{*+} is randomly sampled from the luminous region in the center of the VELO. The D^{*+} kinematics are randomly sampled either from histograms created from generated pp collisions or flat within given ranges for the components of the momentum. The D^{*+} is then decayed and simulated as before. The performance of this method is evaluated by comparing a privately produced sample of one million events to four million events, which have been produced on the central production

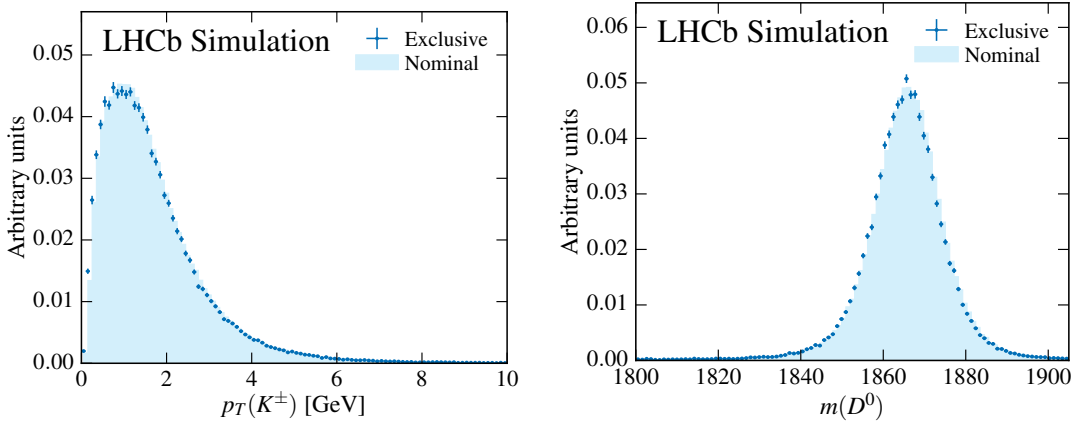


Fig. 4.4: Normalised distributions of K^+ transverse momentum and D^0 invariant mass comparing the nominal sample to the sample containing only the exclusive final state.

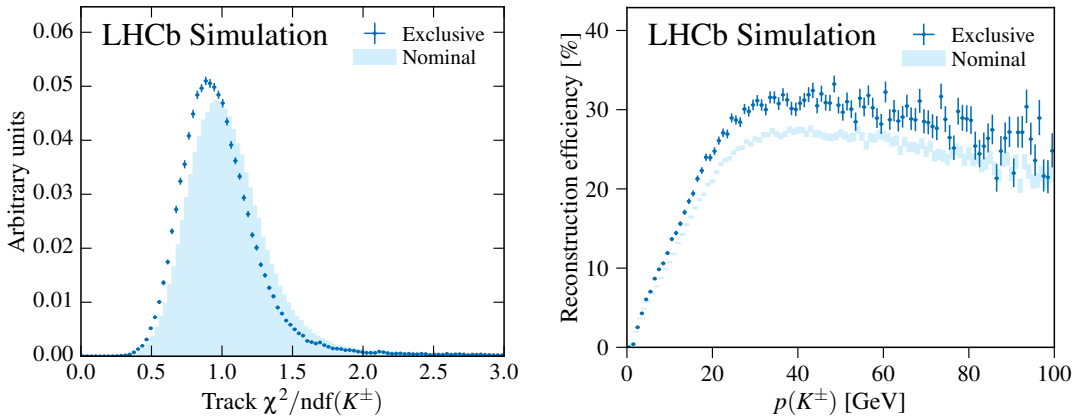


Fig. 4.5: Normalised distributions K^+ track fit χ^2 per degree-of-freedom and the reconstruction efficiency for $D^0 \rightarrow K^- \pi^+$ as a function of the K^+ momentum.

system. Both samples have the exact same conditions. Comparing the time needed to produce one event, reducing the simulation to just the final state results in a speed-up by a factor of 50.

Figure 4.4 compares kinematic distributions between the two samples and shows good agreement. This is hardly surprising as these distributions mainly depend on the ones generated, which are set up such that they are identical. However, discrepancies are seen when looking at variables that highly depend on the reconstruction performance. Figure 4.5 depicts the χ^2 per degree-of-freedom from the track fit and the efficiency to reconstruct the $D^0 \rightarrow K^- \pi^+$ decay as a function of $K^\pm$ momentum. As seen, the performance of the track fit is significantly better in the exclusive simulation compared to the nominal one. This is caused by the very low occupancy of the tracking detectors and the practical elimination of any ambiguities when associating hits to tracks. This also leads to a substantial improvement in the efficiencies. The other subdetectors, especially those strongly dependent on the occupancies like the RICH detectors, display the same behaviour where the exclusive simulation shows a better performance than the full simulation.

Besides this mis-modelling, several other quantities cannot be obtained using exclusive samples, some of which are very commonly used in LHCb to select signal events. For example, χ^2_{IP} , which can be intuitively understood as the squared significance of the impact parameter with respect to the primary vertex, is used to distinguish between prompt and non-prompt particles. It is computed by performing the primary vertex fit with and without the particle in question and taking the absolute difference of the obtained χ^2 . Fitting the primary vertex in an event requires tracks from the primary vertex to be reconstruct; however, with the exception of cases where the signal particle is an instantaneously decaying resonance, no particles from the true primary vertex are even simulated and thus no primary vertex fit can be performed in exclusive simulations.

Both of these issues reduce the range of possible applications of an exclusive simulation to studies of systematic uncertainties or physical background contributions, similar to the parametric approaches but with a more realistic detector simulation.

4.4 ReDecay – An LHCb fast simulation solution

In the previous section, the increases in speed of about a factor of 50 when only simulating the exclusive final state came at a cost: disagreement in resolution and efficiencies due to the low occupancy of the detector. This section will present ReDecay, a new addition to the LHCb simulation, which addresses the drawbacks of a Particle Gun simulation while reaching similar increases in speed.

4.4.1 Idea and procedure

The main lesson learned from comparing a Particle Gun simulation to the full simulation is the realisation that the underlying event², which substantially increases the occupancy, is important to obtain a realistic description of the real data. However, simulating the underlying event takes much longer than the simulation of the actual signal decay. In the ReDecay approach, this is mitigated by re-using the underlying event multiple times instead of generating a new one for every event. The signal particle is kept identical in all sets of events that use the same underlying event to preserve the correct correlation with the underlying event but is independently decayed again each time. To be more precise, in each redecay event, the origin vertex and kinematics of the signal particle are identical, while the decay time and thus decay vertex as well as the final state particles are different.

The following procedure has been implemented in GAUSS:

² Here, underlying event refers to all constituents of an event which do not belong to the studied signal decay.

1. A full event is generated as described in Section 4.2.1. In order for the re-use of the underlying event to be sensible, only event types with a specific signal particle can be used. This event will be referred to as the original in the following discussions.
2. Before the detector simulation, the signal and its decay products are removed from the event and the origin-vertex position and momentum of the signal particle are stored.
3. The remaining event is simulated as usual and the entire output, the information on the true particle and the energy depositions in the detector, is kept.
4. A signal decay is generated and simulated using the stored origin vertex position and momentum similar to a Particle Gun event.
5. The persisted underlying event and the signal decay are merged and written out to disk as a full event.
6. Repeat N_{ReDecay} times from step 4.

The thus produced events, *i.e.* the combination of an underlying event with a signal decay, are then independently digitised, including different random noise, and passed through the remaining LHCb data flow. The individual events are indistinguishable from normal events. This approach has multiple benefits:

- Each event obtained could have been produced by chance in a nominal simulation as ReDecay just reorders the already factorised approach: Hadrons are decayed independently and the quasi-stable tracks are propagated through the detector individually. Therefore, the efficiencies and resolutions are, by construction, identical to those found in a nominal simulation.
- With increasing N_{ReDecay} , the average time per event on disk approaches those of Particle Gun, resulting in a potential increase in speed by a factor of 50.
- The approach is fully compatible with fast simulation options that modify the detector simulation.

However, different events stemming from the same original event are correlated, *e.g.* they all have the same number of primary vertices. This potentially complicates applications of these samples as the correlations need to be taken into account. The statistical treatment of these samples is discussed in Section 4.4.4, after a description of different variants and validation studies that have been performed.

4.4.2 Variants and implementation

Early studies revealed a potential issue with the ReDecay approach when looking at decays of particles with relatively high cross-sections and a large branching fraction. In this case, it is possible that multiple signal particles are found in a generated minimum bias event. Applying the procedure in Section 4.2.1, one of these particles

is then randomly selected and forced to decay to the desired final state. All remaining particles are free to decay to any final state, including the one of the signal decay. If that happens, multiple signal decays exist in the event and consequently need to be removed from the original event, redecaied and reattached to the simulated underlying event to build the new events. The severity of this effect depends on the probability that the hadronisation yields multiple signal particles and the branching fraction for the specific decay of interest. Users can choose between two different variants currently available:

1. Only the signal particle that was forced to decay to a certain final state is taken out and decayed multiple times. The rest of the event is used as the underlying event.
2. Instead of redecaied the signal particle, all hadrons that are at least as heavy as the signal particle are removed. In order to generate the signal portion in the redecaied step, a fake event is created, which consists of those heavy particles maintaining their position of origin and momentum (basically mimicking the generation step typically done by PYTHIA). The procedure presented in Section 4.2.1 is then applied to find and decay a signal particle. After simulating the detector response, the results are merged with the previously simulated underlying event on the heavy hadrons.

The first method is significantly faster but suffers from the aforementioned problems if the signal decay can occur multiple times, while the second resolves this issue. The heavy quarks produced in pairs in the simulation are typically hadronised to two heavy quark resonances which then decay to the actual signal particles of interest. The second variant of ReDecay has the additional benefit that these two heavy flavour sides are redecaied in every event, giving more variation to the signal particle and hence reduces the correlation between different events.

The implementation of ReDecay in GAUSS is modular, without making any assumptions on which generators or decay tools are used. Therefore, it can be applied in a wide variety of configurations. Existing algorithms and tools are adopted as much as possible to ease maintaining the code base. The code is set up to handle any number of particles from the original event to be redecaied and adding variants or extending the two discussed above is simple. The only configuration that needs to be set by the user is the number of redecaies N_{ReDecay} and the variant to be used, everything else is automatically extracted from the settings of the nominal simulation.

4.4.3 Validation and performance

A sample of four million events of $D^{*+} \rightarrow D^0 \pi^+$ decays with $D^0 \rightarrow K^- \pi^+$, created in the nominal way and identical to the one used in the Particle Gun comparisons, is

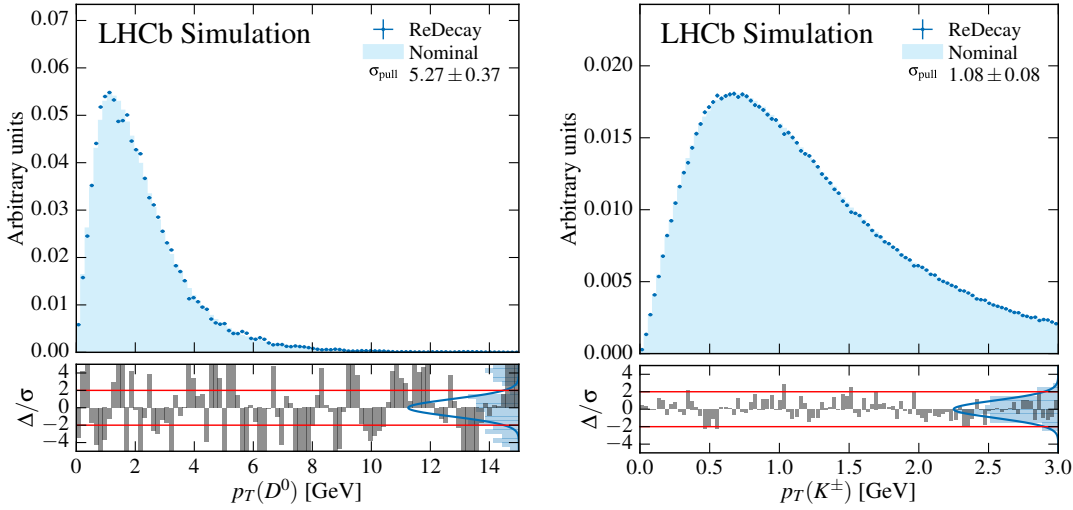


Fig. 4.6: Normalised distributions of transverse momenta of the true D^0 and K^+ at the level of the event generator. Shown pulls are calculated from simple $\sqrt{N_i}$ uncertainties in each bin i . A projection of the distribution of the pulls compared to a unit normal distribution is also given.

compared to a ReDecay sample. This sample contains 10,000 original events with $N_{\text{ReDecay}} = 100$ redecays each. The D^{*+} mesons are very frequently produced in the hadronisation of the $c\bar{c}$ pair in the minimum bias event. Additionally, the decay to the aforementioned final state has a branching fraction of about 2.5%. Using the nominal sample, 1.1% of the events are found to have at least two candidates for the decay of interest. Therefore, the ReDecay sample is configured to redecaay all particles at least as heavy as the signal.

Comparisons are performed for a wide range of observables but in the interest of keeping this section concise, only the most relevant ones are shown here. Figure 4.6 compares the true momenta of the generated D^0 and K^+ as produced by PYTHIA and EVTGEN. The pull plots underneath each distribution have been obtained using simple $\sqrt{N}$ uncertainties for the number of events in each bin:

$$\frac{\Delta}{\sigma} = \frac{N_{\text{rd}}/N_{\text{rd}}^0 - N_{\text{nom}}/N_{\text{nom}}^0}{\sqrt{(\sigma_{\text{rd}}/N_{\text{rd}}^0)^2 + (\sigma_{\text{nom}}/N_{\text{nom}}^0)^2}}, \quad (4.3)$$

where N_{rd} and N_{nom} are the numbers in the individual bins for the ReDecay and nominal sample respectively, N_{rd}^0 and N_{nom}^0 are the total numbers of generated candidates in the respective samples in order to normalise the samples, and σ_{rd} and σ_{nom} are the respective uncertainties, which are assumed to be $\sqrt{N_{\text{rd}}}$ and $\sqrt{N_{\text{nom}}}$ for now. A good agreement between the ReDecay and nominal sample can be seen, with the exception of the pulls for the D^0 transverse momentum. These do not follow a unit normal distribution and have a standard deviation much larger than one. This is caused by the correlation between different events and will be further discussed in Section 4.4.4.

All distributions of reconstructed quantities are normalised to the respective numbers of generated candidates $N_{\text{rd/nom}}^0$ and hence differences in the overall normalisation between the depicted ReDecay and nominal distributions indicate different efficiencies in both samples. Figure 4.7 compares several reconstructed quantities and good agreement is seen for all of them. Especially the distributions of quantities very sensitive to the detector occupancy, like the track fit χ^2 or the RICH based kaon versus pion likelihood $\text{DLL}_{K-\pi}$, show good agreement unlike comparisons performed using Particle Gun samples. No difference is seen for the overall normalisation of the distributions, implying a good agreement in the reconstruction efficiencies, which is confirmed by the plots shown in Figure 4.8. The good agreement is hardly surprising as the ReDecay events are by construction identical to those in the nominal simulation. As mentioned previously, the correct statistical treatment in the presence of inter-event correlations will be discussed in the Section 4.4.4.

The ReDecay sample has been produced on the LHCb computing grid again in an uncontrolled environment. Therefore, any timing studies are just indicative as no attempt has been made to normalise the individual worker nodes. The following numbers are based on the samples used in the validation studies with $N_{\text{ReDecay}} = 100$. The average time taken per event, averaged over all events, is 3.9 seconds, 3.3 seconds of which are taken by the redecays of the signal. Only 0.6 seconds are used for the simulation of the original events. Comparing this timing to the nominal simulation, ReDecay is about 15 times faster. This is slower than the previously stated factor 50, which is caused by the more time consuming ReDecay variant being used as more particles need to be propagated through the simulation of the detector.

4.4.4 Statistical treatment

The visual agreement between the ReDecay and nominal sample is generally good. However, several pull plots in Figures 4.6 and 4.7 reveal problems in the assignment of uncertainties to the ReDecay sample. This is caused by the correlations between different events in the sample if those events stem from the same original event. The effect is more pronounced in cases where the observable is similar in all redecays events (*e.g.* the D^0 transverse momentum). Correlations between random numbers X can be studied using autocorrelation:

$$R(\tau) = \frac{1}{(N - \tau)\sigma^2} \sum_{t=0}^{N-\tau} (X_t - \mu)(X_{t+\tau} - \mu). \quad (4.4)$$

This computes the correlation coefficient for a list of numbers with itself where N is the length of the list, and μ and σ are mean and the standard deviation of all entries in the list. The list is shifted by an offset τ with respect to itself when computing the coefficient. Based on this definition, if a quantity is truly random

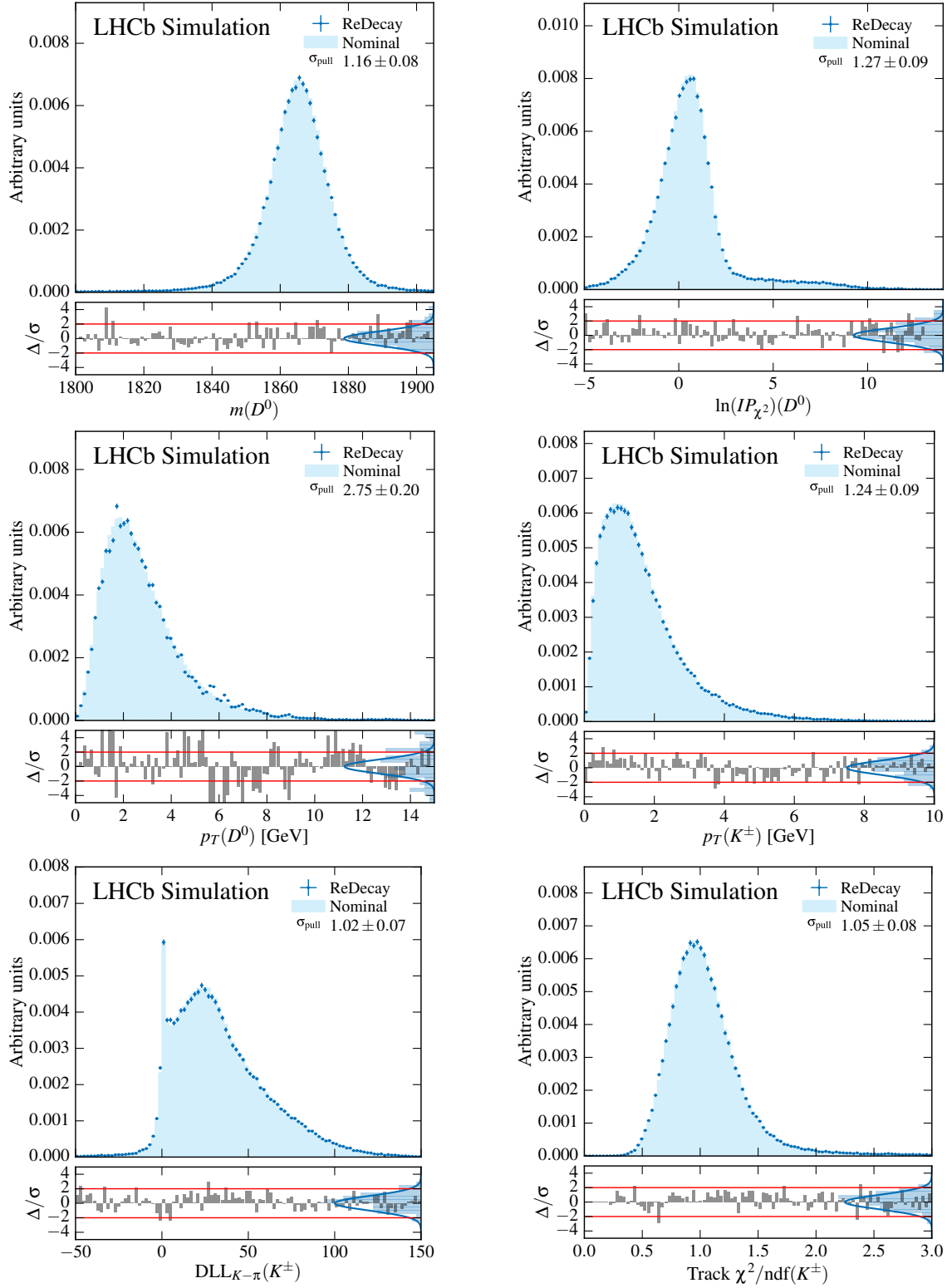


Fig. 4.7: Distributions of reconstructed quantities normalised to the number of generated candidates. Shown pulls are calculated from simple $\sqrt{N_i}$ uncertainties in each bin i . A projection of the distribution of the pulls compared to a unit normal distribution is also given.

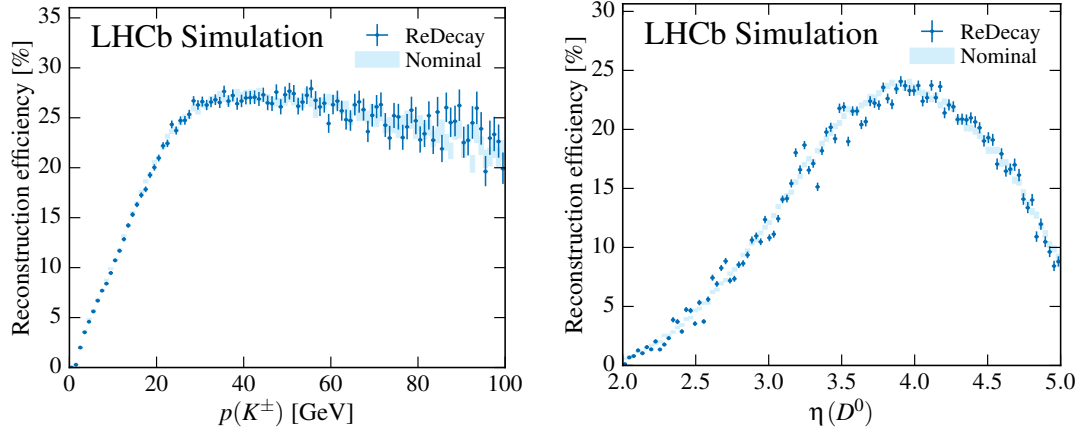


Fig. 4.8: Reconstruction efficiencies for the $D^0 \rightarrow K^- \pi^+$ decay as a function of the K^+ momentum and the pseudorapidity of the D^0 .

and independent, its autocorrelation is zero. The events in the ReDecay sample are sorted such that all events stemming from the same original event are in a block. Therefore, one expects the autocorrelation to reduce with increasing offset as the overlap of the block of events with itself decreases. After offsetting by N_{ReDecay} , the autocorrelation reduces to zero. Four examples of autocorrelations as a function of the offset are given in Figure 4.9. As expected, the nominal sample does not show any sign of correlation between the events. However, for ReDecay, large correlations are seen for the D^0 transverse momentum, decreasing with increasing offset. At around an offset of 40, the correlation practically disappears. The cut-off lower than $N_{\text{ReDecay}} = 100$ is caused by the reconstruction efficiencies as not all of the redecayed events from one original event can be reconstructed. A similar though less pronounced correlation is seen for the transverse momentum of the K^+ . This reduced correlation is caused by the larger degree of freedom in different redecay events. Figure 4.10 illustrates this by comparing the relative standard deviation seen for events stemming from the same underlying event where much larger variations are seen for the K^+ compared to the D^0 . This is not surprising: due to the small phase-space available in the decay $D^{*+} \rightarrow D^0 \pi^+$, the D^0 has very similar kinematics in each redecay event. Furthermore, Figure 4.9 shows that for observables that are mostly based on detector hits which are independent for every signal particle, the correlation almost vanishes; this is seen for the track fit χ^2 due to the track being made mostly by hits from the K^+ which are newly generated in each redecay event. This is also true for quantities that combine a larger number of input observables like the vertex fit χ^2 .

In summary, events are correlated though the amount of correlation depends on the observable studied. This explains the effect seen in the pulls where the approximation of $\sqrt{N}$ for the uncertainties breaks down in the presence of large correlations. In the following, an alternative approach will be discussed to better model the statistical uncertainty in the presence of such correlations. The trivial approach for dealing

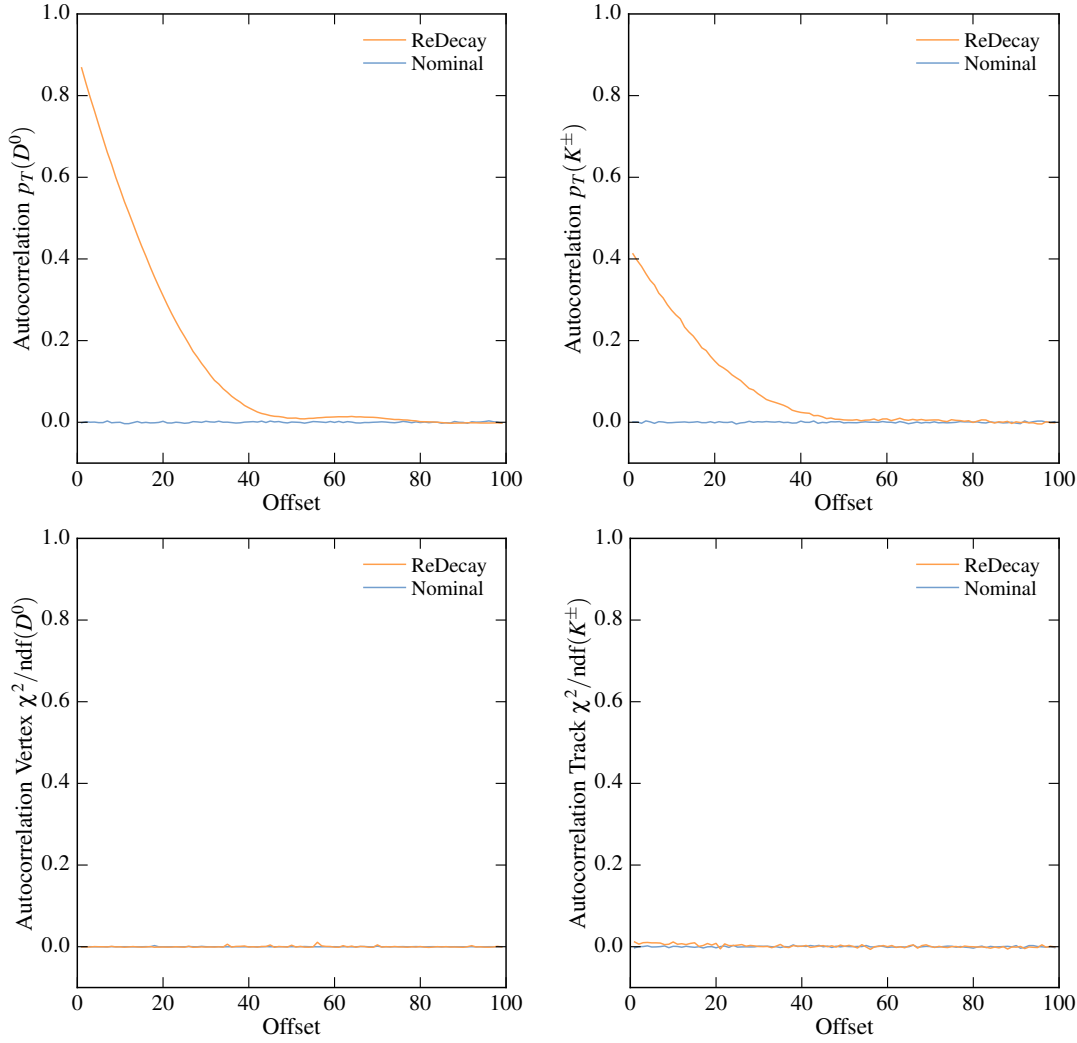


Fig. 4.9: Observed autocorrelations for the reconstructed D^0 and K^+ transverse momentum as well as the D^0 vertex and the K^+ track fit χ^2 . A definition of autocorrelation is given in Equation (4.4). The trivial Offset = 0 point is not included.

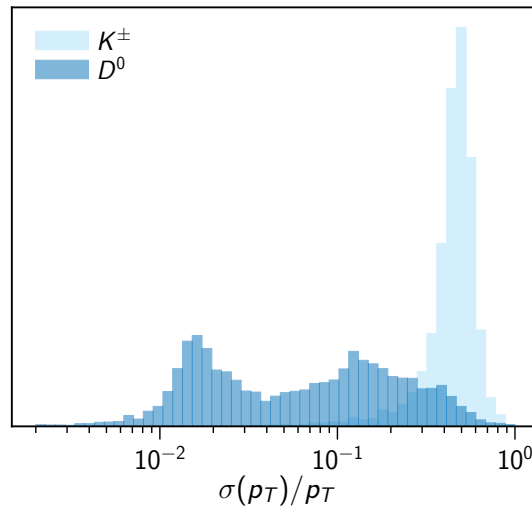


Fig. 4.10: Distribution of the standard deviation $\sigma(p_T)$ computed for all events stemming from the same original event over the average p_T of the same events, separately for D^0 and K^+ . As can be seen, the variation of the transverse momentum between different events re-using the same underlying event is, on average, much larger for the K^+ .

with this problem in an analysis is to make use of the increased speed, and generate so much simulated data that the systematic uncertainty due to the size of the simulated sample is negligible. Obviously, this is not a satisfying solution and a proper treatment of the statistical uncertainties is required. The severity of the effect highly depends on how the sample is used and which variables are of interest and hence deriving a general, analytical solution is difficult if not impossible. A possible solution is given by bootstrapping [101], which is a method for generating pseudo-samples by resampling. These pseudo-samples can then be used to estimate the variance of complicated estimators, for which analytic solutions are impractical.

Starting from a sample with N entries, a pseudo-sample can be obtained according to the following procedure: first, a new total number of events N' is randomly drawn from a Poisson distribution of mean N . Then, entries are randomly drawn with replacement from the original sample N' times to fill the pseudo-sample. As sampling with replacement is employed, the obtained pseudo-sample can contain some entries of the original sample multiple times while others are not present in the pseudo-sample at all. Unfortunately, this cannot directly be applied to the ReDecay samples because it assumes statistically independent entries. In the study of time series, data points ordered in time where each data point can depend on the previous points, different extensions of the bootstrapping algorithm have been developed to preserve the correlations in the pseudo-samples [102]. A common approach is the so-called block bootstrapping where the sample is divided into blocks. To obtain a pseudo-sample, the procedure above is slightly modified: instead of sampling N' individual entries from the original sample, entire blocks are drawn with replacement and all entries constituting a drawn block are filled into the pseudo-sample until the pseudo-sample has reached a size of N' entries. In general, the definition of the blocks, which are also allowed to overlap, is non-trivial and can depend on the actual dataset analysed. Fortunately, the ReDecay samples are a special kind of dependent data: they are grouped in clusters, where dependence only exists for entries stemming from the same original event. This provides a natural choice for the blocks being resampled: the clusters of events stemming from the same original event. For the following studies, $B = 10000$ pseudo-samples are created and each is used to fill one copy of the histograms. These B histograms are then condensed into one depiction by calculating the mean and standard deviation in each bin i :

$$N_{\text{rd},i}^{BS} = \frac{1}{B} \sum_{j=1}^B N_{\text{rd},i}^j \quad \text{and} \quad \sigma_{\text{rd},i}^{BS} = \sqrt{\frac{1}{B} \sum_{j=1}^B \left(N_{\text{rd},i}^j - N_{\text{rd},i}^{BS} \right)^2}, \quad (4.5)$$

where $N_{\text{rd},i}^j$ is the yield of the i -th bin in the j -th bootstrap sample and the superscript indicates that the respective quantities have been obtained from the ensemble of bootstrapped samples. Having B values for each bin, the statistical correlation

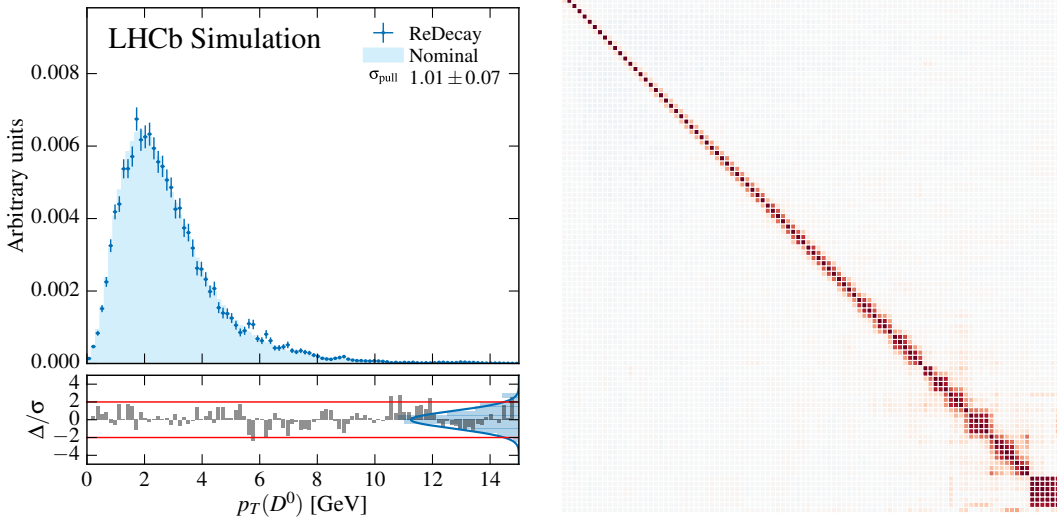


Fig. 4.11: Transverse momentum of the D^0 . For the ReDecay sample, the bootstrapped values (see Equation (4.5)) are plotted and used in the computation of the pull plot. The correlation matrix between the individual bins is shown on the right. A projection of the distribution of the pulls compared to a unit normal distribution is also given.

between different bins m and n can also be computed:

$$\text{corr}_{m,n}^{BS} = \frac{1}{\sigma_{\text{rd},m}^{BS} \sigma_{\text{rd},n}^{BS}} \cdot \frac{1}{B} \sum_{j=1}^B \left(N_{\text{rd},m}^j - N_{\text{rd},m}^{BS} \right) \left(N_{\text{rd},n}^j - N_{\text{rd},n}^{BS} \right). \quad (4.6)$$

Figure 4.11 shows the transverse momentum distribution of the D^0 candidate where bootstrapping has been used to compute the central value and uncertainties in each bin of the ReDecay histogram. The error bars are visually substantially larger than those seen in the naive treatment, depicted in Figure 4.7. The resulting pull plot follows the expectation and the pulls are now approximately distributed according to a unit Gaussian distribution. The correlation matrix, shown in Figure 4.11 as well, reveals very large correlations between neighbouring bins, which is caused by all events only populating a small range in D^0 transverse momentum if they stem from the same original event (see Figure 4.10). On the contrary, the same distribution for the K^+ , shown in Figure 4.12, is visually very similar to the one using naive uncertainties. The correlation matrix shows that instead of having very large correlations with neighbouring bins as seen for the D^0 , small correlations to a wide range of bins is seen (see again Figure 4.10). The overall importance of such small correlations is generally low, making the ReDecay approach especially powerful if the observables comprise final state quantities.

Having the bootstrapped uncertainty, it is possible to compute the fraction of nominally produced events that would be necessary to achieve the same statistical precision as the bootstrapped ReDecay sample. Again, this number will depend on the application. Thus, the following is just a discussion for the simple case of filling

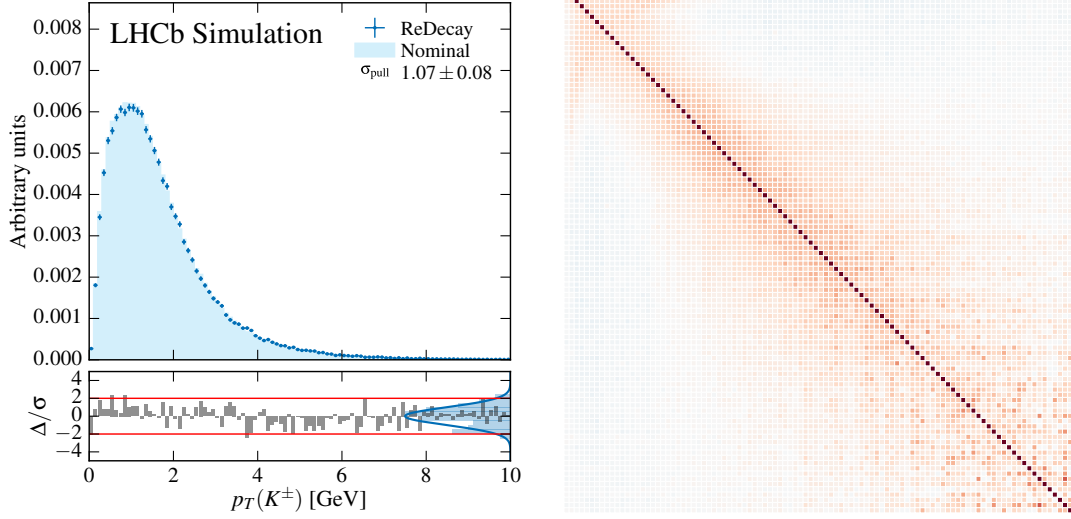


Fig. 4.12: Transverse momentum of the K^+ . For the ReDecay sample, the bootstrapped values (see Equation (4.5)) are plotted and used in the computation of the pull plot. The correlation matrix between the individual bins is shown on the right. A projection of the distribution of the pulls compared to a unit normal distribution is also given.

the simulated data into histograms (where the equivalent number is sensitive to the chosen binning scheme). This fraction f_{eq} can be computed by requiring the same relative precision in ReDecay and equivalent nominal sample in each bin:

$$\frac{\sqrt{N_{\text{eq},i}}}{N_{\text{eq},i}} = \frac{\sigma_{\text{rd},i}^{BS}}{N_{\text{rd},i}^{BS}} \Rightarrow f \equiv \frac{\sum N_{\text{eq},i}}{\sum N_{\text{rd},i}^{BS}} = \frac{\sum (N_{\text{rd},i}^{BS}/\sigma_{\text{rd},i}^{BS})^2}{\sum N_{\text{rd},i}^{BS}} \quad (4.7)$$

The equivalent fraction multiplied by the increase in speed can be understood as a measure for the effective increase in speed, factoring in the lower statistical power of ReDecay events. Equivalent nominal event fractions are given in Table 4.2 for several of the variables already shown before. For the first four variables, which are commonly used in event selections or as the independent variable when calculating efficiencies, values above 70% are found, meaning that an equivalent nominal sample would be required to have almost the same size as the ReDecay sample while being 15 times slower to produce. Even for the transverse momentum of the D^0 , the equivalent fraction is slightly larger than the necessary 1/15 required to benefit from the ReDecay approach. This is caused by the slight variation in the D^0 kinematics in each decay.

4.4.5 Summary and impact

The ReDecay approach has been implemented and validated. Extensive studies into the treatment of statistical uncertainties due to the correlation among different events have been performed and reveal that while the problem is manageable through the use of pseudo-samples, it is expected to be of minor importance for many measurements,

Tab. 4.2: Fraction of nominally produced events f_{eq} required to achieve the same statistical precision as a ReDecay sample (*i.e.* if the ReDecay sample contains N events, $f_{\text{eq}} N$ events produced without ReDecay are required for reaching the same statistical precision). Multiplying this fraction with the increase in speed of a factor 15, an measure for the effective increase is also given (which accounts for the lower statistical power of ReDecay samples). As described in the text, f_{eq} and hence the effective speed-up depend on the application, example values are given for different variables and for 100 and 200 bins separately.

Variable	f_{eq} [%]		Effective speed-up	
	100 bins	200 bins	100 bins	200 bins
$m(D^0)$	82	90	$12.3\times$	$13.5\times$
$\ln \chi^2_{\text{IP}}(K^+)$	81	89	$12.2\times$	$13.3\times$
Track χ^2	80	89	$12.0\times$	$13.3\times$
$p_{\text{T}}(K^+)$	72	83	$10.8\times$	$12.4\times$
$p_{\text{T}}(D^0)$	8	11	$1.2\times$	$1.6\times$

which mostly rely on final state quantities in the simulation. One such case where ReDecay is already successfully used is the estimation of efficiencies as a function of phase-space variables in multibody decays: an analysis determining the amplitude model (see Section 2.4) for $D^0 \rightarrow K^+ K^- \pi^+ \pi^-$ uses a ReDecay sample to determine the reconstruction and selection efficiencies as a function of five invariant mass combinations of the final state particles, which together describe the full decay. For further validation, a nominal sample with otherwise identical settings has been generated as well. Each of the samples contains 9,000,000 selected events, many more than the samples used in the studies presented above. Comparisons are given in Figure 4.13. Despite the use of naive $\sqrt{N}$ errors, the pulls follow a unit normal distribution, underlining that ReDecay is a powerful and almost cost-free method to generate samples for analyses that need to obtain something from simulated data as a function of variables depending on final state particles.

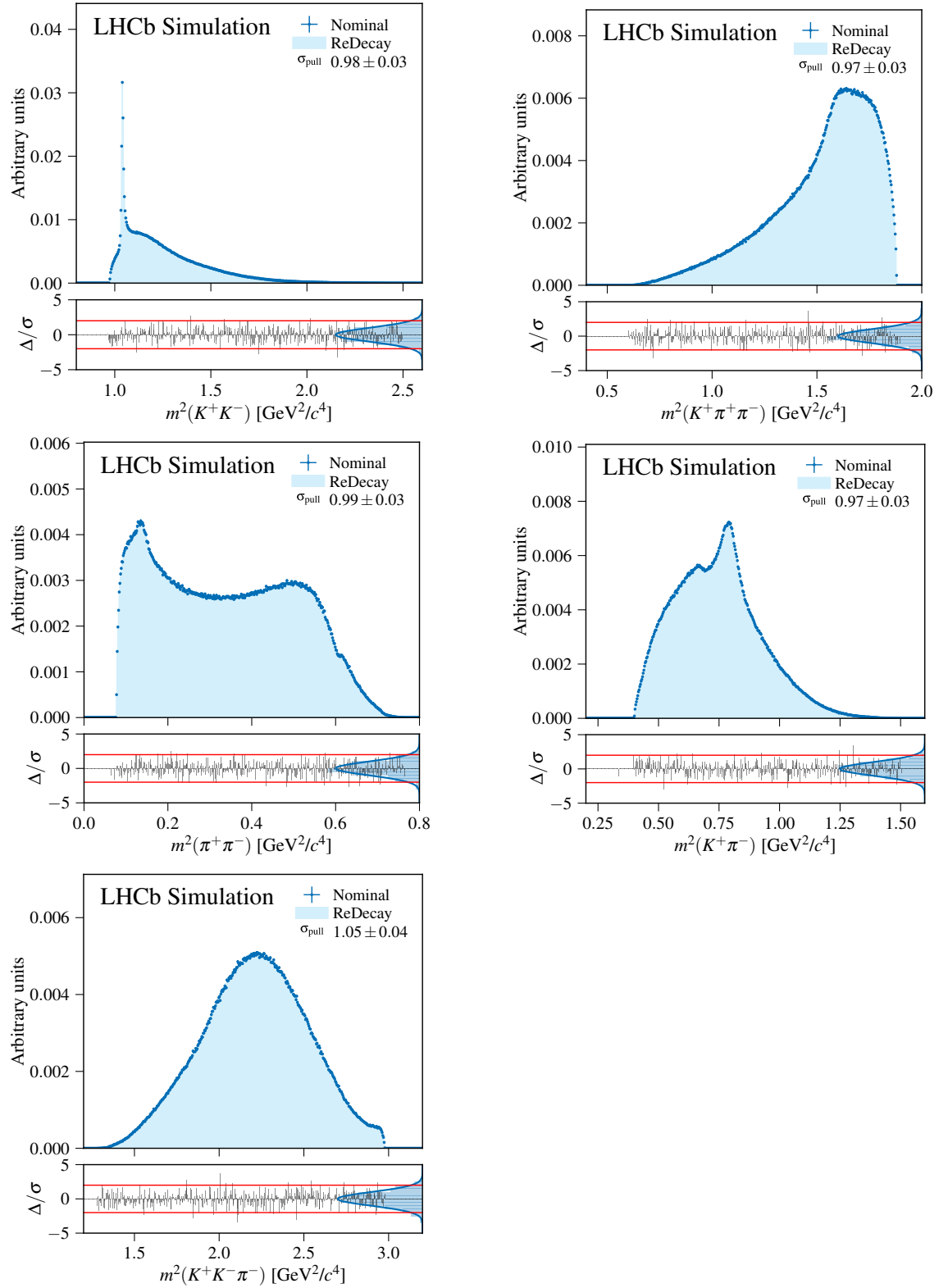


Fig. 4.13: Distributions of the five variables describing the phase-space of $D^0 \rightarrow K^+ K^- \pi^+ \pi^-$ decays. Uncertainties are computed using the naive $\sqrt{N}$ approach. Also shown is a projection of the distribution of the pulls compared to a unit normal distribution.

5. Charm production at 5 and 13 TeV

5.1	Analysis overview	96
5.2	The datasets	98
5.2.1	Simulated samples	100
5.3	The event selection	102
5.3.1	Selection variables	102
5.3.2	Online selection in the trigger	104
5.3.3	Offline selection	105
5.4	Fitting – Extracting the prompt yield	107
5.4.1	The fit procedure	107
5.4.2	Separating combinatorial background – The $m/\delta m$ fit . . .	108
5.4.3	Separating secondary charm – The $\ln \chi^2_{\text{IP}}$ fit	109
5.4.4	Mode-specific details – D^0 ain't D^+	111
5.4.5	Results	112
5.5	Reconstruction and selection efficiencies	122
5.5.1	Efficiencies from simulation	122
5.5.2	Particle identification efficiencies	127
5.6	Systematic uncertainties	135
5.6.1	Uncertainties from external input	135
5.6.2	Tracking correction uncertainties	135
5.6.3	Choice of fit model	136
5.6.4	Monte Carlo modelling	137
5.6.5	Particle identification	141
5.6.6	Summary	146
5.7	Results	148
5.7.1	Ratios between different $\sqrt{s}$	148
5.7.2	Integrated cross-sections	155
5.8	Summary and impact	159

A motivation for the measurement of charm meson production cross-sections, especially at different centre-of-mass energies, has already been given in Section 2.2. This chapter presents the measurements of these cross-sections at two different centre-of-mass energies: $\sqrt{s} = 5$ TeV [2] and $\sqrt{s} = 13$ TeV [1]. While both measurements follow a similar approach, they constitute separate measurements performed about one year apart. In the interest of readability, this chapter gives a unified description making a distinction between the two measurements only where necessary. To this end, all shown $\sqrt{s} = 13$ TeV figures visually deviate from the published ones as they have been recreated using the styling of those at $\sqrt{s} = 5$ TeV to achieve visual uniformity.

Tab. 5.1: Branching ratios for the different modes. Quoted uncertainties are propagated as systematic uncertainties in this measurement.

Mode	$\mathcal{B}$ [%]	Ref.
$D^0 \rightarrow K^- \pi^+$	3.88 ± 0.05	[77]
$D^+ \rightarrow K^- \pi^+ \pi^+$	9.13 ± 0.19	[77]
$D_s^+ \rightarrow \phi \pi^+$	2.24 ± 0.13	[103]
$D^{*+} \rightarrow D^0 \pi^+$	67.7 ± 0.5	[77]

In the following, a general overview of the analysis is presented, highlighting the individual components entering a cross-section measurement.

5.1 Analysis overview

Charm mesons produced at the pp collision point, either directly or as decay products of excited charm and charmonium resonances, are referred to as promptly produced. The goal of this analysis is the measurement the double-differential cross-sections of various promptly produced charm meson in bins of transverse momentum and rapidity, (p_T, y) , using data recorded by the LHCb experiment. The analysis reconstructs and selects the various charm decays, extracts the size of the prompt signal component N , and then computes the efficiency ϵ of each step, from the geometrical limitations imposed by the LHCb detector acceptance and the event selection applied in the trigger to the selection requirements. Combined with a measurement of the collected integrated luminosity $\mathcal{L}$ and the branching fraction $\mathcal{B}$ of the decay mode f in question, the cross-section for a single charm meson H_c in a given (p_T, y) bin i can be computed as:

$$\sigma_i(H_c) = \frac{N_i(H_c \rightarrow f)}{\epsilon_i(H_c \rightarrow f) \mathcal{B}(H_c \rightarrow f) \mathcal{L}}. \quad (5.1)$$

Double differential cross sections can then be determined using the width of the bin in p_T , Δp_T , and the width of the bin in rapidity, Δy , in which the yield and efficiency measurements are made:

$$\left. \frac{d^2 \sigma(H_c)}{dp_T dy} \right|_i = \frac{\sigma_i(H_c)}{\Delta p_T \Delta y}. \quad (5.2)$$

Both the 5 and 13 TeV analyses measure this double-differential cross-section for D^0 , D^+ , D_s^+ , and D_s^+ mesons using the final states $D^0 \rightarrow K^- \pi^+$, $D^+ \rightarrow K^- \pi^+ \pi^+$, $D_s^+ \rightarrow (\phi \rightarrow K^+ K^-) \pi^+$ and $D^{*+} \rightarrow (D^0 \rightarrow K^- \pi^+) \pi^+$ respectively. The respective branching fractions are given in Table 5.1 and taken from Ref. [77] with the exception of the D_s^+ measurement, which instead of determining the resonant contribution to the amplitude, the fraction of $D_s^+ \rightarrow K^- K^+ \pi^+$ decays with a $K^- K^+$ invariant

mass in the range $1000 < m_{K^-K^+} < 1040 \text{ MeV}/c^2$ is used and its branching fraction is taken from Ref. [103].

The width of the rapidity bins used in the measurement is 0.5 units in the range $2.0 < y < 4.5$. The p_T range is $0 < p_T < 10 \text{ GeV}/c$ for the 5 TeV and $0 < p_T < 15 \text{ GeV}/c$ for the 13 TeV measurement, separated into $1 \text{ GeV}/c$ -wide bins with the exception of the $1 < p_T < 4 \text{ GeV}/c$ range for 13 TeV where a width of $0.5 \text{ GeV}/c$ is used. In addition to the double-differential cross-sections, which form the main result, several derived results are obtained, most importantly estimates for the integrated cross-sections and the ratios $R_{13/5}$ and $R_{13/7}$ between different centre-of-mass energies.

Combination and propagation of uncertainties

When computing the uncertainty on a function $f(a)$, the assumptions of linear error propagation may not hold, namely that a linear approximation is sufficient to describe the function over the range covered by the uncertainty. In this analysis, the uncertainties on the quantities in the denominator of Equation 5.1 can be large (exceeding a relative uncertainty of 10%) in some (p_T, y) bins, thereby making a linear approximation of σ_i insufficient.

Instead of using this analytical approach, these measurements employ a Monte Carlo error propagation technique. With this, ten thousand random numbers are sampled from a suitable probability density function (PDF) for each quantity appearing in the analysis, and operations are performed for each of these numbers row-by-row. Statistics such as the mean and standard deviation can then be computed on the resulting set. For inputs common to multiple bins or mesons, *e.g.* the luminosity, only one set of random numbers is sampled, persisted and reused, rendering the correct treatment of arising correlations due to common inputs trivial.

Unless stated otherwise, each quantity to which an uncertainty is assigned is modelled as a normal distribution, $N(\mu, \sigma)$, whose mean μ is the value found by the appropriate procedure, and whose standard deviation σ is taken as the uncertainty on that value. In the case of the prompt signal yield, for example, the mean of the normal distribution is the fitted value, and the width is the uncertainty returned by the fitter. An illustration for the obtained posterior distributions is given in Figure 5.1. When converting these posterior distributions back into values for the cross-section, the central value is given by the most probable value estimated by building a kernel density and the lower and upper uncertainties are given by the 16% and 84% percentiles, respectively.

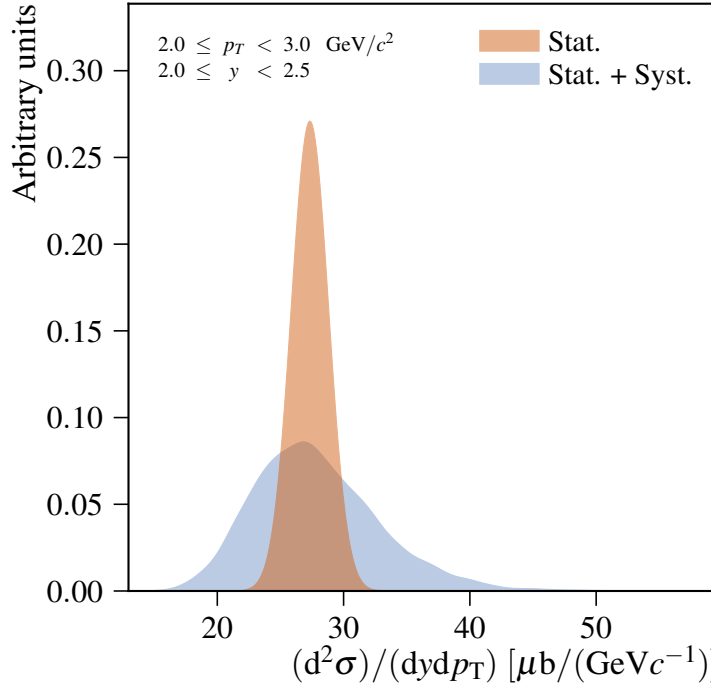


Fig. 5.1: Distribution of the posterior probability for one (p_T, y) bin in the differential $D_s^+ \rightarrow (\phi \rightarrow K^+ K^-) \pi^+$ measurement, illustrating the asymmetric distribution due to large uncertainties in the denominator.

Chapter structure

The structure of the remaining chapter is given by the ingredients of the right-hand side of Equation (5.1). The used datasets and thus the integrated luminosity $\mathcal{L}$ are discussed in Section 5.2, including details on the running conditions of the LHC, the trigger strategy, and the simulated samples that are used. The selection applied to the candidates follows in Section 5.3, including a brief description of how the charm hadron candidates are reconstructed. The fitting framework that determines the yields of the various signal and background components after the full selection are shown in Section 5.4, and in Section 5.5 the efficiencies of each reconstruction and selection stage are described and given. The sources of systematic uncertainty in the analysis are listed and evaluated in Section 5.6. Finally, the measured cross-sections are discussed in Section 5.7, followed by a description of the impact in Section 5.8.

5.2 The datasets

The datasets used for the measurements were recorded during 2015 using the same trigger strategy with the detailed implementation and selection criteria being the topic of the following section. Nonetheless, the used L0 trigger is relevant for the description of the datasets. The studied decays are solely comprised of hadronic particles and thus the only viable trigger options have to either rely on the calorimeter information or select events randomly. The former option requires difficult data-driven

approaches when evaluating the absolute efficiencies due to inaccurate simulation of the calorimeter based hardware trigger. Therefore, the latter option was chosen by selecting events randomly on bunch crossings at a constant rate of about 300 kHz. This kind of trigger simply selects a subsample of all available data in an unbiased way and is thus referred to as a NoBias trigger. The known LHC fill structure determines the total number of beam-beam crossings (the number of colliding bunches) $N_{\text{Coll. bunches}}$, which is then combined with the NoBias rate R_{NoBias} to give the effective efficiency of the L0 trigger ϵ_{L0} for a given fill;

$$\epsilon_{\text{L0}} = \frac{R_{\text{NoBias}}}{N_{\text{Coll. bunches}} f_{\text{LHC}}}, \quad (5.3)$$

where f_{LHC} is the bunch revolution frequency, taken to be 11.246 kHz. The overall efficiency of the L0 trigger is then determined as the luminosity-weighted average of the efficiencies for the individual fills :

$$\bar{\epsilon}_{\text{L0}} = \frac{\sum_{\text{fills}} \epsilon_{\text{L0},i} \mathcal{L}_i}{\sum_{\text{fills}} \mathcal{L}_i}. \quad (5.4)$$

The effective luminosity used is then given by:

$$\mathcal{L}_{\text{eff.}} = \bar{\epsilon}_{\text{L0}} \cdot \sum_{\text{fills}} \mathcal{L}_i. \quad (5.5)$$

The luminosity for both datasets described in the following has been measured using a beam-gas imaging technique [104]. Additionally, the data were recorded with the magnetic field of the LHCb magnet pointing downwards (and thus bending particles along the LHC curvature).

The 13 TeV dataset

The 13 TeV dataset was recorded over a two week period in the summer of 2015 that is commonly referred to as the early measurements period. During this period, the LHC was slowly being ramped up after two years of upgrading and maintenance and thus running at a comparably low number of colliding bunches which were spaced 50 ns apart. Unfortunately, the LHC did not perform as well as hoped with only five fills being usable for analysis resulting in about a factor of ten less recorded data than anticipated while planning this measurement and tuning the event selection. These five fills are summarised in Table 5.2 and sum to a luminosity of:

$$\mathcal{L} = (4.98 \pm 0.19) \text{ pb}^{-1}. \quad (5.6)$$

Tab. 5.2: List of fills used in the $\sqrt{s} = 13$ TeV analysis, the integrated luminosity (“lumi.”) collected in the fill, the number of colliding bunches at LHCb $N_{\text{Coll. bunches}}$ during the fill, the L0 NoBias rate during the fill, and the corresponding effective L0 efficiency (“eff.”).

Fill	Lumi. collected	$N_{\text{Coll. bunches}}$	L0 NoBias rate	L0 NoBias eff.
3976	$(628 \pm 24) \text{ nb}^{-1}$	108	250 kHz	20.58 %
3981	$(1101 \pm 43) \text{ nb}^{-1}$	246	300 kHz	10.84 %
3988	$(931 \pm 36) \text{ nb}^{-1}$	246	300 kHz	10.84 %
3992	$(1310 \pm 51) \text{ nb}^{-1}$	397	350 kHz	7.84 %
3996	$(999 \pm 39) \text{ nb}^{-1}$	397	350 kHz	7.84 %

Using the information on the fill structure $\bar{\epsilon}_{\text{L0}} = 10.67\%$ is obtained, resulting in an effective luminosity of:

$$\mathcal{L}_{\text{eff.}} = (531 \pm 21) \text{ nb}^{-1} . \quad (5.7)$$

For reference, the cross-section measurements performed at $\sqrt{s} = 7$ TeV [27] only used around 15 nb^{-1} , *i.e.* $\sim 3\%$ of the 13 TeV luminosity, allowing for a far more precise measurement.

The 5 TeV dataset

The 5 TeV dataset was recorded over a weekend in November 2015 in preparation for a proton-lead run. Compared to the ‘early measurements’ period, the nominal Run 2 conditions were used with a bunch spacing of 25 ns and a much larger number of bunches circulating in the machine. A total of three fills were performed, which are summarised in Table 5.3. In total, these three fills have a combined luminosity of:

$$\mathcal{L} = (8.60 \pm 0.33) \text{ pb}^{-1} . \quad (5.8)$$

However, with a NoBias efficiency of just $\bar{\epsilon}_{\text{L0}} = 1.86\%$, this results in an effective luminosity of:

$$\mathcal{L}_{\text{eff.}} = (161 \pm 6) \text{ nb}^{-1} , \quad (5.9)$$

which is about three times less than what is available for the 13 TeV measurement but still ten times more than was used to determine the cross-sections at 7 TeV.

5.2.1 Simulated samples

Simulated samples are necessary to obtain most of the efficiencies in the measurement. Therefore, samples for $D^+ \rightarrow K^- \pi^+ \pi^+$, $D_s^+ \rightarrow K^- K^+ \pi^+$ and $D^{*+} \rightarrow (D^0 \rightarrow K^- \pi^+) \pi^+$ decays have been generated. A sample for the decay $D_s^+ \rightarrow (\phi \rightarrow$

Tab. 5.3: List of fills used in the $\sqrt{s} = 5$ TeV analysis, the integrated luminosity (“lumi.”) collected in the fill, the number of colliding bunches at LHCb $N_{\text{Coll. bunches}}$ during the fill, the L0 NoBias rate during the fill, and the corresponding effective L0 efficiency (“eff.”).

Fill	Lumi. collected	$N_{\text{Coll. bunches}}$	L0 NoBias rate	L0 NoBias eff.
4640	$(1.80 \pm 0.07) \text{ pb}^{-1}$	1248	300 kHz	2.138 %
4643	$(4.49 \pm 0.17) \text{ pb}^{-1}$	1486	300 kHz	1.795 %
4647	$(2.31 \pm 0.09) \text{ pb}^{-1}$	1486	300 kHz	1.795 %

$K^+K^-\pi^+$ is then obtained by selecting the window in the di-kaon mass analogous to the real-time data selection. Furthermore, no dedicated sample is needed for $D^0 \rightarrow K^-\pi^+$: the $D^{*+} \rightarrow D^0\pi^+$ sample already contains the decay, and potential differences in the D^0 kinematics are irrelevant due to the double-differential nature of the measurement. Four million events were generated per decay and per centre-of-mass energy. To populate the high p_T region in the 13 TeV measurement up to $15 \text{ GeV}/c$, an additional two million MC events were generated per decay where the generated meson is required to have $p_T > 10 \text{ GeV}/c$. In all samples, all decay products were required to be in the acceptance of LHCb, which is defined as $10 < \theta < 400 \text{ mrad}$. As described in Section 4.2.1, this requirement is imposed to save computing resources by reducing the number of events that are needlessly passed through the expensive detector simulation as they cannot be reconstructed due to one of their decay products being outside of the detector.

In the following, unless otherwise stated, the simulated decays that are used to study the efficiencies are truth-matched. Truth-matching refers to the use of the true particles produced by the generator to identify true candidates in the simulated sample, which can have backgrounds just like real data, albeit at a much lower rate. LHCb performs the truth-matching by searching for track and generated particle combinations, where the overlap between the hits used in the track fit and the hits produced by the true particles is at least 70%. Three possible outcomes of the truth-matching procedure exist for the candidates, illustrated in Figure 5.2:

1. Candidates where the child tracks are matched to truth-particles of the correct type and the matched true particles all stem from the same and correct parent and in turn comprise the full set of true children (not counting radiative photons). Such candidates constitute the vast majority of all reconstructed candidates in the sample and are referred to as truth-matched signals.
2. Candidates where the child tracks are correctly matched to true particles that stem from the same parent. However, this parent then either does not match the correct one or has additional true decay products not associated to the reconstructed decay. Decays such as this are referred to as partial backgrounds,

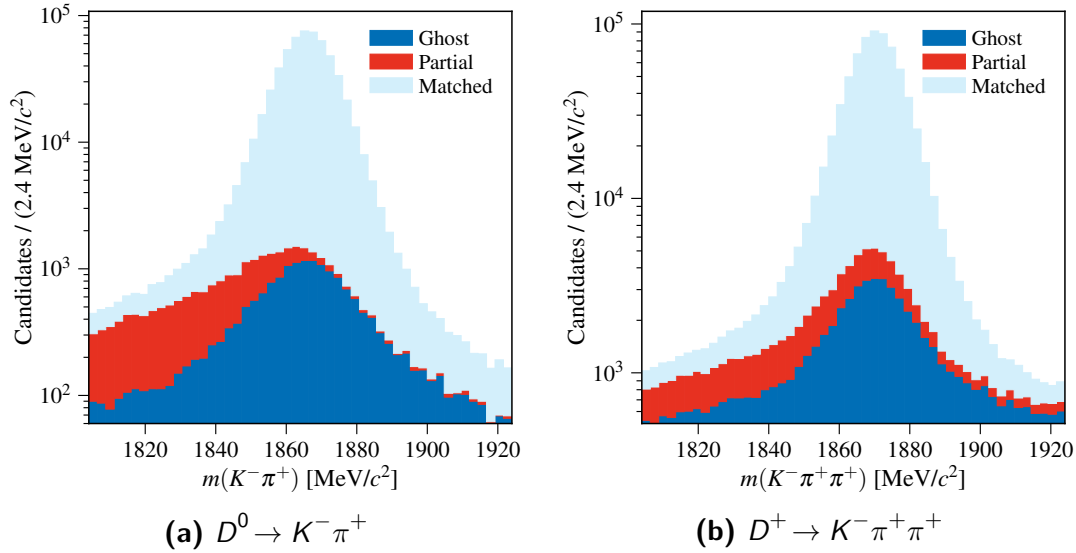


Fig. 5.2: Illustration of the three dominant types of candidates in the simulated samples as described in the text based on truth-matching. The different components in these plots are stacked. To enhance the clarity of these plots, the simulated samples for 5 TeV and 13 TeV are merged.

typically characterised by a shape with negative slope in the mass window of interest due to missing energy.

3. Candidates where no matching true particle for a reconstructed track is found, in which case the candidate is referred to as a ghost.

However, given the truth-matching procedure, it is possible that true signal candidates fail the matching by chance and are incorrectly rejected. This can be seen clearly in Figure 5.2 where the ghost contribution is dominated by true decays with a shape very similar to that of the signal. When using the simulated samples to compute efficiencies based on truth-matched candidates, a correction factor ρ_{Truth} needs to be computed to account for and correct this effect.

5.3 The event selection

The presented measurements are the first charm measurements using the novel TurboStream approach, which allows the direct analysis of candidates from the trigger. Therefore, a large part of the selection criteria that are employed to select the signal as cleanly as possible is implemented in the HLT2 lines of the trigger. Before those criteria are discussed, an overview of the used variables and their definition is given.

5.3.1 Selection variables

In the following, definitions for selection variables commonly used by LHCb and utilised in this analysis are given.

Child and parent χ_{IP}^2

The impact parameter (IP), with respect to the primary vertex (PV), is the distance of closest approach of the track, extrapolated to the PV, as shown in Figure 3.3. Prompt stable particles produced directly from the pp collision will have a low impact parameter, whereas particles produced in long-lived charm decays will not generally have trajectories pointing back to the PV, and will therefore have large impact parameters.

The ‘ χ^2 ’ is computed as the absolute difference of a vertex fit with and without the particle in question. If the particle truly stems from the vertex, the fit χ^2 only changes due to the added degree of freedom and thus the difference, quoted as χ_{IP}^2 , is expected to be around one.

Kaon and pion PID

Particle identification information is obtained as the combined response of the calorimeters, muon stations, and RICH detectors. In this analysis the DLL set of PID variables is used. A definition of these variables is given in Section 3.3.2.

Pairwise child distance of closest approach (DOCA)

For a set of tracks physically originating from a common vertex, the largest distance between any two tracks, the DOCA (for distance of closest approach), should be small.

Pairwise child DOCA χ^2

This variable is used in one trigger line and relates to DOCA as χ_{IP}^2 relates to IP.

Parent vertex displacement

Displacement of the decay vertex from the production vertex, assumed to be the PV, as either the reconstructed decay time $c\tau$ or the vertex displacement χ^2 , where the latter is computed in the same way as the track χ_{IP}^2 . The ‘ χ^2 ’ is computed as $\Delta^T \Sigma^{-1} \Delta$, where Δ is the difference in the spatial positions of the two vertices, and Σ is the sum of the covariance matrices of the two vertex fits.

Parent direction angle (DIRA)

If the decay products are chosen correctly, and the decay vertex position is correctly reconstructed, then the momentum vector of the parent should point in the same direction as the displacement vector between the decay vertex and the assumed production vertex. The angle between the momentum and displacement vectors, called the direction angle (DIRA), should be small for a correctly reconstructed signal decay. (It is the cosine of the direction angle that is usually called “DIRA”, which should be very close to 1 for a correctly reconstructed signal decay.)

These variables, together with the information on the four-momenta of the particles, are available in the trigger reconstruction which this analysis relies on. The following

Tab. 5.4: Requirements made on the track that fires the HLT1 trigger line.

Variable	Requirement
p_T	$> 800 \text{ MeV}$
p	$> 3 \text{ GeV}$
Track χ^2	< 3
VELO hits	> 9
χ_{IP}^2	> 10

section will describe how they are used to select the four different decays in the trigger.

5.3.2 Online selection in the trigger

For this measurement, dedicated triggers have been developed for the HLT1 and HLT2 step, which have been used exclusively for the discussed cross-section measurements and have been inactive during the nominal Run 2 data taking. As previously mentioned, the L0 trigger is replaced by a random triggering on bunch crossings.

The procedure applied in HLT1 has already been discussed in Section 3.3.5. The selection criteria for this trigger are given in Table 5.4 and their goal is to identify events with a track that could potentially stem from one of the studied decays. Requirements on p and p_T are made since higher momentum kaons and pions, which comprise the final state in this analysis, are more likely to stem from the decay of a heavy particle like a charm meson. The requirements on the track χ^2 and the number of hits in the VELO are imposed to suppress ghost tracks. Lastly, the studied charm mesons all have a long-lived component in their decay chain, either themselves directly in the case of the D^0 , D^+ and D_s^+ or the D^0 in $D^{*+} \rightarrow D^0 \pi^+$ decays. Therefore, the tracks are required to have a large χ_{IP}^2 with respect to all primary vertices, implying they are incompatible with having been produced directly in the pp collision.

Once an event has been triggered by the HLT1 line, several HLT2 lines are run to build and select the signal decays. Three distinct lines are used. One of these lines constructs $D^0 \rightarrow K^- \pi^+$ decays and the applied selection criteria are given in Table 5.5; while the two three-body decays $D^+ \rightarrow K^- \pi^+ \pi^+$ and $D_s^+ \rightarrow K^- K^+ \pi^+$ are, due to their similarity, selected by a common HLT2 line that selects all $D_s^+ \rightarrow h^- h^+ h^+$ decays, split into the different invariant mass windows of the D^+ and D_s^+ and final state particle hypothesis. The selection criteria are listed in Table 5.6. As described before, this measurement only uses those $D_s^+ \rightarrow K^- K^+ \pi^+$ decays where $1000 < m_{K^- K^+} < 1040 \text{ MeV}/c^2$. This requirement is not included in the trigger

Tab. 5.5: Selection requirements made in the $D^0 \rightarrow K^- \pi^+$ HLT2 trigger line.

Particle	Variable	Requirement
$\pi^\pm, K^\pm$	p_T	$> 250 \text{ MeV}$
	p	$p > 2 \text{ GeV}$
	Track χ^2	< 3
	χ_{IP}^2	> 16
$\pi^\pm$	$DLL_{K-\pi}$	< 5
$K^\pm$	$DLL_{K-\pi}$	> 5
D^0	$m(K^- \pi^+)$	$1784 \text{ MeV} < m < 1944 \text{ MeV}$
	$K^\pm$ to $\pi^\pm$ DOCA	$< 0.1 \text{ mm}$
	Vertex fit χ^2	< 10
	Direction angle	> 0.99985
	Vertex displacement χ^2	> 49

lines but applied manually later on. Lastly, D^{*+} candidates are built by taking $D^0 \rightarrow K^- \pi^+$ candidates that have been selected and then adding additional tracks of opposite charge with respect to the kaon. The sum of pion and D^0 masses is only $5 \text{ MeV}/c^2$ smaller than the mass of the D^{*+} , and thus the pion in the $D^{*+} \rightarrow D^0 \pi^+$ decay tends to be very soft and is often referred to as the slow pion. At this low momentum the performance of the particle identification system is poor. Therefore, even though the added track is referred to as a pion track, no actual PID requirements are imposed. The ‘correct’ pion is later identified when estimating the signal yield, as only the correct choice will lead to a peak in the $m(D^{*+}) - m(D^0)$ distribution. The additional selection criteria imposed onto the π^+ and D^0 combination are given in Table 5.7.

5.3.3 Offline selection

Additional selection criteria are applied on top of those in the trigger to improve the purity. This is especially necessary for the three-body decays where a large combinatorial background remains after the trigger, illustrated in Figure 5.3 for $D^+ \rightarrow K^- \pi^+ \pi^+$. The final-state particles of the D^0 , D^+ and D_s^+ candidates are required to be within $3 < p < 100 \text{ GeV}/c$ and $2 < \eta < 5$ to achieve a good performance of the PID system. Furthermore, the PID requirement for pions is tightened to $DLL_{K-\pi} < 3$. To further reduce the background in the three-body decays, the vertex fit χ^2 is required to be less than 6 and their DIRA must exceed 0.9999.

Figure 5.3 illustrates the power of these additional selection criteria. While $D^0 \rightarrow K^- \pi^+$ barely changes and displays a very clean sample online and offline, the additional requirements substantially clean up $D^+ \rightarrow K^- \pi^+ \pi^+$. With the full

Tab. 5.6: Selection requirements made in the three-body HLT2 lines. The lines are split into $D^+ \rightarrow K^- \pi^+ \pi^+$ and $D_s^+ \rightarrow K^- K^+ \pi^+$ decays according to the mass window. Criteria of the form $x > x_1, x_2, x_3$ require that all particles satisfy $x > x_1$, at least two satisfy $x > x_2$, and at least one satisfies $x > x_3$.

Particle	Variable	Requirement
$\pi^\pm, K^\pm$	p_T	$> 200, 400, 1000 \text{ MeV}$
	p	$p > 2 \text{ GeV}$
	Track χ^2	< 3
	χ_{IP}^2	$> 4, 10, 50$
$\pi^\pm$	$\text{DLL}_{K-\pi}$	< 5
$K^\pm$	$\text{DLL}_{K-\pi}$	> 5
D^+	$m(h^- h^+ h^+)$	$1789 \text{ MeV} < m < 1949 \text{ MeV}$
D_s^+	$m(h^- h^+ h^+)$	$1889 \text{ MeV} < m < 2049 \text{ MeV}$
D^+/D_s^+	Vertex fit χ^2	< 25
	Direction angle	> 0.9994
	Vertex displacement	$VD\chi^2 > 16 \text{ AND } \tau > 0.150 \text{ ps}$

Tab. 5.7: Selection requirements made on HLT2 D^{*+} candidates.

Particle	Variable	Requirement
$\pi^\pm$	p_T	$> 100 \text{ MeV}$
	Track χ^2	< 3
D^{*+}	$m(D^{*+}) - m(D^0)$	$130 \text{ MeV} < m < 160 \text{ MeV}$
	Vertex fit χ^2	< 25

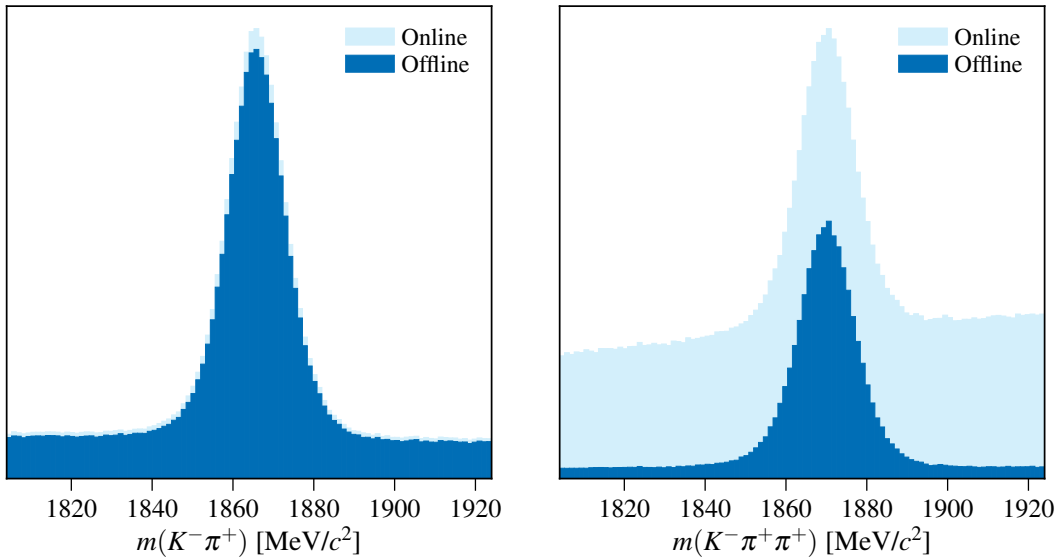


Fig. 5.3: Illustration of the online and offline event selection in a two-body decay ($D^0 \rightarrow K^- \pi^+$, left) and three-body decay ($D^+ \rightarrow K^- \pi^+ \pi^+$, right) in 5 TeV data.

selection defined, the next section discusses the estimation of the prompt signal yield.

5.4 Fitting – Extracting the prompt yield

The number of signal events in each (p_T, y) bin needs to be estimated by distinguishing it from two distinct backgrounds: combinatorial background and secondary charm decays, which originate from the decay of a heavy b -hadron. These are two very different backgrounds that cannot be resolved by a simple one-dimensional fit to one quantity. The invariant mass of the decay products is ideally suited to distinguish signal and combinatorial background while having no sensitivity to separate prompt and secondary signal decays. It has already been mentioned that the impact parameter or derived quantities like χ_{IP}^2 are well equipped to distinguish prompt and secondary production; however, the combinatorial background distribution is hard to discern. Therefore, a multi-step fit procedure is employed that is described in the following.

5.4.1 The fit procedure

The number of fully selected prompt signal candidates, for a given charm hadron in a given (p_T, y) bin, is determined by a one-dimensional binned extended maximum likelihood fit to the charm hadron $\ln \chi_{\text{IP}}^2$ distribution in a signal region in the meson invariant mass of the fully selected dataset. To control the combinatorial background in this fit, its shape is constrained using mass sidebands and its yield is constrained from a prior fit to the invariant mass distribution. In case of the D^{*+} , instead of its invariant mass, the difference between the D^{*+} and D^0 mass $\delta m = m(K^- \pi^+ \pi^+) - m(K^- \pi^+)$ is used.

The ‘signal region’ in each long-lived charm hadron mass distribution is defined as a $40 \text{ MeV}/c^2$ -wide window centred on the nominal rest mass of the respective meson, taken to be $m_{D^0} = 1864.84 \text{ MeV}/c^2$, $m_{D^+} = 1869.61 \text{ MeV}/c^2$, and $m_{D_s^+} = 1968.30 \text{ MeV}/c^2$ [77]. The ‘lower sideband’ is a $20 \text{ MeV}/c^2$ -wide window centred $50 \text{ MeV}/c^2$ below the centre of the signal window, and the ‘upper sideband’ is a $20 \text{ MeV}/c^2$ -wide window centred $50 \text{ MeV}/c^2$ above the centre of the signal window. For the D^{*+} measurements, the ‘signal region’ is defined as a window of $6 \text{ MeV}/c^2$ width, centred on the nominal δm value of $145.43 \text{ MeV}/c^2$ [77]. The ‘upper sideband’, or just ‘sideband’, as a ‘lower sideband’ region is not defined, is defined as the region from $4.5 \text{ MeV}/c^2$ to $9 \text{ MeV}/c^2$ above the centre of the signal region. As an illustration, the sidebands and signal region are highlighted in Figure 5.4 for each of the decays.

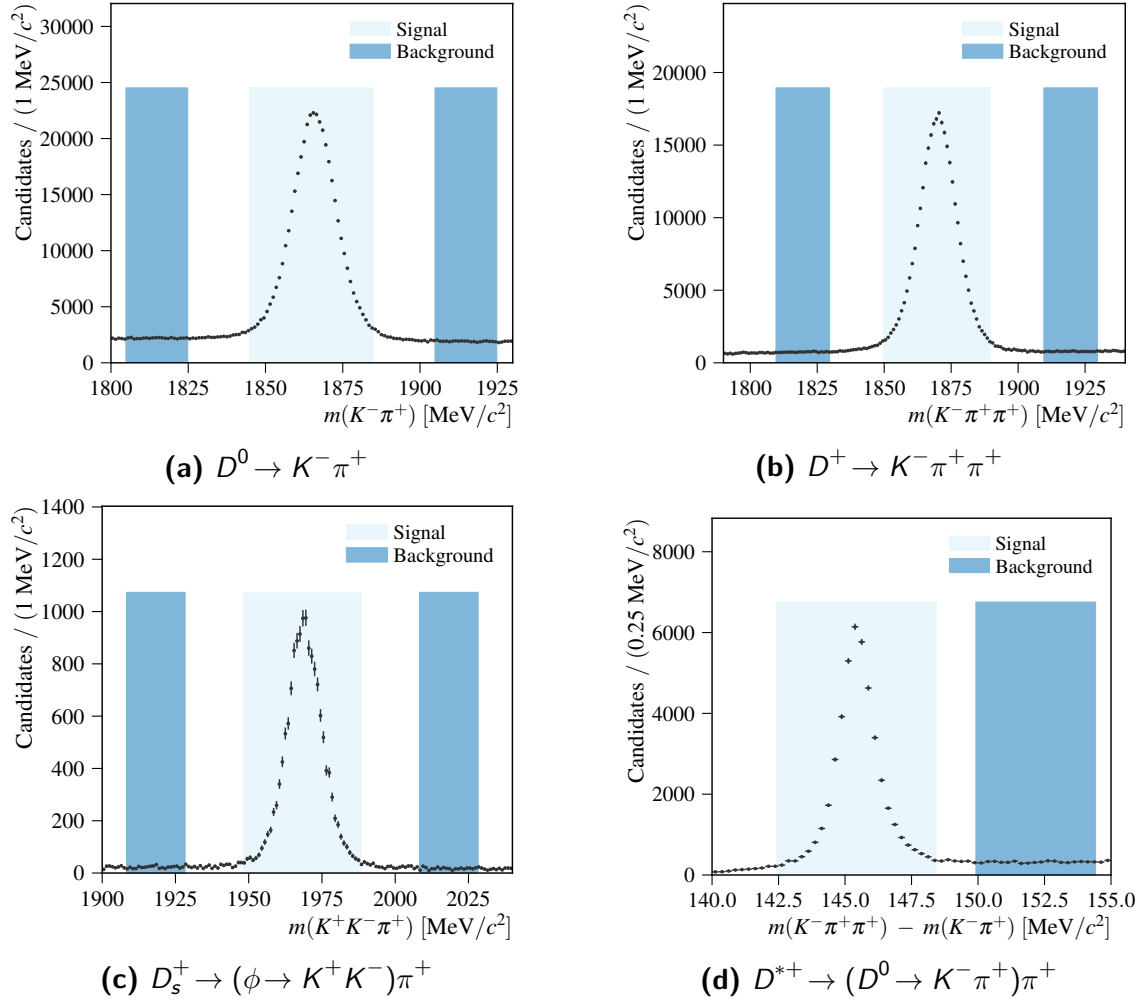


Fig. 5.4: Illustration of the signal and background regions for the different decays.

For the fit, different probability density functions (PDF) are identified for the different species and the total PDF is obtained as their sum, weighted by the respective yields:

$$f(m \text{ or } \delta m) = \sum_{\text{species}} N_{\text{species}} \cdot f_{\text{species}}(m \text{ or } \delta m). \quad (5.10)$$

5.4.2 Separating combinatorial background – The $m/\delta m$ fit

The main goal of this first fit is the separation of the combinatorial background as promptly and secondarily produced mesons have the same distribution and cannot be distinguished. Therefore, they are treated as one species in the fit. Additionally, the two different backgrounds in the δm fit, combinatorial background and background where a true D^0 is paired with a random pion, are assumed to be indistinguishable and no effort is made to separate them.

For every D^0 , D^+ and D_s^+ fit, the background PDF is taken to be a first-order polynomial, and the signal PDF is decay-dependent and given in Section 5.4.4. In

the D^{*+} fit, the background in the δm fits is an empirical function of the form:

$$R(\delta m; \delta m_0, A, B) \propto e^{B \cdot \delta m} \cdot (\delta m - \delta m_0)^A, \quad (5.11)$$

which has been found to describe the background well, but is not motivated by any physics reason.

One such PDF, $f(m)$ or $f(\delta m)$, is constructed per (p_T, y) bin, and the likelihood is formed as the product of these such that each bin is fitted simultaneously when the negative log-likelihood is minimised.

It can be seen in Figure 5.4 that the signal window does not cover all of the visible peak in the m or δm distributions and thus the requirement is not 100% efficient for signal events. Therefore, its efficiency $\varepsilon_{\text{signal}}$ is evaluated by integrating the fitted m or δm PDF over the signal region in every (p_T, y) bin. The uncertainty on the parameters of the PDF obtained from the fit is taken into account by performing pseudo-experiments. In each of these experiments, new values for the parameters of the PDF are randomly drawn from a multinomial distribution that uses the central values obtained from the fits as the means and the full covariance matrix to describe the correlations between the parameters. This results in one value for the integral over the signal region per pseudo-experiment. The central value and the uncertainty of $\varepsilon_{\text{signal}}$ are then obtained by taking the mean and standard deviation over all samples. As the shape of the signal PDF is allowed to vary between different (p_T, y) bins, values are obtained for each bin ranging from 97% to 99% for the m and 91% to 99% for the δm fits with relative uncertainties ranging from 0.1% to 0.9%. These uncertainties are later propagated as systematic uncertainties on the measured cross-sections.

5.4.3 Separating secondary charm – The $\ln \chi_{\text{IP}}^2$ fit

In the second fit, performed in the signal region as defined above, all three species can be distinguished and hence are described as individual contributions to the total PDF. While the signal is expected to peak at zero in $\ln \chi_{\text{IP}}^2$ and the secondary charm generally populates the large $\ln \chi_{\text{IP}}^2$ region, it is *a priori* not clear what shape is appropriate to describe the combinatorial background. Therefore, a data-driven approach is used: the data in the mass sidebands are used to construct templates for each (p_T, y) bin. This non-parametric distribution is then fitted with only the yield being left free to float in the $\ln \chi_{\text{IP}}^2$ fit. This approach necessitates that the $\ln \chi_{\text{IP}}^2$ distribution in the sidebands is representative for that in the signal region. Figure 5.5 shows two exemplary distributions, comparing the higher and lower sidebands for two (p_T, y) bins for $D^0 \rightarrow K^- \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$. The two sidebands are in good agreement, thus supporting the assumption that the region between the

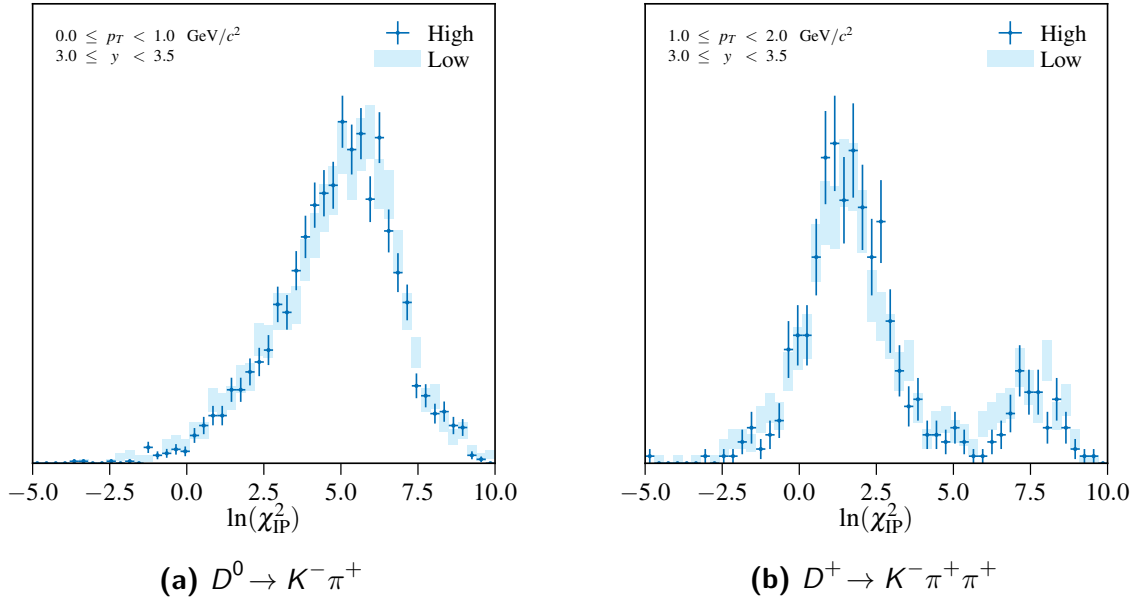


Fig. 5.5: Comparison of the distribution of $\ln \chi_{\text{IP}}^2$ in the high and low mass sidebands for two exemplary (p_T, y) bins of $D^0 \rightarrow K^- \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$. This data is used to construct a template for the combinatorial background shape in the $\ln \chi_{\text{IP}}^2$ fit.

sidebands also shows a similar behaviour. Furthermore, Figure 5.5b motivates the non-parametric data-driven approach as it reveals a complicated two peak structure, which would be challenging to model with analytic shapes. To construct the templates for each bin, the events in the sidebands are used to build a smooth distribution by creating kernel density estimates (KDE) [105]. The constraint on the background yield for each bin is applied by multiplying the likelihood by a normal distribution whose mean is the number of background candidates N measured by the mass fit and whose width is the uncertainty on N returned by the fitter with an additional term accounting for fluctuations of N itself.

The prompt signal component in the $\ln \chi_{\text{IP}}^2$ fit is modelled by a modified normal distribution H . This modified distribution is allowed to have an asymmetric width around the mean and its tails are described by exponential functions:

$$H(x; \mu, \sigma, \epsilon, \rho_L, \rho_R) \propto \begin{cases} \exp\left(\frac{\rho_L^2}{2} + \rho_L \frac{x-\mu}{(1-\epsilon)\sigma}\right) & x < \mu - (\rho_L \sigma (1-\epsilon)) \\ \exp\left(-\left(\frac{x-\mu}{\sqrt{2}\sigma(1-\epsilon)}\right)^2\right) & \mu - (\rho_L \sigma (1-\epsilon)) \leq x < \mu \\ \exp\left(-\left(\frac{x-\mu}{\sqrt{2}\sigma(1+\epsilon)}\right)^2\right) & \mu \leq x < \mu + (\rho_R \sigma (1+\epsilon)) \\ \exp\left(\frac{\rho_R^2}{2} - \rho_R \frac{x-\mu}{(1+\epsilon)\sigma}\right) & x \geq \mu + (\rho_R \sigma (1+\epsilon)) \end{cases} \quad (5.12)$$

where x represents $\ln \chi_{\text{IP}}^2$. This function differs from a standard normal distribution by the asymmetry parameter ϵ , which skews the peak of the distribution, and $\rho_{L,R}$, which describes the position where the exponential tail begins in units of the width on each side of the mean. The secondary distribution is described by a normal

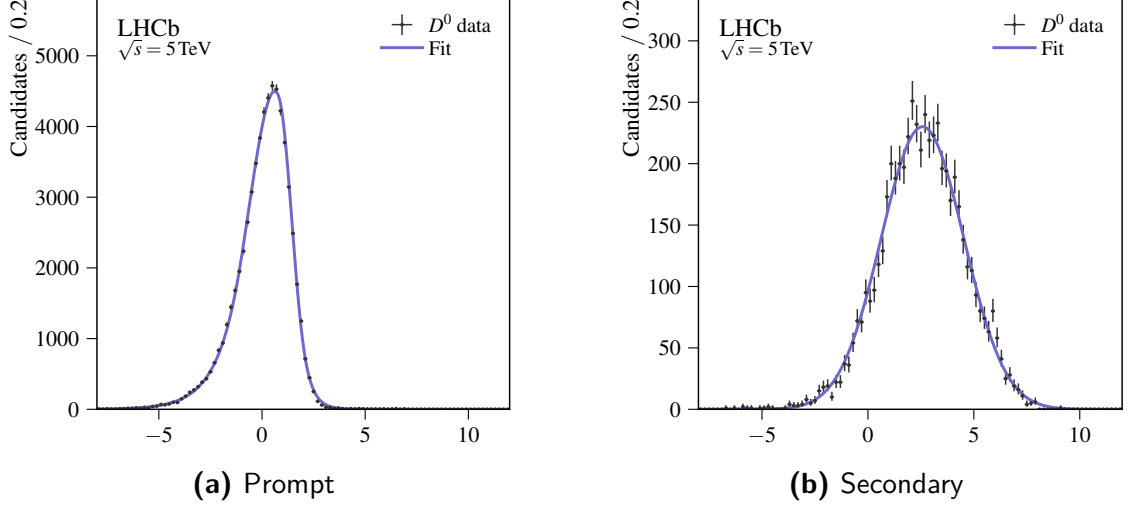


Fig. 5.6: Prompt and secondary signal from a simulated sample for D^0 at 5 TeV. The plotted curve is the sum of the individual prefits performed in each (p_T, y) bin.

distribution. In the interest of a stable fit it is important to note that the signal shape reduces to the secondary shape for $\varepsilon = 0$ and large values for $|\rho_{L,R}|$. This introduces ambiguities in the fit as prompt and secondary contributions might swap and thus some of the parameters need to be kept constant to achieve a converging fit. Therefore, these parameters are determined in ‘prefits’ to simulated data that have been split into prompt and secondary events. An example of these prefits is given in Figure 5.6.

5.4.4 Mode-specific details – D^0 ain’t D^+

The four decay modes studied in this measurement differ in their m , δm and $\ln \chi^2_{\text{IP}}$ distributions and thus require slightly different settings for the shapes and especially the parameters, which can either be split or shared in different (p_T, y) bins as well as floating or fixed to their prefit value. The settings have been determined by trial-and-error, testing different combinations until the fit for the specific decay converges and describes the data well. Fortunately, the same settings can be used for fitting the 5 TeV and 13 TeV data.

Details for $D^0 \rightarrow K^- \pi^+$

In the mass fit, the signal shape is the sum of a normal and a Crystal Ball [106] distribution, with a common mean but different widths. The mean and width of the signal mass PDF are split across (p_T, y) bins, as is the slope of the linear background PDF. For the $\ln \chi^2_{\text{IP}}$ fit, the mean and width of both the prompt and secondary signals are split between different (p_T, y) bins and allowed to float in the fit. All remaining parameters of the prompt $\ln \chi^2_{\text{IP}}$ shape are fixed to the values obtained from the prefits.

Details for $D^+ \rightarrow K^- \pi^+ \pi^+$

The setup is near-identical to that used for the $D^0 \rightarrow K^- \pi^+$ fit, with the only change being the parameters split between different (p_T, y) bins in the $\ln \chi_{\text{IP}}^2$: Only the means of the prompt and secondary signal distributions are split.

Details for $D_s^+ \rightarrow (\phi \rightarrow K^+ K^-) \pi^+$

Compared to the previous two modes, it is sufficient to use the sum of two normal distributions, sharing a common mean but allowed to have different widths, to describe the signal contribution in the mass fit. In this fit, all parameters are split between different (p_T, y) bins. The mean and width of the signal mass PDF and the slope of the background mass PDF, along with the mean of the $\ln \chi_{\text{IP}}^2$ prompt signal PDF, are split across (p_T, y) bins. During the final $\ln \chi_{\text{IP}}^2$ fit, only the mean of the prompt signal $\ln \chi_{\text{IP}}^2$ PDF and all parameters of the secondary PDF are floated, everything else is fixed to the values obtained in the prefits.

Details for $D^{*+} \rightarrow (D^0 \rightarrow K^- \pi^+) \pi^+$

For the δm fit, the signal PDF is the sum of three normal distributions, sharing a common mean but allowed to have varying widths. To enhance the stability of the fit and avoid ambiguous solutions, the widths are configured in cascading order. Only the width of the widest component of the signal model and the parameter A for the background are split across (p_T, y) bins. As the D^{*+} is an instantaneously decaying particle, its χ_{IP}^2 cannot be used to distinguish prompt and secondary signals. Therefore, the $D^0 \ln \chi_{\text{IP}}^2$ distribution is fitted, where the mean of the prompt signal $\ln \chi_{\text{IP}}^2$ PDF are split across different (p_T, y) bins. The same parameters as in the $D_s^+ \rightarrow \phi \pi^+$ fit are floated.

5.4.5 Results

The starting parameters of the fits have been set to sensible values close to the expected values in the mass fits and prefits. Starting values for the yields have been estimated from the number of events in the mass sidebands, extrapolated to the signal region and a 5% contamination from secondary decays is assumed. The fits are required to converge with a good quality covariance matrix and inspected visually to ensure a good description of the data. A selection of fits is shown in Figures 5.7 to 5.10 for the 5 TeV data and in Figures 5.11 to 5.14 for the 13 TeV data.

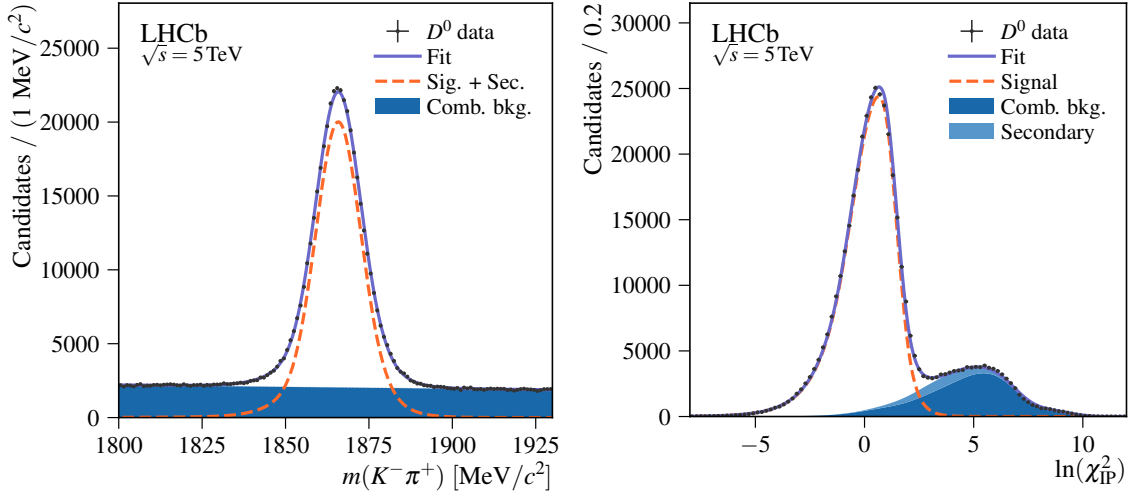
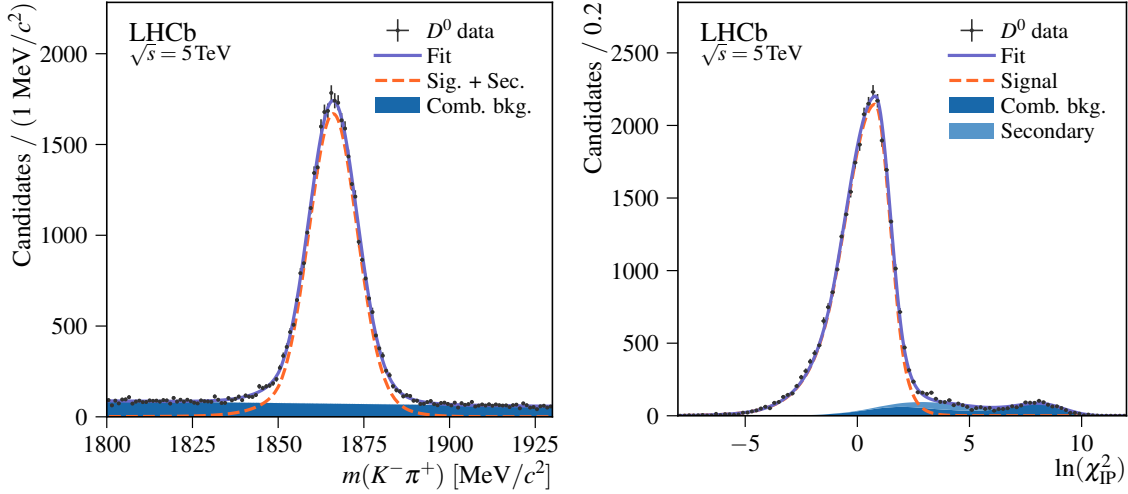
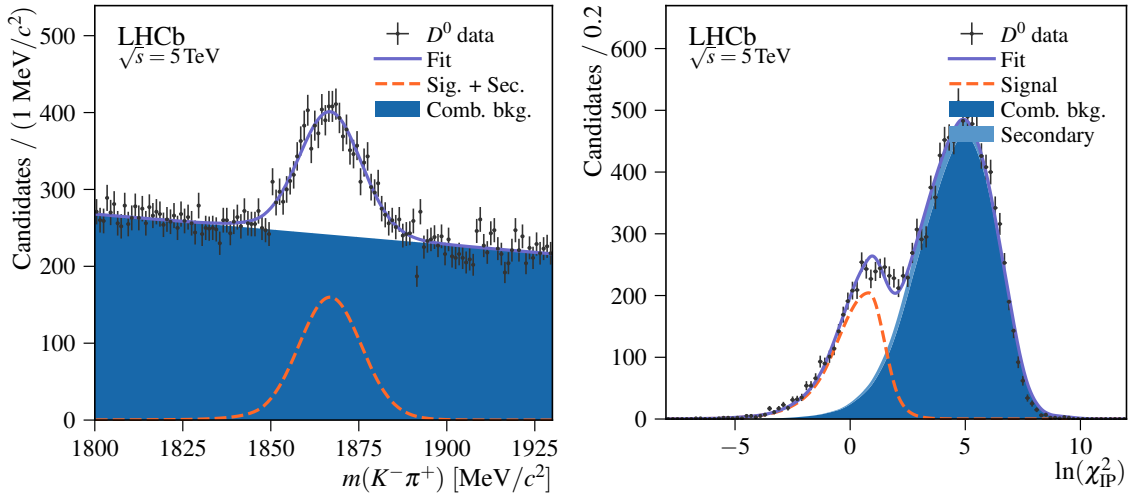
The individual bins have been chosen to illustrate the versatility of the fit, giving the results for a (p_T, y) bin with a large signal as well as one with a large background contribution. The integrated fits shown are, in fact, not obtained by a single fit to

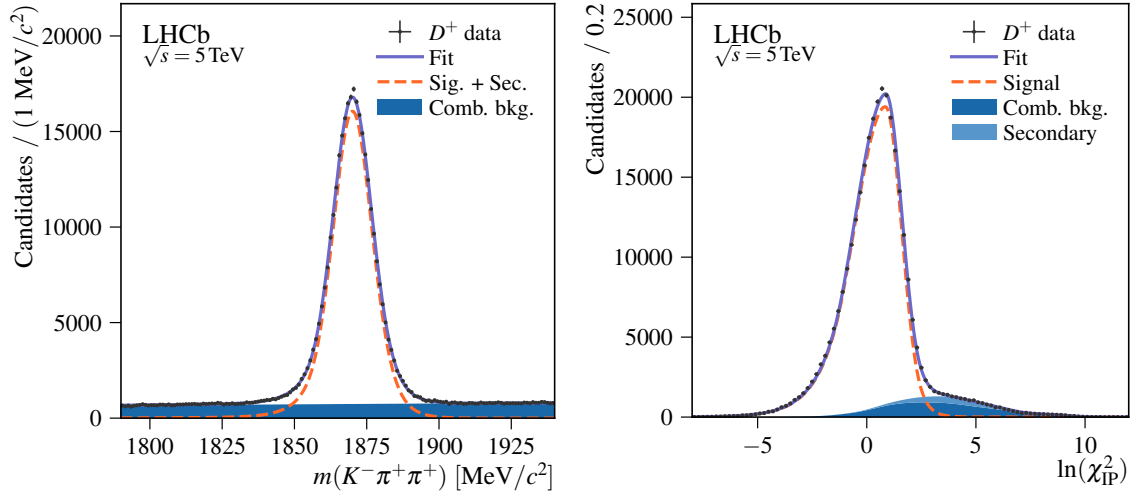
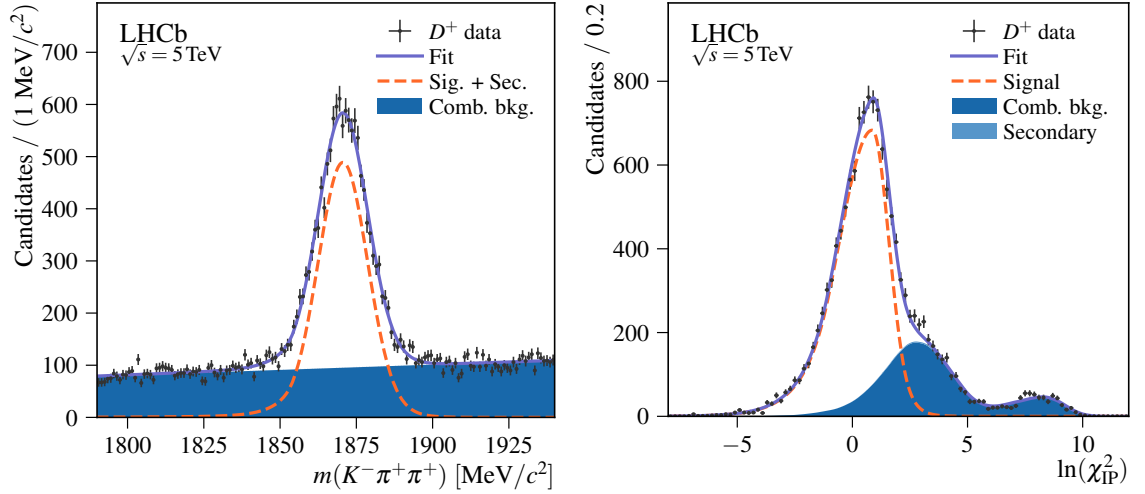
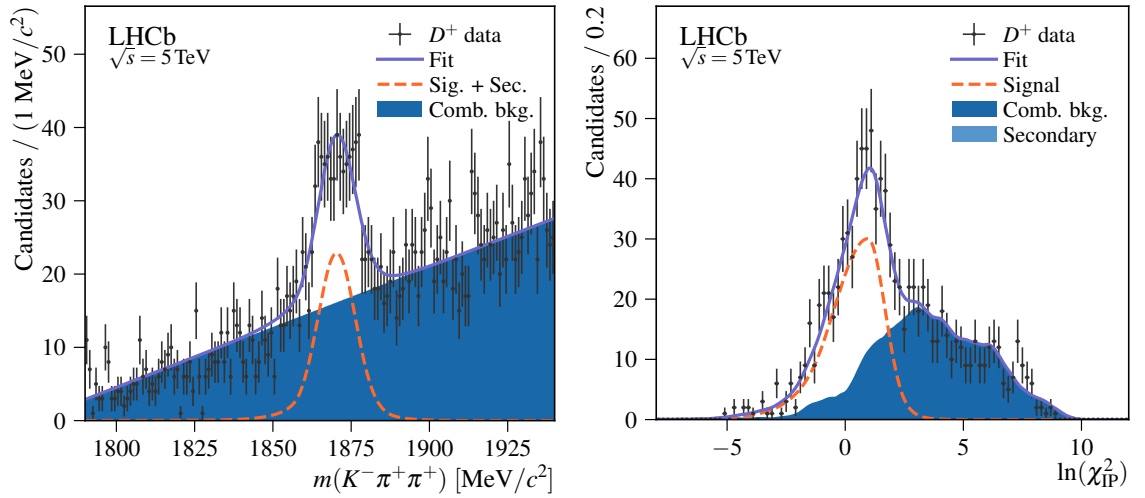
Tab. 5.8: Prompt signal yields in the fully selected dataset, summed over all (p_T, y) bins in which a measurement is made. Only statistical uncertainties are given.

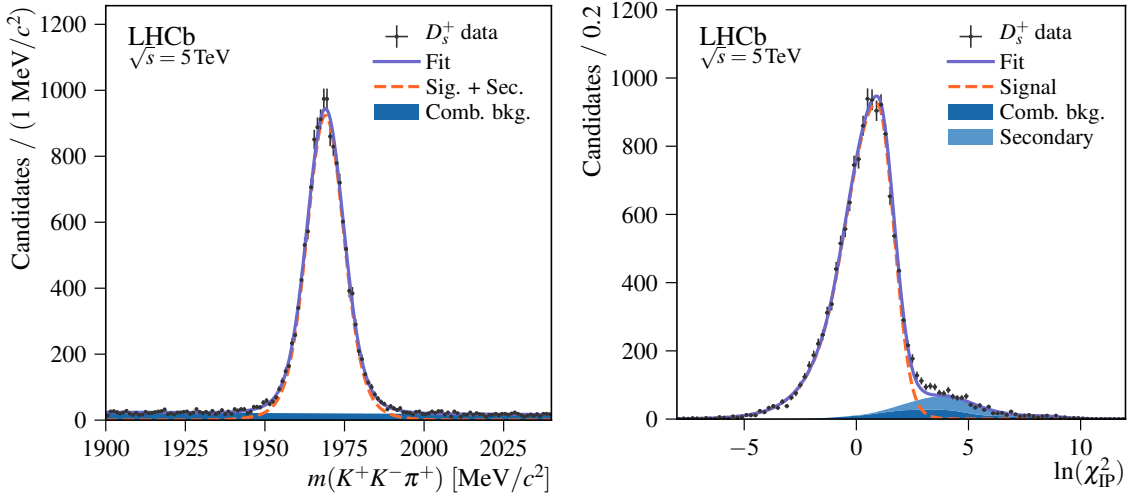
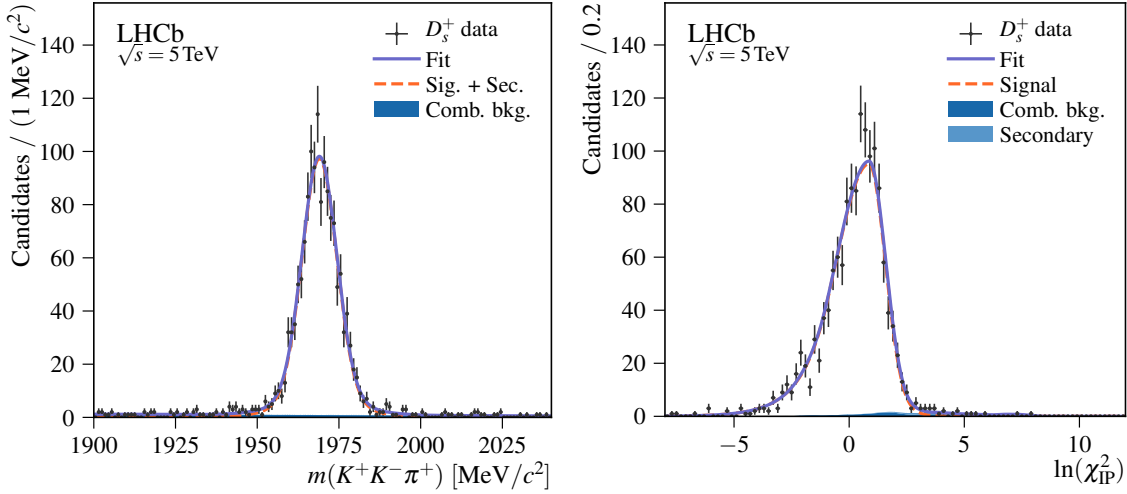
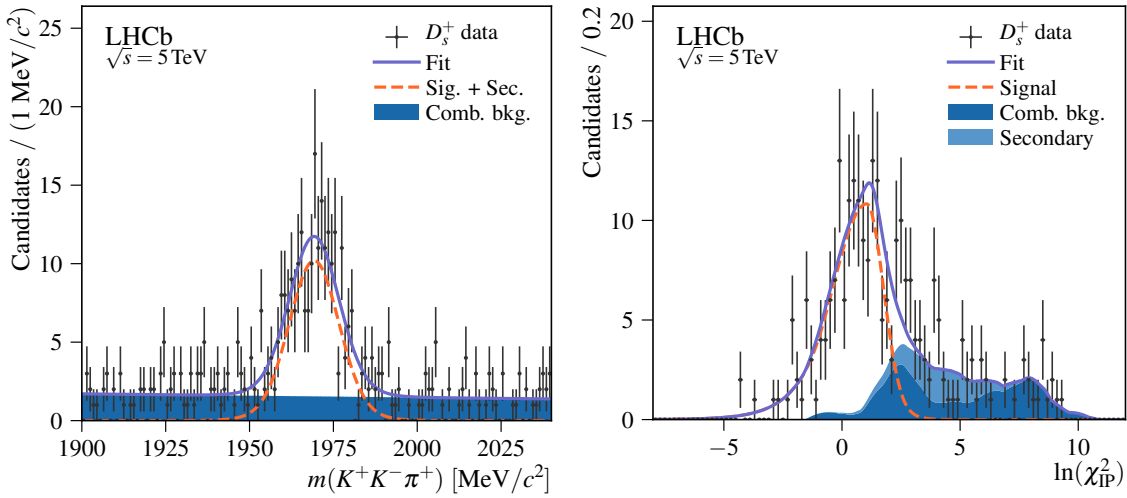
Decay	Prompt signal yield	
	13 TeV	5 TeV
$D^0 \rightarrow K^- \pi^+$	$(25.77 \pm 0.02) \times 10^5$	$(34.4 \pm 0.7) \times 10^4$
$D^+ \rightarrow K^- \pi^+ \pi^+$	$(19.74 \pm 0.02) \times 10^5$	$(27.6 \pm 0.6) \times 10^4$
$D_s^+ \rightarrow (\phi \rightarrow K^+ K^-) \pi^+$	$(11.32 \pm 0.04) \times 10^4$	$(13.2 \pm 0.1) \times 10^3$
$D^{*+} \rightarrow (D^0 \rightarrow K^- \pi^+) \pi^+$	$(30.12 \pm 0.06) \times 10^4$	$(39.0 \pm 0.2) \times 10^3$

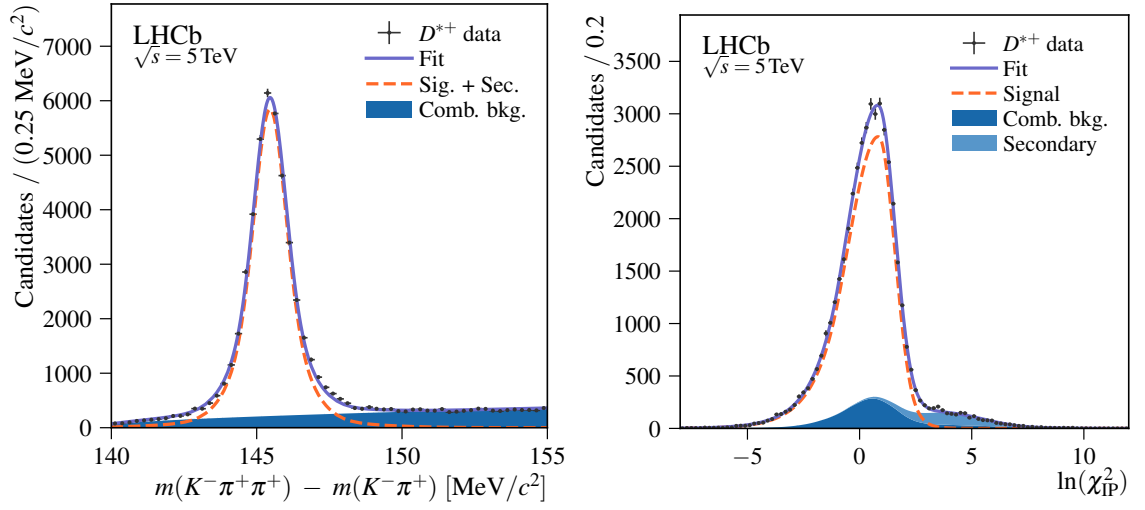
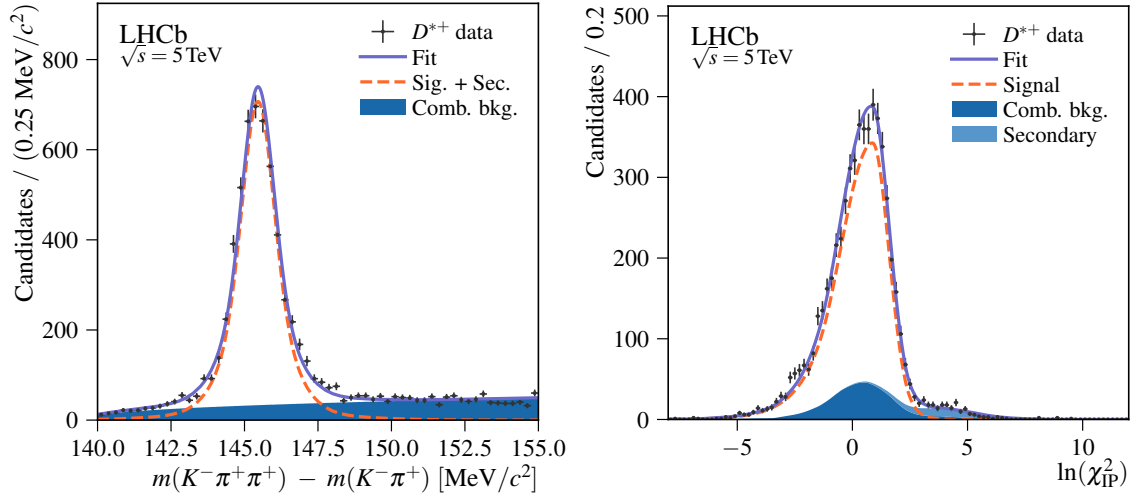
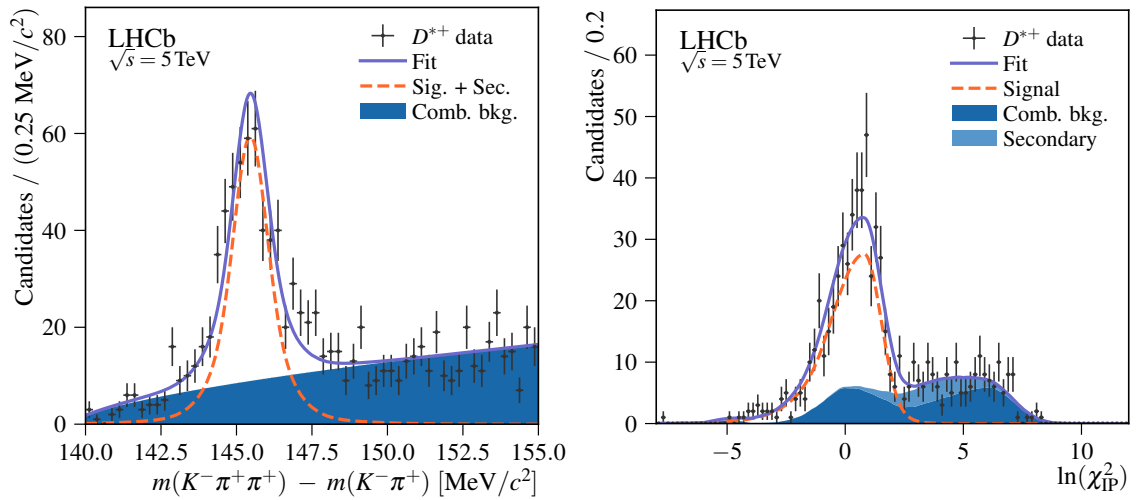
the entire dataset, but rather display the sum of the individual per-bin fits, scaled with the respective yields. While there is visible mis-modelling, especially for the D_s^+ fits in Figures 5.9 and 5.13 in the integrated fit, these are not statistically significant when looking at the individual binned fits. Nonetheless, a systematic uncertainty will later be assigned to account for this.

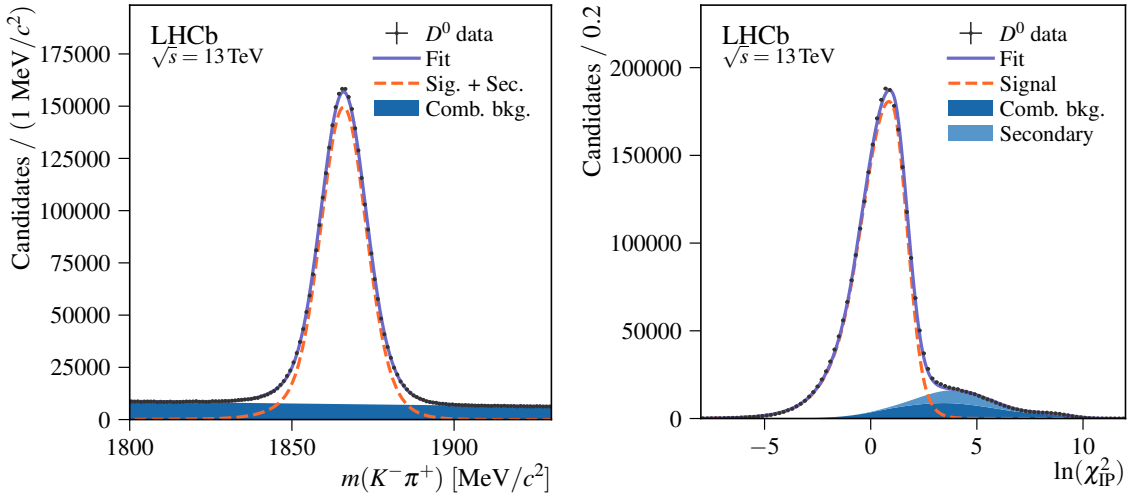
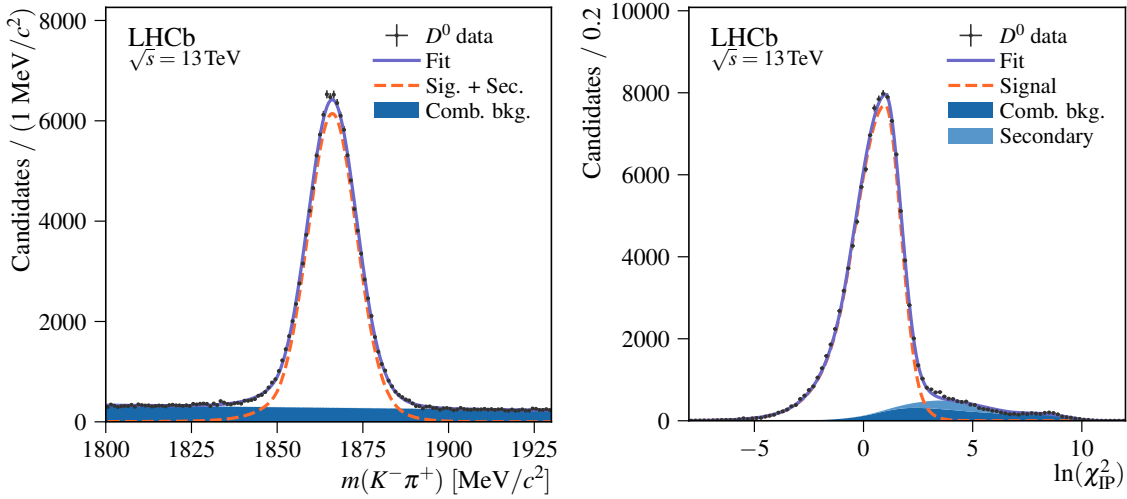
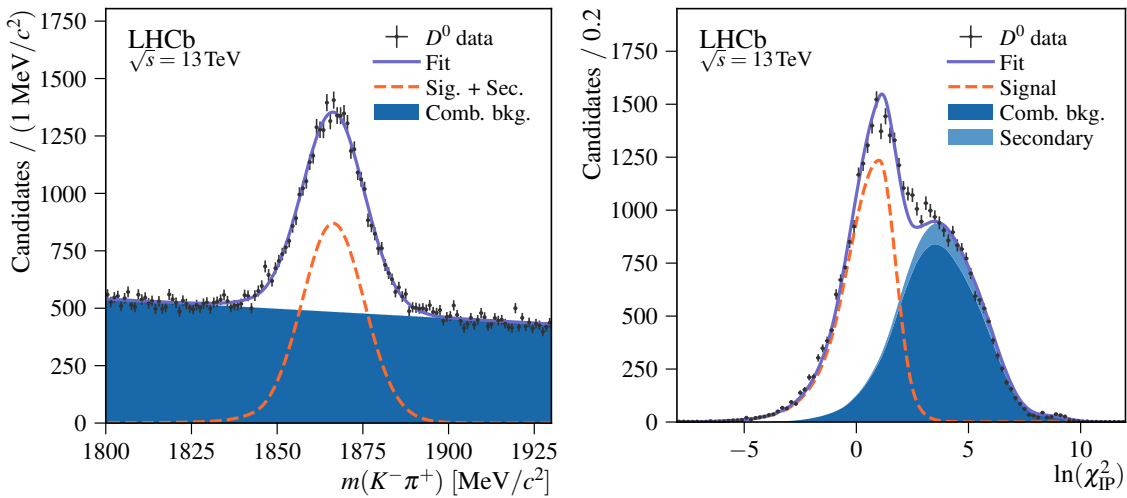
The total signal yield obtained by summing the individual per-bin yields is given in Table 5.8. As can be seen, D^0 and D^+ contribute about one order of magnitude more data to the analysis than D_s^+ and D^{*+} . Furthermore, there is also a factor of ten difference between the 13 TeV and the smaller 5 TeV dataset. This is hardly surprising: assuming roughly the same efficiencies, one expects a factor of 3.3 difference due to the lower effective luminosity and a factor of $13/5 \approx 2.6$ difference due to lower cross-sections. In the next section, those efficiencies are investigated and determined.

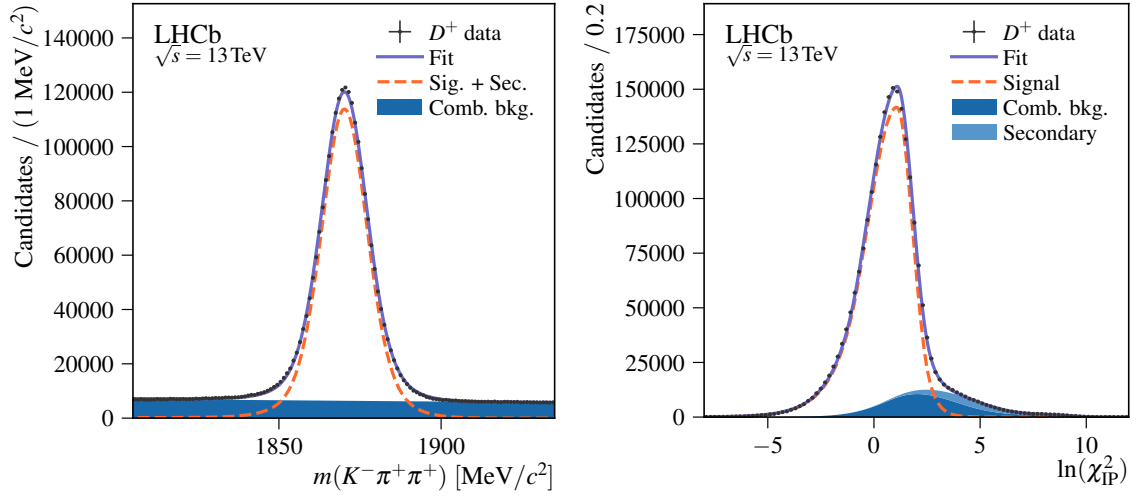
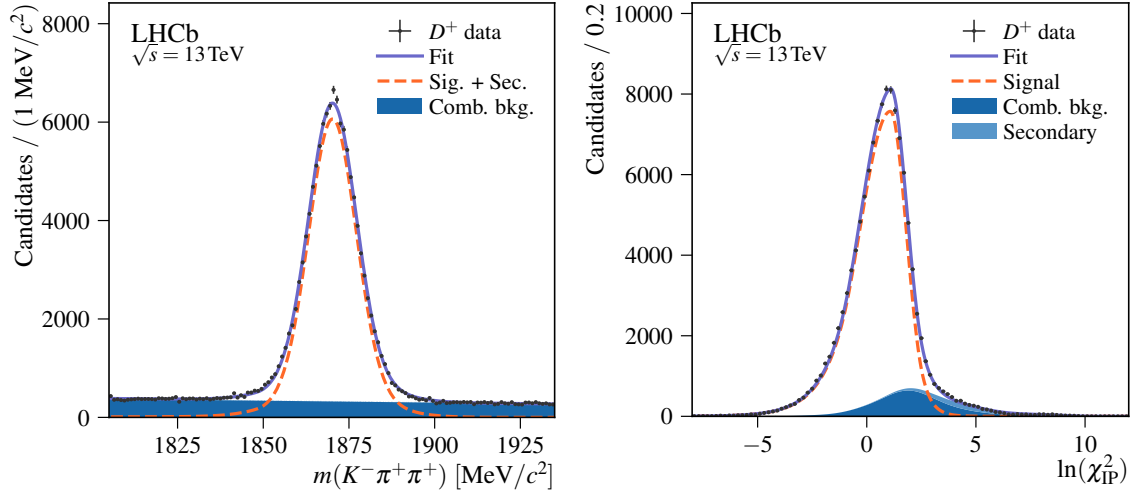
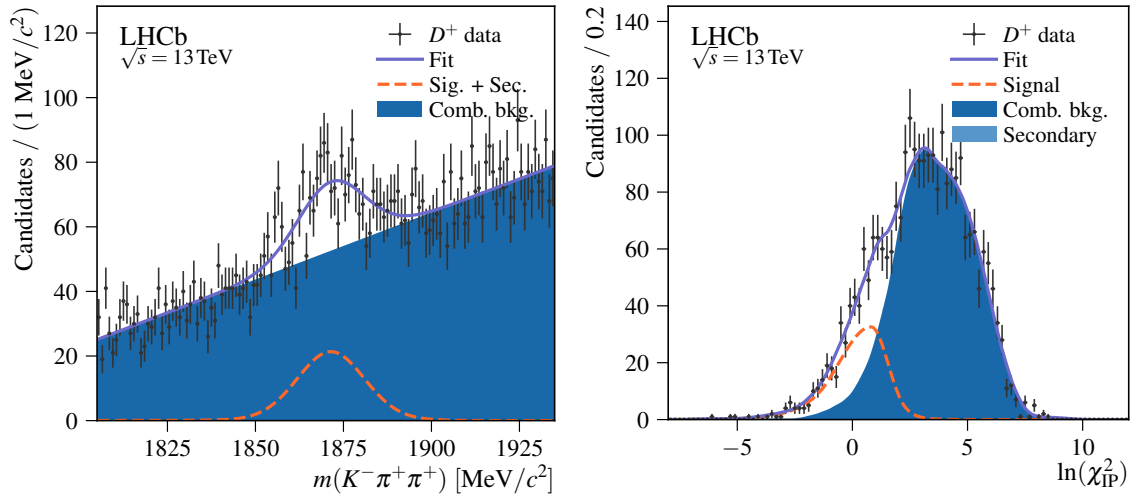
(a) Integrated fit with sum of individual (p_T, y) bin PDFs.(b) The bin with a large signal $3 \leq y < 3.5$; $1 \leq p_T < 2 \text{ GeV}/c$ (c) The bin with a large relative background $4 \leq y < 4.5$; $0 \leq p_T < 1 \text{ GeV}/c$ **Fig. 5.7:** Fits for the $D^0 \rightarrow K^- \pi^+$ fit at $\sqrt{s} = 5 \text{ TeV}$.

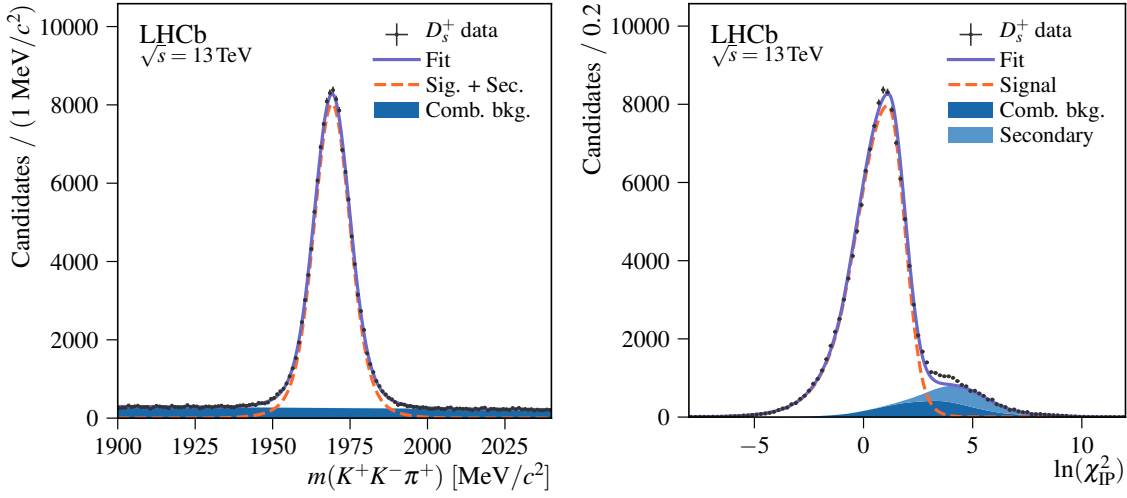
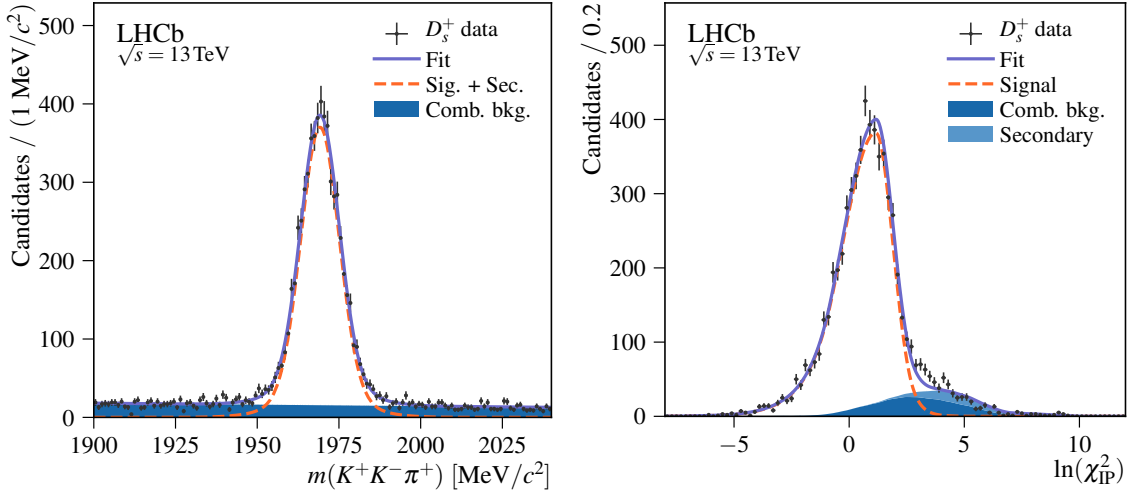
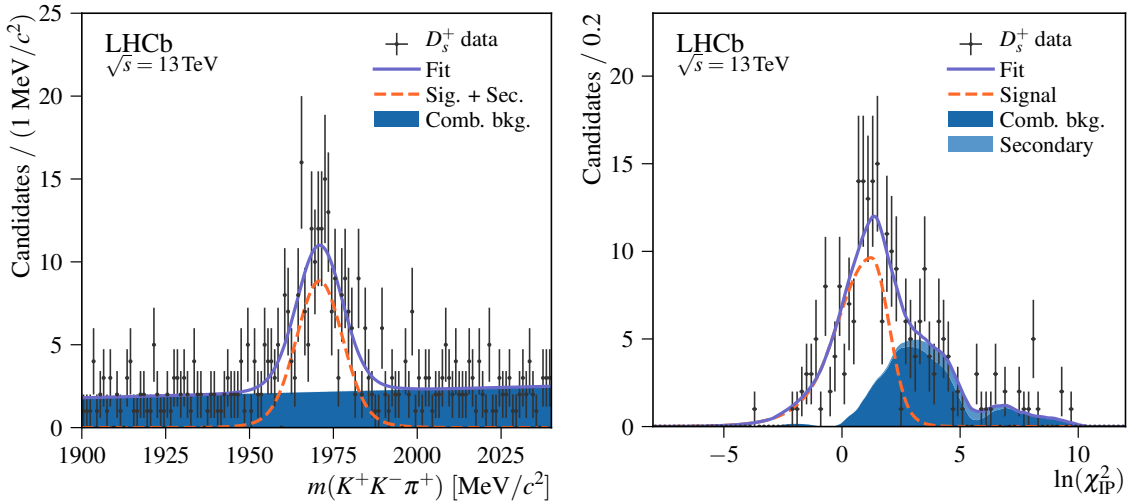
(a) Integrated fit with sum of individual (p_T, y) bin PDFs.(b) The bin with a large signal $3 \leq y < 3.5$; $2 \leq p_T < 3 \text{ GeV}/c$ (c) The bin with a large relative background $3 \leq y < 3.5$; $0 \leq p_T < 1 \text{ GeV}/c$ **Fig. 5.8:** Fits for the $D^+ \rightarrow K^- \pi^+ \pi^+$ fit at $\sqrt{s} = 5 \text{ TeV}$.

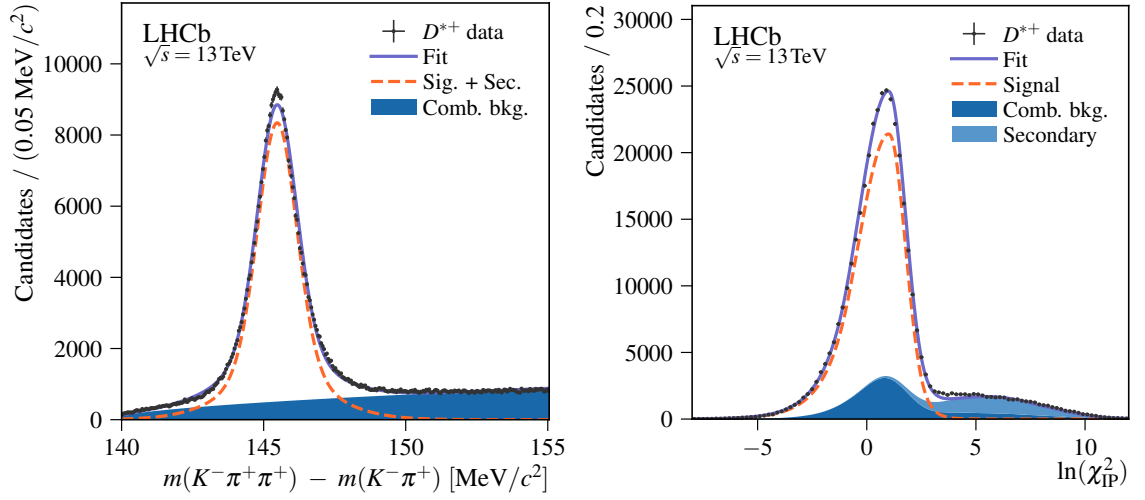
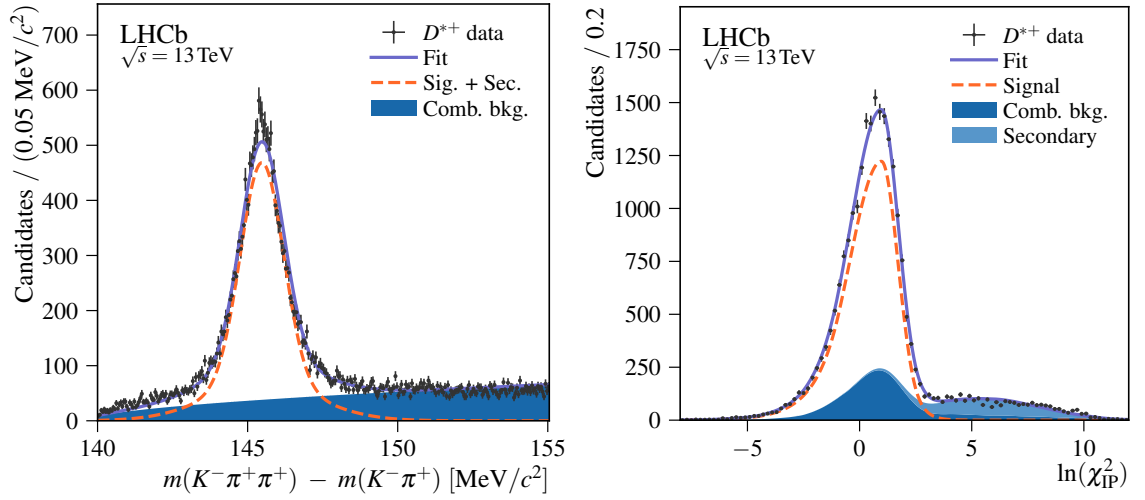
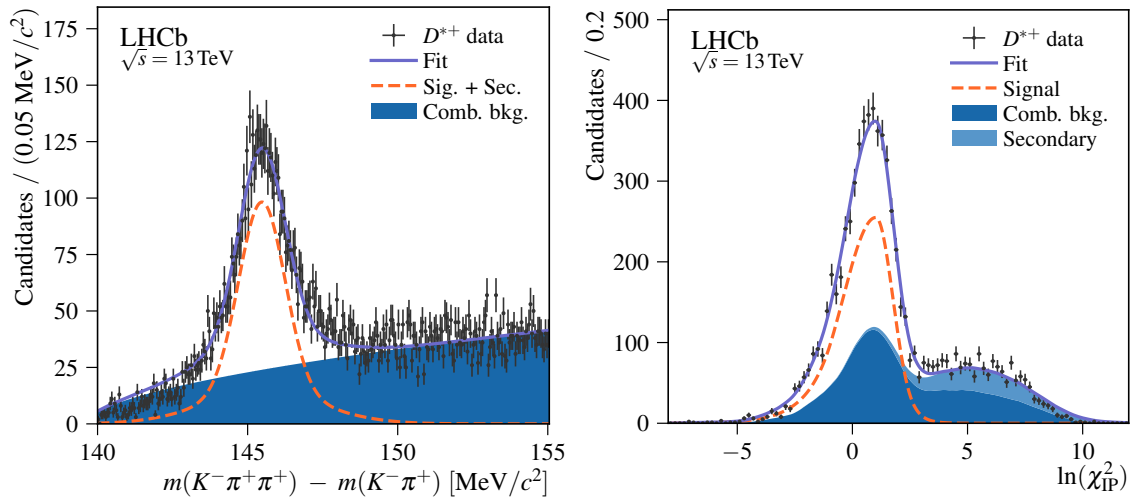
(a) Integrated fit with sum of individual (p_T, y) bin PDFs.(b) The bin with a large signal $2.5 \leq y < 3$; $2 \leq p_T < 3 \text{ GeV}/c$ (c) The bin with a large relative background $3.5 \leq y < 4$; $1 \leq p_T < 2 \text{ GeV}/c$ **Fig. 5.9:** Fits for the $D_s^+ \rightarrow (\phi \rightarrow K^+K^-)\pi^+$ fit at $\sqrt{s} = 5 \text{ TeV}$.

(a) Integrated fit with sum of individual (p_T, y) bin PDFs.(b) The bin with a large signal $2.5 \leq y < 3$; $2 \leq p_T < 3$ GeV/c(c) The bin with a large relative background $4 \leq y < 4.5$; $1 \leq p_T < 2$ GeV/c**Fig. 5.10:** Fits for the $D^{*+} \rightarrow (D^0 \rightarrow K^- \pi^+) \pi^+$ fit at $\sqrt{s} = 5$ TeV.

(a) Integrated fit with sum of individual (p_T, y) bin PDFs.(b) The bin with a large signal $3 \leq y < 3.5$; $1.5 \leq p_T < 2 \text{ GeV}/c$ (c) The bin with a large relative background $4 \leq y < 4.5$; $0 \leq p_T < 1 \text{ GeV}/c$ **Fig. 5.11:** Fits for the $D^0 \rightarrow K^- \pi^+$ fit at $\sqrt{s} = 13 \text{ TeV}$.

(a) Integrated fit with sum of individual (p_T, y) bin PDFs.(b) The bin with a large signal $3 \leq y < 3.5$; $2.5 \leq p_T < 3 \text{ GeV}/c$ (c) The bin with a large relative background $4 \leq y < 4.5$; $0 \leq p_T < 1 \text{ GeV}/c$ **Fig. 5.12:** Fits for the $D^+ \rightarrow K^- \pi^+ \pi^+$ fit at $\sqrt{s} = 13 \text{ TeV}$.

(a) Integrated fit with sum of individual (p_T, y) bin PDFs.(b) The bin with a large signal $2.5 \leq y < 3$; $2.5 \leq p_T < 3 \text{ GeV}/c$ (c) The bin with a large relative background $3.5 \leq y < 4$; $1 \leq p_T < 2 \text{ GeV}/c$ **Fig. 5.13:** Fits for the $D_s^+ \rightarrow (\phi \rightarrow K^+K^-)\pi^+$ fit at $\sqrt{s} = 13 \text{ TeV}$.

(a) Integrated fit with sum of individual (p_T, y) bin PDFs.(b) The bin with a large signal $2.5 \leq y < 3$; $2.5 \leq p_T < 3 \text{ GeV}/c$ (c) The bin with a large relative background $3.5 \leq y < 4$; $1.5 \leq p_T < 2 \text{ GeV}/c$ **Fig. 5.14:** Fits for the $D^{*+} \rightarrow (D^0 \rightarrow K^- \pi^+) \pi^+$ fit at $\sqrt{s} = 13 \text{ TeV}$.

5.5 Reconstruction and selection efficiencies

The last remaining ingredient in Equation (5.1) that is necessary to compute a cross-section is the efficiency ε_i in each (p_T, y) bin. This number combines all sources that affect how many of the true candidates that are produced in all pp collisions in a given data taking period enter the fit in the previous section. One such effect, the signal window efficiency, has already been presented. This section will discuss the remaining contributions, which can be broadly classified into two categories: efficiencies that are evaluated from the simulated sample and those obtained using data-driven techniques. The last category only comprises the efficiencies of the PID selection as the modelling of the RICH detectors in simulation does not reach sufficient precision.

5.5.1 Efficiencies from simulation

The efficiency from simulation can be split into multiple, factorising components:

$$\varepsilon_{\text{Simulation}} = \varepsilon(\text{acc.}|\text{gen.}) \times \varepsilon(\text{sel.}|\text{acc.}) \times \rho_{\text{truth}} \times \rho_{\text{tracking}} , \quad (5.13)$$

where the first term, $\varepsilon(\text{acc.}|\text{gen.})$, accounts for the efficiency of the generator level requirement, which has been applied before the simulation of the detector to save computing resources. The second component, $\varepsilon(\text{sel.}|\text{acc.})$, encapsulates the efficiency to reconstruct a candidate, the fraction of those candidates that pass the trigger selection, and the final off-line selection. As the efficiencies for the PID system are evaluated separately later on, any requirements made on PID quantities are removed here. The third term, ρ_{truth} , corrects the previous term for inefficiencies in the truth-matching of the simulated candidates (see Section 5.2.1) and, finally, the last term, ρ_{tracking} , corrects for a mis-modelling in the track reconstruction efficiencies. In the following, each of these components is discussed. While the next chapter presents the treatment of more involved systematic uncertainties on the measured cross-sections due to the efficiency determination from simulation, contributions due to the limited sample size are already presented alongside the evaluation of the efficiencies themselves.

The acceptance efficiency $\varepsilon(\text{acc.}|\text{gen.})$

To save computing resources, Section 5.2.1 already introduced a requirement of $10 < \theta < 400 \text{ mrad}$, which all final state particles in the true signal need to fulfil. While this requirement is trivial for charm mesons with a transverse momentum and rapidity in the central region of the measurement, significant fractions of the mesons in the (p_T, y) bins at the edge of the acceptance fail it and thus need to be

corrected. While the requirement is added to save precious computing resources due to the expensive detector simulation, the evaluation of its efficiency only requires simulated samples at generator level, which can be produced very quickly. Therefore, large samples have been generated and the efficiency is obtained as the fraction of the true candidates that pass the selection.

The obtained acceptance efficiencies increase with increasing p_T towards the central rapidity region, reaching values compatible with 100%. In the lowest and highest rapidity bin, the variation with p_T is the largest and the acceptance drops to about 75% in the $2 \leq y < 2.5$ and $4 \leq y < 4.5$ bin at $0 \leq p_T < 1 \text{ GeV}/c$.

When propagating the uncertainty to the cross-section, a binomial model is used instead of a simple binomial approximation. This is necessary because the approximation breaks down for very large or very small efficiencies when the number of events is small, a scenario that is present at the edges of the p_T and y region studied in this analysis. Given an efficiency ε and total number of events N , the probability to observe k events is given by the binomial distribution. Assuming a flat prior in the efficiency, the PDF describing the efficiency becomes:

$$P(\varepsilon) \propto \varepsilon^k \cdot (1 - \varepsilon)^{N-k} ,$$

which is also commonly referred to as the beta distribution $B(\alpha, \beta)$ with $\alpha = k + 1$ and $\beta = N - k - 1$. When performing the MC error propagation, this PDF is used to randomly sample values for the efficiency.

The selection efficiency $\varepsilon(\text{sel.}|\text{acc.})$

The reconstruction and selection efficiencies are estimated using the simulated samples with a full simulation of the LHCb detector. The denominator of the efficiency is given by all generated events in the sample (which have to pass the generator level selection) while the numerator is given by all reconstructed and fully selected candidates (excluding the PID requirements). The resolution on the transverse momentum of the charm meson in the simulated samples, obtained by comparing the true and reconstructed momentum, is 0.6% while the same number for the rapidity is only 0.2%. Therefore, the amount of migration between different (p_T, y) bins when going from generator to reconstruction level quantities is assumed to be negligible.

The selection efficiency shows the same general behaviour as the acceptance efficiency described above, where the efficiency is lowest in the $2 \leq y < 2.5$ and $4 \leq y < 4.5$ bin at small p_T , reaching values in the sub-per-mille range, and increases to 20% in other regions of phase-space for the D^0 and 15% for the remaining mesons at 13 TeV. Owing to the generally softer particles at 5 TeV, the efficiencies are about 10% smaller relative to those at 13 TeV. Uncertainties on those efficiencies are computed

as described above, with relative uncertainties ranging from less than 1% in the central p_T and y region to values significantly exceeding 10% towards the edges of the acceptance. As independent samples are used, this uncertainty is uncorrelated to the acceptance efficiencies discussed before. Obviously, using the simulated samples in this way requires a reliable modelling of the detector effects in the simulation. While the two correction factors discussed in the following already address potential disagreements, a more substantive validation of the simulated samples is performed in Section 5.6.4.

The truth-matching correction factor ρ_{truth}

Section 5.2.1 already discussed the truth-matching procedure and demonstrated that the truth-matching is not fully efficient for signal candidates. This is mostly caused by events in which the track cannot be matched to a true particle. Any true signal candidates that are falsely rejected, due to failing the truth-matching, bias the efficiencies estimated from these samples to lower values. Therefore, a correction factor ρ_{truth} is derived.

Figure 5.15 shows the distribution of candidates that could not be matched to the signal or constitute a partial decay chain. A clear peak consistent with true signal can be seen in the rejected candidates. To estimate the number of signal events, N_{fail} , the distributions are fitted with the same setup as described in Section 5.4.2. Only the δm fit for D^{*+} candidates has been modified to use only the sum of two normal distributions to model the peak. The correction factor is then obtained:

$$\rho_{\text{truth}} = \frac{N_{\text{pass}} + N_{\text{fail}}}{N_{\text{pass}}}, \quad (5.14)$$

where N_{pass} is the number of candidates matched as signal. In the Monte Carlo error propagation, N_{pass} is modelled as a Poisson distribution while N_{fail} is a normal distribution where the mean and width are set to the values from the fit. A normal distribution is used to allow for non-poissonian uncertainties from the fit. Due to the limited sample size of the simulated sample, the correction is assumed to be uniform across all (p_T, y) bins. The correction factor differs from unity by about 1.3% per final state track with a relative uncertainty due to the size of the simulated sample of about 0.2%.

The tracking correction factor ρ_{tracking}

The tracking efficiency itself, which is the efficiency of reconstructing a certain track and usually depends on the track kinematics, is part of the selection efficiency $\varepsilon(\text{sel.}|\text{acc.})$. However, simulated samples are known to show differences in tracking efficiencies with respect to real data [80], caused by slight differences in the detector

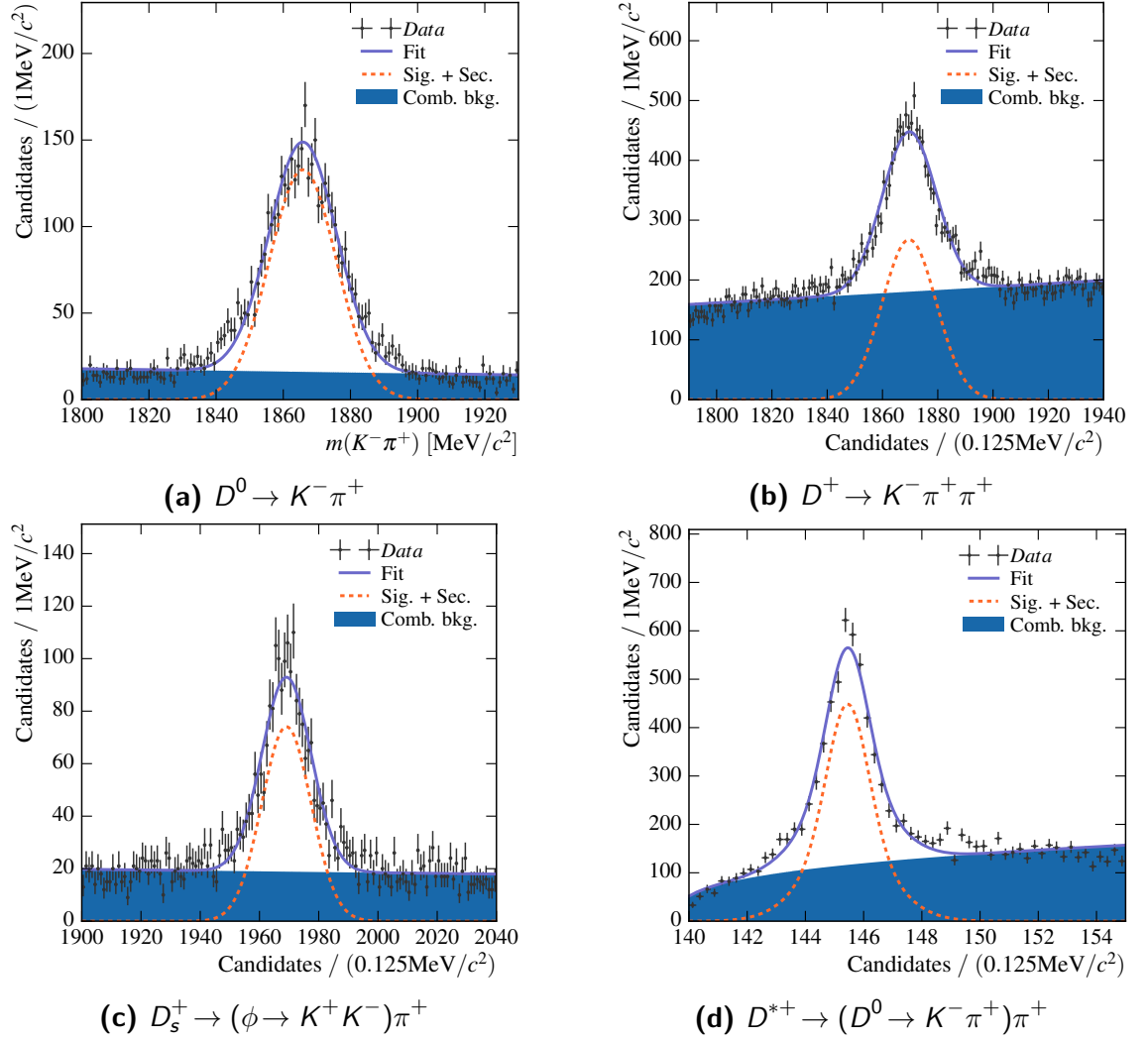


Fig. 5.15: Fits to the candidates that fail the truth-matching to extract the correction factor ρ_{truth} . The data shown are comprised of all events in the simulated samples except those that have been matched to the signal decay or are associated with a partially reconstructed background. The plots use the 5 TeV simulated samples.

occupancy and hit efficiencies. This difference needs to be accounted for in a measurement of absolute cross-sections. In order to measure this correction factor, it is first necessary to estimate the tracking efficiencies themselves. A detailed description for Run 1 is provided in Ref. [80] and a short summary is given in the following.

All tracks used in this analysis are so called long tracks, which are built from and are required to have hits in both the VELO and the T stations. Additional hits in the TT may be added to improve the momentum resolution; however, this is not a requirement. To estimate the long track efficiency, $J/\psi \rightarrow \mu^+ \mu^-$ decays are used in a tag-and-probe technique. In this approach, one muon is used as a tag, which is reconstructed like a normal long track. The probe muon is constructed primarily using the information recorded by the muon stations in a way similar to the muon-based L0 trigger. A track constructed from muon hits alone is extrapolated upstream and

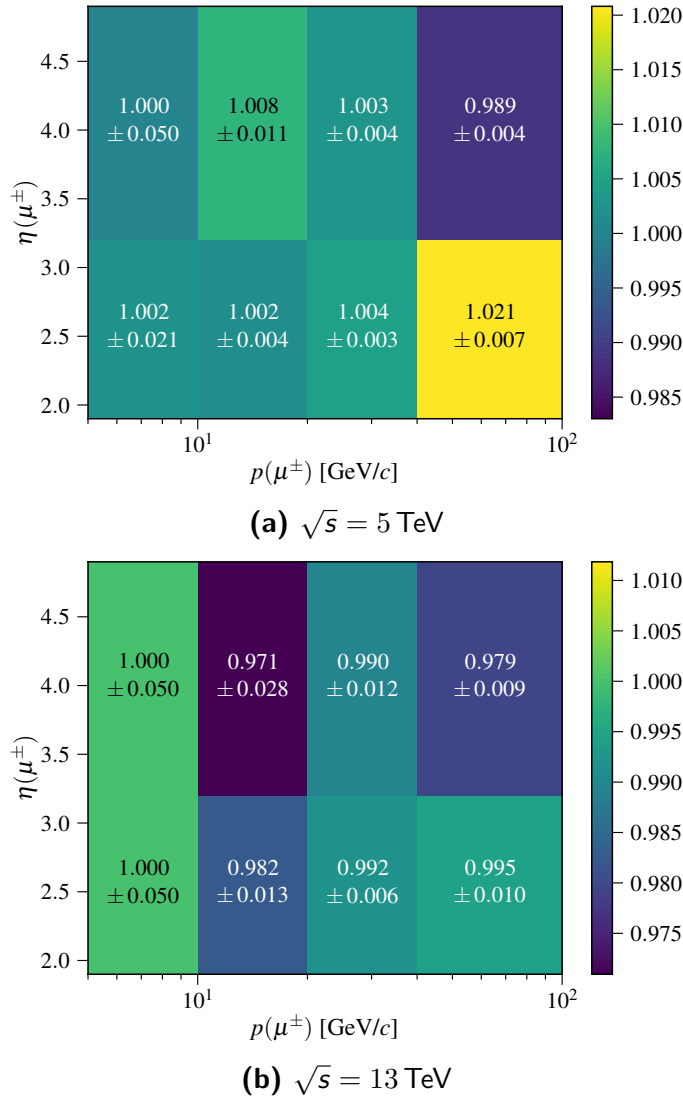


Fig. 5.16: Tracking correction tables for the two running periods, obtained using a tag-and-probe approach with $J/\psi \rightarrow \mu^+ \mu^-$ as described in the text and in more detail in Ref. [80]. The uncertainties given are due to the limited sample size of the long and muon track calibration sample. In the lowest p bin in (b) and lowest p high η bin, no calibration could be performed so the correction factor is set to 1 and is assigned as the uncertainty.

a ‘kick’ is applied at the plane of the magnet such that the track points to the same PV as the tag track. The tracking efficiency is then estimated by determining the number of $J/\psi \rightarrow \mu^+ \mu^-$ signal events from a likelihood fit, once to the entire sample and once to only those events where the probe track was successfully matched to a reconstructed long track. Tracking correction tables are then obtained by applying this method to both real data and simulated $J/\psi \rightarrow \mu^+ \mu^-$ decays and their ratio is taken. As the tracking efficiency itself and thus the corrections also depend on the kinematics of the tracks, the procedure is performed in bins of momentum p and pseudorapidity η of the muon. The resulting tables are given in Figure 5.16 for the two data-taking periods. In the following, a description is given how these tables are used in cross-section measurement to calibrate the simulated samples.

Weights are computed for every fully selected candidate in the simulated samples to align the distribution in the number of tracks with that observed in data. For each simulated and selected candidate, the tracking correction factor is computed as the product of the final state particles' correction factors taken from the tables above. The final correction factor ρ_{tracking} is then calculated as the weighted average of the individual candidate factors using the weights aligning the track multiplicity to data. As before, a value is computed for each (p_T, y) bin and the obtained correction factors are in the range of 1.02–0.97 (5 TeV) and 1–0.95 (13 TeV) for the three-body decays, with slightly smaller corrections for $D^0 \rightarrow K^- \pi^+$ as it is comprised of only two tracks. The propagation of the statistical uncertainty in the correction tables as well as additional uncertainties arising due to the application of a table derived using muons to hadrons are discussed in Section 5.6.2.

5.5.2 Particle identification efficiencies

This analysis applies requirements on the PID response of the tracks comprising the reconstructed candidate. As previously discussed, requirements are applied on the $DLL_{K-\pi}$ property whose computation is described in Section 3.3.2. While these requirements are quite loose and mainly intended to reduce the amount of background making them efficient for signal events, their efficiency needs to be determined nonetheless. However, the simulation of the RICH detectors is non-trivial: the RICH response depends on the particles kinematics, the number of particles traversing the detector in an event, and experimental conditions like temperature and gas pressure, which may change over time during a data taking period. Therefore, a data-driven technique is applied to estimate the efficiency of the requirements imposed on pions and kaons.

The calibration method

The data-driven method employed in the cross-section measurement is described in more detail in Ref. [107]. The starting point is a calibration sample containing positively identified kaon and pion tracks that do not have any PID requirements applied. The sample is then partitioned in variables that the PID response depends on: the track's momentum p and pseudorapidity η , and the number of hits in the SPD detector N_{SPD} as a measure of the detector occupancy. In each of these bins, the PID efficiency of a selection is given by:

$$\varepsilon_i = \frac{C'_i}{C_i}, \quad (5.15)$$

where C_i is the total number of tracks and C'_i is the number of tracks passing the requirement in the respective bin. The bins form a look-up table, which gives the

probability for a track with given p and η as well as N_{SPD} to pass the PID selection. Assuming the partition is fine enough such that the value of ε_i does not depend on the distributions in the calibration sample, the constructed efficiency look-up table can be applied to any other sample (henceforth called the reference sample) to compute the efficiency:

$$\varepsilon_{\text{PID}} = \frac{1}{R} \sum_j R_j \varepsilon_j , \quad (5.16)$$

where R_j is the number of events in the reference sample whose track is in bin j , and R is the total number of events in said sample. However, the analysis presented here has multiple final state particles with PID requirements. Therefore, the application of the efficiency is done in a similar way to that of the tracking correction factor by applying the calibration table per event:

$$\varepsilon_{\text{PID}} = \frac{1}{R} \sum_j w_j \prod_i \varepsilon(p_i, \eta_i, N_{\text{SPD},i}) , \quad (5.17)$$

with the sum running over all events, w_j is the weight of the j -th event in the reference sample, and the product is taken over all decay products. This approach includes the implicit assumption that the tracks PID responses are independent from each other. Typically, the simulated sample is used as the reference sample. This requires a good modelling of p and η as well as N_{SPD} in the simulation. As this is not guaranteed, this analysis uses the fully selected (including the PID requirements) data sample that is also used when extracting the signal yield from the $\ln \chi^2_{\text{IP}}$ fit in Section 5.4. This is a valid approach if no region is completely excluded by the PID selection [107]. As the PID selection has already been applied, the inverse of the efficiency from the calibration tables is used to recover the number of events in the reference sample before the cut and compute the efficiency:

$$\varepsilon_{\text{PID}} = \frac{R}{\sum_j w_j \prod_i \varepsilon(p_i, \eta_i, N_{\text{SPD},i})^{-1}} . \quad (5.18)$$

As the data sample does not contain pure signal events, the weight w_j is extracted for every event using the *sPlot* technique [108] applied to the $\ln \chi^2_{\text{IP}}$ fit. As the kinematics of the decay products significantly depend on the kinematics of the parent, PID efficiencies are computed separately for each (p_T, y) bin. Systematic uncertainties introduced into the measurement by this data-driven approach can be substantial and are discussed in detail in Section 5.6.5.

Creation of the calibration tables

Calibration samples for kaons and pions are extracted from $D^{*+} \rightarrow (D^0 \rightarrow K^- \pi^+) \pi^+$ decays which have been selected with a dedicated trigger line. Unlike the HLT2 line discussed in Section 5.3.2, which is used in the measurement, this dedicated line accepts all events that have passed at least one L0 and one HLT1 line. This leads to a substantially larger number of events compared to the dataset used for the cross-section measurement and a negligible overlap between the two samples. To extract the number of signal events in this sample (and at the same time identify which of the two D^0 tracks is the kaon and which one is the pion), a two-dimensional likelihood fit is performed in $m(K^- \pi^+)$ and δm . In this fit, both possible mass hypotheses for the child tracks are considered. Due to the excellent resolution of the LHCb detector, only the correct assignment produces the expected peak in the distribution while the wrong assignment, if it even ends up in the fit range, cannot be distinguished from true combinatorial background. The signal is modelled by the sum of two normal distributions in both dimensions. The background consisting of real $D^0 \rightarrow K^- \pi^+$ associated with the wrong pion is modelled with the same shape in $m(K^- \pi^+)$ while the δm part is modelled using:

$$B(\delta m) \propto \begin{cases} \sqrt{\frac{\delta m}{b}} - 1 \cdot \exp\left(\frac{a}{b} \delta m\right) & \delta m \geq b \\ 0 & \delta m < b \end{cases}, \quad (5.19)$$

where $b = 139.57 \text{ MeV}/c^2$ is fixed to the mass of the $\pi^\pm$ and a is allowed to float in the fit. Lastly, the combinatorial background in δm uses the same shape (with independent parameters) and a linear polynomial for the background in $m(K^- \pi^+)$. From this fit and using the *sPlot* technique, weights are computed which allow the number of signal events to be determined. The yields C'_i and C_i in Equation (5.15) are then given by the sum over these ${}_s\mathcal{W}$ for all events in the i -th partition of the calibration sample, before and after any PID requirements are imposed respectively.

Figures 5.17 and 5.18 show the one-dimensional efficiencies in the calibration sample as a function of p , η and N_{SPD} for kaons and pions, and for 5 TeV and 13 TeV respectively. The observed variation is strongest for kaons, which show a strong dependence on p and η with a drop in efficiency towards larger values as the kaon and pion hypothesis become harder to distinguish (and hence reduce the difference of there likelihoods, $\text{DLL}_{K-\pi}$). The pion efficiency shows only a small variation as a function of the kinematics, which is hardly surprising because the selection requirement of $\text{DLL}_{K-\pi} < 3$ accepts all tracks that do not look too much like a kaon, and especially at large values of p or η , the difference in velocity and hence Cherenkov angle for pion and kaons of identical momentum is small. Comparing the two datasets shows slightly different behaviour, especially in the low p and η region

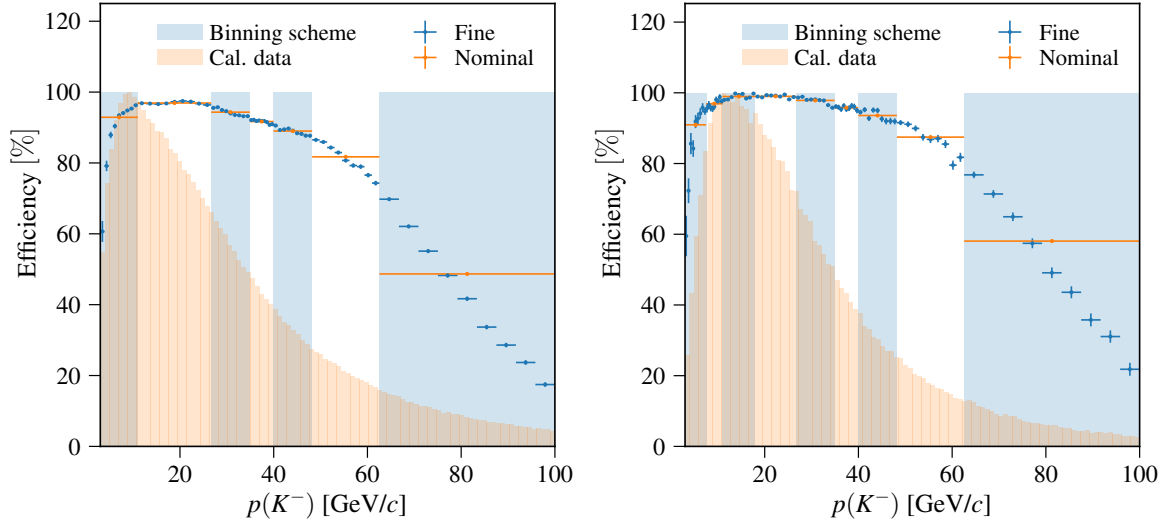
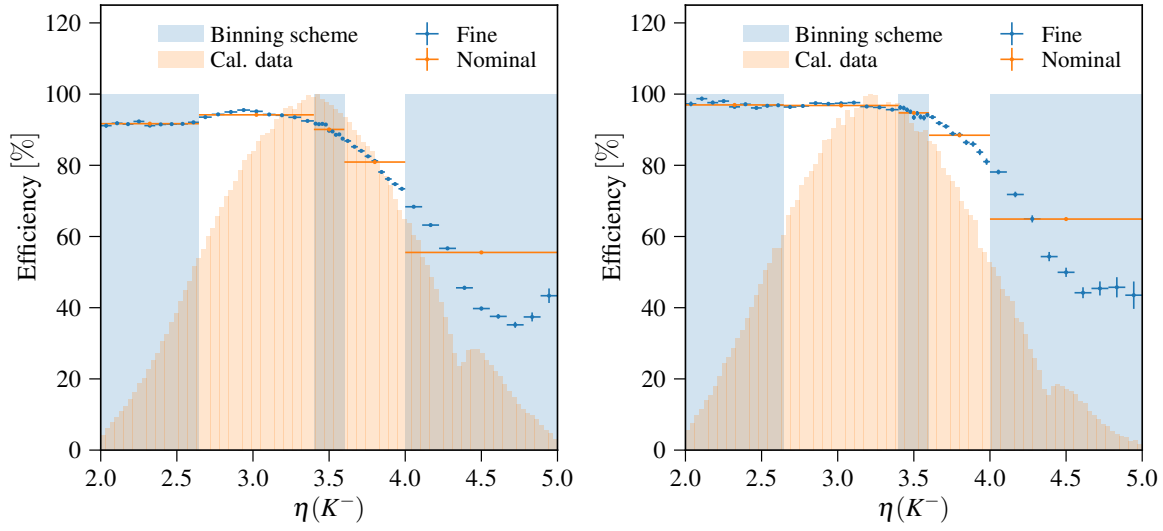
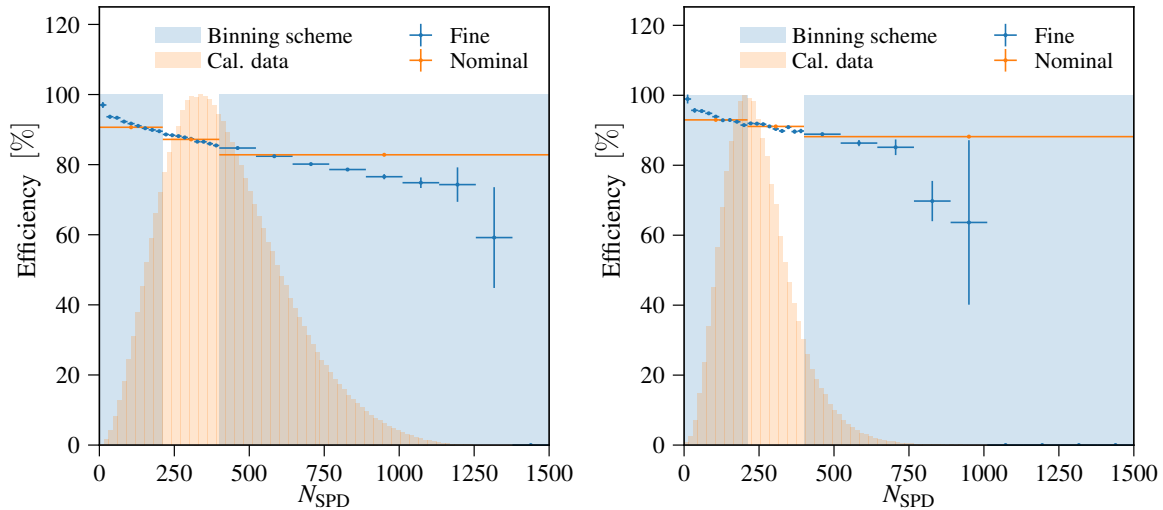
(a) p dependence at 13 TeV (left) and 5 TeV (right).(b) η dependence at 13 TeV (left) and 5 TeV (right).(c) N_{SPD} dependence at 13 TeV (left) and 5 TeV (right).

Fig. 5.17: One dimensional PID efficiency for kaons for $\text{DLL}_{K-\pi} > 5$ in the calibration sample. The alternating transparent and shaded area show the binning scheme used for the three-dimensional efficiencies with the efficiency within those bins shown in orange. The efficiency with a much finer binning is shown in blue, revealing the insufficient precision of the nominal binning scheme.

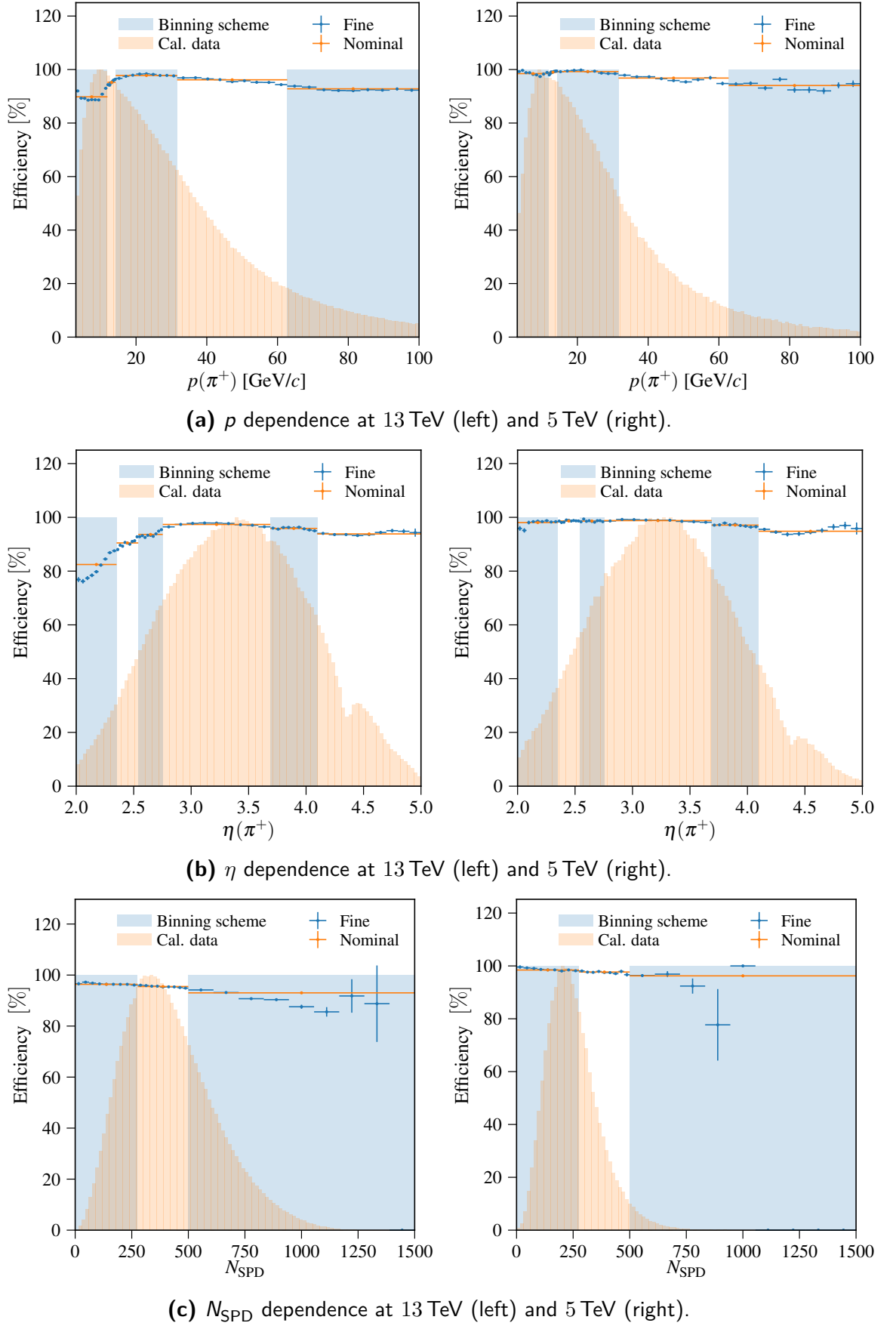


Fig. 5.18: One dimensional PID efficiency for pions for $\text{DLL}_{K-\pi} < 3$ in the calibration sample. The alternating transparent and shaded area show the binning scheme used for the three-dimensional efficiencies with the efficiency within those bins shown in orange. The efficiency with a much finer binning is shown in blue, revealing the insufficient precision of the nominal binning scheme.

with a better performance seen for 5 TeV. This is caused by a non-optimal calibration and alignment of the RICH detectors being used during the early measurement data-taking period. The data recorded during that period was used to optimise the calibration for the remainder of the 2015 data-taking, including the 5 TeV period. Also indicated in the figure is the binning scheme in the respective dimension, which is used to create the calibration tables. Bin boundaries have been placed to best utilise the available calibration data, by adding fewer bins where the efficiency is constant and adding more bins where the variation is large.

The actual definition of bin boundaries and how the efficiencies are computed in each bin differs between the two measurements. In the classical approach used for the 13 TeV results, the weights ${}_s\mathcal{W}_j$ obtained from the *sPlot* technique are used to determine the numbers before and after applying the PID selection. The statistical uncertainty in each bin is then given by a modified binomial uncertainty based on Chapter 8.5 in Ref. [109]:

$$\sigma_{\varepsilon_i}^2 = \frac{1}{C_i^2} \left[\left(1 - 2 \frac{C'_i}{C_i} \right) \sum_{j \in C'_i} {}_s\mathcal{W}_j^2 + \frac{C_i'^2}{C_i^2} \sum_{j \in C_i} {}_s\mathcal{W}_j^2 \right]. \quad (5.20)$$

Due to the nature of *sPlot* weights, which can be negative to subtract background events, sparsely populated $(p, \eta, N_{\text{SPD}})$ bins may have unphysical efficiencies if the PID requirement removed more positively than negatively weighted events. Additionally, binomial uncertainties do not correctly model the statistical uncertainty in this case. To counter this behaviour, the binning scheme is required to keep a sufficient number of events in each bin, enlarging bins especially in the low and high momentum region. Following the idea of Monte Carlo error propagation when calculating the PID efficiency using Equation (5.18), the statistical uncertainty in every calibration table bin is propagated by generating a large number of toy calibration tables by randomly sampling a new value for every $(p, \eta, N_{\text{SPD}})$ bin from a normal distribution. The mean and width of this normal distribution are those obtained from the nominal calibration sample as described above. This approach ensures the correct statistical correlation of tracks falling into the same $(p, \eta, N_{\text{SPD}})$ bin.

For the 5 TeV measurement, the approach for filling the calibration tables for kaons and pion has been modified. Owing to a significantly smaller calibration sample size, which would require even larger bins using the classical approach, the weights ${}_s\mathcal{W}$ obtained from a global fit to the sample are not used. Instead, the same likelihood fit is performed in every $(p, \eta, N_{\text{SPD}})$ bin separately. To avoid having to perform two fits in each bin, one before and one after applying the PID cut, the likelihood fit is performed in three dimensions. These are, as before, the invariant mass and mass difference with the addition of the binary cut decision f which is either zero or one. As this is nothing but a simple coin toss process, it is modelled using a

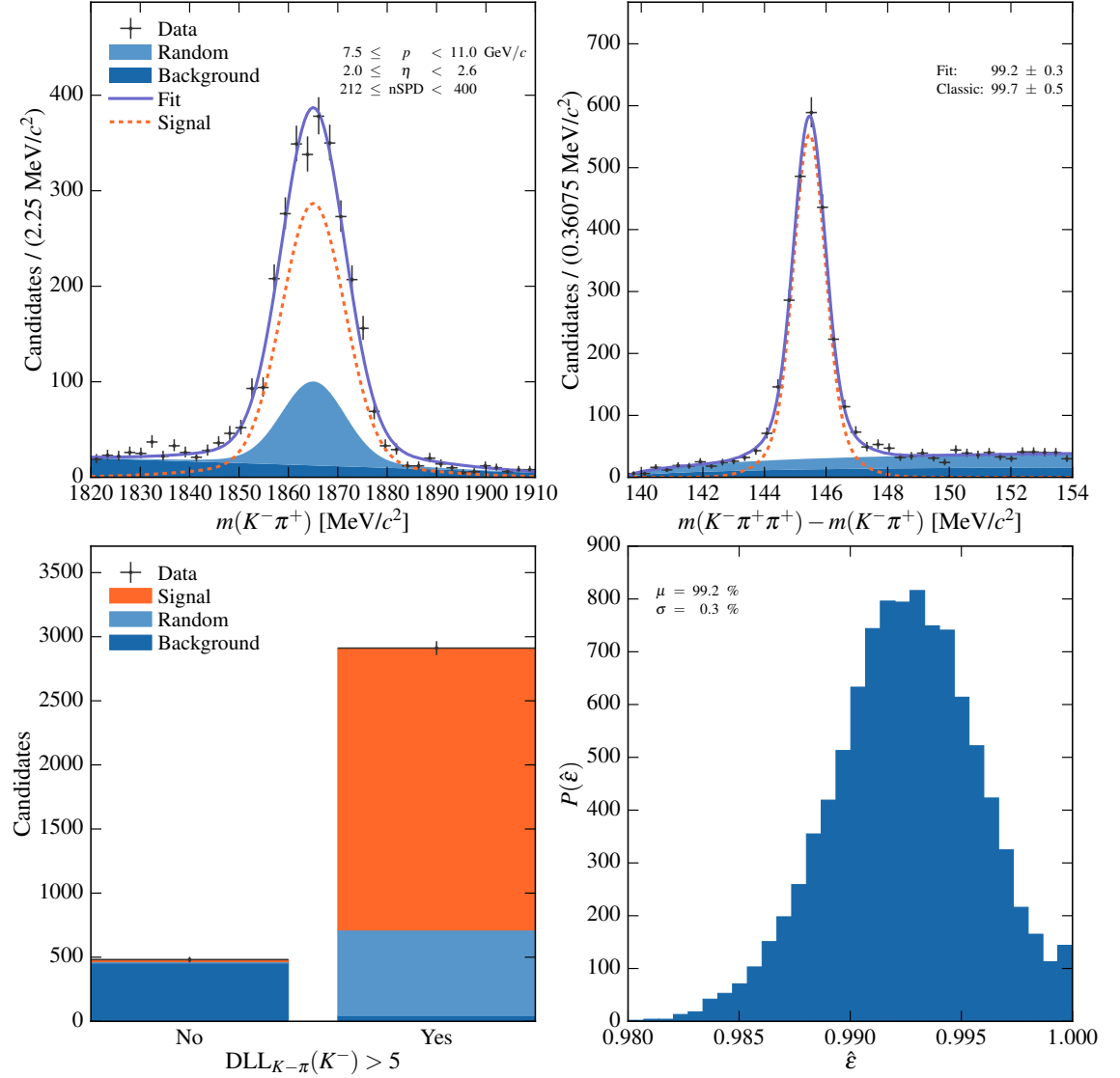


Fig. 5.19: Illustration of three-dimensional likelihood fits to the 5 TeV kaon calibration sample used to extract the PID efficiency in a specific $(p, \eta, N_{\text{SPD}})$ bin. Shown are the projection on the invariant mass (top left), mass difference between the D^{*+} and D^0 candidate (top right), and the PID cut decision (bottom left). Also shown is the obtained distribution of the estimated efficiency after bootstrapping the fit 10,000 times (bottom right).

Bernoulli PDF:

$$F(f) \propto \varepsilon^f \cdot (1 - \varepsilon)^{1-f} . \quad (5.21)$$

Each component of the fit is modelled with an independent Bernoulli PDF hence allowing signal and background to have different efficiencies. An example for such a fit is given in Figure 5.19, displaying the fit results in one $(p, \eta, N_{\text{SPD}})$ bin for the kaon calibration sample. The values for the efficiency in a specific bin are compatible with those obtained using the classical approach if it was possible to apply. Using those fits, it is also possible to improve the determination and propagation of the uncertainty due to the finite size of the calibration sample. In the example given in Figure 5.19, the classical approach yields an efficiency $99.7 \pm 0.5\%$ which raises the question of how to correctly propagate it, as the assumption of it being normally distributed would lead to unphysical efficiencies leaking past one. Instead, the calibration sample is bootstrapped and the fit is repeated in each of the bootstrapped samples, yielding one value for the efficiency, and it is this list of efficiencies that is used in the Monte Carlo error propagation. An example of the distribution of the obtained efficiencies is also given in Figure 5.19. Lastly, applying this bootstrapping approach allows checking for possible statistical correlations between the pion and kaon calibration tables as both are filled from the same calibration sample of $D^{*+} \rightarrow (D^0 \rightarrow K^- \pi^+) \pi^+$ decays. In those decays, correlations exist between the kaon and pion kinematics and the number of hits in the SPD is obviously identical. All possible combinations of two bins, one from each table, is investigated but no significant correlation is found. As a cross-check, the new approach has also been applied to the 13 TeV measurement and yields efficiencies compatible with the classic approach.

The obtained PID efficiencies

A PID efficiency is computed for every charm meson (p_T, y) bin. As expected from Figures 5.17 and 5.18, the obtained efficiencies are similar between the 5 TeV and 13 TeV measurements and driven by the dependence of the PID requirement imposed on the kaon. Therefore, ε_{PID} is largest in the low y , central p_T (caused by the loss of efficiency for low momentum kaons) and reduces gradually towards the high y , high p_T corner. Typical values range from 55% to 95% for 5 TeV and are approximately 10% lower for 13 TeV due to the much larger number of SPD hits impacting the performance. The relative uncertainty due to the size of the calibration sample is 0–2%. This, however, is not the only systematic effect due to PID and is in fact dwarfed by the uncertainty due to the chosen binning scheme in the calibration tables in many (p_T, y) bins. Looking at Figure 5.17, one can hardly say that the chosen binning scheme is good at capturing and describing the shape of the efficiency

dependence for the kaon and the associated uncertainty is dedicatedly discussed in the following section.

5.6 Systematic uncertainties

In the previous sections, the estimation and propagation of potential systematic uncertainties for several sources have already been discussed where appropriate. This covered the uncertainty on the signal window efficiency when measuring the prompt signal yield, the finite size of the simulated samples giving rise to an uncertainty on the selection and acceptance efficiencies as well as the truth-matching correction factor, and the uncertainty on the PID efficiency tables due to a finite size of the PID calibration samples. This section discusses additional sources of uncertainty that affect the measured cross-section.

5.6.1 Uncertainties from external input

The branching ratios for the studied final states are taken as the current best averages as summarised in Table 5.1. The uncertainties are propagated through the Monte Carlo error propagation assuming they follow a normal distribution. As the same set of branching ratios are used in the 5 TeV and 13 TeV measurements, they are fully correlated. The second external input with its own set of uncertainties is the luminosity measured for the two data-taking periods. The components of the uncertainty are described in Ref. [104]. In total, the luminosity contributes 3.8% uncertainty to the cross-sections measured at 5 TeV and 3.9% to those at 13 TeV. The uncertainty is fully correlated between different (p_T, y) bins and mesons and are 32% correlated between different data-taking periods.

5.6.2 Tracking correction uncertainties

The tracking correction factors, which are obtained in a separate analysis as described in Ref. [80], come with their own set of systematic uncertainties that need to be propagated as systematic uncertainties on the cross-sections. The propagation of the statistical uncertainties in the procedure used to obtain the tracking correction tables given in Figure 5.16 is similar to that used for the PID calibration tables at 13 TeV: toy tracking correction tables are generated by randomly sampling new values for each entry in the tracking correction table and repeating the correction procedure for every table. This ensures the correct treatment of any correlations.

Aside from the size of the calibration sample, systematic uncertainties arise when applying the tracking correction table in this measurement: the tracking efficiency and subsequently the correction factors are derived using a muon probe track but are

applied to hadron tracks. This is not a problem if the change in tracking efficiency when comparing hadrons and muons is identical in real data and simulation. However, the material budget in simulation is not precisely known and an conservative estimate for the uncertainty of 10% is assumed. Using simulated events, it is determined that 11% (14%) of the kaons (pions) cannot be reconstructed because they participate in hadronic interactions before the last T station. Combining this with the uncertainty on the material budget leads to a 1.1% and 1.4% systematic uncertainty on the tracking correction factors, which is applied per track. Lastly, the variation of the tracking correction tables has been studied when changing the observable used to weight the multiplicity in simulation to match that in data. This leads to a maximal deviation of 0.4%, which is conservatively assigned as a systematic uncertainty per track.

The obtained uncertainty on the cross-section varies substantially between different (p_T, y) bins of the measurement. It is highest in the low p_T , low y region caused by the large statistical uncertainty in the correction tables at low track momentum, and decreases with increasing p_T and y . While this behaviour is the same for all mesons at both centre-of-mass energies, the actual ranges differ at the different energies and for the different mesons where $D^0 \rightarrow K^- \pi^+$ has a generally lower uncertainty due to having only two tracks instead of three. At 5 TeV, the relative uncertainty for D^0 covers a range of 3–5% while the three-body modes lie in the 4–7% range. At 13 TeV due to both low p bins in the correction tables having a large uncertainty, D^0 covers a range of 3–10% while the three-body modes extend to the 4–14% range.

The tracking correction factor is one of the dominant sources of systematic uncertainty affecting the measured cross-sections. This is partly remedied when taking ratios between measurements at different $\sqrt{s}$, where the contributions due to the material budget uncertainty and occupancy reweighting are assumed to be fully correlated.

5.6.3 Choice of fit model

The PDFs used in the $\ln \chi^2_{\text{IP}}$ fit in Section 5.4.3 to extract the number of prompt signal candidates have been chosen somewhat arbitrarily. With the exception of the background model, which is based on data-driven methods, the shapes for prompt and secondary signal decays have no physical meaning but have been chosen solely based on their ability to describe the data. Therefore, the difference in measured prompt signal yield when using different models is assigned as an uncertainty for each (p_T, y) bin.

One alternative shape for the prompt signal contribution is obtained by modifying the nominal PDF given in Equation (5.12) to use the sum of two normal distributions for the central part instead of a single one. A second shape is obtained by modifying

the normal χ^2 distribution to allow for asymmetric widths on either side of the most probable value and transforming the result into a PDF for $\ln \chi^2$ [110]. For the secondary shape, all prompt signal shapes are tested. To compute the uncertainty, the $\ln \chi^2_{\text{IP}}$ fit is performed again with each combination of alternate shapes for prompt and secondary signals. The uncertainty in each (p_T, y) bin is assigned as the largest difference in the obtained prompt signal yield with respect to the nominal fit. The relative uncertainty introduced by this can reach a few percent, though tends to be much smaller than the statistical uncertainty in each bin.

5.6.4 Monte Carlo modelling

If the simulated samples do not model the data correctly, the selection efficiencies that are evaluated using these simulated samples could be incorrect. Small deviations of the MC from the distributions in data can lead to small deviations from the true efficiencies, and these deviations need to be evaluated and assigned as systematic uncertainties. As a first step, distributions of selection variables are compared between data and simulation. Obviously, this can only be performed using fully selected candidates, especially in the measurements presented here where most of the selection is integrated into the trigger. As before, the prompt signal component is extracted from data using signal weights ${}_s\mathcal{W}$ [108]. To account for the effect of the PID efficiencies, the data are corrected by using the inverse of the PID efficiency as an additional weight similar to Equation (5.18). In a similar manner, the simulated sample is corrected using the tracking correction factors. In both cases, only the central value is used as a correction factor and all involved uncertainties are ignored. Furthermore, it is paramount to point out that the *sPlot* technique, which allows the signal distribution to be obtained in some variable, can only be applied if this variable is independent (not linearly correlated) to the discriminating variable used to extract the weights ($\ln \chi^2_{\text{IP}}$ in this analysis).

Examples are given in Figures 5.20 and 5.21 for $D^0 \rightarrow K^- \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$ at 5 TeV. Overall, good agreement is seen when comparing the simulated distributions to data.

To assign an additional uncertainty to encapsulate potential mis-modelling in the simulated sample, the following procedure is performed for each selection variable. The procedure is not performed for selection variables that are practically 100% efficient for signal events. As described above, no data with a looser selection are available and thus, the effect of mis-modelling is studied using a tighter selection, assuming the agreement between data and simulation is similar for events being accepted and those being rejected by the selection.

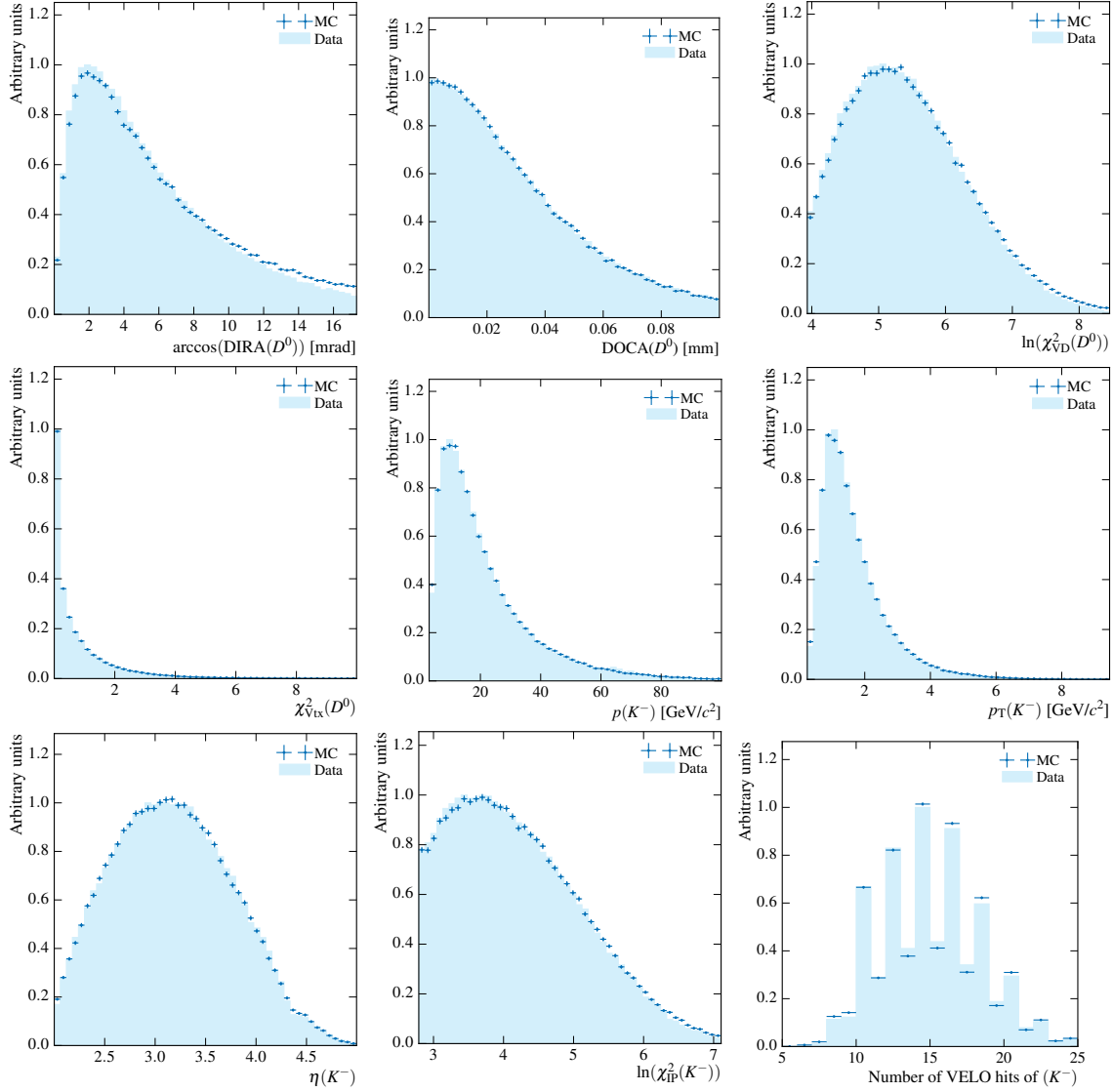


Fig. 5.20: Comparison of several quantities used in the event selection between simulated and real data for $D^0 \rightarrow K^- \pi^+$ at 5 TeV. Tracking correction factors are applied to the simulated sample and the real data is corrected for the PID efficiency. To reflect the double differential nature of the measurement, simulated events are reweighted using p_T and y in the (p_T, y) binning scheme used for the measurement.

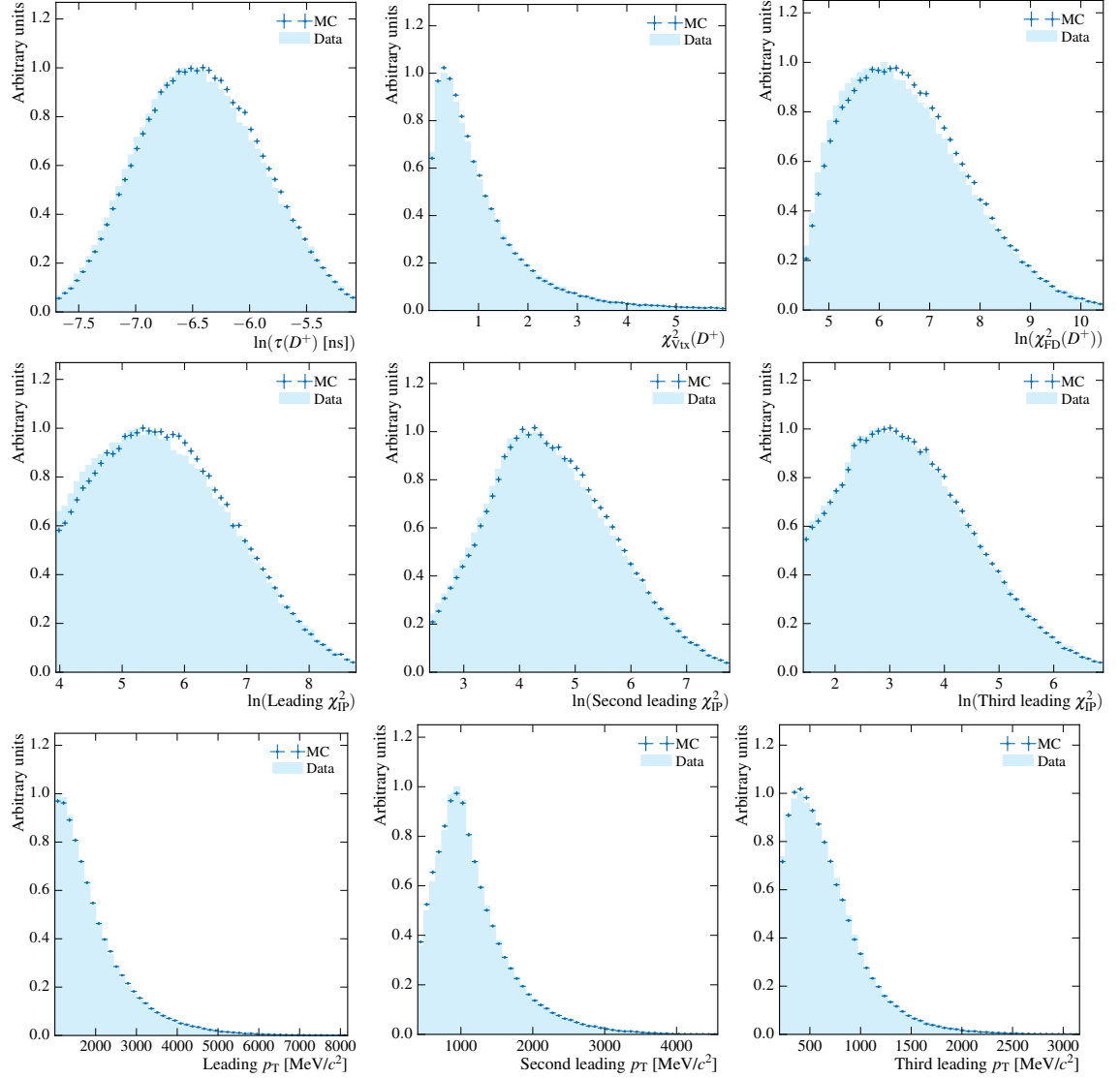


Fig. 5.21: Comparison of several quantities used in the event selection between simulated and real data for $D^+ \rightarrow K^- \pi^+ \pi^+$ at 5 TeV. Tracking correction factors are applied to the simulated sample and the real data is corrected for the PID efficiency. To reflect the double differential nature of the measurement, simulated events are reweighted using p_T and y in the (p_T, y) binning scheme used for the measurement.

1. For a given selection variable, find the cut value that rejects half of the fully selected simulated data. When defining this cut, the tracking correction factors applied to the simulated sample are taken into account. Depending on the size of the sample, the efficiency of this cut, ε_{MC} , is close to 50%.
2. Apply the cut value to the *real data* and count the number of signal candidates. The number of signal events is given by the sum over the sWeights after correcting for the PID selection, which is not applied in simulation:

$$N_{\text{Passed}} = \sum_{\text{passed}}^i \frac{s \mathcal{W}_i}{\epsilon_{\text{PID},i}}. \quad (5.22)$$

3. The efficiency-corrected signal yield corresponding to the requirement is $N_{\text{Passed}}/\varepsilon_{\text{MC}}$. It should be equal to the number of signal events seen in the entire sample, if the MC simulation describes the data. The quantity:

$$\Delta = \frac{N_{\text{Passed}}}{\varepsilon_{\text{MC}}} - \sum_{\text{all}}^i \frac{w_{i,\text{signal}}}{\epsilon_{\text{PID},i}}, \quad (5.23)$$

then allows the agreement between simulated and real data to be judged, as it should be zero.

The procedure is performed in every (p_{T}, y) bin for every selection variable separately, resulting in one Δ value per bin and variable. To account for the statistical uncertainty due to the amount of available data and to include correlations between selection variables, the data and simulated samples are bootstrapped and the procedure above is repeated ten thousand times using those bootstrapped samples. This ensures that the correlations between variables are correctly taken into account when combining the deviations per variable into a single number. For example, if one final state particle passes a p_{T} requirement, the others are more likely to do so as well due to correlations between their transverse momenta, which needs to be taken into account in the combination and is ensured by the bootstrapping. Furthermore, each bootstrapped sample uses modified values for the tracking corrections and PID efficiencies randomly creating new tables as described before. Therefore, the spread of the ten thousand obtained values for Δ includes contributions from the amount of simulated and real data as well as the available calibration statistics and the uncertainties in the tracking correction tables. An overall Δ is then obtained by summing over the individual Δ values per variable, yielding one value per (p_{T}, y) bin and bootstrap toy.

Due to the limited amount of available data and simulated events, most Δ values per bin are not significantly different from zero. To search for an overall trend, assuming mis-modelling would have a similar effect in all (p_{T}, y) bins, the sum of

the individual bin Δ 's is computed. Due to the bootstrapping, ten thousand values for this sum are obtained whose distribution resembles a normal distribution with mean μ and width σ . If $|\mu| < \sigma$, no uncertainty is assigned to the mis-modelling in simulation. If $|\mu| - \sigma > 0$, a relative uncertainty of $(|\mu| - \sigma)/N_{\text{total}}$ is assigned, where N_{total} is the total number of signal events computed in Equation (5.23). Subtracting the width of the observed distribution is deliberately done to reduce the impact of double-counting the uncertainties making up said width as they are already included as systematic uncertainties on the cross-sections. Applying the procedure, about 1% of uncertainty is observed.

5.6.5 Particle identification

Two sources of systematic uncertainty are considered for the estimation of the PID efficiencies, which was described in Section 5.5.2. The first is due to the size of the calibration sample itself, leading to statistical uncertainties on the PID efficiencies in each bin of the $(p, \eta, N_{\text{SPD}})$ look-up table. Their propagation, in essence identical to the same uncertainty on the tracking correction factor, is described in Section 5.5.2.

A second source is given by the chosen binning scheme which can be incapable of properly removing the dependence on the calibration sample kinematics from the $(p, \eta, N_{\text{SPD}})$ PID efficiency look-up tables. The general cause for such a behaviour can be easily illustrated: let x be an observable on which the efficiency $\varepsilon(x)$ depends. Let $R(x)$ and $C(x)$ be the probability density function of x in the reference and calibration sample respectively and let x_i and x_{i+1} be the chosen boundaries of the i -th bin in the efficiency table. When computing the efficiency for this bin $\varepsilon_i = C'_i/C_i$, one provides an estimate whose expectation value is given by:

$$\langle \varepsilon_i \rangle_C = \frac{\int_{x_i}^{x_{i+1}} \varepsilon(x) \cdot C(x) dx}{\int_{x_i}^{x_{i+1}} C(x) dx},$$

where the subscript C indicates that the expectation is with respect to $C(x)$. The same expectation value for the reference sample is hence given by:

$$\langle \varepsilon_i \rangle_R = \frac{\int_{x_i}^{x_{i+1}} \varepsilon(x) \cdot R(x) dx}{\int_{x_i}^{x_{i+1}} R(x) dx}.$$

When applying the values in the efficiency tables to the reference sample to obtain ε_{PID} , the obtained ε_i must be an estimate of $\langle \varepsilon_i \rangle_R$ in order to yield an unbiased estimate of ε_{PID} :

$$\frac{\sum \langle \varepsilon_i \rangle_R R_i}{R} = \frac{\sum \frac{\int_{x_i}^{x_{i+1}} \varepsilon(x) R(x) dx}{\int_{x_i}^{x_{i+1}} R(x) dx} \cdot R \cdot \frac{\int_{x_i}^{x_{i+1}} R(x) dx}{\int_X R(x) dx}}{R}$$

$$= \frac{\sum \int_{x_i}^{x_{i+1}} \varepsilon(x) R(x) dx}{\int_X R(x) dx} = \frac{\int_X \varepsilon(x) R(x) dx}{\int_X R(x) dx} = \langle \varepsilon_{\text{PID}} \rangle .$$

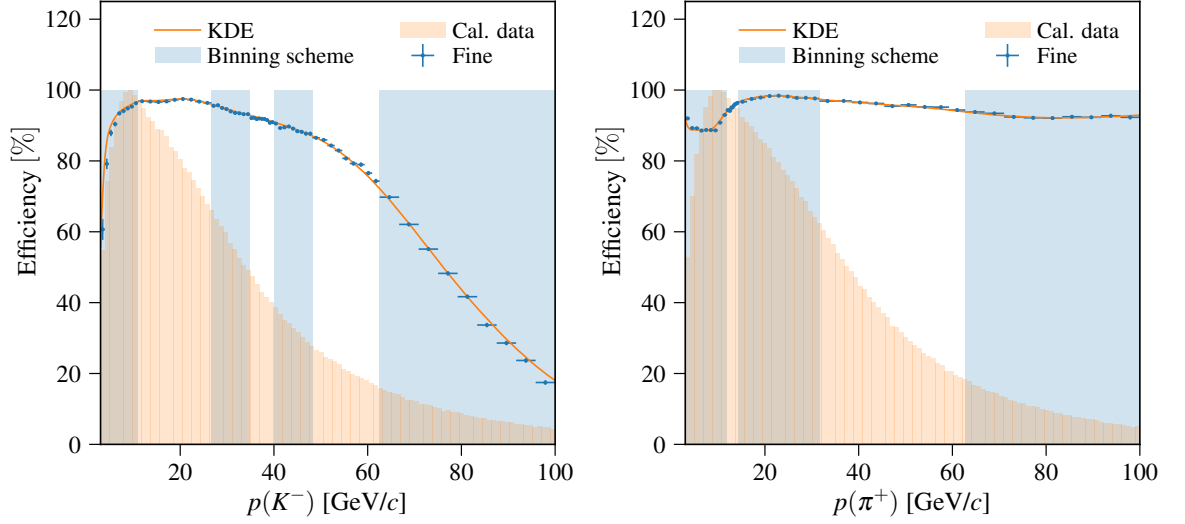
However, the expectation value of ε_i is $\langle \varepsilon_i \rangle_C$, which in general is not identical to $\langle \varepsilon_i \rangle_R$, thus introducing a bias to ε_{PID} . The magnitude of the bias (and thus the systematic uncertainty) significantly depends on the chosen binning scheme's approximation of the assumption $\langle \varepsilon_i \rangle_R = \langle \varepsilon_i \rangle_C$.

To estimate the introduced bias, knowledge of both the reference sample distributions as well as the efficiency $\varepsilon(p, \eta, N_{\text{SPD}})$ (as a function, not as the binned tables) is necessary. While the former is given by the actual data sample using ${}_s\mathcal{W}$ to extract the prompt signal contribution, the latter, the efficiency as an actual function of p , η and N_{SPD} is not known. To estimate a systematic uncertainty in this analysis, different approaches are used in the 5 TeV and 13 TeV measurements, based on newly available techniques.

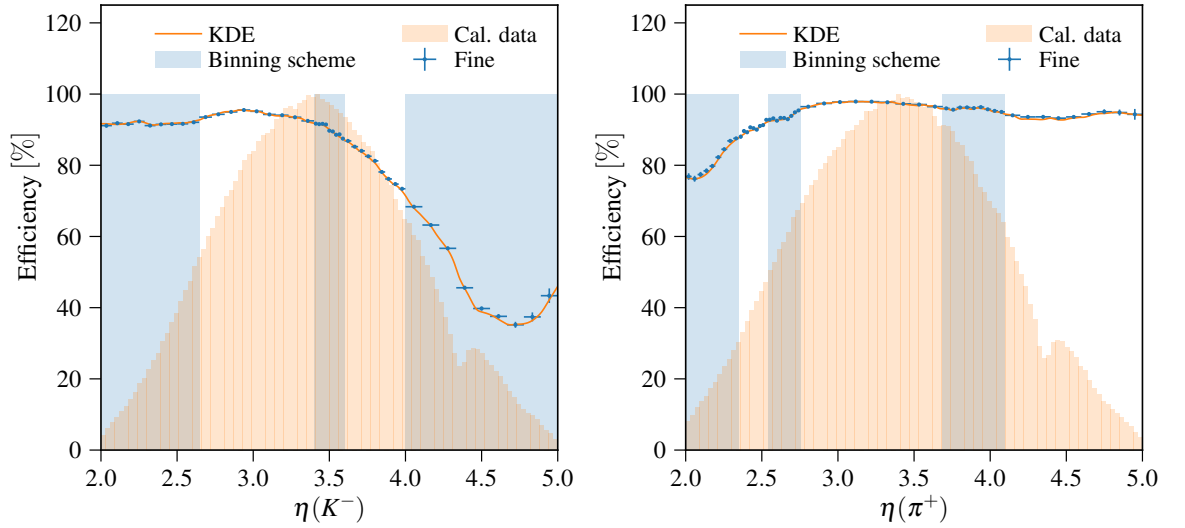
In the 13 TeV measurement, based on the dependence of the one-dimensional efficiencies on the dependent variables as shown in Figures 5.17 and 5.18, effects due to the binning in N_{SPD} are assumed to be negligible and an approximation of the efficiency is derived as a function of p and η . Similar to the derivation of the shapes describing the combinatorial background in the $\ln \chi_{\text{IP}}^2$ fit, two-dimensional kernel density estimates are derived to describe the kinematics in the calibration samples separately for kaons and pions. The signal weights in the calibration sample are used to build the KDEs for the signal component. For each species, two KDEs are constructed, one before and one after applying the PID cut on the track in the calibration sample. Using both KDEs, the efficiency can be computed for a (p, η) bin with any bin boundaries:

$$\varepsilon_i = \frac{C'}{C} \cdot \frac{\int_{\text{bin}} \text{KDE}'(p, \eta) dp d\eta}{\int_{\text{bin}} \text{KDE}(p, \eta) dp d\eta} , \quad (5.24)$$

where the kernel density estimates KDE and KDE' are normalised to unity. Figure 5.22 shows a projection of the PID efficiencies in the calibration sample on to p and η , overlayed with the projection of the KDE model. It is obtained using Equation (5.24) by splitting the range of the plotted variable into 500 equally sized bins, each stretching the full range of the variable that is not plotted. A good agreement between the actual calibration data and the KDE estimate is seen. To derive an uncertainty based on the cross-sections based on this model, two-dimensional calibration tables are created with a progressively finer binning in (p, η) , starting from the nominal binning scheme for these two variables. Figures 5.23 and 5.24 give an example of these two-dimensional calibration tables with progressively finer binning. The calibration procedure (see Equation (5.18)) is repeated multiple times while using increasingly finer binned calibration tables for pions and kaons. The obtained

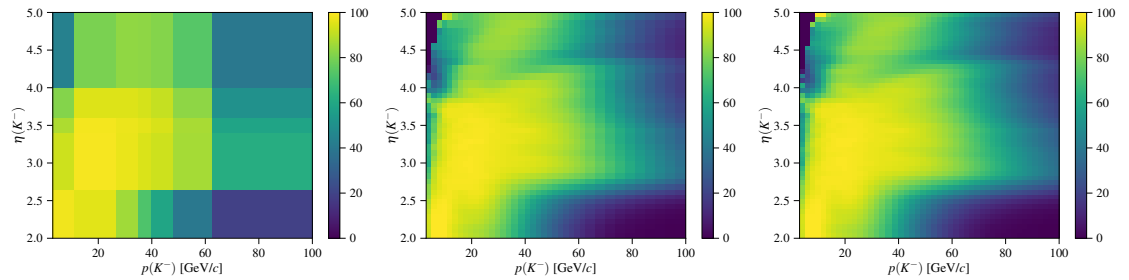


(a) p dependence $DLL_{K-\pi} > 5$ for kaons (left) and $DLL_{K-\pi} < 3$ for pions (right).



(b) η dependence $DLL_{K-\pi} > 5$ for kaons (left) and $DLL_{K-\pi} < 3$ for pions (right).

Fig. 5.22: One dimensional PID efficiency for 13 TeV comparing the data and the projections from a two-dimensional kernel density estimate as described in the text. The variable not used in the projection is integrated over its full range.



(a) 35 bins (nominal)

(b) 1221 bins

(c) 2256 bins

Fig. 5.23: Two-dimensional toy calibration tables for the kaon, generated from KDE description of the true calibration sample with increasing granularity.

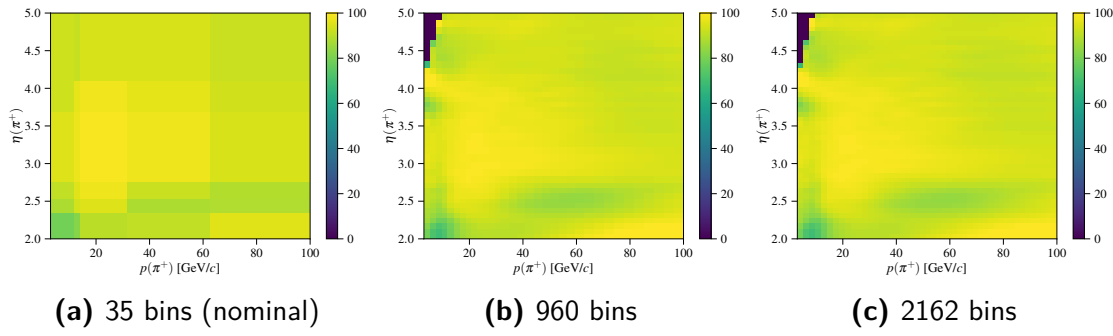


Fig. 5.24: Two-dimensional toy calibration tables for the pion, generated from KDE description of the true calibration sample with increasing granularity.

efficiency ε_{PID} as a function of the number of bins in the calibration tables shows a convergence towards a higher number of bins and it is the difference between the efficiency at convergence and the one obtained using the nominal tables (estimated using the KDE approach to avoid biases in the difference due to the missing N_{SPD} dimension) that is taken as the systematic uncertainty. This difference is computed in each (p_T, y) bin separately. The relative uncertainty obtained is lowest in the central region of meson kinematics, increasing towards higher values in meson p_T with values ranging from as little as 1.2% to well above 20% depending on the meson. Large uncertainties usually coincide with bins with sizeable statistical uncertainty, which are more prone to significantly change with altered binning schemes applied to the calibration sample.

For the 5 TeV result, a simplified approach was chosen based on the idea that an efficiency is nothing but the weight transforming a distribution before a certain requirement is imposed to the one after the requirement is applied, and therefore equating the problem of finding an efficiency description to a reweighting problem. Recent developments in machine-learning-based reweighting techniques [111, 112] allow for efficient reweighting of higher-dimensional problems compared to obtaining the weights from the ratio of histograms. Therefore, an alternative efficiency description is derived with a reweighting algorithm using boosted decision trees (BDT) [112], which is trained using the calibration data. The trained reweighter then allows to obtain an efficiency for any $(p, \eta, N_{\text{SPD}})$ point and thus acts like an approximation of the function $\varepsilon(p, \eta, N_{\text{SPD}})$. To train the reweighter on the signal in the calibration sample alone, signal weights ${}_s\mathcal{W}$ are used, which have been extracted from a fit to the total calibration sample, similar to those in the 13 TeV sample. Figure 5.25 compares the efficiency description from the BDT to that obtained directly from the calibration data and reveals a good agreement between the two. An uncertainty on the measured cross-sections is then obtained in a similar way to the KDE based approach for 13 TeV where the difference between the nominal efficiencies and those obtained using the BDT based approach is assigned as uncertainty. The obtained relative uncertainties show the same behaviour and ranges as those for the 13 TeV

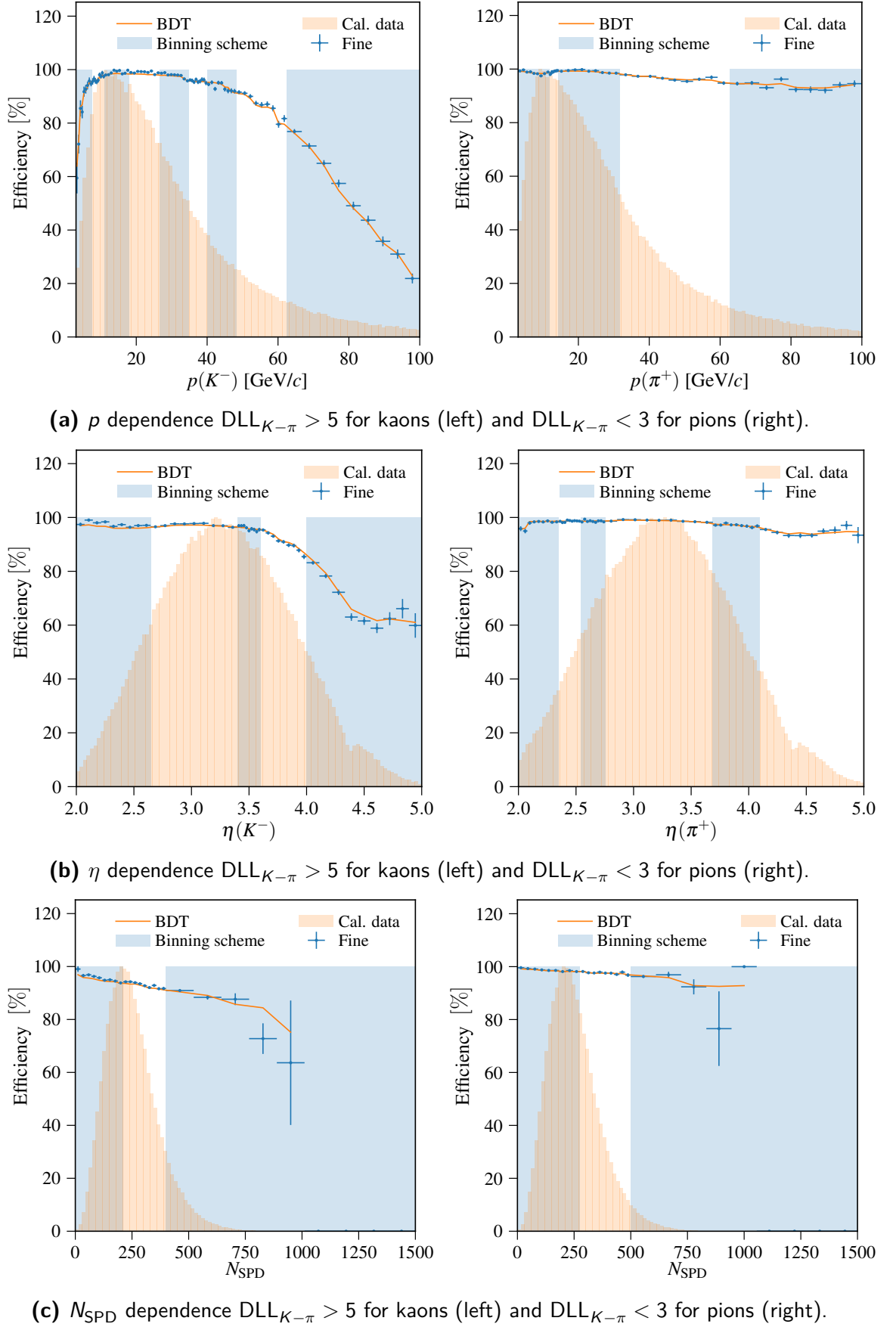


Fig. 5.25: One dimensional PID efficiency for 5 TeV comparing the data and the projections averages in each bin of the efficiencies obtained from the BDT reweighter. For stability, all plots have been created using the fiducial range of $2 \leq \eta < 5$ and $3 \leq p < 100$ GeV/c for the track kinematics.

measurement. However, due to the additional bins in p , their total effect on an integrated cross-section obtained as the sum over all bins is only about 1%, compared to 2–3% for 13 TeV. As a cross-check, a BDT based efficiency description has also been obtained for the 13 TeV measurement and the resulting uncertainties on the cross-section are consistent with the ones quoted based on the KDE approach.

At this point, an obvious question might arise, namely what the reason is for not using these two techniques to derive the nominal efficiencies and avoiding the uncertainty due to the binning systematic. This decision was made due to practical considerations especially concerning the treatment of statistical uncertainties caused by the finite size of the calibration sample. While the propagation in the nominal approach using binned PID efficiency tables, is rather straightforward, this is not true anymore in the case of unbinned descriptions such as the KDE and BDT reweighter approach. Probably the only viable approach would require the creation of bootstrap samples from the calibration sample. In each of these samples, a fit is performed to obtain the signal weights ${}_s\mathcal{W}$ and the KDE or BDT procedure is repeated, yielding one KDE description or trained BDT classifier per bootstrap sample. Unfortunately, the creation of the non-parametric descriptions is time-consuming and persisting them to ensure a correct treatment of correlations when applying them to different charm mesons requires large amounts of disk space. With this in mind, using the simplified binned tables is deemed sufficient in this measurement.

5.6.6 Summary

All sources of systematic uncertainty are summarised in Tables 5.9 and 5.10 where the individual ranges are given. As the quoted ranges for the different sources of systematic uncertainty and different charm mesons can be rather sizeable, often ranging from values negligibly small to $\mathcal{O}(10\%)$, the resulting uncertainty on the sum of (p_T, y) bins in the range $0 \leq p_T < 8 \text{ GeV}/c$ and $2 \leq y < 4.5$ is also given to allow a clearer judgement on the dominant sources of systematic uncertainty on the measured cross-sections. The top two contenders for the most dominant systematic uncertainty are the uncertainty on the measured luminosity, and the uncertainty involved in the tracking correction procedure, both trading places depending on the considered meson. Depending on the considered (p_T, y) bin, the uncertainties assigned to the PID calibration procedure due to the chosen binning scheme of the PID calibration tables and the uncertainty due to the limited size of the simulated sample give contributions significantly greater than those dominating the integrated cross-sections. Overall, the individual measurements presented in the next section are completely dominated by the systematic uncertainty with only marginal contributions due to the statistical uncertainty, the uncertainty assigned to the prompt signal yields extracted from the $\ln \chi^2_{\text{IP}}$ fit described in Section 5.4.3.

Tab. 5.9: Fractional systematic uncertainties for 5 TeV, in percent and rounded to one significant digit. Uncertainties that are computed bin-by-bin are expressed as ranges giving the minimum to maximum values. As these can vary significantly between different bins, the uncertainties on the integrated over the range $0 \leq p_T < 8 \text{ GeV}/c$ and $2 \leq y < 4.5$ is given in round braces.

Source	Uncertainties (%)			
	D^0	D^+	D_s^+	D^{*+}
Luminosity	3.8			
$\mathcal{B}(H_c \rightarrow f)$	1.2	2.1	5.8	1.5
MC/truth	0.043	0.085	0.15	0.52
MC modelling	0.28	0.65	0.56	2.0
Tracking	3.0–4.9 (3)	4.6–7.4 (6)	4.3–7.1 (5)	4.8–7.2 (6)
MC sample size	1–11 (0.3)	1–12 (1)	1.9–9.4 (1)	1.1–9.7 (0.9)
PID sample size	0.1–1.4 (0.2)	0.1–1.2 (0.3)	0.1–1.7 (0.5)	0.1–1.6 (0.1)
PID binning	0–25 (0.7)	0–10 (1)	0–17 (2)	0–20 (0.6)
Fit model shapes	0.0–2.7 (0.2)	0.1–3.3 (0.2)	0.1–2.8 (0.2)	0.0–1.0 (0.1)
Signal window	0.2–1.3(0.06)	0.1–1.2(0.09)	0.2–2.5(0.07)	0.2–1.7(0.06)
Stat.	1–14 (0.2)	1–10 (1.0)	3–20 (2)	2–17 (1)

Tab. 5.10: Fractional systematic uncertainties for 13 TeV, in percent and rounded to one significant digit. Uncertainties that are computed bin-by-bin are expressed as ranges giving the minimum to maximum values. As these can vary significantly between different bins, the uncertainties on the integrated over the range $0 \leq p_T < 8 \text{ GeV}/c$ and $2 \leq y < 4.5$ is given in round braces.

Source	Uncertainties (%)			
	D^0	D^+	D_s^+	D^{*+}
Luminosity	3.9			
$\mathcal{B}(H_c \rightarrow f)$	1.2	2.1	5.8	1.5
MC/truth	0.037	0.057	0.11	0.099
MC modelling	1.0	1.2	0.20	0.91
Tracking	2.8–10.0 (4)	4–14 (8)	4–14 (7)	5–11 (7)
MC sample size	1–26 (0.4)	1–55 (2)	1–20 (2)	1–23 (2)
PID sample size	0.1–2.0(0.09)	0.1–1.2 (0.2)	0.1–2.7 (0.3)	0.1–1.1(0.09)
PID binning	0–45 (2)	0–11 (3)	0–18 (4)	0–14 (0.6)
Fit model shapes	0.0–6.3(0.05)	0.0–6.9(0.08)	0.0–1.4(0.05)	0–46 (0.8)
Signal window	0.0–2.0(0.05)	0.0–2.2(0.04)	0.1–1.2(0.03)	0.2–6.8 (0.3)
Stat.	0–17 (0.09)	0–10 (0.5)	1–17 (0.8)	1–18 (0.8)

5.7 Results

Results are presented for the four different meson species and at the two studied centre-of-mass energies. The main result is given by the double-differential cross-sections, shown in Figures 5.26 to 5.29, which are compared to up to three sets of theoretical predictions described in Section 2.2.4. Overall, the measured cross-sections are in agreement with the different sets of predictions albeit tending to lie at the upper edge of the theoretically predicted ranges. Generally, the best agreement is seen for the GMVFNS predictions; however, these predictions are only available for $p_T > 3 \text{ GeV}/c$.

Based on the double-differential cross-section, additional results are derived, most notably the ratios of measured cross-sections between different centre-of-mass energies and integrated cross-section which are described in the following.

5.7.1 Ratios between different $\sqrt{s}$

Based on the double-differential cross-sections, ratios between different centre-of-mass energies are computed. The ratio $R_{13/5}$ between the measurements at 13 TeV and 5 TeV solely relies on the cross-section measurements presented in this thesis while the ratio of 13 TeV to 7 TeV uses previous measurements of the production cross-sections by LHCb, which are discussed in Ref. [27]. Ratios are measured for the kinematic range of $0 \leq p_T < 8 \text{ GeV}/c$ of the 7 TeV measurement using bins with a width of $1 \text{ GeV}/c$. To accommodate this, the finer bins used in the 13 TeV measurement are combined by computing the average differential cross-section of the bins covering a wider $1 \text{ GeV}/c$ bin. The obtained ratios are presented in Figures 5.30 to 5.33. These cross-section ratios benefit from heavily reduced uncertainties due to cancellation of both experimental uncertainties affecting the measurement as well as theoretical uncertainties. For the measured values, this is mostly caused by the correlated luminosity and branching ratios used in the measurements as well as a cancellation of uncertainties associated to the material budget in the tracking efficiency corrections. Likewise, uncertainties due to the charm quark mass and the QCD scale are heavily reduced in these ratios, resulting in a significantly more powerful test of the theoretical predictions.

Overall, the measured values are found to be in good agreement with the predictions for all meson species, indicating a good modelling of the cross-section scaling with increasing centre-of-mass energy. Furthermore, the measured ratios for the D^0 and D^+ mesons shown in Figures 5.30 and 5.31 have uncertainties that are still much smaller than the predictions and thus provide valuable input to constrain and reduce theoretical uncertainties.

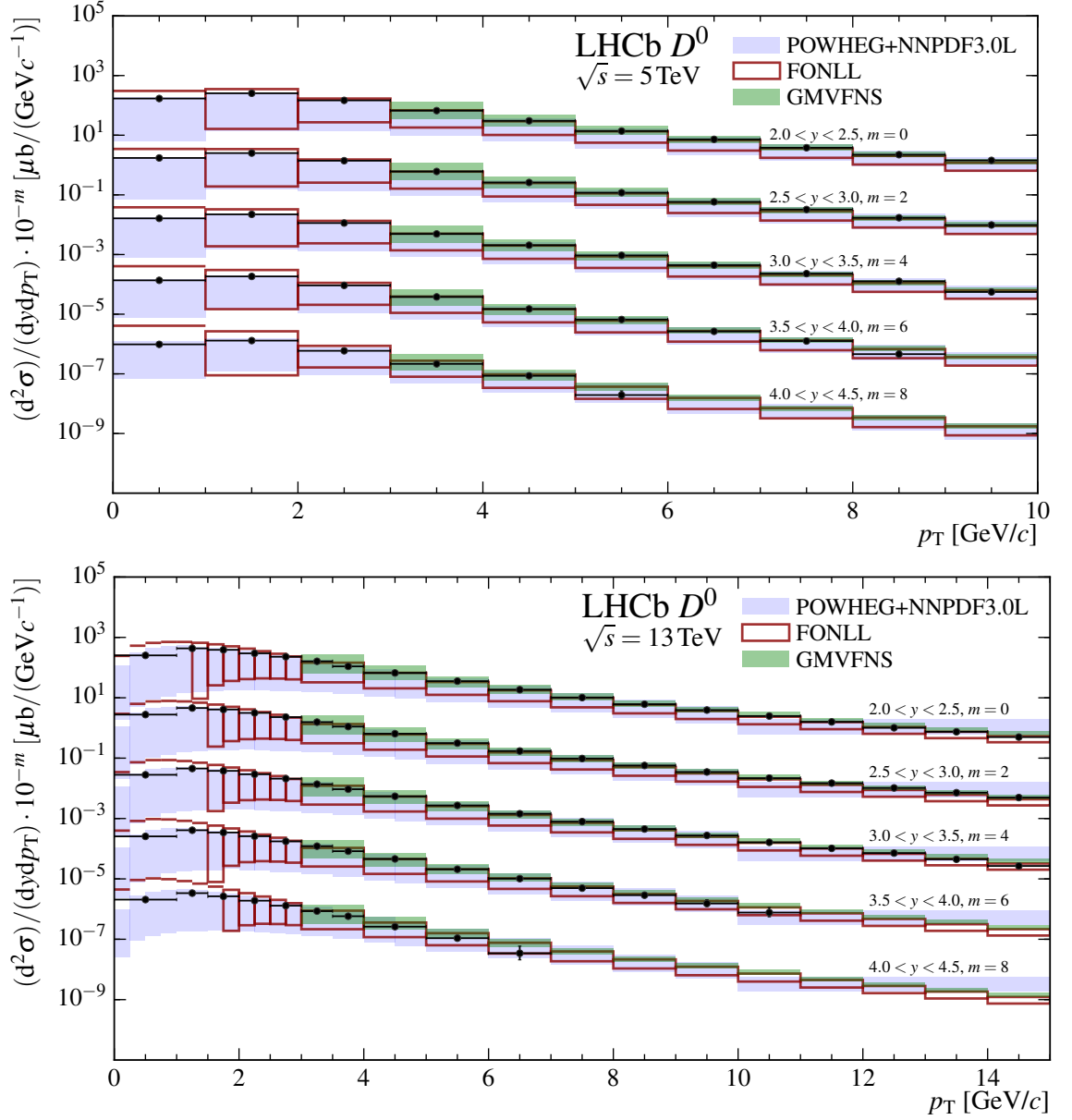


Fig. 5.26: Double-differential cross-sections at 5 TeV (top) and 13 TeV (bottom) for D^0 . The theoretical predictions shown have been provided by the respective theorist groups and are described in Section 2.2.4.

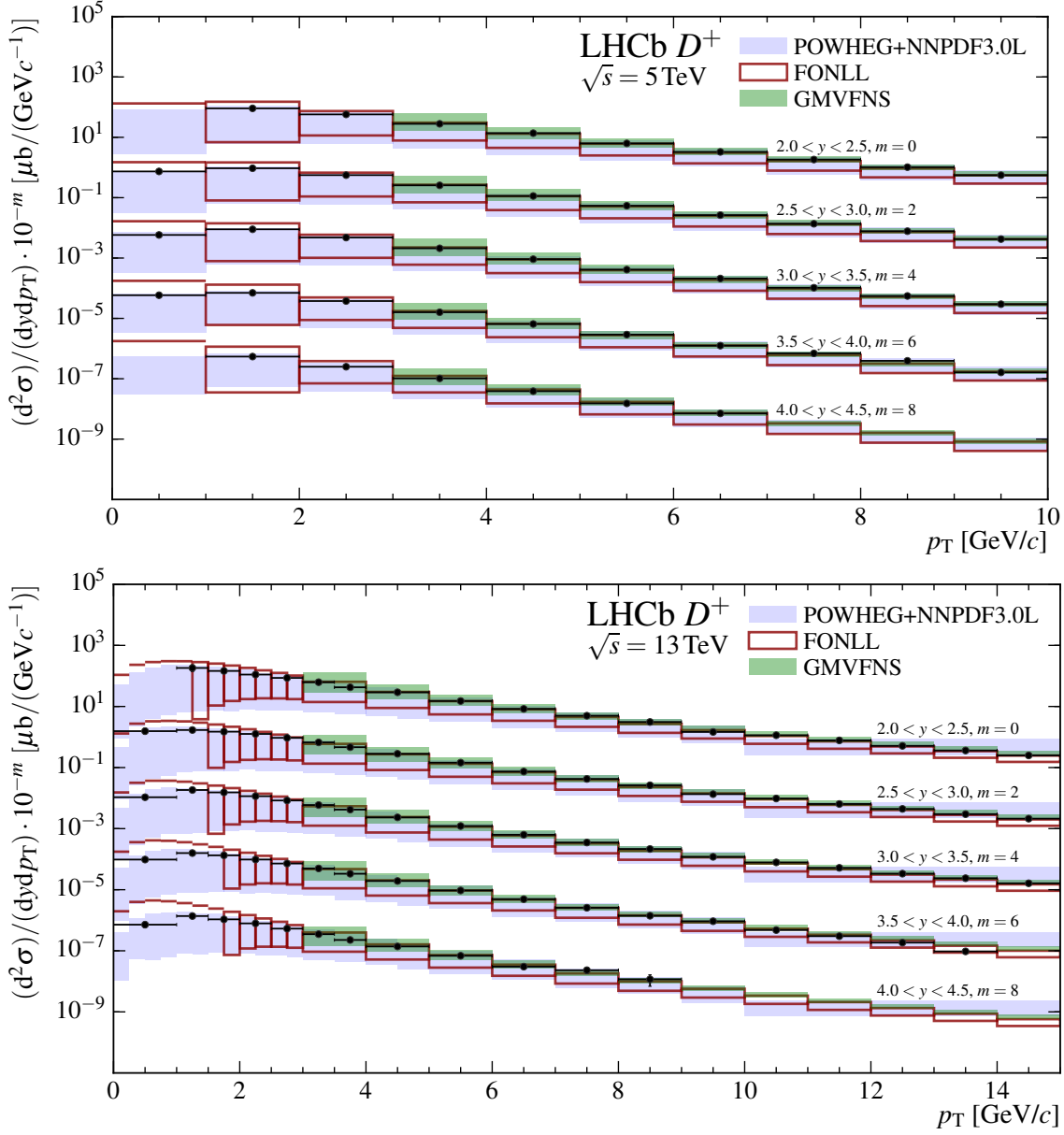


Fig. 5.27: Double-differential cross-sections at 5 TeV (top) and 13 TeV (bottom) for D^+ . The theoretical predictions shown have been provided by the respective theorist groups and are described in Section 2.2.4.

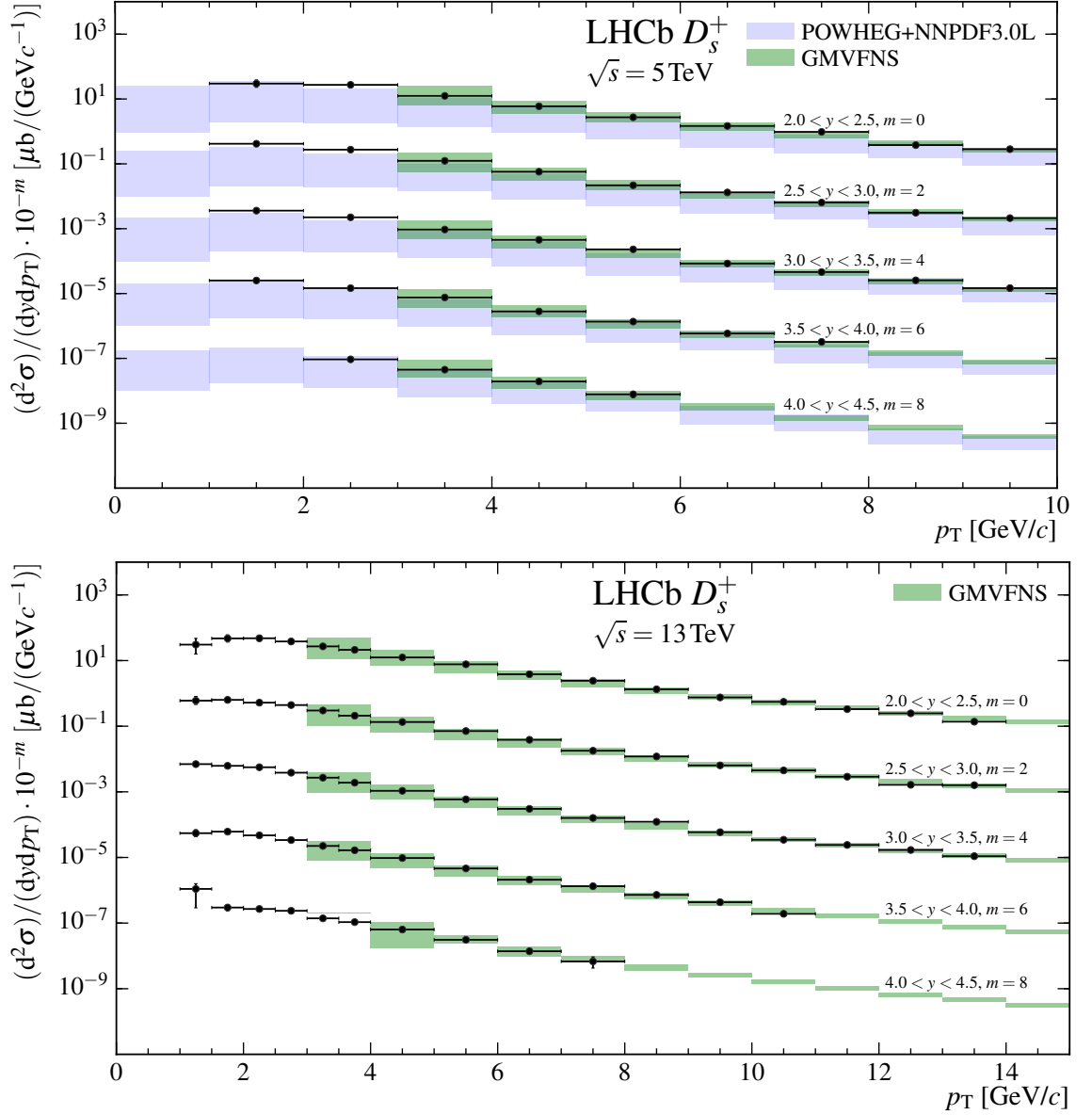


Fig. 5.28: Double-differential cross-sections at 5 TeV (top) and 13 TeV (bottom) for D_s^+ . The theoretical predictions shown have been provided by the respective theorist groups and are described in Section 2.2.4.

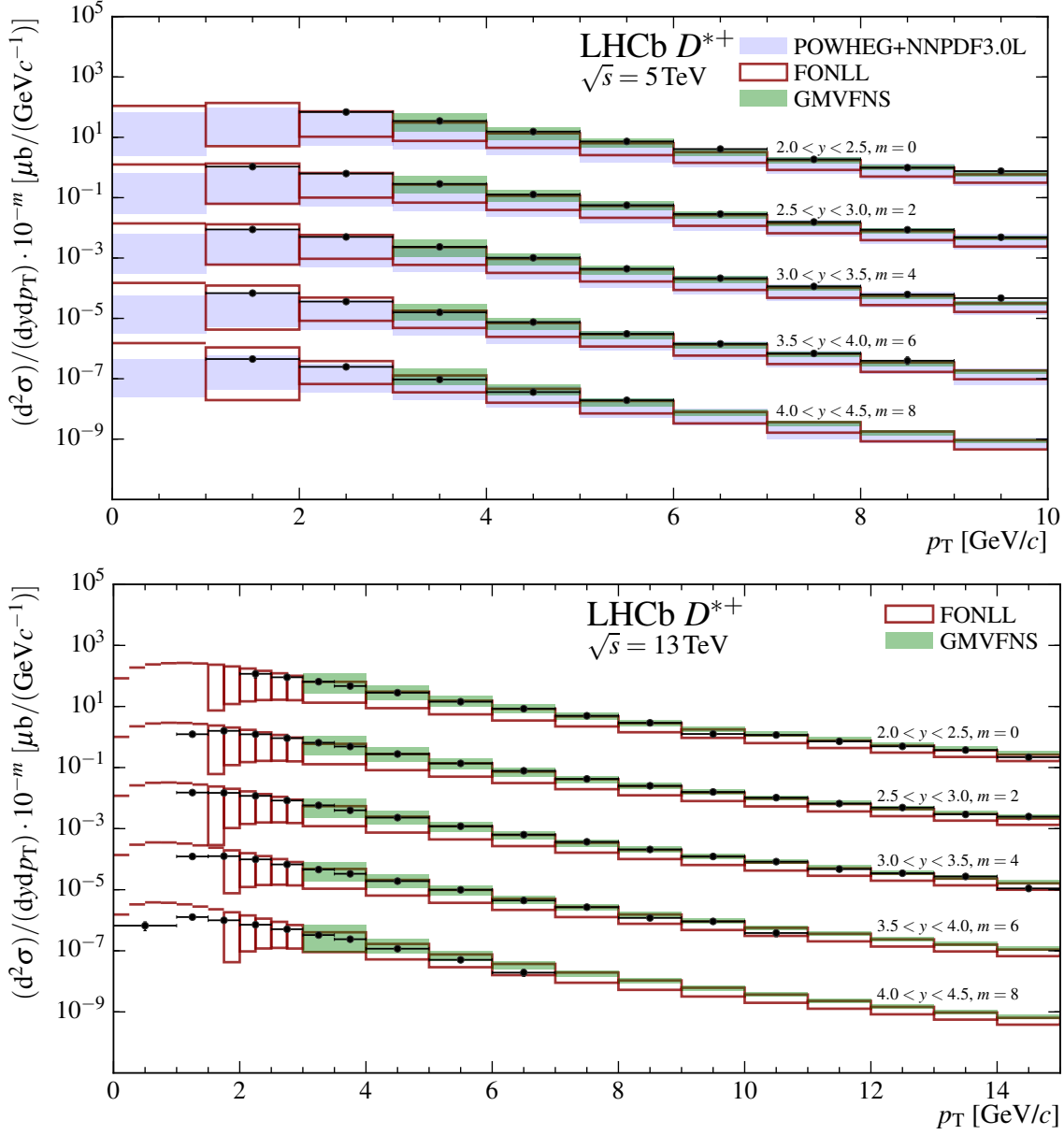


Fig. 5.29: Double-differential cross-sections at 5 TeV (top) and 13 TeV (bottom) for D^{*+} . The theoretical predictions shown have been provided by the respective theorist groups and are described in Section 2.2.4.

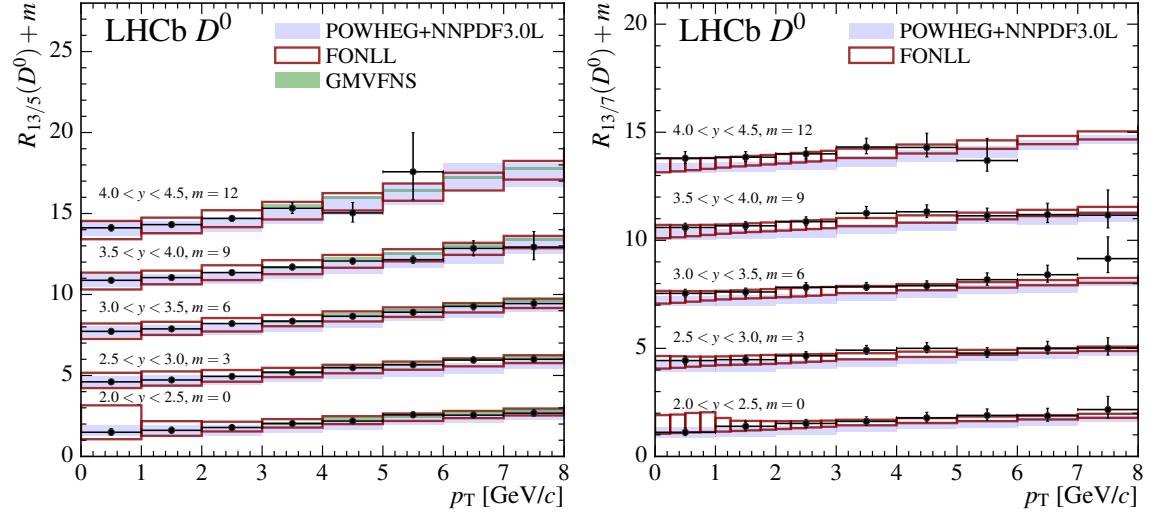


Fig. 5.30: Cross-section ratio $R_{13/5}$ between 13 TeV and 5 TeV (left) and $R_{13/7}$ between 13 TeV and 7 TeV (right) for D^0 . The theoretical predictions shown have been provided by the respective theorist groups and are described in Section 2.2.4.

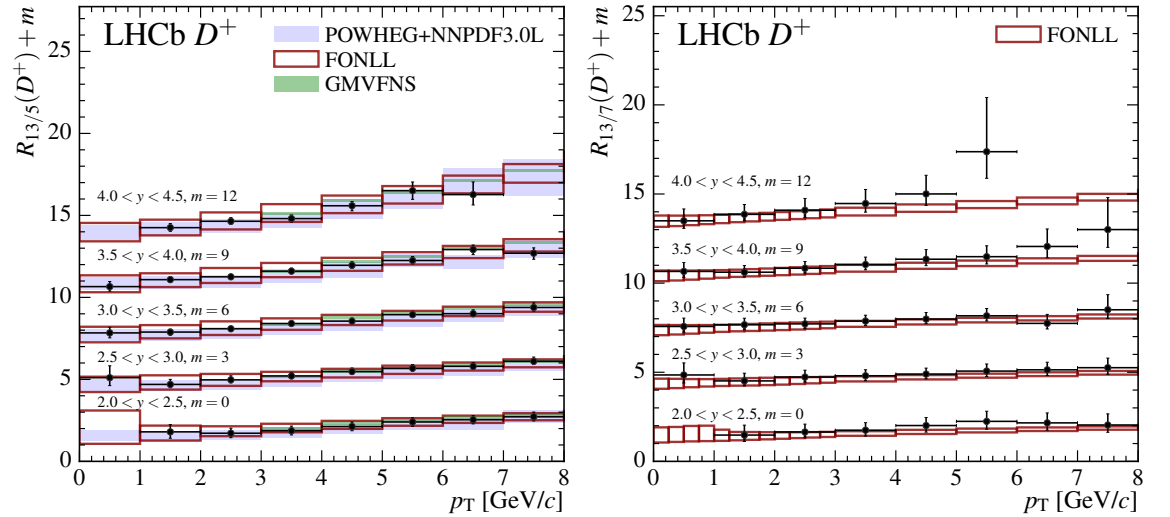


Fig. 5.31: Cross-section ratio $R_{13/5}$ between 13 TeV and 5 TeV (left) and $R_{13/7}$ between 13 TeV and 7 TeV (right) for D^+ . The theoretical predictions shown have been provided by the respective theorist groups and are described in Section 2.2.4.

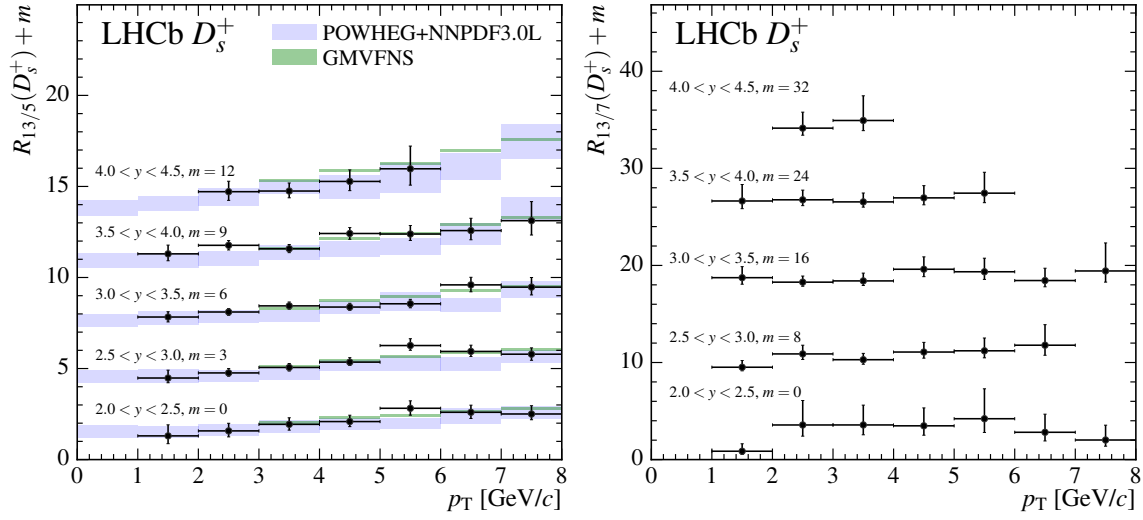


Fig. 5.32: Cross-section ratio $R_{13/5}$ between 13 TeV and 5 TeV (left) and $R_{13/7}$ between 13 TeV and 7 TeV (right) for D_s^+ . The theoretical predictions shown have been provided by the respective theorist groups and are described in Section 2.2.4.

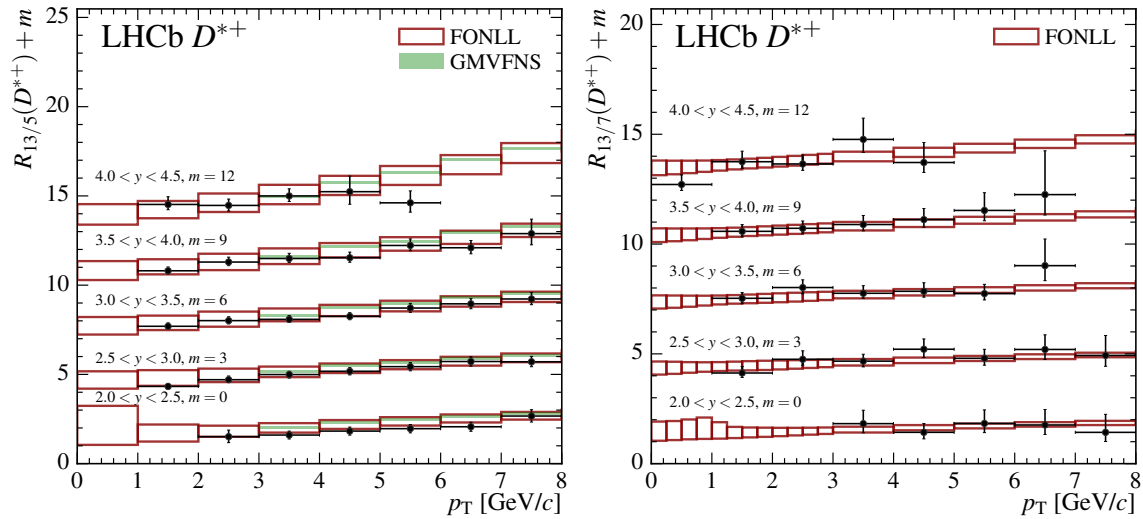


Fig. 5.33: Cross-section ratio $R_{13/5}$ between 13 TeV and 5 TeV (left) and $R_{13/7}$ between 13 TeV and 7 TeV (right) for D^{*+} . The theoretical predictions shown have been provided by the respective theorist groups and are described in Section 2.2.4.

Tab. 5.11: Prompt D -meson production cross-sections at $\sqrt{s} = 5$ TeV in the kinematic ranges given. The first uncertainty on the cross-section is statistical, the second is systematic and includes the contribution from the extrapolation factor.

			Extrapolation factor	Cross-section (μb)		
D^0	$0 < p_T < 8 \text{ GeV}/c$	$2 < y < 4.5$	1.0013 ± 0.0019	1374 ± 3	± 74	
D^+	$0 < p_T < 8 \text{ GeV}/c$	$2 < y < 4.5$	1.108 ± 0.049	551 ± 5	± 48	
D^0	$1 < p_T < 8 \text{ GeV}/c$	$2 < y < 4.5$	1.0017 ± 0.0020	1004 ± 3	± 54	
D^+	$1 < p_T < 8 \text{ GeV}/c$	$2 < y < 4.5$	1.00062 ± 0.00099	402 ± 2	± 30	
D_s^+	$1 < p_T < 8 \text{ GeV}/c$	$2 < y < 4.5$	1.0734 ± 0.0080	170 ± 4	± 16	
D^{*+}	$1 < p_T < 8 \text{ GeV}/c$	$2 < y < 4.5$	1.122 ± 0.046	421 ± 5	± 36	

Tab. 5.12: Prompt D -meson production cross-sections at $\sqrt{s} = 13$ TeV in the kinematic ranges given. The first uncertainty on the cross-section is statistical, the second is systematic and includes the contribution from the extrapolation factor.

			Extrapolation factor	Cross-section (μb)		
D^0	$0 < p_T < 8 \text{ GeV}/c$	$2 < y < 4.5$	1.0014 ± 0.0024	2709 ± 2	± 165	
D^+	$0 < p_T < 8 \text{ GeV}/c$	$2 < y < 4.5$	1.049 ± 0.031	1102 ± 5	± 111	
D^0	$1 < p_T < 8 \text{ GeV}/c$	$2 < y < 4.5$	1.0018 ± 0.0025	2072 ± 2	± 124	
D^+	$1 < p_T < 8 \text{ GeV}/c$	$2 < y < 4.5$	—	834 ± 2	± 78	
D_s^+	$1 < p_T < 8 \text{ GeV}/c$	$2 < y < 4.5$	—	353 ± 9	± 76	
D^{*+}	$1 < p_T < 8 \text{ GeV}/c$	$2 < y < 4.5$	1.102 ± 0.081	784 ± 4	± 87	

5.7.2 Integrated cross-sections

Integrated cross-sections are computed for the different charm mesons at the two centre-of-mass energies in the range $2 \leq y < 4.5$ and $p_T < 8 \text{ GeV}/c$. This upper limit is chosen to coincide with the limit used in the 7 TeV measurement. However, due to the exponential fall-off of the cross-section with increasing p_T , the region past this upper limit does not contribute significantly to the integrated cross-sections. Furthermore, insufficient data in the low transverse momentum region for D_s^+ and D^{*+} requires a lower limit of $1 \text{ GeV}/c$ for these mesons. To account for (p_T, y) bins in the integration range for which no measurement was possible, a correction factor is derived, which is based on the theoretical prediction. The factor is computed as the ratio between the sum of the predicted cross-sections in the considered kinematic region, divided by the sum over all bins in this range for which a measurement is possible. An uncertainty on this factor is taken as the difference between the factors obtained using the upper and lower bounds of the theoretical predictions, respectively. The obtained integrated cross-sections are given in Tables 5.11 and 5.12. Ratios of the integrated cross-sections in the range $1 \leq p_T < 8 \text{ GeV}/c$ between 13 TeV and 5 TeV are given in Figure 5.30 and show good agreement with POWHEG+NNPDF3.0L predictions in Figure 5.34

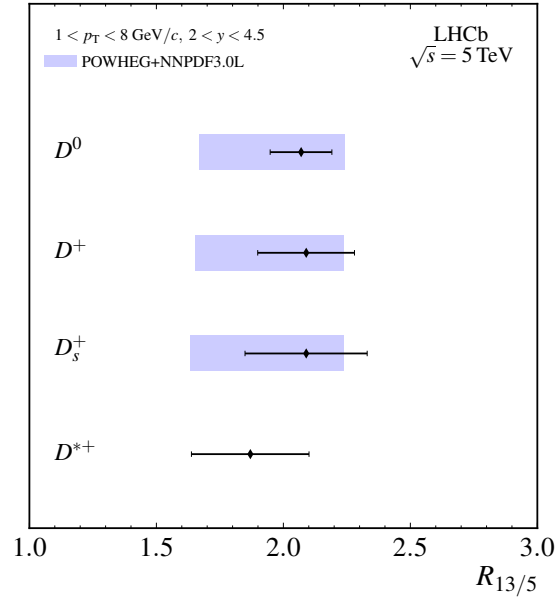


Fig. 5.34: Ratio of integrated cross-sections measured at 5 TeV and 13 TeV compared to theoretical predictions as described in 2.2.4.

Estimates for the production cross-section of $c\bar{c}$ pairs, $\sigma(pp \rightarrow c\bar{c}X)$, where at least one of the c quarks hadronises to a meson in the fiducial acceptance, can also be computed. Assuming that other contributions to the production mechanism of the measured, integrated cross-sections $\sigma(D + \bar{D})$ are negligible, an estimate is given by:

$$\sigma(pp \rightarrow c\bar{c}X) = \frac{\sigma(D + \bar{D})}{2f(c \rightarrow D)},$$

where the term $f(c \rightarrow D)$ is the quark-to-hadron transition probability, and the factor 2 accounts for the inclusion of charge-conjugate states in the measurement. Measurements at e^+e^- colliders operating at a centre-of-mass energy close to the $\Upsilon(4S)$ resonance are used to determine the transition probabilities [113] $f(c \rightarrow D^0) = 0.565 \pm 0.032$, $f(c \rightarrow D^+) = 0.246 \pm 0.020$, $f(c \rightarrow D_s^+) = 0.080 \pm 0.017$, and $f(c \rightarrow D^{*+}) = 0.224 \pm 0.028$. The fragmentation fraction $f(c \rightarrow D^0)$ has an overlapping contribution from $f(c \rightarrow D^{*+})$. While one estimate can be computed per studied charm meson, the obtained values should be independent of the used meson and be in agreement. Figure 5.35 shows good agreement of the $\sigma(pp \rightarrow c\bar{c}X)$ estimates both in the full and reduced fiducial region. To improve the precision, a combined value is computed by means of a weighted average that minimises the variance of the estimate [114], which is dominated by the estimates obtained from the D^0 cross-section and also shown in Figure 5.35. The combination of the D^0 and D^+ measurements using the full fiducial range yields:

$$\sigma(pp \rightarrow c\bar{c}X)_{p_T < 8 \text{ GeV}/c, 2.0 < y < 4.5}^{5 \text{ TeV}} = 1193 \pm 3 \pm 67 \pm 58 \text{ } \mu\text{b},$$

for 5 TeV and:

$$\sigma(pp \rightarrow c\bar{c}X)_{p_T < 8 \text{ GeV}/c, 2.0 < y < 4.5}^{13 \text{ TeV}} = 2369 \pm 3 \pm 152 \pm 118 \text{ } \mu\text{b} \text{ ,}$$

for 13 TeV. In both cases, the uncertainties are due to statistical and systematic uncertainties as well as the uncertainties related to the fragmentation fractions.

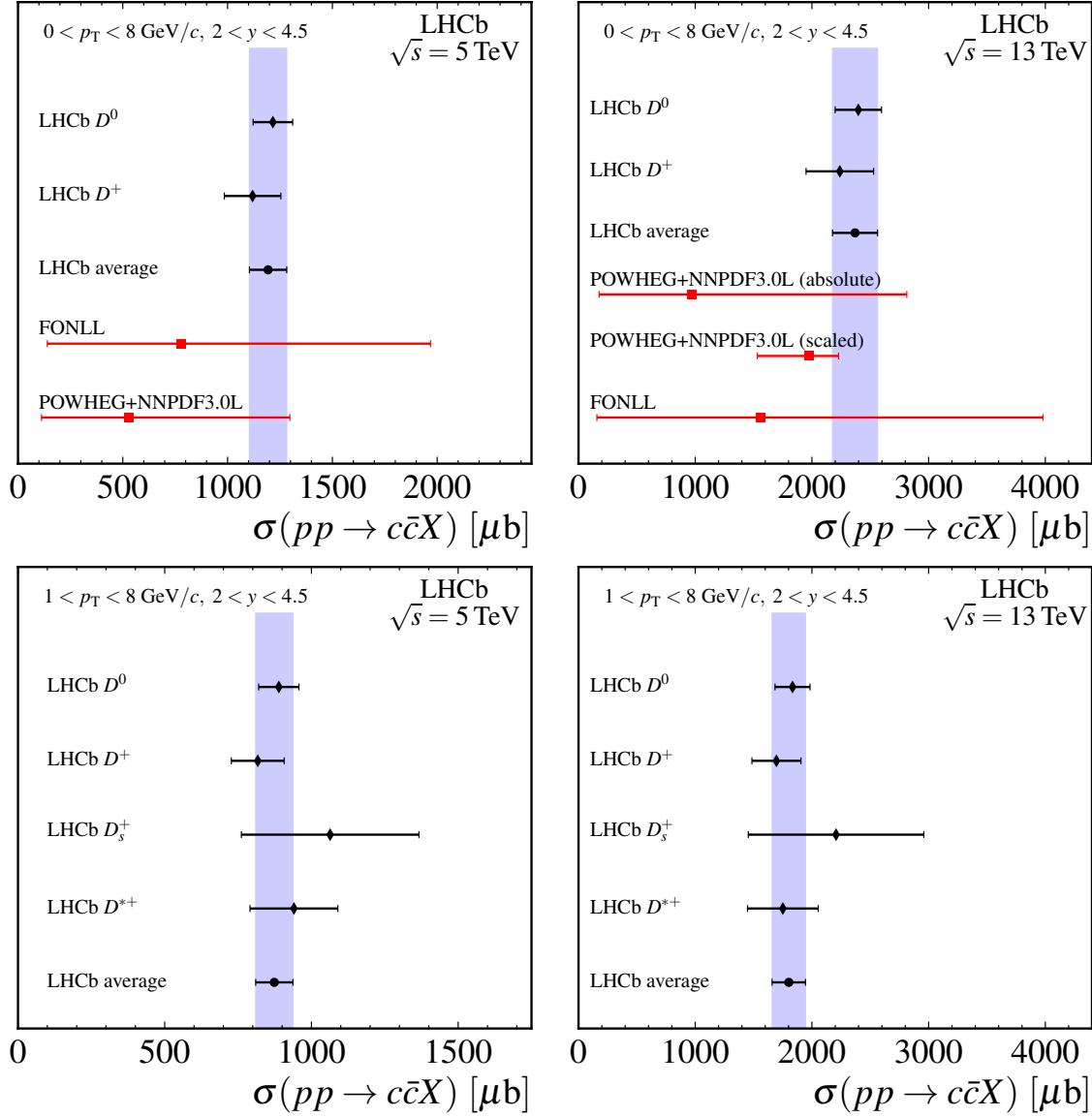


Fig. 5.35: Estimates for $\sigma(pp \rightarrow c\bar{c}X)$ in the full and reduced fiducial region for the different mesons. Also shown is the LHCb average computed as the weighted average of the different mesons and theory predictions where available. Theoretical predictions shown have been provided by the respective theorist groups and are described in Section 2.2.4.

5.8 Summary and impact

Measurements of production cross-sections are obtained for D^0 , D^+ , D_s^+ and D^{*+} mesons at centre-of-mass energies of 5 TeV and 13 TeV. Integrated cross-sections are obtained in the transverse momentum range $1 \leq p_T < 8 \text{ GeV}/c$ and range $2 \leq y < 4.5$ at $\sqrt{s} = 5 \text{ TeV}$ to be ;

$$\begin{aligned}\sigma(pp \rightarrow D^0 X) &= 1004 \pm 3 \pm 54 \mu\text{b}, \\ \sigma(pp \rightarrow D^+ X) &= 402 \pm 2 \pm 30 \mu\text{b}, \\ \sigma(pp \rightarrow D_s^+ X) &= 170 \pm 4 \pm 16 \mu\text{b}, \\ \sigma(pp \rightarrow D^{*+} X) &= 421 \pm 5 \pm 36 \mu\text{b},\end{aligned}$$

and at $\sqrt{s} = 13 \text{ TeV}$:

$$\begin{aligned}\sigma(pp \rightarrow D^0 X) &= 2072 \pm 2 \pm 124 \mu\text{b}, \\ \sigma(pp \rightarrow D^+ X) &= 834 \pm 2 \pm 78 \mu\text{b}, \\ \sigma(pp \rightarrow D_s^+ X) &= 353 \pm 9 \pm 76 \mu\text{b}, \\ \sigma(pp \rightarrow D^{*+} X) &= 784 \pm 4 \pm 87 \mu\text{b}.\end{aligned}$$

Additionally, double-differential cross-sections are found to be in good agreement with theoretical predictions. This extends to very precise double-differential cross-section ratios between different centre-of-mass energies which are very well described by theoretical calculations.

A main motivation for measuring the charm meson production cross-sections at LHCb is their potential to be used to constrain uncertainties in QCD calculations. This is mainly driven by LHCb's unique ability to measure the production in the forward region, allowing to probe very small values of gluon momentum fraction x . While this has been discussed before as the general motivation, Gauld *et al.* [115]

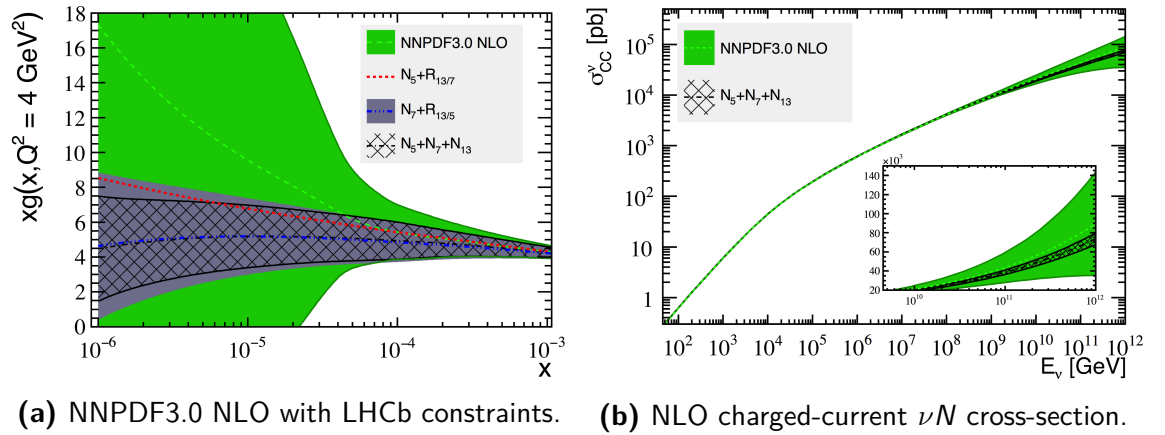


Fig. 5.36: Impact of LHCb charm cross-section data on the low- x gluon parton density function (a) and resulting improvement on predicted neutrino-nucleus interaction at high neutrino energies (b). Both plots are taken from Ref. [115]

have recently published updated NNPDF3.0 PDF sets, which are constrained using the data presented in this thesis together with the LHCb measurement at 7 TeV. They construct five different datasets, three normalised cross-sections N_5 , N_7 and N_{13} based on the double-differential cross-sections, where the differential cross-sections in every (p_T, y) bin are normalised to the result in the $3 \leq y < 3.5$ bin for the respective p_T range. In addition to these three, the two ratios $R_{13/5}$ and $R_{13/7}$ are used. Figure 5.36a compares the results of the PDF fits for different combinations of those datasets. All combinations result in consistent constraints, which are substantially more precise, exceeding one order of magnitude improvements for certain regions in x .

These improvements have a wide range of applications. As an example, Gauld *et al.* demonstrate the substantial reduction in the charged-current neutrino-nucleus interaction cross-section for very high-energy neutrinos as studied by the IceCube experiment (see Figure 5.36b). A substantial improvement is seen, showing the impact that the cross-section measurements have in other areas of particle physics which rely on calculations involving low- x gluons.

6. Mixing in charm four-body decays

6.1	Selection of the dataset	162
6.1.1	The dataset	162
6.1.2	The trigger and offline selection	163
6.1.3	Boosted decision tree	167
6.1.4	Removal of incorrect candidates	170
6.1.5	Number of signal and background events	172
6.2	The mixing fit	175
6.2.1	The signal model	176
6.2.2	The background model	178
6.2.3	A description for the efficiencies	181
6.2.4	External constraints in the fit	184
6.2.5	A few implementation details	185
6.3	The sensitivity study	187
6.3.1	Fixed amplitude models	188
6.3.2	Floating DCS amplitude model	189
6.3.3	Systematic uncertainties	192
6.4	Summary – impact on world average of x and y	194

Section 2.3 has already introduced the formalism of neutral meson mixing and compared two different approaches of extracting the parameters x and y that describe the oscillation between a D^0 and a $\bar{D}^0$ meson. Here, the potential of a model-dependent measurement using wrong-sign (WS)¹ $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ decays, exploiting the very large dataset of charm mesons recorded during Run 2 by the LHCb experiment, is presented. These sensitivity studies are implemented such that they already use high precision, realistic descriptions of experimental backgrounds and acceptance effects such that the estimated statistical uncertainties on the mixing parameters give a very realistic picture of the prospects of such a measurement.

Section 6.1 presents the definition of the used dataset and the selection of candidates for the WS decay, using the corresponding right-sign (RS) $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ decay as a control channel to tune the selection due to its much higher purity. Following this, Section 6.2 presents the used amplitudes and the structure of the fit setup, including a discussion on the construction of an efficiency description, a crucial ingredient for the aforementioned fit. Furthermore, the results of the sensitivity studies are presented in the concluding section, which also evaluates their relation to the current world average.

¹ It is useful here to recall the purposeful distinction between the labels WS and DCS as the former includes both contributions from the DCS and CF decay amplitudes due to mixing and thus its level of suppression is subject to the amount of mixing.

6.1 Selection of the dataset

Chapter 5 presents cross-section results that confirm the expectations from theoretical calculations that the cross-sections of D^0 mesons nearly doubled at $\sqrt{s} = 13$ TeV with respect to Run 1 of the LHC (at 7 TeV and 8 TeV). Additionally, the offline-quality reconstruction in the trigger as well as an optimised trigger efficiency further justify to solely focus on the data recorded during Run 2 of the LHC. This section details the used dataset and the applied selection, both online and offline.

6.1.1 The dataset

The studies in this chapter rely on data recorded during 2015 and 2016 with an integrated luminosity of 0.33 fb^{-1} and 1.67 fb^{-1} respectively. At the time of writing, data recording is ongoing with an integrated luminosity of 0.74 fb^{-1} having been recorded during 2017 so far². However, the 2017 data is not yet available for analysis and thus the following discussion uses the combined 2015 and 2016 dataset, which totals an integrated luminosity of 2 fb^{-1} . Furthermore, only a randomly selected subset of 65 pb^{-1} is used for the data-driven parts of the study as this amount is sufficient for the scope of this sensitivity study.

It is not possible to directly reconstruct and use $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ decays as they are experimentally indistinguishable from $\bar{D}^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ decays. Thus it is impossible to determine which of the reconstructed four-body decays is a WS decay. To differentiate WS and RS decays, a tagged sample is used by requiring the D^0 to stem from the D^{*+} decay $D^{*+} \rightarrow D^0 \pi^+$. As this is a strong decay, the decay $D^{*+} \rightarrow \bar{D}^0 \pi^+$ is forbidden and thus the flavour of the D^0 meson can be determined from the charge of the produced pion. Additionally, the selected $D^{*+} \rightarrow D^0 \pi^+$ with $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ or $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ candidates must be consistent with being produced in the primary vertex, exploiting the much higher production cross-sections for prompt charm mesons. In the previous chapter, the production cross-section in the general LHCb acceptance is measured to be 784 mb (see Table 5.12). Using the known branching fractions [77], this yields about $8.5 \cdot 10^{10}$ RS and $2.8 \cdot 10^8$ WS decays in the LHCb acceptance in the combined 2015 and 2016 dataset. While these numbers do not directly reflect the respective numbers of reconstructed and fully selected decays, which can have typical efficiencies on the per-mille level, they still illustrate LHCb's potential as a charm factory. In the following, the trigger and offline selection requirements, which are applied to select the candidates, are discussed.

² Until 6th of September 2017.

6.1.2 The trigger and offline selection

Similar to the previously presented cross-section measurements using Run 2 data, the events for this measurement also rely on dedicated HLT2 lines. While the RS and WS decays are separated into two different lines, both use a common container of $K^-\pi^+\pi^-\pi^+$ combinations, which then get the corresponding D^0 flavour assigned based on the charge of the additional tagging pion. Therefore, the same four-body combination can be present in both samples of selected decays, albeit with a different flavour and tagging pion, where at most one being a correctly assigned tagging pion. Unlike the cross-section measurement, no requirements are imposed on the used L0 and HLT1 lines. However, the events selected by the HLT2 lines are dominated by events that have been triggered by the calorimeter based triggers and been further selected by HLT1 lines looking for high p_T , good quality charged tracks. Table 6.1 lists the actual selection criteria used in the trigger. They are mostly very similar to those used for the cross-section measurement with one important exception: a requirement is imposed on the transverse momentum of the D^0 which is very effective in suppressing combinatorial background. This requirement is also the cornerstone of a further cut-based selection, which is applied after the trigger, where it is tightened from 2 to 4 GeV/ c .

In general, the additional requirements applied during the offline selection aim to reduce the amount of background under the peak in both the $m(K^-\pi^+\pi^-\pi^+)$ and the mass difference between the D^{*+} and D^0 candidates. Therefore, the PID requirements applied to the $K^\pm$ and $\pi^\pm$ are significantly tightened by using the more effective ProbNN variables. This family of variables, for which a variant exists for each particle hypothesis ($K^\pm$, $\pi^\pm$, $\mu^\pm$, $e^\pm$, $p/\bar{p}$), are obtained by training an artificial neural network on simulated tracks to distinguish the respective hypothesis from all other hypotheses using the DLL variables, the momentum p and p_T , and the track fit quality in both the entire detector and each subdetector separately. The output of the neural network is then simply scaled to be between zero and one. It therefore does not represent the probability that a specific particle is of the respective type. Furthermore, the requirements on the D^0 vertex are further narrowed by requiring a smaller χ^2 per degree-of-freedom and a much shorter distance-of-closest-approach of the four tracks. When dealing with events selected to be prompt, the contamination of secondary decays needs to be understood. Section 5.4 already discussed ways to discriminate prompt and secondarily produced particles in the context of the cross-section measurements. For simplicity, the sensitivity study discussed here applies a very tight cut of $\ln \chi_{\text{IP}}^2 < 1$ to remove essentially all secondary candidates from the selected samples. However, this cut also removes about 30% to 40% of signal events (see for example Figure 5.7) and an extension of the presented sensitivity

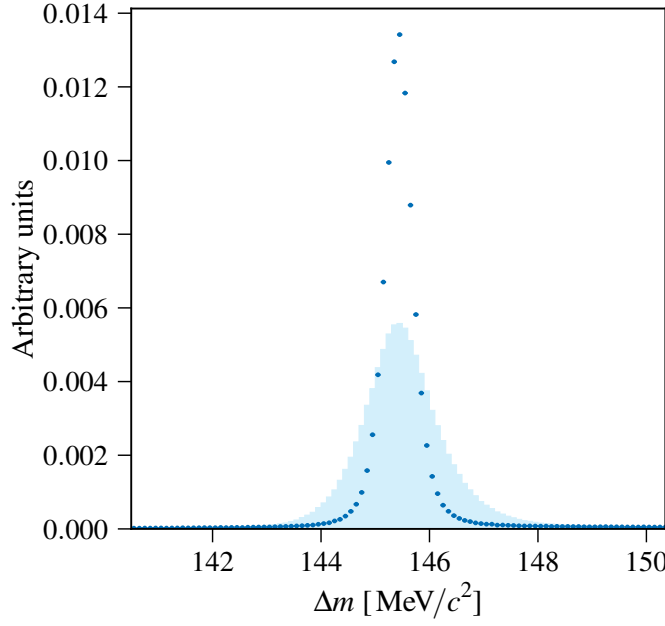


Fig. 6.1: Illustration of the improved resolution after applying a decay tree fit for the mass difference between reconstructed D^{*+} and D^0 candidates.

study might change to a more sophisticated treatment of secondary contamination to boost the statistical sensitivity.

All reconstructed particles are affected by resolution effects that lead to the reconstructed momenta being smeared with respect to the true quantities. This effect is propagated to observables constructed from these momenta such as the set of variables used to discriminate signal and background as well as the variables to describe the phase-space of the decay. The resolution can be improved by using the knowledge about the decay chain and performing a decay tree fit (DTF) [116]. In such a fit, the vertex positions and momenta of the particles are allowed to float within known resolution functions and constrained such that the D^{*+} decay products stem from one common point consistent with the PV. Additionally, the D^0 decay products are required to stem from a common point in space with momenta such that the four-body invariant mass is exactly equal to the known D^0 mass. An example illustrating the improvement when applying a DTF is given in Figure 6.1, where the mass differences Δm between the D^{*+} and D^0 candidate with and without the DTF are compared. Only events whose DTF converges are selected and a very loose upper limit is imposed on the fit χ^2 . Lastly, an upper and lower limit on the decay time t is imposed whose cuts have been chosen such that they remove one permille in the RS sample after the full selection has been applied. This cut is enforced to remove very sparsely populated areas when obtaining the data-driven efficiencies in Section 6.2.3.

Figures 6.2 and 6.3 depict the obtained mass distributions after applying all the selection criteria listed in Table 6.1 for the RS and WS sample respectively. The distribution in $m(D^0)$ reveals that the contribution of combinatorial background,

Tab. 6.1: Selection criteria for $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ and $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ decays tagged to stem from a $D^{*+} \rightarrow D^0 \pi^+$ decay. Due to a very different kinematical distribution, the pion produced in the decay of the D^{*+} is labelled as $\pi_s^\pm$ with different sets of cuts applied than to the pions stemming from the D^0 decay. Cuts applied in the HLT2 line are given separately to those which are applied offline.

Particle	Variable	HLT2 selection	Offline selection
$\pi^\pm, K^\pm$	p_T	$> 250 \text{ MeV}/c$	$> 250 \text{ MeV}/c$
	p	$> 1 \text{ GeV}/c$	$3 < p[\text{GeV}/c] < 100$
	η	—	$2 < \eta < 5$
	Track χ^2	< 3	—
	χ_{IP}^2	> 3	> 4
$\pi^\pm$	$\text{DLL}_{K-\pi}$	< 5	—
	ProbNN_π	—	> 0.5
	ProbNN_K	—	< 0.7
	ProbNN_μ	—	< 0.2
$K^\pm$	$\text{DLL}_{K-\pi}$	> 5	—
	ProbNN_K	—	> 0.7
	ProbNN_π	—	< 0.7
	ProbNN_μ	—	< 0.2
D^0	p_T	$> 2 \text{ GeV}/c$	$> 4 \text{ GeV}/c$
	p	$> 30 \text{ GeV}/c$	—
	$\sum p_T$	$> 1.8 \text{ GeV}/c$	—
	Decay time t	$> 0.1 \text{ ps}$	$0.17 < t[\text{ps}] < 3.25$
	DOCA	$< 100 \text{ mm}$	$< 0.2 \text{ mm}$
	$\text{DOCA } \chi^2$	< 12	—
	$\arccos(\text{DIRA})$	$< 20 \text{ mrad}$	—
	Vertex fit χ^2/ndf	< 12	< 4
	$\ln \chi_{\text{IP}}^2$	—	< 1
	Vertex fit χ^2/ndf	< 10	—
D^{*+}	$m(D^{*+}) - m(D^0) - m(\pi_s)$	$< 30 \text{ MeV}/c^2$	—
	DTF	—	converged
	DTF χ^2	—	< 100
$\pi_s^\pm$	p_T	$> 100 \text{ MeV}/c$	—
	p	—	$3 < p[\text{GeV}/c] < 100$
	ProbNN_π	—	> 0.3
	ProbNN_K	—	< 0.7
	ProbNN_p	—	< 0.15
	$\text{ProbNN}_{\text{ghost}}$	—	< 0.15
	ProbNN_μ	—	< 0.1

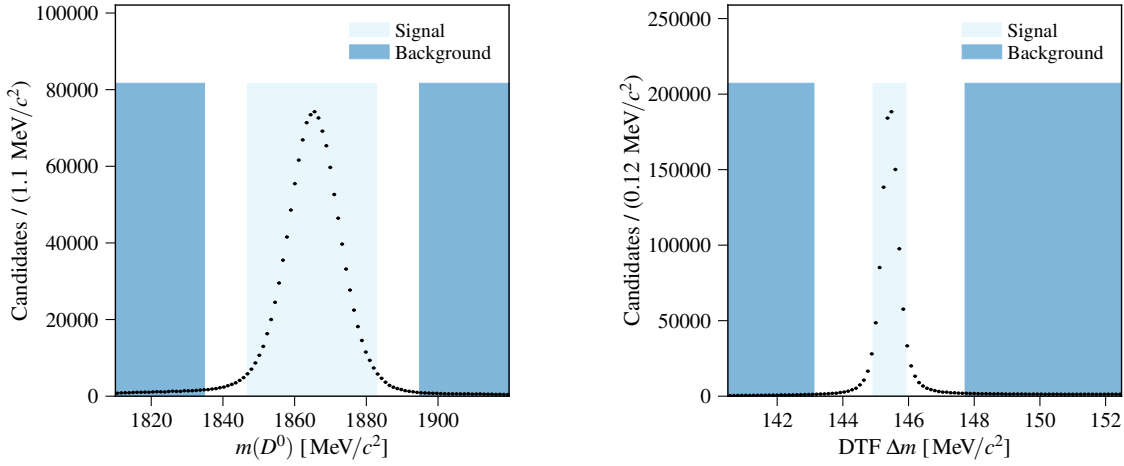


Fig. 6.2: Invariant mass distribution of the D^0 candidate (without DTF) and the Δm distribution (with DTF) in the RS sample. All selection criteria listed in Table 6.1 have been applied. Also indicated are the signal and background regions.

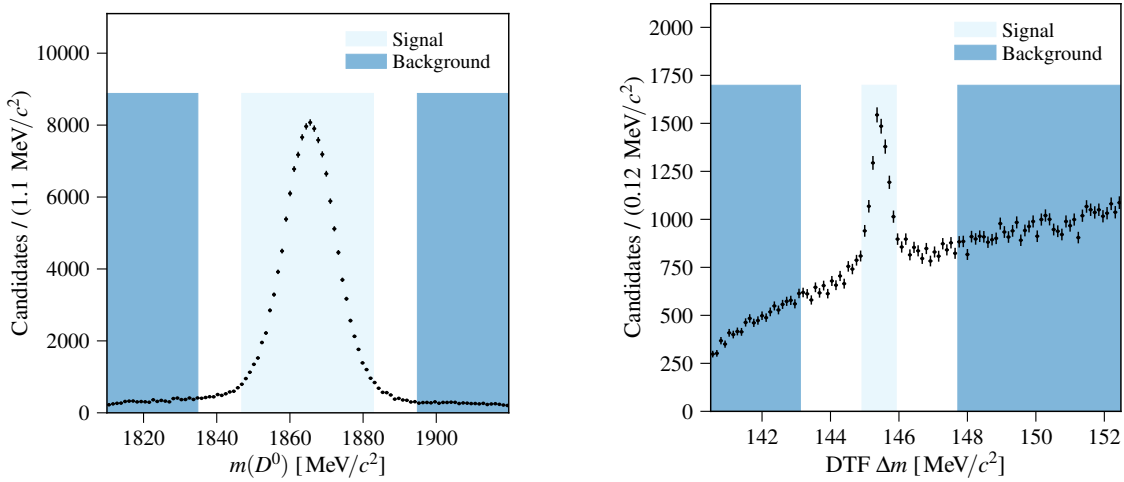


Fig. 6.3: Invariant mass distribution of the D^0 candidate (without DTF) and the Δm distribution (with DTF) in the WS sample. All selection criteria listed in Table 6.1 have been applied. Also indicated are the signal and background regions.

which is characterised by a practically flat distribution in the mass window, is small compared to the size of the peak. While this signal peak is replicated in the mass difference for the RS sample, indicating that the sample already has a high purity, this is not the case for the WS sample that is dominated by true $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ candidates associated with an incorrect pion when building the D^{*+} candidate (often referred to as random slow pion background). In the following, an additional, optimised selection is developed to further reduce the background contributions in the two samples.

6.1.3 Boosted decision tree

A boosted decision tree (BDT) is used to further reduce the backgrounds after the cut based selection given in Table 6.1. The trained BDT focuses on suppressing the random slow pion background³. In order to develop the selection, the RS sample is used, which is assumed to be identical to the WS sample as far as variables commonly used in selections are concerned. A training sample consisting of mainly signal events is constructed from the signal regions in both mass variables and a sample rich in random slow pion background is formed by combining the signal region in $m(D^0)$ with the background region in Δm .

In general, the BDT optimises the selection to achieve the best separation between different signal and background categories by giving a greater weight to events which have been mis-classified in each iteration of the boosting. Therefore, even if the BDT was not provided with the specific input variable, due to correlations it is capable of picking up underlying relations, e.g. that the measured decay time t is a variable with strong separation power. Taking the decay time as an example, imposing a selection criterion on the BDT discriminant might result in a complicated efficiency dependence on the decay time and a suppression in the low decay-time region. However, as discussed in Section 2.3.4, the sensitivity to the mixing parameters x and y comes from the decay-time dependence and the best sensitivity is achieved if the covered decay time range is as large as possible. Therefore, a modified BDT implementation is used that on the one hand still applies a boosting algorithm penalising misclassification while on the other hand also using a cost function in each step of the boosting to improve uniformity of the efficiencies for a given set of variables [117]. The BDT is trained using the transverse momenta p_T of both the D^0 and slow pion $\pi_s^\pm$ candidate, the quality of the D^{*+} candidate vertex fit, the vertex distance significance χ_{VD}^2 , and the χ^2 of the decay tree fit. Figure 6.4 compares the distributions of the input variables between the signal and background samples as well as the produced output distribution of the BDT classifier, which shows good

³ The potential improvement with a second BDT, dedicated to reject combinatorial background, was investigated but no improvements were found.

Tab. 6.2: Optimised selection criteria obtained from applying a Bayesian optimisation procedure to the Punzi figure of merit.

Variable	Cut value
BDT	> 0.45
$\pi_s^\pm \text{ ProbNN}_\pi$	> 0.75
$\pi_s^\pm \text{ ProbNN}_K$	< 0.25
$\pi_s^\pm \text{ ProbNN}_p$	< 0.15
$\pi_s^\pm \text{ ProbNN}_{\text{ghost}}$	< 0.15

agreement between the distributions obtained in the training and testing samples (30% of the signal and background samples are used for training, respectively).

The receiver operating characteristic (ROC) curve of the trained BDT, evaluated on the test sample, is shown in Figure 6.5a and reveals an improved performance compared to a cut-based selection on the individual input variables and a near identical performance when comparing the two different boosting algorithms. In addition, Figure 6.5b investigates the decay-time dependence of the BDT discriminant by applying four different cuts corresponding to different global cut efficiencies and the obtained efficiency dependence is flat in all four cases. As stated above, the random slow pion background is dominating the overall background in the selected WS sample and thus including $\pi_s^\pm$ variables, especially PID variables, in the BDT seems like a natural decision. However, while not applied in this sensitivity study, the estimation of the efficiencies of requirements imposed on PID variables is typically done using a data-driven technique relying on calibration samples (see Section 5.5.2), which is substantially more complicated to apply to cases where PID variables are included in the BDT. Therefore, the PID requirements imposed on the slow pion candidate are optimised in conjunction with the cut on the BDT discriminant as described in the following.

The cuts applied on the BDT discriminant as well as ProbNN_π , ProbNN_K , ProbNN_p , and $\text{ProbNN}_{\text{ghost}}$ of the slow pion are optimised using Bayesian parameters optimisation (following the description in Ref. [118]). The Punzi figure of merit is used, given by:

$$P = \frac{N_{\text{sig}}/N_{\text{sig},0}}{\frac{1}{2} + \sqrt{N_{\text{bkg}}}}, \quad (6.1)$$

where N_{sig} and $N_{\text{sig},0}$ are the number of events in the RS Δm signal window before and after the optimised selection divided by 300, and N_{bkg} is the number of events in the WS Δm sidebands after the optimised selection divided by 20. Both scaling factors are applied to approximate the number of signal and background events found in the WS Δm signal window, the sample of interest.

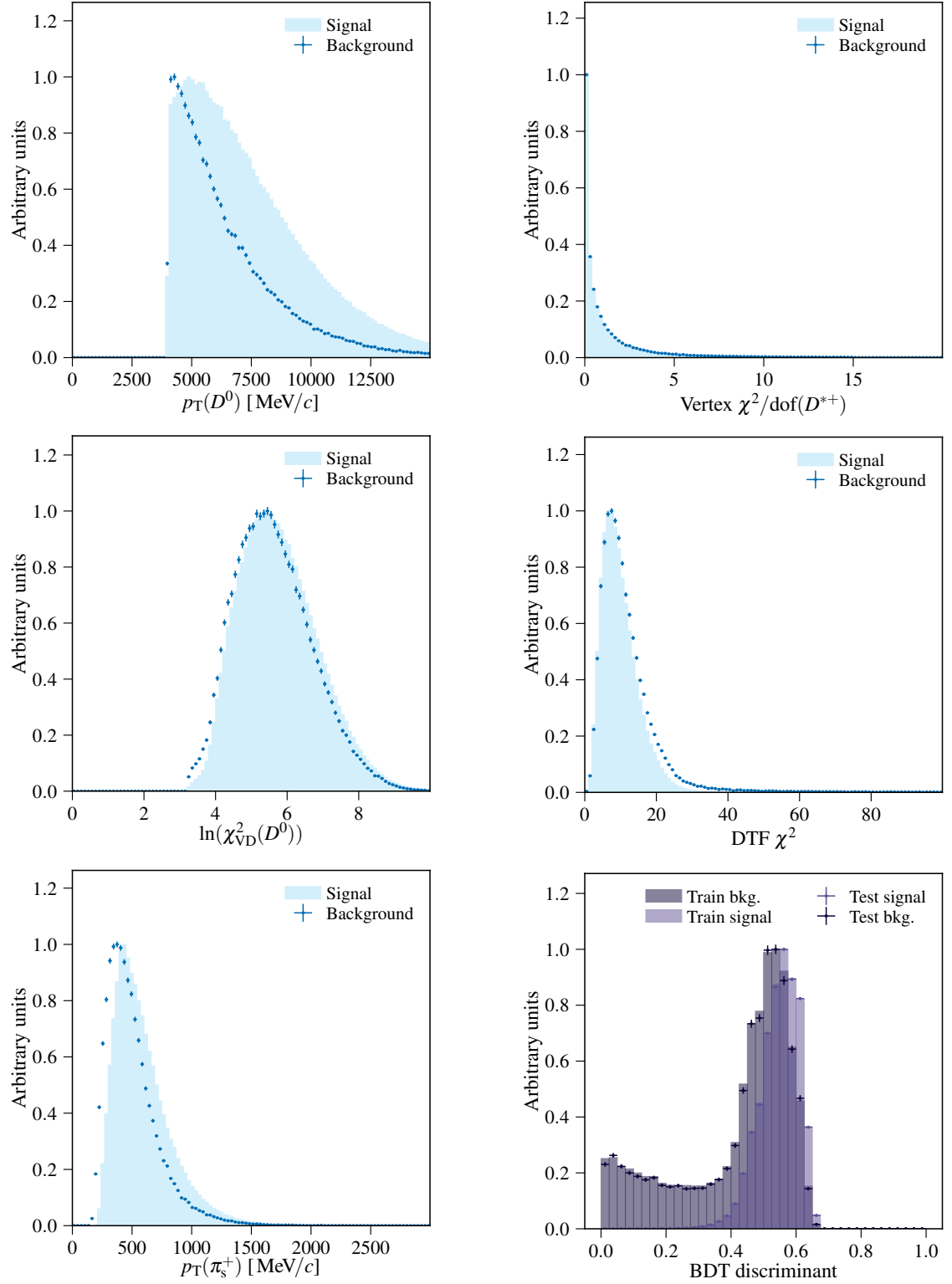


Fig. 6.4: Input variables for the BDT trained to distinguish signal from random slow pion background. Data plotted is taken from the signal window in $m(D^0)$ and Δm (for signal) and the $m(D^0)$ signal combined with Δm background window (for random slow pion background). Also shown is the distribution in the BDT discriminant after training for the classifier.

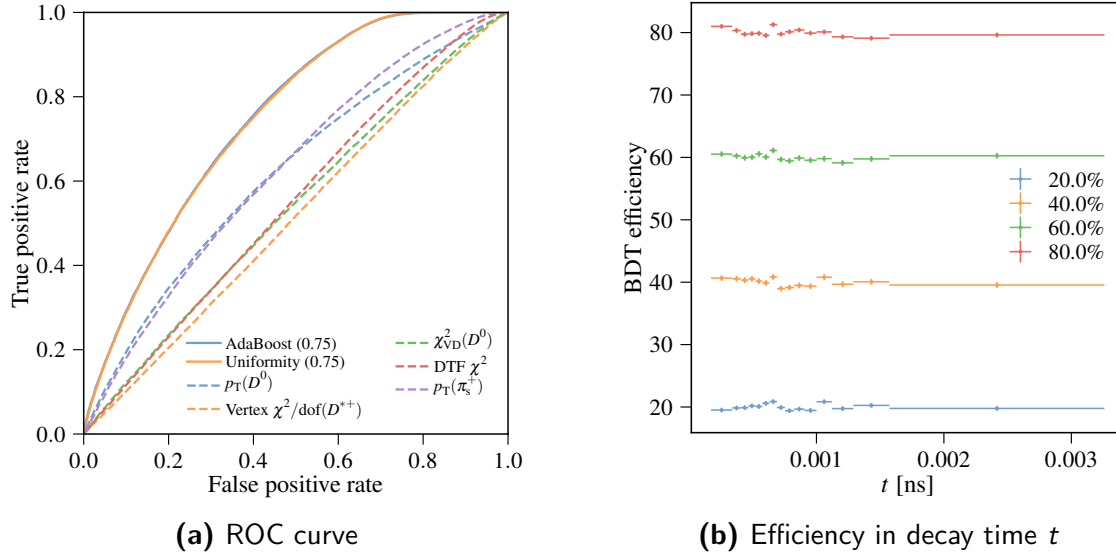


Fig. 6.5: Performance of the BDT: (a) ROC curve comparison between the uniform BDT and the classical AdaBoost implementation with simple cuts on the input variables. The area under the ROC curve is given in the legend as well; (b) efficiency for different cuts on the BDT discriminant (with total efficiencies as stated in the legend) as a function of the decay time t .

6.1.4 Removal of incorrect candidates

While the selection presented above focused on removing general backgrounds stemming mostly from incorrectly selected slow pion candidates, the purity of the WS sample can be further enhanced by explicitly studying the contamination from certain misidentified or incorrectly reconstructed candidates. In the following, the studies are performed with respect to the selection criteria described above.

Double mis-identification refers to the candidates which are built from two tracks that have the incorrect particle hypothesis assigned. This is especially concerning in the case where a true kaon is misidentified as a pion of opposite charge and vice versa as this allows true RS decays to be accepted as WS candidates while being tagged by the correct slow pion. While the selection summarised in the previous section applies tight requirements to the PID variables, which reduces the likelihood of this happening, exchanging the hypothesis in such a way roughly preserves the overall energy of the decay and thus it is possible to obtain mis-identified candidates still entering the mass range considered in this analysis. In order to remove any contamination from RS decays that have been doubly mis-identified and thus tagged as a WS candidate, a $60 \text{ MeV}/c^2$ exclusion window centered around the nominal D^0 mass of $1864.83 \text{ MeV}/c^2$ is defined on the invariant mass constructed after exchanging the kaon and oppositely charged pion (both oppositely charged pions are considered). Figure 6.6a shows the Δm distribution in the WS sample, indicating the number of events that are removed by this requirement. To gauge the efficiency of true, correctly identified decays to pass the requirement, the same cut is applied to the RS

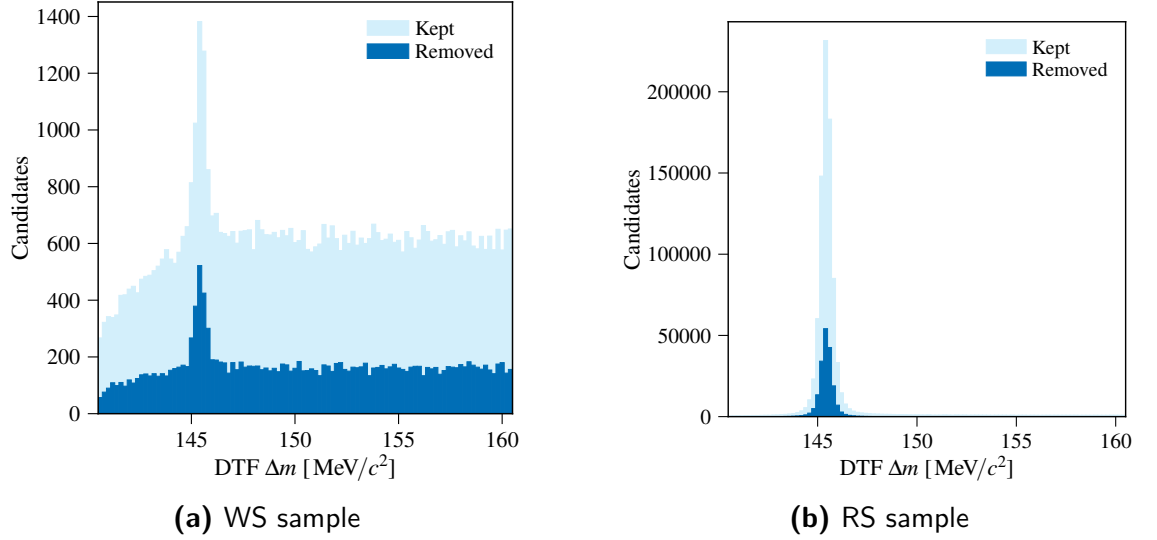


Fig. 6.6: Distribution of Δm in the WS and RS sample after excluding events where any of the two D^0 invariant masses fall into $\pm 30 \text{ MeV}/c^2$ window around $1864.83 \text{ MeV}/c^2$ after exchanging the kaon's and oppositely charged pion's mass hypotheses.

sample and the respective distribution is shown in Figure 6.6b. In the RS sample, the contamination of doubly mis-identified WS decays falsely added to the RS sample is assumed to be negligible and thus all decays removed by the requirement are true RS decays. Comparing the numbers in the signal window before and after the application of the cut, the efficiency is found to be 81.2%. Comparing the number of events in the signal peak of the WS sample before and after the cut (using the fit described in Section 6.1.5), 67.1% of the events are removed by the cut. Assuming that the requirement removes all of the doubly mis-identified events and the efficiency for true WS candidates is identical to the one obtained for the RS sample, 17.4% of the WS candidates in the signal peak are mis-identified candidates before this rejection cut is applied.

As stated before, both RS and WS samples are constructed from the same set of four-body D^0 decays as only the additional π_s^+ can tag the flavour of the D^0 . Additionally, after the selection, the background in the WS sample is dominated by true D^0 four-body decays that have not been matched with the correct slow pion. Due to the much larger branching fraction of the CF decays, these decays are very likely to be RS decays that have been incorrectly tagged as WS candidates. Therefore, events are removed from the WS sample if a RS candidate built from the same D^0 candidate is found in the Δm signal window. Figure 6.7a depicts the distribution of the accepted and rejected WS candidates, showing that about 10% of the background is removed without affecting the number of signal events in the WS sample.

A further source of combinatorial background is introduced by partial clones, which are sets of tracks that share parts of their track segments, most commonly the segment

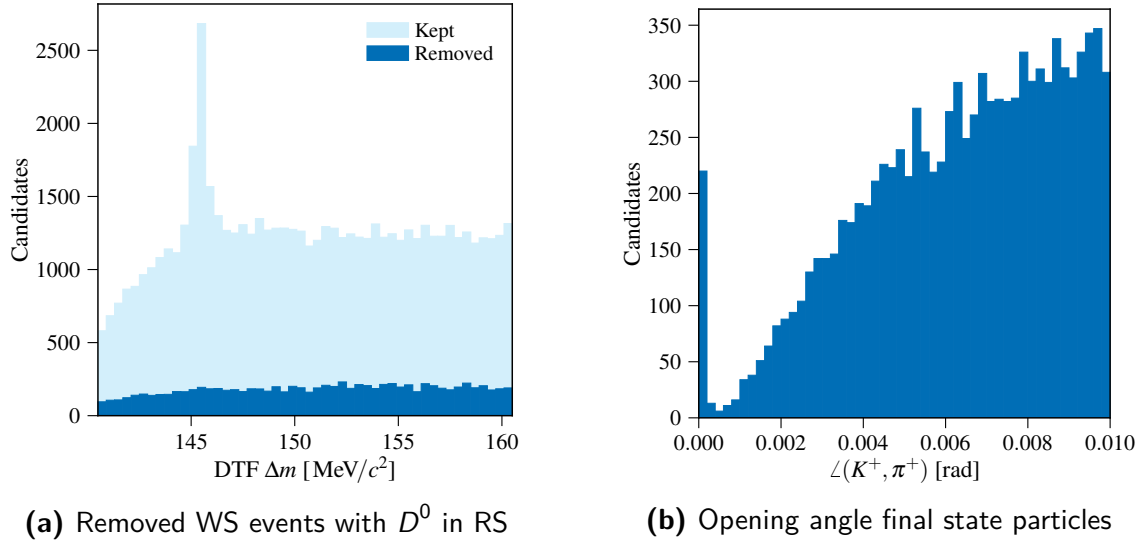


Fig. 6.7: Overlap removal between WS and RS samples (shown distributions are not stacked) (a) and clone track identification by looking for track pairs with tiny opening angle (b).

in the VELO: during the track reconstruction procedure, different tracks in the T stations are matched to the same VELO track, resulting in tracks with identical direction but different momentum and PID properties. To identify candidates containing final state particles that are partial clones, the opening angle between every possible pair of particles is studied (using the information before the decay tree fit). One such example is shown in Figure 6.7b where a clear increase is seen for an angle of practically zero. To remove these candidates, the opening angle between every possible combination of final state particles comprising the candidate is required to exceed 0.03° .

After applying all selections discussed, 0.3% of the events (pp bunch crossings) in the WS sample have more than one selected candidate. Due to the highly suppressed decay, it is practically impossible that two or more genuine candidates exist in the same event and hence at most one of the candidates per event can be an actual signal decay. Furthermore, additional candidates are not necessarily independent as they might share a set of common particles, with the most prominent example being a D^0 candidate that is matched with multiple $\pi_s^\pm$ resulting in multiple candidates⁴. To avoid any adverse effects from those multiple candidates and in light of the overall very small fraction of events with multiple candidates, one of the candidates is chosen randomly in events with multiple selected candidates.

6.1.5 Number of signal and background events

For a realistic estimate of the precision that can be expected on the mixing parameters x''' and y''' , the number of signal and background candidates in the signal window of

⁴ In fact, this accounts for about half of the additional candidates in the sample.

the WS sample needs to be determined. Similar to the cross-section measurement, these numbers are obtained from extended maximum likelihood fits, this time to the two-dimensional distribution in $m(D^0)$ and the DTF Δm .

In principle, three different contributions can be disentangled in a two-dimensional fit to $m(D^0)$ and Δm : signal, random slow pion, and combinatorial background. However, the amount of combinatorial background is negligible and the inclusion of a component describing it in the fit is found to cause instabilities in the fit and is therefore ignored in the following. For the $m(D^0)$, both signal and random slow pion contributions are modelled using a single Johnson S_U distribution given by:

$$S_U(m|\mu, \sigma, \nu, \tau) = \frac{e^{-\frac{1}{2}r^2}}{2\pi c \cdot \sigma \cdot \tau \sqrt{z^2 + 1}}, \quad (6.2)$$

with the substitutions

$$\begin{aligned} r &= -\nu + \frac{\arcsin(z)}{\tau}, \\ z &= \frac{m - (\mu + c \cdot \sigma \cdot \sqrt{\omega} \sinh(\omega))}{c \cdot \sigma}, \\ c &= \frac{h - 1}{2\sqrt{h \cosh(2\omega) + 1}}, \\ h &= e^{\tau^2}, \text{ and} \\ \omega &= -\nu\tau. \end{aligned}$$

In this parameterisation, μ , σ , ν and τ correspond to the mean, standard deviation, skewness and kurtosis of the distribution, respectively. As the expected distribution in $m(D^0)$ is identical for both signal and random slow pion background, a common shape with one set of parameters is used. Due to its great flexibility when modelling skewed, asymmetric peaks, a Johnson S_U is also used to model the signal peak in Δm while the background is described using:

$$B(\Delta m|\Delta m^0, a, c) \propto (1 - e^{-\frac{\Delta m - \Delta m^0}{c}}) \cdot \left(\frac{\Delta m}{\Delta m^0}\right)^a, \quad (6.3)$$

where the parameter Δm^0 is fixed to $139.57 \text{ MeV}/c^2$. The fits are performed in the range $1810 < m(D^0) < 1920 \text{ MeV}/c^2$ and $140.5 < \Delta m < 160.5 \text{ MeV}/c^2$.

Figure 6.8 gives the projections on the two dimensions of the fit for the WS. Generally, a good description of the data is seen with a clear WS signal above a large background. While not actually needed, the figure also depicts the corresponding RS fit, showing a very pure RS sample with practically no background under the peak. The results of the WS fit for all parameters describing the fit shapes as well as the yields are given in Table 6.3. The number of signal and background events in the full dataset

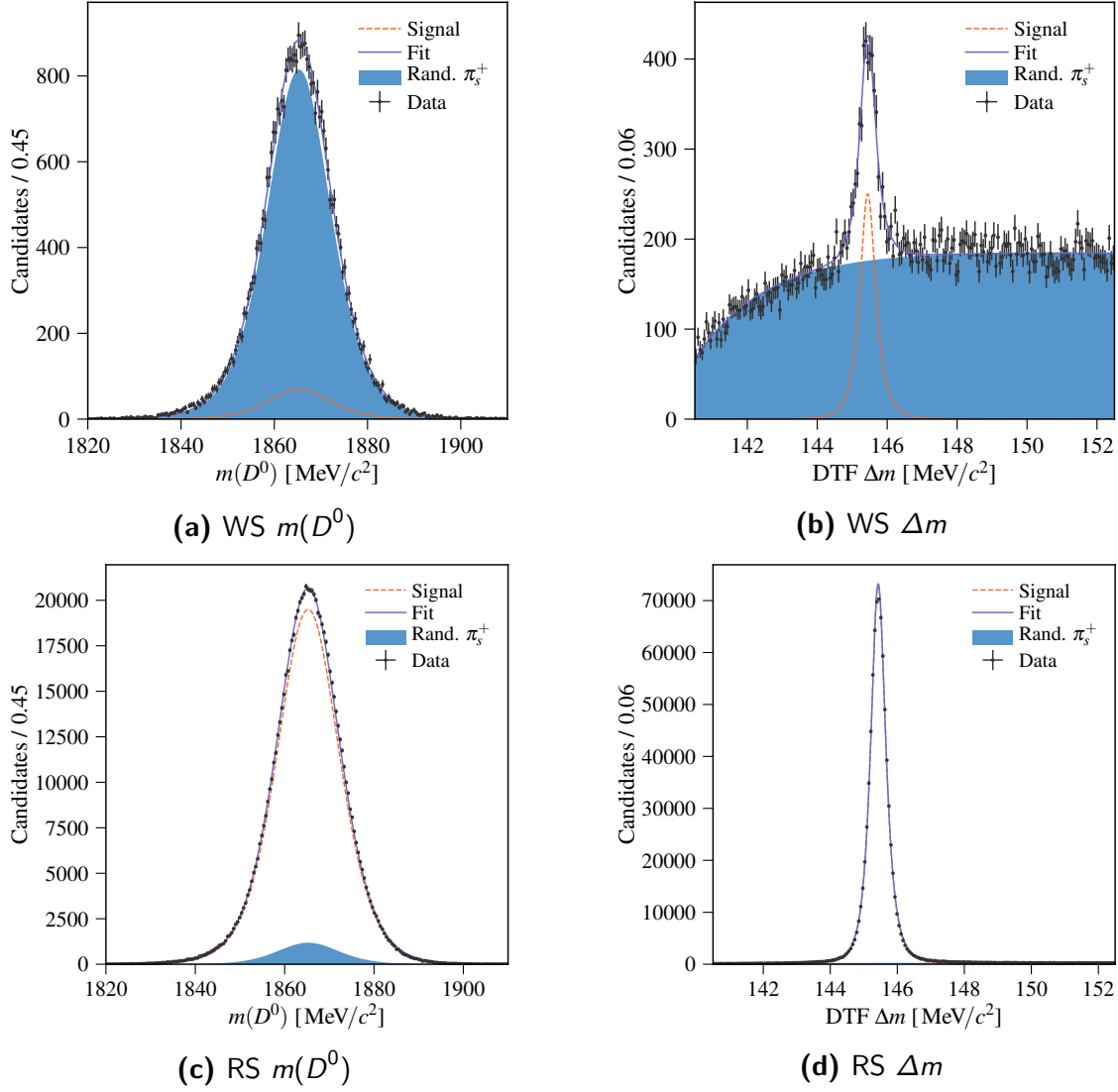


Fig. 6.8: Projections of the two-dimensional fit to $m(D^0)$ and Δm for the RS and WS sample. Only 65 pb^{-1} of the data is used. Also note that the plots show only a reduced area of the full fit range.

(the combination of data recorded in 2015 and 2016) is then estimated by multiplying the result for the yield by a factor 30 ($\approx 2 \text{ fb}^{-1}/65 \text{ pb}^{-1}$). The mixing fit, which is described in the following section, is only performed in the Δm signal window as indicated in Figure 6.3, covering a $\pm 0.5 \text{ MeV}/c^2$ window around the peak at $\Delta m = 145.42 \text{ MeV}/c^2$. By integrating the signal and background PDFs one obtains that 85.1% and 5.1% of the signal and background events respectively within the fit range are situated in the signal region. Assuming Poisson uncertainties on the expected number of events in the entire fit range, the expected number of signal and background events in the relevant signal region (SR) is:

$$N_{\text{Sig}}^{\text{SR}} = 71170 \pm 250, \text{ and}$$

$$N_{\text{Bkg}}^{\text{SR}} = 87900 \pm 80.$$

Tab. 6.3: Fit results for the two dimensional fit to $m(D^0)$ and Δm in the WS sample. All parameters are given dimensionless and hence rely on the masses being given in MeV/c^2 .

τ_m	0.5117 ± 0.0082	$\mu_{\Delta m}$	145.455 ± 0.029
σ_m	8.375 ± 0.032	$\sigma_{\Delta m}$	0.415 ± 0.090
μ_m	1865.309 ± 0.032	$\nu_{\Delta m}$	0.04 ± 0.15
Δm_{sp}^0	139.57	$\tau_{\Delta m}$	0.89 ± 0.18
$a_{\Delta m, sp}$	-0.34 ± 0.23	N_{Sig}	2790 ± 150
$c_{\Delta m, sp}$	2.116 ± 0.096	N_{Rnd}	57580 ± 280

6.2 The mixing fit

Having defined the selection and the dataset which is intended to be used in the measurement, this section deals with the construction of a fit model to describe the overall data in the WS signal region and extract the mixing parameters. To provide structure to the remaining section, the basic ingredients necessary are neatly explained by starting from the extended likelihood that will be minimised:

$$\mathcal{L} = \left\{ \prod_i^N \frac{1}{N_{\text{Sig}} + N_{\text{Bkg}}} \left[N_{\text{Sig}} \cdot \mathcal{P}'_{\text{Sig}}(\Phi_{n,i}, t_i | \vec{\theta}) + N_{\text{Bkg}} \cdot \mathcal{P}'_{\text{Bkg}}(\Phi_{n,i}, t_i | \vec{\theta}) \right] \right\} e^{-(N_{\text{Sig}} + N_{\text{Bkg}})} \cdot \frac{(N_{\text{Sig}} + N_{\text{Bkg}})^N}{N!} \cdot \mathcal{C}(\mathcal{N}_{\text{Sig}}, \mathcal{N}_{\text{Bkg}}, \vec{\theta}) , \quad (6.4)$$

where N is the number of observed events in the dataset, N_{Sig} and N_{Bkg} the yields for the individual species – signal and background – in the fit and $\mathcal{P}'_{\text{Sig}}$ and $\mathcal{P}'_{\text{Bkg}}$ are normalised models describing the six-dimensional distribution in phase-space Φ_n and decay time t that depend on a potentially very large number of parameters encoded in $\vec{\theta}$. Furthermore, a Poisson probability ensures that the sum of N_{Sig} and N_{Bkg} is close to the observed number of events and an additional term $\mathcal{C}$ collects all additional constraints, which are added to the likelihood to include external input on the free parameters. Additionally, the signal and background models factorise as:

$$\begin{aligned} \mathcal{P}'_{\text{Sig}}(\Phi_n, t | \vec{\theta}) &\propto \varepsilon_{\text{sig}}(\Phi_n, t) \cdot \mathcal{P}_{\text{Sig}}(\Phi_n, t | \vec{\theta}) \text{ and} \\ \mathcal{P}'_{\text{Bkg}}(\Phi_n, t | \vec{\theta}) &\propto \varepsilon_{\text{bkg}}(\Phi_n, t) \cdot \mathcal{P}_{\text{Bkg}}(\Phi_n, t | \vec{\theta}) , \end{aligned}$$

where $\varepsilon_{\text{sig}}(\Phi_n, t)$ and $\varepsilon_{\text{bkg}}(\Phi_n, t)$ describe the acceptance and reconstruction effects, which might differ for signal and background events, and $\mathcal{P}_{\text{Sig}}$ and $\mathcal{P}_{\text{Bkg}}$ give the actual distribution parameterising the physics. While the description of the background is not necessarily separable into an efficiency and a model describing the physics, *e.g.* if the background is mostly combinatorial, it is intentionally done here, and the reasoning is discussed below. Furthermore, one assumes that given a fine enough discretisation in the dependent variables, efficiency descriptions are independent of the used decay model.

Usually, the $\mathcal{P}'_{\text{Sig}}$ and $\mathcal{P}'_{\text{Bkg}}$ should also be convolved with a function describing the resolution effects. However, resolution effects are not investigated on the scope of this study. Consequently, a proper treatment of these resolution effects is a necessity when expanding this study to a measurement. Nonetheless, due to the excellent momentum resolution of LHCb, as discussed in Section 3.3.1, the impact of the resolution effects are likely small, especially as a DTF is already applied.

After decomposing the different components of the likelihood, the following section discusses the construction of the signal model $\mathcal{P}_{\text{Sig}}$, followed by a discussion of the modelling of background contributions, a derivation of an efficiency description, a presentation of the constraints applied in the fit and concluded by an overview on the implementation of the fit.

6.2.1 The signal model

The signal model $\mathcal{P}_{\text{Sig}}$ is given by the time and phase-space dependent decay amplitude given in Equation (2.53), which is repeated here for convenience:

$$\begin{aligned} \mathcal{P}_{\text{Sig}}(t, \Phi_n) = \mathcal{N}_f \cdot e^{-\Gamma t} & \left[r^2 |A_{\text{DCS}}(\Phi_n)|^2 \right. \\ & + |\bar{A}_{\text{CF}}(\Phi_n)|^2 \cdot \frac{y'''^2 + x'''^2}{4} (\Gamma t)^2 \\ & - r |A_{\text{DCS}}(\Phi_n)| |\bar{A}_{\text{CF}}(\Phi_n)| \Gamma t \\ & \left. \cdot \{ y''' \cos(\delta'(\Phi_n)) - x''' \sin(\delta'(\Phi_n)) \} \right]. \end{aligned} \quad (6.5)$$

As a PDF, $\mathcal{N}_f$ is the normalisation factor necessary to force the integral of $\mathcal{P}_{\text{Sig}}$ over all six dimensions to be unity. However, due to the additional term describing the efficiencies $\varepsilon_{\text{sig}}(\Phi_n, t)$, there is no need to determine this factor explicitly as only the product of $\mathcal{P}_{\text{Sig}}$ and $\varepsilon_{\text{sig}}(\Phi_n, t)$ is required to be normalised. Details on the integration itself are going to be discussed in Section 6.2.5. The additional factor r is the same factor as in Equation (2.45) and introduced here to ensure the correct overall difference in magnitude. This is necessary due to the way the amplitude models are defined in the following but also very useful in practice as it allows to easily constrain the overall difference in the decay rate between the two amplitudes using external input (*e.g.* from the model-independent measurement discussed in Section 2.3.3).

While the parameterisation of the time component is trivially done by simply using the decay time, multiple choices exist to parameterise the phase-space component Φ_n . For this study, a combination of two invariant mass pairs and three angles is utilised that is inspired by the decay $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ proceeding via a $K^+ \pi^-$ and $\pi^+ \pi^-$

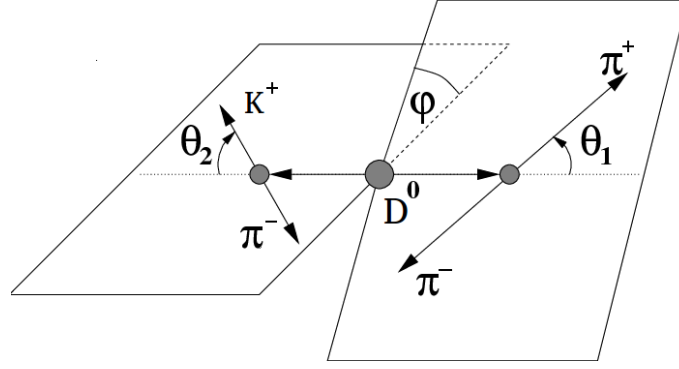


Fig. 6.9: Definition of the observables used to describe the five-dimensional phase-space. Figure based on Ref. [119].

resonance. Therefore, the invariant masses are given by $m(K^+\pi^-)$ and $m(\pi^+\pi^-)$. The angle ϕ is given by the angle between the decay planes of these two resonances in the rest frame of the D^0 and the two remaining angles, θ_1 and θ_2 , are the helicity angles of the π^+ and K^+ and are defined as the angle between the particle's momentum measured in the resonance rest frame and the flight distance of the resonance determined in the D^0 rest frame. An illustration is provided in Figure 6.9.

Section 2.4 discusses the general construction approach of models for decay amplitudes and the different components entering that model. The actual models, which are used in the following studies, are taken from preliminary studies performed within the LHCb collaboration (which are internal at the time of this thesis). The model for $A_{\text{DCS}}(\Phi_n)$ is given in Table 6.4 while Table 6.5 lists all entering resonances with their complex coefficients for $\bar{A}_{\text{CF}}(\Phi_n)$. Additionally, Table 6.6 provides the masses and widths used for the respective lineshape. While the sensitivity study discussed in this chapter will not delve into the construction of the models themselves, *i.e.* which individual resonances are included, a discussion of the general way these models are obtained is necessary. Both are constructed from fits to decay-time integrated phase-space distributions of RS and WS events, using the full Run 1 dataset of secondarily produced D^0 decays. Therefore, the model given in Table 6.4 is not, strictly speaking, a model for the DCS decay but rather a model for WS decays as it already includes the contribution from mixing. Nonetheless, due to the small effect of mixing, the model is assumed to be a good enough approximation for a sensitivity study. In an actual measurement though, it would be necessary to extract $A_{\text{DCS}}(\Phi_n)$ in conjunction with the mixing parameters.

It was already discussed previously that the overall phase of the amplitudes cannot be determined in the time-integrated fits as the PDF in those fits is proportional to the amplitude squared. Furthermore, PDFs are, by definition, normalised and hence there is also no sensitivity to the overall magnitude of the amplitudes $A_{\text{DCS}}(\Phi_n)$ and $\bar{A}_{\text{CF}}(\Phi_n)$ as any such factor is absorbed in the normalisation of the PDF. For both rea-

Tab. 6.4: Decay chains considered in the amplitude $A_{\text{DCS}}(\Phi_n)$ describing the decay $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ and their corresponding complex coefficients. The model is a preliminary result of the work that culminated in Ref. [120]. The shown spin factor enumerations refer to the indices in Table 2.2.

Decay Chain	Spin Factor	$\Re(c_i)$	$\Im(c_i)$	Fit Fraction
$K^*(892)^0[K^+\pi^-]\rho(770)[\pi^+\pi^-]$	1	1.000	0.000	0.0786
$[K^*(892)^0[K^+\pi^-]\rho(770)[\pi^+\pi^-]]^{\mathcal{P}\text{-wave}}$	2	-0.283	1.697	0.0360
$[K^*(892)^0[K^+\pi^-]\rho(770)[\pi^+\pi^-]]^{\mathcal{D}\text{-wave}}$	3	23.046	7.441	0.1128
$\rho(1450)[\pi^+\pi^-]K_0^*(1430)^*[K^+\pi^-]$	4	1.853	0.106	0.0742
$K_1(1270)^+[K^*(892)^0[K^+\pi^-]\pi^+]\pi^-$	7	-0.419	-1.590	0.0578
$K_1(1270)^+[\rho(770)[\pi^+\pi^-]K^+]\pi^-$	7	2.294	2.669	0.1334
$K_1(1270)^+[K_0^*(1430)^0[K^+\pi^-]\pi^-]\pi^-$	9	-0.740	0.375	0.0233
$K_1(1400)^+[K^*(892)^0[K^+\pi^-]\pi^+]\pi^-$	7	-2.783	-2.295	0.2557
$K^+\pi^-\pi^-\pi^+(\text{non-resonant})$	-	-0.407	-0.125	0.1767

sons, the complex coefficient of a large resonance, here $K^*(892)^0[K^+\pi^-]\rho(770)[\pi^+\pi^-]$ in the $\mathcal{S}$ -wave configuration, is set to one. However, the overall magnitude of a decay amplitude describes how likely a certain process is and the difference in overall magnitude between the DCS and CF amplitude is crucial for the extraction of the mixing parameters. Therefore, the additional parameter r was added to Equation (6.5) and in the fit, $A_{\text{DCS}}(\Phi_n)$ is normalised to have the same phase-space integrated magnitude as $\bar{A}_{\text{CF}}(\Phi_n)$. In the following, the true value of r is approximated as $1/\sqrt{300}$.

6.2.2 The background model

The construction of a background model for the fit usually involves data-driven methods to extract the shape from sidebands as the parameterisations for some potential contributions are hard to define. However, as noted before, the background in the WS sample is completely dominated by true D^0 decays that have been associated with the incorrect slow pion. Additionally, simply due to the large difference in branching fraction of the CF to DCS, this background must consist of mostly RS decays. Figure 6.10 shows a comparison between the WS sample sidebands in Δm as a proxy for the shape of the background in the signal region, with the signal region in the RS sample for all variables used in the fit. A very good agreement is seen for all phase-space variables with the decay time t being the only one showing small differences. This is expected: the D^0 candidates in the WS sidebands are much softer than those in the RS signal region, leading to slightly reduced efficiencies in the decay time as the same decay time leads, on average, to less spatial displacement between the production and decay vertex. This difference

Tab. 6.5: Decay chains considered in the amplitude $\bar{A}_{CF}(\Phi_n)$ describing the decay $\bar{D}^0 \rightarrow K^+\pi^-\pi^+\pi^-$ and their corresponding complex coefficients. Details on the given information is given in Table 6.4.

Decay Chain	Spin Factor	$\Re(c_i)$	$\Im(c_i)$	Fit Fraction
$K^*(892)^0[K^+\pi^-]\rho(770)[\pi^+\pi^-]$	1	1.000	0.000	0.0356
$[K^*(892)^0[K^+\pi^-]\rho(770)[\pi^+\pi^-]]^{\mathcal{P}-wave}$	2	1.531	-0.886	0.0172
$[K^*(892)^0[K^+\pi^-]\rho(770)[\pi^+\pi^-]]^{\mathcal{D}-wave}$	3	20.513	20.398	0.0726
$[K^*(1410)^0[K^+\pi^-]\rho(770)[\pi^+\pi^-]]$	2	4.001	-2.620	0.0093
$K^*(892)^0[K^+\pi^-]f_0(600)[\pi^+\pi^-]$	4	-0.770	-1.530	0.0468
$\rho(1450)[\pi^+\pi^-]K_0^*(1430)^0[K^+\pi^-]$	4	-0.110	1.814	0.0318
$K(1460)^+[K^*(892)^0[K^+\pi^-]\pi^+]\pi^-$	5	-0.696	0.326	0.0131
$K(1460)^+[f_0(1370)[\pi^+\pi^-]K^+]\pi^-$	6	-0.849	0.972	0.1135
$K_1(1270)^+[K^*(892)^0[K^+\pi^-]\pi^+]\pi^-$	7	0.601	-0.182	0.0038
$K_1(1270)^+[\rho(770)[\pi^+\pi^-]K^+]\pi^-$	7	-1.523	1.244	0.0190
$K_1(1270)^+[K^*(1430)^0[K^+\pi^-]\pi^+]\pi^-$	9	0.248	-0.088	0.0011
$K_1(1400)^+[K^*(892)^0[K^+\pi^-]\pi^+]\pi^-$	7	-0.808	-0.358	0.0070
$1^+[K^*(892)^0[K^+\pi^-]\pi^+]\pi^-$	7	-15.322	-12.089	0.0209
$K_2^*(1430)^+[K^*(892)^0[K^+\pi^-]\pi^+]\pi^-$	10	17.008	-5.014	0.0022
$K_2^*(1430)^+[\rho(770)[\pi^+\pi^-]K^+]\pi^-$	10	13.039	-1.935	0.0007
$a_1(1260)^-[\rho(770)[\pi^+\pi^-]\pi^-]K^+$	7	-0.639	-6.801	0.3417
$a_1(1260)^-[f_0(600)[\pi^+\pi^-]\pi^-]K^+$	9	-0.062	2.872	0.0621
$a_1(1260)^-[\rho(770)[\pi^+\pi^-]\pi^-]^{\mathcal{D}-wave}K^+$	8	-9.465	15.390	0.0122
$K^+\pi^-\pi^-\pi^+(\text{non-resonant})$	-	-0.265	-0.003	0.0312

Tab. 6.6: Necessary resonance properties for the model building. The listed mass, width, and spin of the resonances are taken from Ref. [121] with the exception of $K(1460)^-$ and $a_1(1260)^+$ whose properties are taken from Ref. [120].

Resonance	Mass [GeV]	Width [GeV]	Spin
$\rho(770)^0$	0.775	0.1478	1
$K^*(892)^0$	0.896	0.0474	1
$a_1(1260)^+$	1.237	0.5260	1
$K_1(1270)^+$	1.272	0.0900	1
$K_1(1400)^+$	1.403	0.1740	1
$K^*(1410)^0$	1.414	0.2320	1
$K^*(1430)^0$	1.425	0.2700	0
$K_2^*(1430)^+$	1.426	0.0985	2
$K(1460)^-$	1.351	0.2810	0

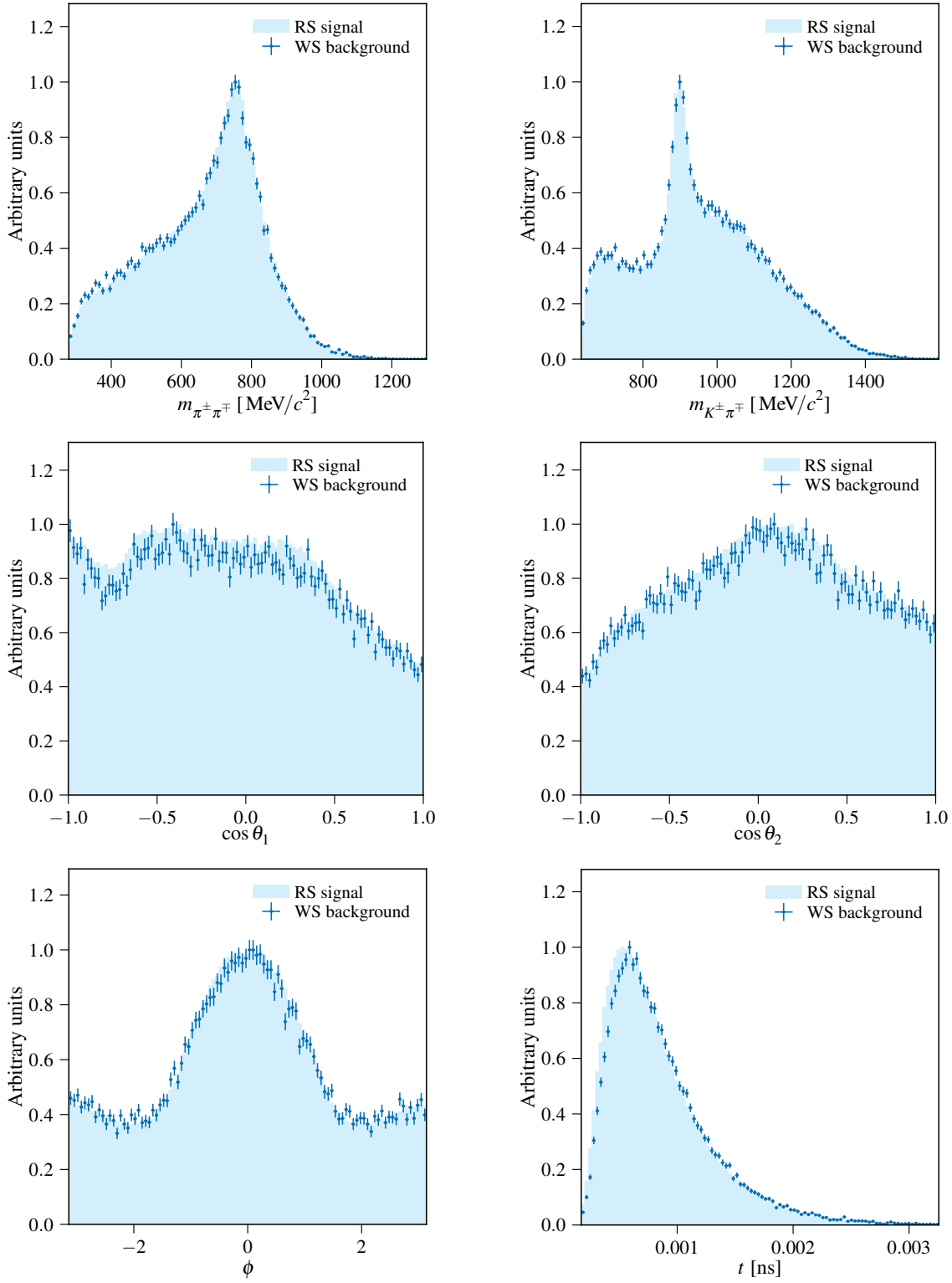


Fig. 6.10: Comparison of the variables describing the decay between the WS sample sidebands (as a proxy for the background in the signal region) and the signal region in the RS sample.

is caused by the wrong slow pion being used. The true slow pion's momentum is highly correlated with that of the D^0 due to the small available energy in D^{*+} decays and the selection applied to the slow pion candidate highly favours harder particles. Therefore, D^0 mesons matched to correct pion average a higher momentum than their incorrectly tagged counterparts whose momentum acceptance is not limited by the slow pion.

Due to the very good agreement, the fit will model the background contribution $\mathcal{P}_{\text{Bkg}}$ using the RS decay rate as given, for example, in Equation (2.39). Likewise, in pseudo-experiments events are sampled from this RS model. Lastly, the tiny differences in the decay time acceptance compared to actual RS decays are neglected for the scope of the sensitivity study. Therefore, the efficiencies $\varepsilon_{\text{bkg}}(\Phi_n, t)$ and $\varepsilon_{\text{sig}}(\Phi_n, t)$ are assumed to be identical. A model for these efficiencies is derived in the following.

6.2.3 A description for the efficiencies

The task in this section is the derivation of an efficiency model, parameterising it as a six-dimensional function of the five phase-space variables and the decay time. A classical approach to construct the efficiencies has already been discussed in the description of the cross-section measurements in Section 5.5: simulated samples are used to obtain reconstruction and most of the selection efficiencies with data driven corrections for tracking effects as well as a data-driven calibration technique for PID selection requirements. Especially concerning the simulated sample, the measurement discussed here is an ideal candidate to benefit from ReDecay to significantly reduce the required computing resources. Unfortunately, due to technical problems with the production of simulated samples for the latest LHC running conditions combined with strict prioritisation within the collaboration, no simulated samples could be obtained in time for this thesis.

This measurement does not need absolute efficiencies as any global factor again disappears in the normalisation of $\mathcal{P}'_{\text{Sig}}$. Therefore, only the shape differences between the distributions in the numerator and denominator of the efficiency are relevant. For the numerator, a basically background-free sample describing the shape is given by the events in the signal window of the RS sample. For the denominator, events generated from the RS model given in Table 6.5 are used, which implicitly includes the measurement which constructed the RS decay amplitude model as a prior in this work.

Figure 6.11 depicts the obtained efficiencies as a function of the different observables and in bins of another observable. While the decay time shows an expected increase with increasing decay time and only a small dependence on other variables (and

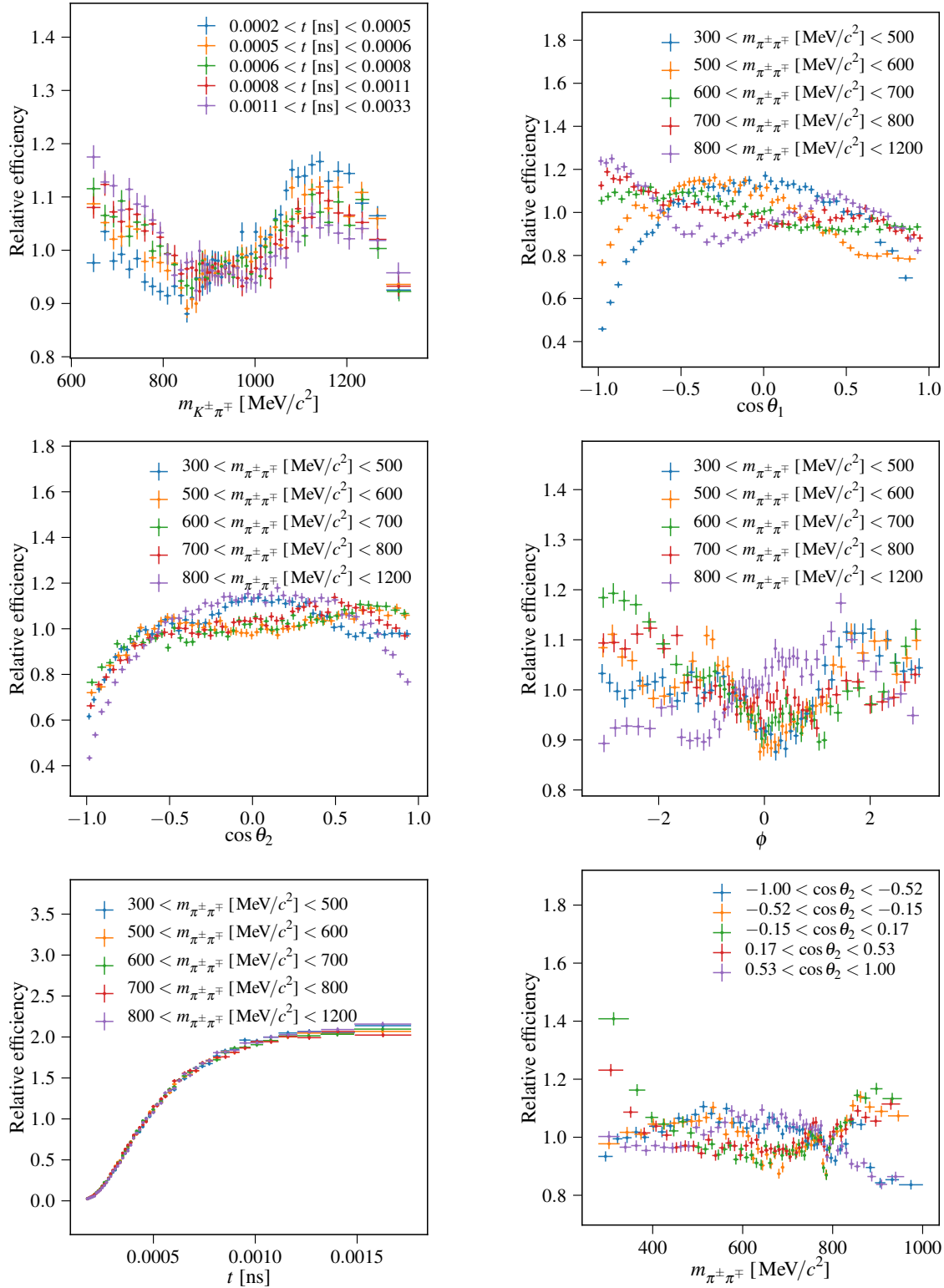


Fig. 6.11: Example selection of plots illustrating the shape of the relative efficiency as a function of the observables and in bins of another observable to illustrate that the efficiency does not factorise.

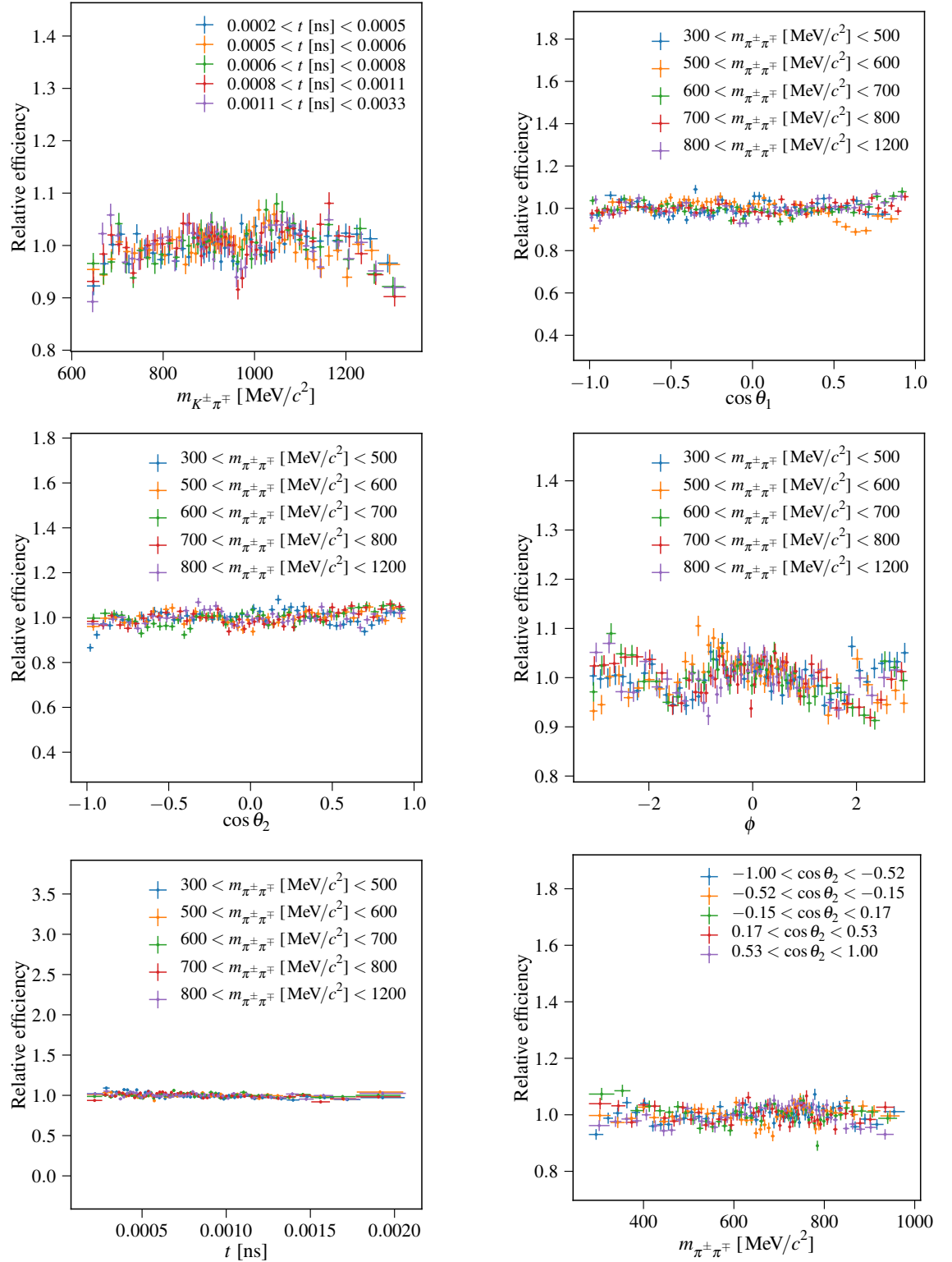


Fig. 6.12: Example selection of plots illustrating the shape of the relative efficiency as a function of the observables. For these plots the events in the denominator have been weighted using the result of the BDT reweighter, showing the very good quality of this non-parametric description when correcting the six-dimensional efficiency.

vice versa), the variations as a function of the phase-space variables lead to larger variations especially at the edges of their respective ranges where those variations are usually largest with an additional strong dependence on other phase-space variables. In order to obtain a six-dimensional efficiency model that can capture all this dependence, the same non-parametric approach as in Section 5.6.5 is used by training a BDT reweighter [112] to reweight events generated from the RS model to achieve a distribution that looks like the events in the signal region of the RS sample. For this training, an independent set of generated events is used for the denominator as well as only 20% of the respective RS data. A special treatment is necessary for the phase-space variable ϕ in order to ensure a consistent result for the efficiency, *i.e.* enforce identical efficiencies for $\phi = \pi$ and $\phi = -\pi$, the BDT is trained using the sine and cosine of ϕ instead of the angle itself. Figure 6.12 shows the result of this reweighting after training the BDT by showing the same plots as in Figure 6.11, but with the obtained efficiencies applied as weights to the generated events used in the denominator of the efficiency. As expected, after applying those weights the dependence of the relative efficiencies has essentially disappeared, pointing to a good description of the six-dimensional efficiency by the BDT.

6.2.4 External constraints in the fit

Adding additional constraints to the likelihood, while optional, can greatly improve the performance of the fit itself and reduce uncertainties on the parameters of interest. Furthermore, adding a constraining term for a secondary parameter, such as r , that is given by an external measurement, acts as an error propagation, projecting this uncertainty on the axis of the parameter of interest in the likelihood parameter space. The studies presented here will use two constraints, which, at their core, are implemented to tackle the same underlying ambiguity when extracting the mixing parameters with finite samples.

The first constrained parameter in the fit is r , which is the parameter describing the overall difference in magnitude of the CF and DCS amplitude. When looking at the terms with different powers in decay time in the time-dependent amplitude given in Equation (2.53), it is possible to separately absorb any increase or decrease of the relative magnitude between the CF and DCS amplitude in a redefinition of the mixing parameter. This is not true anymore when considering the entire time-dependent amplitude instead of the individual terms as a different transformation is needed for the quadratic and linear term. Nonetheless, due to the linear part being more sensitive, this still leads to a flattening of the $\ln \mathcal{L}$ over the x - y plane around the minimum which, in turn, leads to an increase in the reported uncertainties. In fact, tests using the sensitivity study set up presented in the next section reveal that, while the minimisation still converges, the uncertainties are increased by up to one order

of magnitude when no constraint is applied. The constraint itself is implemented as a normal distribution centered at $\mu = 1/\sqrt{300}$ with a width σ corresponding to a relative 2% uncertainty. The uncertainty corresponds to the one obtained in the model-independent measurement [46]. Similar to this constraint, another constraint is added to limit the freedom of the total background yield N_{Bkg} to the expected amount of background in the signal window, estimated from a fit to the full mass range as given in Section 6.1.5. This is necessary due to the nature of the background, which is comprised of RS decays. Therefore, the effect of mixing in the WS signal in the signal window could be compensated by the fitter by slightly increasing the amount of background and thus adding more RS likeness to the distribution similar to the effect of mixing itself. Consequently, this constraint is added to reduce any occurring bias in the fit.

Adding constraints on the coefficients of the used models for the CF and DCS decay amplitudes is also a possibility. However, they are not included for various reasons: the CF amplitude is obtained from a sample of RS decays recorded during Run 1 of the LHC and the number of events far exceeds the statistics used in the fit to the WS data in this work. Therefore, the CF amplitude is kept fixed in the fits presented in the next section. However, on the contrary, the used DCS amplitude was constructed from a much smaller sample and thus all constraints that could be added would have little effect.

Lastly, when generating pseudo-data during the studies in the following section it is important to remember that constraints added to the likelihood must be treated as measurements [122]. In practice, this implies that in every toy, a new central value for the mean μ of the constraint must be sampled to ensure the correct distribution of any pull statistic constructed from the ensemble of toy samples.

6.2.5 A few implementation details

The fit has been implemented in GooFit [123] to benefit from a substantial increase in speed provided by the use of graphics processing units (GPU) compared to classical CPUs. Details on the extension of GooFit as well as performance studies are discussed in Ref. [124]. The most relevant detail of the implementation in the context here is the handling of the integrals necessary to normalise $\mathcal{P}'_{\text{Sig}}$ and $\mathcal{P}'_{\text{Bkg}}$. By construction, no analytic expression for the integral across the five phase-space dimensions and the decay time can exist due to the non-parametric implementation of the efficiency description. Therefore, Monte Carlo integration is used (see Equation (4.2)) to integrate the six-dimensional function. A common technique to improve the precision of the integration is the use of importance sampling in the integration. This approach is based on the idea that the integration sample used to compute the integral does not necessarily need to follow a uniform distribution as Equation (4.2) can be modified

as follows:

$$\int f(\vec{x})d\vec{x} = \int \frac{f(\vec{x})}{g(\vec{x})} \cdot g(\vec{x})d\vec{x} \approx \overbrace{\int g(\vec{x})d\vec{x}}^{\equiv V} \cdot \frac{1}{N} \sum_{i=0}^N \frac{f(\vec{x}_i)}{g(\vec{x}_i)}, \quad (6.6)$$

where f is the function that is supposed to be integrated by using Monte Carlo integration, and g is any distribution that is non-zero over the entire integration space with an integral V . The entries $\vec{x}_i$ are random points sampled from g . Assuming g has an analytic integral, the precision can be improved if g resembles f as closely as possible. Therefore, the task is to find a distribution that resembles the six-dimensional, time-dependent amplitude model including the efficiencies as close as possible. For the decay-time dimension, g is chosen as a combination of a linear distribution with an exponential tail, whose exponent is chosen such that it falls less quickly than the decay time in the model. The resulting shape is compared to the decay time in the efficiency corrected model in Figure 6.13. For the integration of the phase-space dimensions, a very useful feature of Equation (6.6) in light of the application here, a likelihood fit, is exploited: V solely depends on the sampling distribution, not f . Therefore, the same factor V exists in front of every term that needs to be normalised in Equation (6.4), namely $\mathcal{P}'_{\text{Sig}}$ and $\mathcal{P}'_{\text{Bkg}}$. Therefore, the factor V can be pulled out of the product over all events, leading to a constant global factor V^N that has no effect on the minimisation as it drops out in the derivatives of the log-likelihood. Hence, it does not matter whether the integral of $g(\vec{x})$ is known or not as it only changes the height of the minimum, but not its position or shape. Thus, the importance sampling for the phase-space components uses a distribution that naturally closely resembles the distributions that need to be integrated: the RS model with the parameters as given in Table 6.5.

Nonetheless, the overall precision of the integration still gives rise to a potential systematic uncertainty. While the precision is proportional to $1/\sqrt{N}$, where N is the size of the integration sample, and could thus be eliminated by a sufficiently large sample, this is not applicable in the current setup of the fit. Due to memory limitations of the GPU, only around seven million integration events can be stored on the GPU and on-the-fly generation of integration events is not possible due to the non-parametric description of the efficiencies. The efficiencies can only be evaluated on the CPU and are thus included as a precomputed value together with each integration event that are loaded onto the GPU at the beginning of each fit. Performing these memory transfers is too costly to be performed per minimisation step, and thus it is unfeasible to access more than the seven million events originally loaded onto the GPU at the beginning of each fit. The effect of the chosen integration sample is investigated in the following studies and treated as a systematic uncertainty.

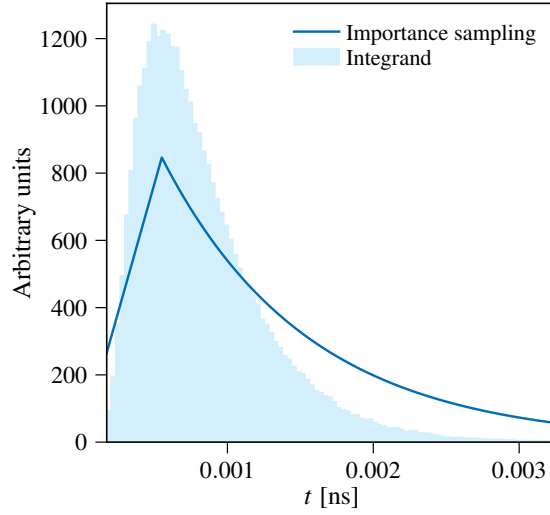


Fig. 6.13: Illustration of the importance sampling shape (orange) compared to the integrand (blue).

6.3 The sensitivity study

Having determined the number of signal and background events and discussed the setup of the mixing fit and the used models and shapes, this section presents results of sensitivity studies to shed light onto the expected statistical precision on the mixing parameters x''' and y''' . For this purpose, pseudo-experiments are performed where each sample is obtained by combining WS (for signal) and RS (for background) events in the correct mixture to achieve the same ratio of signal to background as in the signal window in the WS data sample.

As a first step, ten million events are randomly generated according to the models listed in Tables 6.4 and 6.5, and each event is assigned the respective efficiency from the BDT reweighter. The mixing parameters are set to the current world average values, $x''' = 0.0046$ and $y''' = 0.0062$, and $\tau = 0.4101$ ps is used for the lifetime. Additionally, no CP violation is assumed and hence p/q is set to one. To generate one pseudo-sample, new yields for the two species, N_{Sig} and N_{Bkg} , are randomly sampled from a Poisson distribution centred at the expected values for the combined 2015 and 2016 dataset as obtained in Section 6.1.5. For each of the species, events are then randomly taken out of the respective large sample (with replacement) and a hit-and-miss method using the efficiency is applied to obtain an unweighted sample following the efficiency-corrected model. This process is repeated until the respective, sample-specific number of signal and background events is obtained. Finally, the two species are combined and new values for the two constraints are sampled in each pseudo-experiment.

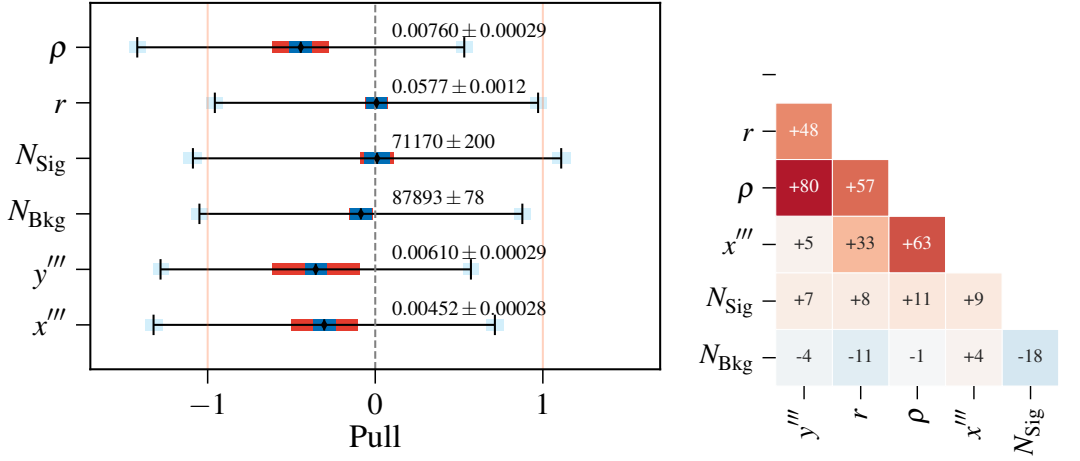


Fig. 6.14: Results of a fit of 200 pseudo-samples where the amplitude models $\bar{A}_{\text{CF}}(\phi_n)$ and $A_{\text{DCS}}(\phi_n)$ are fixed, leaving only five parameters left to float. On the left, a summary of the obtained pull distribution is given: the black marker indicates the mean of the obtained pulls for the respective parameter while the extent of the error bar illustrates the standard deviation. Ideally, the marker should be centred at zero with the standard deviation extending out to ± 1 . In addition, the light and dark blue areas indicate the uncertainty on both the standard deviation and the mean due to the finite number of pseudo-experiments while the red area indicates the spread of the pull mean from variation of the used integration sample. The obtained correlations between the parameters are given on the right in percent. The parameter ρ is defined as $\rho \equiv \sqrt{x'''^2 + y'''^2}$.

6.3.1 Fixed amplitude models

First sensitivity studies are performed by fixing all parameters in the two amplitude models, leaving only the number of signal and background events, N_{Sig} and N_{Bkg} , the ratio of the amplitudes r and the two mixing parameters x''' and y''' free. In total, 200 pseudo-samples are created and fitted. Figure 6.14 depicts the result of this sensitivity test. The pull distributions for the different parameters, where the pull is defined as:

$$\text{Pull} = \frac{\hat{\nu} - \nu^{\text{true}}}{\hat{\sigma}_{\nu}}, \quad (6.7)$$

for a parameter ν , generally show a good agreement with the desired case where the pull distribution is centred at zero with a standard deviation of one. The main parameters of interest, x''' and y''' , show a small bias towards values smaller than the respective true value of 20% and 10% of the statistical uncertainty respectively. By construction, the fit model is at least in theory capable of describing the pseudo-samples perfectly and thus potential sources for this bias need to be investigated. First of all, the fit model only describes the data perfectly if the individual components of the fit are correctly normalised. However, as noted before, this is not strictly true in this case due to the finite number of integration events being used. The effect of the integration sample is investigated by repeating the fit multiple times with an

identical pseudo-sample but using statistically independent sets of integration events. The standard deviation of a parameter is then taken as a measure for accessing the impact of the integration sample choice onto the specific parameter. While only negligible effects are seen for N_{Sig} , N_{Bkg} and r , the deviations correspond to 15% of the total statistical uncertainty on x''' and y''' . This deviation is also indicated in Figure 6.14. It accounts for almost all of the visible bias on the respective parameters and is thus assigned as a systematic uncertainty. The final statistical uncertainties on the mixing parameters are:

$$\sigma_{x'''} = 0.028 \cdot 10^{-2} \text{ and} \quad (6.8)$$

$$\sigma_{y'''} = 0.029 \cdot 10^{-2} . \quad (6.9)$$

The correlation table, which is also given in Figure 6.14, shows a negligible amount of correlation between the two parameters.

6.3.2 Floating DCS amplitude model

Section 6.2.1 already highlighted the fact that, while being a good approximation, the used $A_{\text{DCS}}(\Phi_n)$ is not an actual DCS model but rather a WS model as it includes the effects of mixing. Therefore, its coefficients will need to float in the fit to actually extract the amplitude model at $t = 0$, the true DCS model. To investigate the effect on the expected sensitivity, which might be reduced by strong correlations between mixing parameters and coefficients, all but the coefficient of the $K^*(892)^0[K^+\pi^-]\rho(770)[\pi^+\pi^-]$ resonance in the $\mathcal{S}$ -wave configuration are allowed to float in the fit. This is deliberately done to ensure the fit converges. Floating the magnitude of all coefficients would result in a disambiguity as the overall magnitude of the DCS model is absorbed when normalising it to ensure a consistent definition of r . Additionally, the unlocking of the phase of all coefficients allows a rotation of the entire amplitude by a global phase, which also results in ambiguity with the mixing parameters x''' and y''' as discussed in Section 2.3.4.

The results of these pseudo-experiments are presented in the same format as before in Figures 6.15 and 6.16, which show the statistics of the pull distribution for all relevant parameters as well as the correlation matrix. Most importantly, the general picture for the mixing parameters, x''' and y''' , has slightly improved as they now show good agreement with the true value. Furthermore, the uncertainties obtained from the fit have only marginally changed. This behaviour is confirmed by the correlation pattern as x''' and y''' are only mildly correlated with the additional shape parameters. Overall, the extracted coefficients are in decent agreement with expectations while a much larger dependence on the integration sample is seen. This is consistent with the previous fit where the dependence on the integration sample

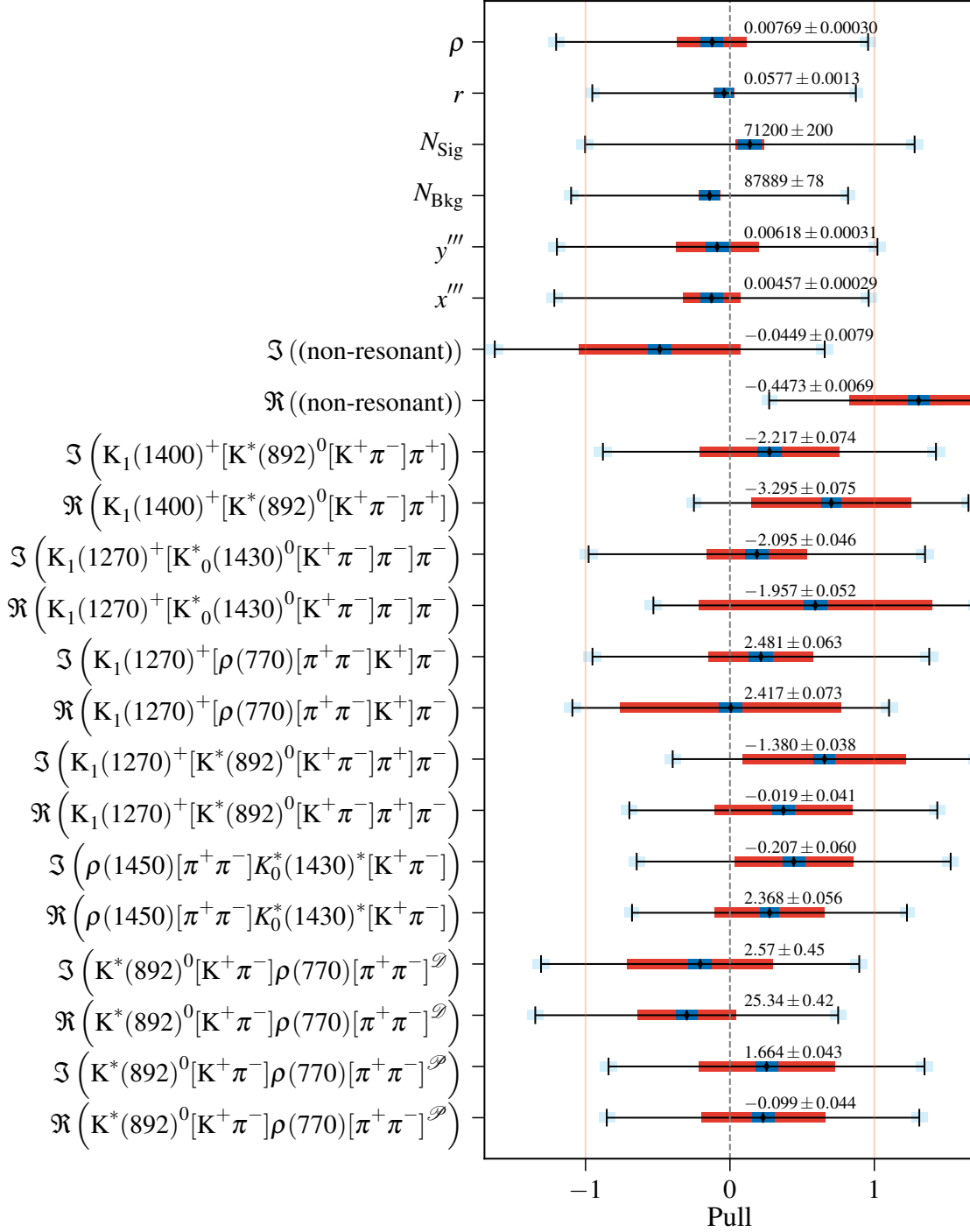


Fig. 6.15: Results of a fit of 200 toys where the amplitude model $\bar{A}_{\text{CF}}(\Phi_n)$ is fixed and all coefficients in $A_{\text{DCS}}(\Phi_n)$ (except the one for $K^*(892)^0[K^+\pi^-]\rho(770)[\pi^+\pi^-]$ in $\mathcal{S}$ -wave configuration as explained in the text) are floating. See Figure 6.14 for a description.

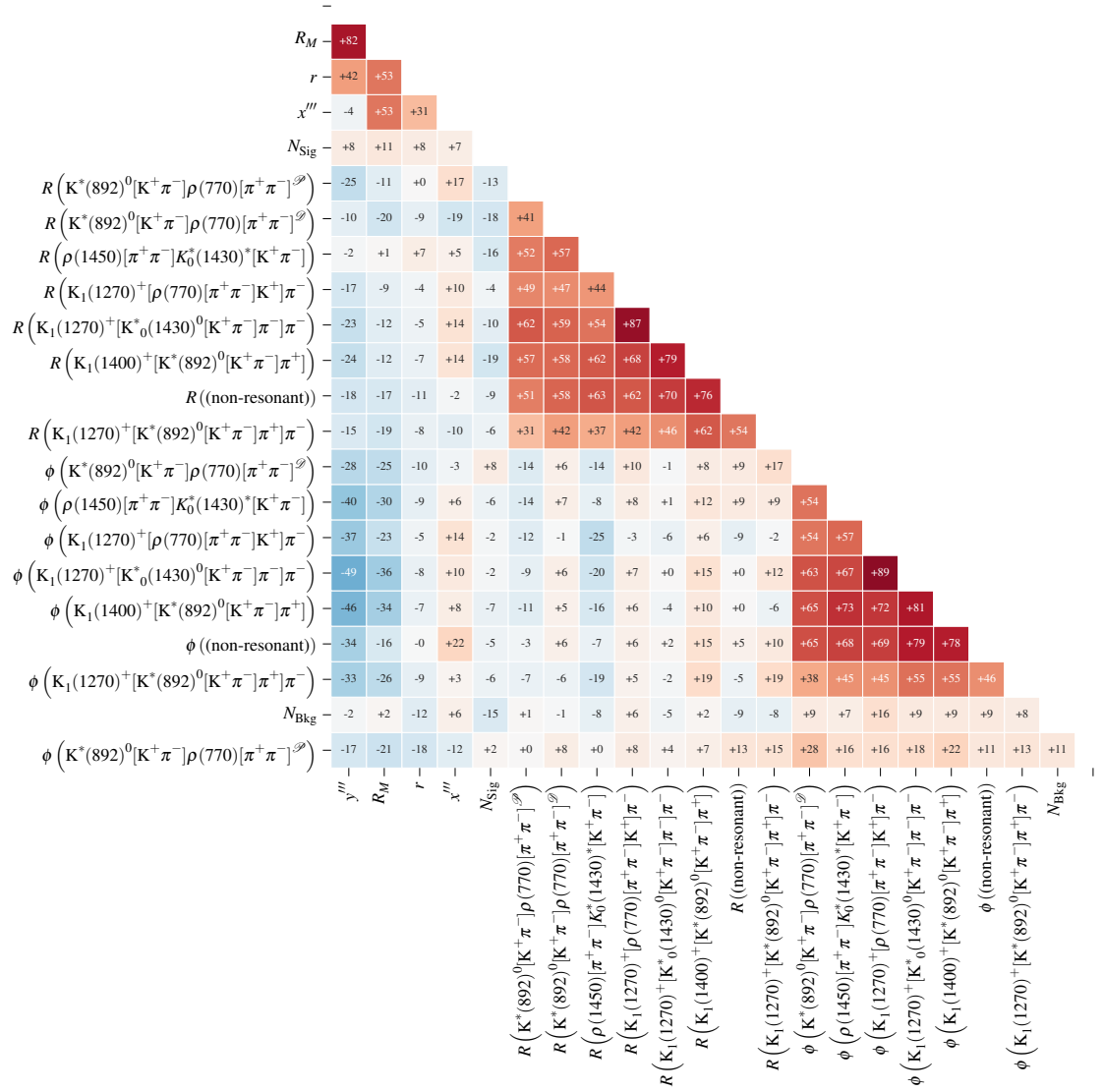


Fig. 6.16: Correlation matrix for a toy fit with all coefficients in the DCS model floating. For better readability, the shown parameters are automatically sorted to create large blocks of similar correlation.

had to be absorbed into the mixing parameters, leading to the bias seen while in this fit the bias from the integration sample is contained in the model coefficients.

In summary, the expected statistical precision on the mixing parameters when floating all coefficients of the $A_{\text{DCS}}(\Phi_n)$ amplitude model are:

$$\sigma_{x'''} = 0.029 \cdot 10^{-2} \text{ and} \quad (6.10)$$

$$\sigma_{y'''} = 0.031 \cdot 10^{-2} . \quad (6.11)$$

6.3.3 Systematic uncertainties

In the following, different sources of systematic uncertainty are discussed, which have been identified as leading sources in other model-dependent mixing measurements [125]. As most of these sources rely on introducing modifications to the amplitude models $A_{\text{DCS}}(\Phi_n)$ and $\bar{A}_{\text{CF}}(\Phi_n)$ used in the fit, one problem arises: modified models lead to a different phase difference, which changes the unknown global phase between $A_{\text{DCS}}(\Phi_n)$ and $\bar{A}_{\text{CF}}(\Phi_n)$. As Section 2.3.4 detailed, this global phase is absorbed in a redefinition of the mixing parameters. However, a modified model leads to a different global phase difference, which implicitly changes the definition of the mixing parameters in the fit. Therefore, the introduction of a systematic variation can change the mixing parameters due to a redefinition of the parameters themselves or due to an actual systematic effect altering their values, where only the latter is of interest when accessing the impact of systematic variations. To estimate this impact, the quantity:

$$\rho \equiv \sqrt{x'''^2 + y'''^2} = \sqrt{x^2 + y^2} , \quad (6.12)$$

which is invariant under a redefinition⁵ of the mixing parameters, is used instead. The statistical precision on this parameter in the scenario where the coefficients of the $A_{\text{DCS}}(\Phi_n)$ model are floating is:

$$\sigma_\rho = 0.029 \cdot 10^{-2} . \quad (6.13)$$

To remove statistical effects, the variation of ρ with systematic effects is studied using the same set of pseudo-experiments as before, producing one deviation per pseudo-sample. The total systematic effect is then taken as the average of these deviations over all pseudo-experiments. The estimates given in the following are intended purely to illustrate the expected sizes and do not aim to be accurate enough for an actual measurement. Furthermore, resolution-related effects cannot be studied

⁵ The redefinition due to an unknown global phase difference between $A_{\text{DCS}}(\Phi_n)$ and $\bar{A}_{\text{CF}}(\Phi_n)$ corresponds to a rotation of x''' and y''' .

due to the lack of a simulated sample to study the resolution on the phase-space variables and the decay time.

Meson radius in Blatt-Weisskopf form factors

The meson radius in the Blatt-Weisskopf form factors of 5 GeV^{-1} for the D^0 and 1.5 GeV^{-1} for the resonances are chosen rather arbitrarily here as no dedicated measurements exist to date. To access the impact this choice can have on the mixing parameters, the previous pseudo-experiments are repeated with fit models where the meson radii of the D^0 meson and the resonances are separately changed by $\pm 4\%$. When varying the D^0 radius, the average absolute change (averaging the up and down variation) of ρ is $0.0028 \cdot 10^{-2}$ which corresponds to 9% of the total statistical precision while a shift of $0.0085 \cdot 10^{-2}$ (28% of the statistical uncertainty) is seen when varying the resonance radius. Taking both values in quadrature, this source of systematic uncertainty is estimated to contribute 30% of the statistical uncertainty.

Modifying the resonances used in the $A_{\text{DCS}}(\Phi_n)$ model

Extending this measurement to use real data requires the construction of the amplitude model that includes the decision on which resonances to use as multiple solutions usually return a similar fit quality. To get an idea of the expected effect on the obtained mixing parameters in the sensitivity study, the pseudo-experiments are repeated after adding additional resonances found in the $\bar{A}_{\text{CF}}(\Phi_n)$ model, which are not part of the $A_{\text{DCS}}(\Phi_n)$ model used when generating the pseudo-samples. Several different resonances are tried one at a time and the largest shift is seen when adding $K^*(892)^0[K^+\pi^-]f_0(600)[\pi^+\pi^-]$, which changes the obtained value for ρ by $0.0076 \cdot 10^{-2}$, corresponding to 25% of the statistical uncertainty.

Efficiency description

The non-parametric efficiency description obtained from a trained BDT reweighter introduces multiple potential sources of systematic uncertainty: a statistical component due to the size of the sample used for the training, an uncertainty due to the chosen inputs for the training, as well as the quality of the training itself. The first source of uncertainty due to the finite size of the training sample is assumed to be negligible: when real RS data is used in the training, the potential sample size far exceeds the WS sample which is fitted. For the second source, this picture changes: while the numerator of the efficiency is correct by construction if real data is used, the assumed model for the denominator comes with its own set of uncertainties (on the coefficients, lineshape parameters *etc.*). However, for the efficiency computation, only the distribution in the six variables is necessary, which is given by the magnitude of the amplitude and is not sensitive to phases or the used resonances as long as the

overall shape remains unchanged. Therefore, the sensitivity to modelling effects is expected to be significantly reduced. Nonetheless, to propagate the uncertainties on the model to an uncertainty on the obtained mixing parameters it is necessary to use pseudo-experiments where different non-parametric descriptions are tested. These were trained using models for the denominator with varied values for the parameters as well as different settings for the BDT itself. However, such a study is beyond the scope of this sensitivity study and therefore, based on the good description seen in Figure 6.12, the effect is expected to be small and assumed to be at most 20% of the statistical uncertainty in the following.

Integration precision

As already discussed earlier, the finite size of the sample used for the normalisation is another potential source of systematic uncertainty. It is accessed by repeating the fit to the same pseudo-sample, *i.e.* with the identical events in it as well as the same central values for the constraints, with statistically independent integration samples. Overall, the deviation of ρ across all integration samples is found to be 15% of the statistical uncertainty. Validation studies reveal that this spread does not depend on the pseudo-sample that is chosen for this study.

Summary

Adding the four different sources of systematic uncertainty in quadrature yields a total systematic uncertainty of 46% on ρ . It is noteworthy here that none of the sources of systematic uncertainty discussed above is linked to the size of the data sample and is expected to decrease with increasing data. Therefore, with more data becoming available, a real measurement is expected to become systematically dominated very quickly.

6.4 Summary – impact on world average of x and y

A realistic estimate is obtained for the statistical uncertainty on the mixing parameters x''' and y''' using $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ decays. These studies rely on extensive work based on real data to estimate the number of signal events in the combination of the 2015 and 2016 dataset recorded by LHCb. Additionally, an appropriate non-parameteric description of the reconstruction and selection efficiency is derived. Furthermore, a discussion on potential systematic uncertainties together with rough estimates is given. In conclusion, the final result of the sensitivity study is:

$$\sigma_{x'''} = 0.029 \cdot 10^{-2}(\text{stat.}) \pm 0.026 \cdot 10^{-2}(\text{syst.}) \text{ and} \quad (6.14)$$

$$\sigma_{y'''} = 0.032 \cdot 10^{-2}(\text{stat.}) \pm 0.029 \cdot 10^{-2}(\text{syst.}) , \quad (6.15)$$

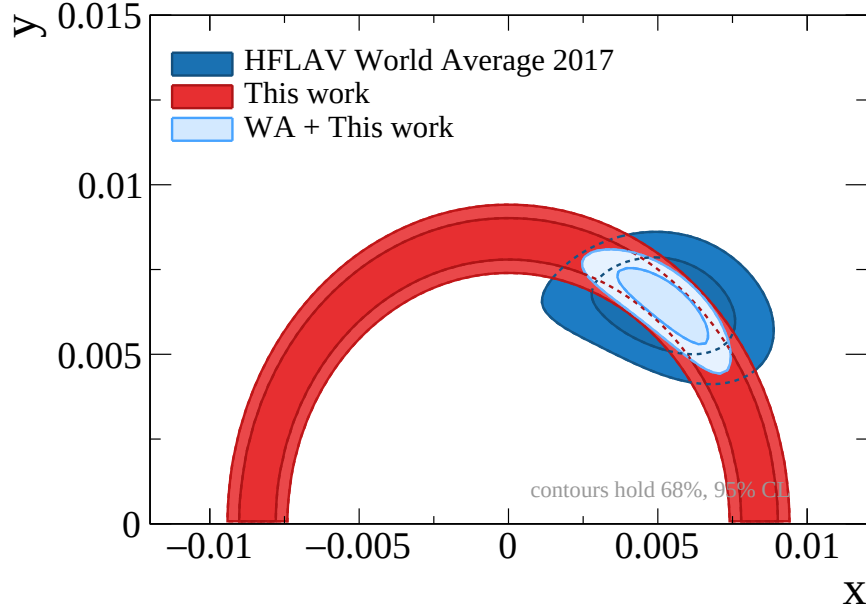


Fig. 6.17: Predicted impact of a measurement of x''' and y''' in $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ [126].

where it is assumed that the uncertainty expressed as a fraction of the statistical uncertainty on ρ can be applied to x''' and y''' . Additionally, the systematic uncertainty is inflated by a factor of two to take into account contributions from potential sources that cannot be investigated on the scope of this sensitivity study. If a careful study of the systematic uncertainties in the actual measurement reveal a similar precision this would constitute the single most precise measurement of the mixing parameters to date. However, as the mixing parameters are only measured with an unknown rotation they will only contribute a ring in the x - y plane to the world average fit, albeit with very high constraining power in the radial direction.

The impact on the world average (see Section 2.3.6) is illustrated in Figure 6.17, where a substantial improvement is seen along the radial direction, as expected. This sensitivity study can also be used to assess the expected sensitivity on p/q , the parameter describing the effect of indirect CP violation, which has been assumed to be one so far. By reintroducing the parameter:

$$\frac{q}{p} = \left| \frac{q}{p} \right| e^{i\phi} \quad (6.16)$$

to Equation (2.50), modified transformations of the mixing parameters can be obtained:

$$x_+''' = \left| \frac{q}{p} \right| x \cos(\delta_0 - \phi) + \left| \frac{q}{p} \right| y \sin(\delta_0 - \phi) , \quad (6.17)$$

$$y_+''' = \left| \frac{q}{p} \right| y \cos(\delta_0 - \phi) - \left| \frac{q}{p} \right| x \sin(\delta_0 - \phi) , \quad (6.18)$$

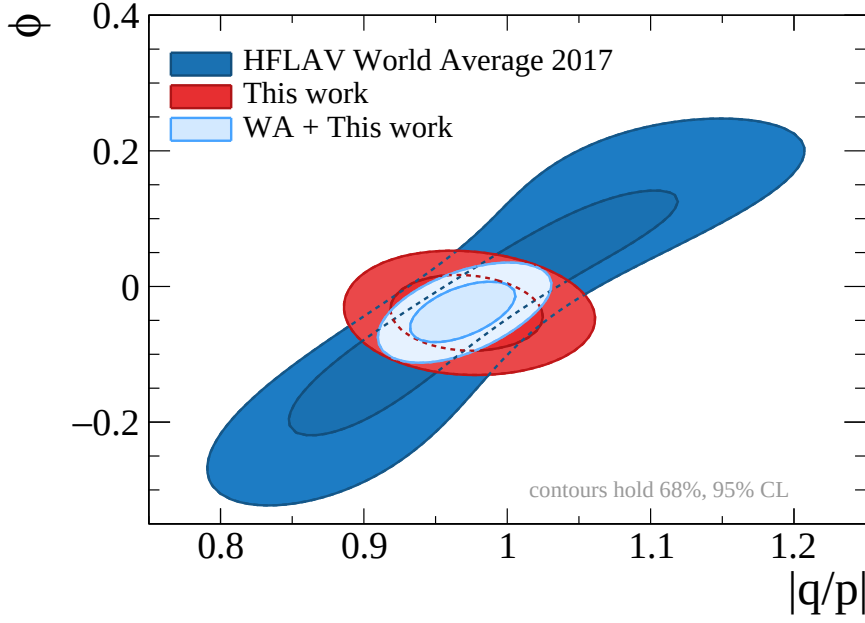


Fig. 6.18: Impact of a measurement of $\left|\frac{q}{p}\right|$ and ϕ as in the studies presented in this work on the world average [126]. The estimate based on this work is constructed as described in the text.

$$x_-''' = \left|\frac{p}{q}\right| x \cos(\delta_0 + \phi) + \left|\frac{p}{q}\right| y \sin(\delta_0 + \phi) , \quad (6.19)$$

$$y_-''' = \left|\frac{p}{q}\right| y \cos(\delta_0 + \phi) - \left|\frac{p}{q}\right| x \sin(\delta_0 + \phi) , \quad (6.20)$$

where x_+''' and y_+''' are the transformations for $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ decays and x_-''' and y_-''' are those for $\bar{D}^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ decays. These transformations absorb not only the unknown strong phase difference between the decay amplitudes but also the weak phase difference and absolute values of q/p in a redefinition of the mixing parameters. It is therefore possible to estimate both $|q/p|$ and ϕ by splitting the WS sample into D^0 and $\bar{D}^0$ decays and performing the fit in each of these samples to extract the respective mixing parameters. For the sensitivity study it is assumed that the sample contains the same amount of D^0 and $\bar{D}^0$, resulting in an increase of the statistical uncertainty by a factor $\sqrt{2}$ with respect to the values given in Equations (6.15) and (6.15). Figure 6.18 shows the results of this sensitivity study, which assumes that the mixing parameters for both the D^0 and $\bar{D}^0$ sub-sample are set such that the resulting two-dimensional likelihood is at its maximal in the best-fit point of the world average. The impact on the world average is again substantial.

In conclusion, the presented sensitivity studies reveal great potential for a future measurement of neutral charm meson mixing using $D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ decays in a model-dependent approach. Such an extension can rely on the data-driven work undertaken for the studies discussed in this work, namely a finalised event selection

procedure, the determination of a suitable background model and a description of the efficiencies in the fit.

7. A glimpse into the future

At the end of each of the previous chapters, a summary and conclusion of the respective results has already been given. Instead of simply reiterating the points already made, this thesis concludes by collecting the different aspects and glimpsing into a hopefully optimistic future for charm physics at the LHC.

The measurements of charm meson production cross-sections in Chapter 5 reveal very large cross-sections for the D^0 production in pp collisions at $\sqrt{s} = 13$ TeV of around 2 mb. To put this into perspective, the total inelastic pp cross-section at the same energy is around 80 mb [127]. This implies that in the current LHC running period, where LHCb sees, on average, one pp collision per bunch crossing, 2% of all bunch crossings in the detector contain a D^0 meson in the LHCb acceptance. Considering the plans for the LHCb upgrade, which involves an increase of the number of pp collisions per bunch crossing to five, the fraction of bunch crossings containing a D^0 in the acceptance will rise to 12%, corresponding to about three million D^0 per second at full intensity. This truly illustrates LHCb's great potential as a charm factory.

This potential is exploited in Chapter 6, which finds an expected statistical precision on the mixing parameters that is a factor three more precise in $\sqrt{x^2 + y^2}$ than the current world average, solely based on the data that were recorded by LHCb during 2015 and 2016, corresponding to an integrated luminosity of 2 fb^{-1} . The entire luminosity recorded during Run 2 is expected to amount to 5 fb^{-1} , which will improve the statistical uncertainty further by a factor of roughly $\sqrt{2}$. Even more so, current projections expect an integrated luminosity of 50 fb^{-1} during Run 3 of the LHC with the LHCb upgrade installed, improving the statistical precision by a factor of five compared to the values discussed in this work.

At such a statistical precision, the measurements are expected to be dominated by systematic uncertainties. Especially amplitude-model related uncertainties, which gave the largest contribution in the studies presented here, will not reduce with more data. Additionally, the limitations of the isobar approach will become more apparent, requiring updated theoretical descriptions to further enhance the precision. Furthermore, the large datasets require very large simulated samples to understand resolution and acceptance effects in an analysis. The ReDecay approach, which is presented in this thesis, is an ideal solution, providing an order of magnitude increase in speed at the full precision of the nominal simulation. This allows the systematic uncertainty due to the size of the simulated samples to be kept small relative to the

statistical precision, a crucial requirement to be able to benefit from the large data samples.

Closing the circle to the introduction of this thesis, the substantial improvement in the precision on the mixing parameters implies a similar increase in the sensitivity to indirect CP violation. Experimental focus is expected to shift from measurements of mixing parameters to those sensitive to differences between D^0 and $\bar{D}^0$. CP violation in the charm sector is the last missing piece in the Standard Model quark CP violation. With the much larger sample and resulting increased sensitivity, there is hope that this missing piece will be determined and will not fit into the puzzle, hence providing a hint at physics beyond the Standard Model. However, only more data can tell if this is the case or not.

A. References

- [1] LHCb collaboration, R. Aaij *et al.*, *Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV*, JHEP **03** (2016) 159, Erratum ibid. **09** (2016) 013, Erratum ibid. **05** (2017) 074, [arXiv:1510.01707](#).
- [2] LHCb collaboration, R. Aaij *et al.*, *Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 5$ TeV*, JHEP **06** (2017) 1, [arXiv:1610.02230](#).
- [3] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. **C40** (2016) 100001.
- [4] R. Davis, D. S. Harmer, and K. C. Hoffman, *Search for neutrinos from the sun*, Phys. Rev. Lett. **20** (1968) 1205.
- [5] SNO collaboration, Q. R. Ahmad *et al.*, *Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **87** (2001) 071301.
- [6] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.
- [7] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964) 168.
- [8] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579 .
- [9] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.
- [10] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.
- [11] ATLAS collaboration, 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*, Phys. Lett. **B716** (2012) 1, [arXiv:1207.7214](#).
- [12] CMS collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, [arXiv:1207.7235](#).
- [13] Belle collaboration, S.-K. Choi *et al.*, *Observation of a Resonancelike Structure in the $\pi^\pm\psi'$ Mass Distribution in Exclusive $B \rightarrow K\pi^\pm\psi'$ Decays*, Phys. Rev. Lett. **100** (2008) 142001.
- [14] LHCb collaboration, R. Aaij *et al.*, *Observation of the resonant character of the $Z(4430)^-$ state*, Phys. Rev. Lett. **112** (2014) 222002, [arXiv:1404.1903](#).
- [15] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \rightarrow J/\psi p K^-$ decays*, Phys. Rev. Lett. **115** (2015) 072001, [arXiv:1507.03414](#).

- [16] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett.**10** (1963) 531.
- [17] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progr. Theoret. Phys. **49** (1973), no. 2 652.
- [18] Belle collaboration, K. Abe *et al.*, *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802, [arXiv:hep-ex/0107061](#).
- [19] BaBar collaboration, B. Aubert *et al.*, *Observation of CP violation in the B^0 meson system*, Phys. Rev. Lett. **87** (2001) 091801, [arXiv:hep-ex/0107013](#).
- [20] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.
- [21] CDF collaboration, D. Acosta *et al.*, *Measurement of prompt charm meson production cross-sections in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. Lett. **91** (2003) 241804, [arXiv:hep-ex/0307080](#).
- [22] ALICE collaboration, B. Abelev *et al.*, *Measurement of charm production at central rapidity in proton-proton collisions at $\sqrt{s} = 2.76$ TeV*, JHEP **07** (2012) 191, [arXiv:1205.4007](#).
- [23] ALICE collaboration, B. Abelev *et al.*, *D_s^+ meson production at central rapidity in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B718** (2012) 279, [arXiv:1208.1948](#).
- [24] ALICE collaboration, B. Abelev *et al.*, *Measurement of charm production at central rapidity in proton-proton collisions at $\sqrt{s} = 7$ TeV*, JHEP **01** (2012) 128, [arXiv:1111.1553](#).
- [25] ALICE collaboration, J. Adam *et al.*, *D-meson production in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV and in pp collisions at $\sqrt{s} = 7$ TeV*, [arXiv:1605.07569](#).
- [26] ATLAS collaboration, G. Aad *et al.*, *Measurement of $D^{*\pm}$, $D^\pm$ and $D_s^\pm$ meson production cross sections in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Nucl. Phys. **B907** (2016) 717, [arXiv:1512.02913](#).
- [27] LHCb collaboration, R. Aaij *et al.*, *Prompt charm production in pp collisions at $\sqrt{s} = 7$ TeV*, Nucl. Phys. **B871** (2013) 1, [arXiv:1302.2864](#).
- [28] NNPDF collaboration, R. D. Ball *et al.*, *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, [arXiv:1410.8849](#).
- [29] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, Sov. Phys. JETP **46** (1977) 641.
- [30] V. N. Gribov and L. N. Lipatov, *Deep inelastic ep scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438.
- [31] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. **B126** (1977) 298.

-
- [32] H.-L. Lai *et al.*, *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, [arXiv:1007.2241](#).
 - [33] R. Gauld, J. Rojo, L. Rottoli, and J. Talbert, *Charm production in the forward region: constraints on the small- x gluon and backgrounds for neutrino astronomy*, JHEP **11** (2015) 009, [arXiv:1506.08025](#).
 - [34] IceCube collaboration, A. Achterberg *et al.*, *First Year Performance of The IceCube Neutrino Telescope*, Astropart. Phys. **26** (2006) 155, [arXiv:astro-ph/0604450](#).
 - [35] IceCube collaboration, M. G. Aartsen *et al.*, *The IceCube Neutrino Observatory - Contributions to ICRC 2015 Part II: Atmospheric and Astrophysical Diffuse Neutrino Searches of All Flavors*, in *Proceedings, 34th International Cosmic Ray Conference (ICRC 2015): The Hague, The Netherlands, July 30-August 6, 2015*, 2015. [arXiv:1510.05223](#).
 - [36] LHCb collaboration, *Study of Cold Nuclear Matter effect with prompt D^0 meson production in pPb collisions at LHCb*, LHCb-CONF-2016-003.
 - [37] M. Cacciari, M. L. Mangano, and P. Nason, *Gluon PDF constraints from the ratio of forward heavy-quark production at the LHC at $\sqrt{s} = 7$ and 13 TeV*, Eur. Phys. J. **C75** (2015) 610, [arXiv:1507.06197](#).
 - [38] B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, *Inclusive charmed-meson production at the CERN LHC*, Eur. Phys. J. **C72** (2012) 2082, [arXiv:1202.0439](#).
 - [39] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: The POWHEG BOX*, JHEP **06** (2010) 043, [arXiv:1002.2581](#).
 - [40] T. Sjöstrand *et al.*, *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, [arXiv:1410.3012](#).
 - [41] T. Kneesch, B. A. Kniehl, G. Kramer, and I. Schienbein, *Charmed-meson fragmentation functions with finite-mass corrections*, Nucl. Phys. **B799** (2008) 34, [arXiv:0712.0481](#).
 - [42] B. A. Kniehl and G. Kramer, *Charmed-hadron fragmentation functions from CERN LEP1 revisited*, Phys. Rev. **D74** (2006) 037502, [arXiv:hep-ph/0607306](#).
 - [43] T. Inami and C. S. Lim, *Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes $K_L^0 \rightarrow \mu^- MIP$, $K^+ \rightarrow \pi^+ \nu \bar{\nu}$ and $K^0 \leftrightarrow \bar{K}^0$* , Prog. Theor. Phys. **65** (1981) 297, [Erratum: Prog. Theor. Phys.65,1772(1981)].
 - [44] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. **D2** (1970) 1285.
 - [45] LHCb collaboration, R. Aaij *et al.*, *Observation of $D^0 - \bar{D}^0$ oscillations*, Phys. Rev. Lett. **110** (2013) 101802, [arXiv:1211.1230](#).
 - [46] LHCb collaboration, R. Aaij *et al.*, *First observation of $D^0 - \bar{D}^0$ oscillations in $D^0 \rightarrow K^+ \pi^+ \pi^- \pi^-$ decays and a measurement of the associated coherence parameters*, Phys. Rev. Lett. **116** (2016) 241801, [arXiv:1602.07224](#).

- [47] Y. Amhis *et al.*, *Averages of b -hadron, c -hadron, and τ -lepton properties as of summer 2016*, [arXiv:1612.07233](https://arxiv.org/abs/1612.07233), and online updates at <http://www.slac.stanford.edu/xorg/hflav/>.
- [48] D. J. Herndon, P. Söding, and R. J. Cashmore, *Generalized isobar model formalism*, Phys. Rev. D **11** (1975) 3165.
- [49] J. M. Blatt and V. F. Weisskopf, *Theoretical nuclear physics*, John Wiley & Sons, 1952.
- [50] F. Von Hippel and C. Quigg, *Centrifugal-barrier effects in resonance partial decay widths, shapes, and production amplitudes*, Phys. Rev. **D5** (1972) 624.
- [51] C. Zemach, *Use of angular-momentum tensors*, Phys. Rev. **140** (1965) B97.
- [52] C. Zemach, *Determination of the spins and parities of resonances*, Phys. Rev. **140** (1965) B109.
- [53] V. Filippini, A. Fontana, and A. Rotondi, *Covariant spin tensors in meson spectroscopy*, Phys. Rev. D **51** (1995) 2247.
- [54] B. S. Zou and D. V. Bugg, *Covariant tensor formalism for partial wave analyses of ψ decay to mesons*, Eur. Phys. J. **A16** (2003) 537, [arXiv:hep-ph/0211457](https://arxiv.org/abs/hep-ph/0211457).
- [55] M. Williams, *Measurement of differential cross sections and spin density matrix elements along with a partial wave analysis for $\gamma p \rightarrow p\omega$ using CLAS at Jefferson Lab*, PhD thesis, Carnegie Mellon U., 2007.
- [56] S. U. Chung, *Spin formalisms; 3nd. updated version*, Brookhaven Nat. Lab., Upton, NY, 2013. This report is an updated version of CERN-71-08.
- [57] P. D'Argent, *Amplitude analysis of $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$ decays at the LHCb experiment*, PhD thesis, Heidelberg U., Dec, 2014.
- [58] D. Aston *et al.*, *A Study of $K^- \pi^+$ Scattering in the Reaction $K^- p \rightarrow K^- \pi^\pm n$ at 11-GeV/c*, Nucl. Phys. **B296** (1988) 493.
- [59] D. V. Bugg, *The Mass of the sigma pole*, J. Phys. **G34** (2007) 151, [arXiv:hep-ph/0608081](https://arxiv.org/abs/hep-ph/0608081).
- [60] H. Kamano, S. X. Nakamura, T. S. H. Lee, and T. Sato, *Unitary coupled-channels model for three-mesons decays of heavy mesons*, Phys. Rev. **D84** (2011) 114019, [arXiv:1106.4523](https://arxiv.org/abs/1106.4523).
- [61] P. C. Magalhaes *et al.*, *Towards three-body unitarity in $D^+ \rightarrow K^- \pi^+ \pi^+$* , Phys. Rev. **D84** (2011) 094001, [arXiv:1105.5120](https://arxiv.org/abs/1105.5120).
- [62] J.-P. Dedonder, R. Kaminski, L. Lesniak, and B. Loiseau, *Dalitz plot studies of $D^0 \rightarrow K_S^0 \pi^+ \pi^-$ decays in a factorization approach*, Phys. Rev. **D89** (2014) 094018, [arXiv:1403.2971](https://arxiv.org/abs/1403.2971).
- [63] J. T. Daub, C. Hanhart, and B. Kubis, *A model-independent analysis of final-state interactions in $\overline{B}_{d/s}^0 \rightarrow J/\psi \pi \pi$* , JHEP **02** (2016) 009, [arXiv:1508.06841](https://arxiv.org/abs/1508.06841).
- [64] CERN, *TE-EPC-LPC in LHC*, <http://te-epc-lpc.web.cern.ch/te-epc-lpc/machines/lhc/general.stm>. Accessed: 25.09.2014.

-
- [65] L. Evans and P. Bryant, *LHC Machine*, JINST**3** (2008) S08001.
- [66] CERN, *The accelerator complex*, <http://home.web.cern.ch/about/accelerators>. Accessed: 24.07.2014.
- [67] K. Wille, *The Physics of Particle Accelerators: An Introduction*, Oxford University Press, Oxford, 5, 2001.
- [68] ATLAS collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
- [69] CMS collaboration, *The CMS experiment at the CERN LHC*, JINST**3** (2008) S08004.
- [70] ALICE collaboration, *The ALICE experiment at the CERN LHC*, JINST**3** (2008) S08002.
- [71] LHCb collaboration, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.
- [72] LHCb collaboration, R. Aaij *et al.*, *Probing matter-antimatter asymmetries in beauty baryon decays*, Nature Physics **13** (2017) 391, [arXiv:1609.05216](#).
- [73] LHCb and CMS collaborations, V. Khachatryan *et al.*, *Observation of the rare $B_s^0 \rightarrow \mu^+ \mu^-$ decay from the combined analysis of CMS and LHCb data*, Nature **522** (2015) 68, [arXiv:1411.4413](#).
- [74] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay using 3 fb^{-1} of integrated luminosity*, JHEP **02** (2016) 104, [arXiv:1512.04442](#).
- [75] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **113** (2014) 151601, [arXiv:1406.6482](#).
- [76] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays*, [arXiv:1705.05802](#), submitted to JHEP.
- [77] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001, and 2015 update.
- [78] LHCb collaboration, R. Aaij *et al.*, *Measurement of b -hadron branching fractions for two-body decays into charmless charged hadrons*, JHEP **10** (2012) 037, [arXiv:1206.2794](#).
- [79] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, [arXiv:1405.7808](#).
- [80] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, [arXiv:1408.1251](#).
- [81] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, [arXiv:1211.6759](#).
- [82] P. Perret and X. Vilasis-Cardona, *Performance of the LHCb calorimeters during the period 2010—2012*, Journal of Physics: Conference Series **587** (2015) 012012.

- [83] G. Dujany and B. Storaci, *Real-time alignment and calibration of the LHCb Detector in Run II*, J. Phys. Conf. Ser. **664** (2015) 082010.
- [84] R. Aaij *et al.*, *Tesla: An application for real-time data analysis in High Energy Physics*, [arXiv:1604.05596](#).
- [85] LHCb collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001. LHCb-TDR-015.
- [86] LHCb collaboration, *LHCb Trigger and Online Technical Design Report*, CERN-LHCC-2014-016. LHCb-TDR-016.
- [87] LHCb collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021. LHCb-TDR-013.
- [88] LHCb collaboration, *LHCb PID Upgrade Technical Design Report*, CERN-LHCC-2013-022. LHCb-TDR-014.
- [89] A. Buckley *et al.*, *General-purpose event generators for LHC physics*, Phys. Rept. **504** (2011) 145, [arXiv:1101.2599](#).
- [90] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, [arXiv:hep-ph/0603175](#).
- [91] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, [arXiv:0710.3820](#).
- [92] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.
- [93] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [94] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97, [arXiv:hep-ph/0506026](#).
- [95] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270; Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [96] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.
- [97] G. Barrand *et al.*, *GAUDI - A software architecture and framework for building HEP data processing applications*, Comput. Phys. Commun. **140** (2001) 45.
- [98] M. Clemencic *et al.*, *Recent developments in the LHCb software framework Gaudi*, J. Phys. Conf. Ser. **219** (2010) 042006.
- [99] G. A. Cowan, D. C. Craik, and M. D. Needham, *RapidSim: an application for the fast simulation of heavy-quark hadron decays*, Comput. Phys. Commun. **C214** (2017) 239, [arXiv:1612.07489](#).
- [100] J. de Favereau *et al.*, *DELPHES 3, A modular framework for fast simulation of a generic collider experiment*, JHEP **02** (2014) 057, [arXiv:1307.6346](#).

-
- [101] B. Efron, *Bootstrap methods: Another look at the jackknife*, Ann. Statist. **7** (1979) 1.
 - [102] J.-P. Kreiss and S. N. Lahiri, *1 - Bootstrap Methods for Time Series*, in *Time Series Analysis: Methods and Applications*, vol. 30 of *Handbook of Statistics*, pp. 3 – 26. Elsevier, 2012. doi: <http://doi.org/10.1016/B978-0-444-53858-1.00001-6>.
 - [103] CLEO collaboration, J. P. Alexander *et al.*, *Absolute measurement of hadronic branching fractions of the $D_{(s)}^+$ meson*, Phys. Rev. Lett. **100** (2008) 161804, arXiv:0801.0680.
 - [104] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay using 3 fb^{-1} of integrated luminosity*, JHEP **02** (2016) 104, arXiv:1512.04442.
 - [105] A. Poluektov, *Kernel density estimation of a multidimensional efficiency profile*, JINST **10** (2015) P02011, arXiv:1411.5528.
 - [106] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.
 - [107] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021.
 - [108] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.
 - [109] F. E. James, *Statistical Methods in Experimental Physics; 2nd ed.*, World Scientific, Singapore, 2006.
 - [110] LHCb collaboration, *CP-violating asymmetries from the decay-time distribution of prompt $D^0 \rightarrow K^+ K^-$ and $D^0 \rightarrow \pi^+ \pi^-$ decays in the full LHCb Run 1 data sample.*, LHCb-CONF-2016-010.
 - [111] G. Louppe, K. Cranmer, and J. Pavez, *Carl: a likelihood-free inference toolbox*, Mar., 2016. doi: 10.5281/zenodo.47798.
 - [112] A. Rogozhnikov, *Reweighting with boosted decision trees*, Journal of Physics: Conference Series **762** (2016) 012036.
 - [113] Particle Data Group, C. Amsler *et al.*, *Fragmentation functions in e^+e^- annihilation and lepton-nucleon DIS, in Review of particle physics*, Phys. Lett. **B667** (2008) 1.
 - [114] L. Lyons, D. Gibaut, and P. Clifford, *How to combine correlated estimates of a single physical quantity*, Nucl. Instrum. Meth. **A270** (1988) 110.
 - [115] R. Gauld and J. Rojo, *Precision Determination of the Small- x Gluon from Charm Production at LHCb*, Phys. Rev. Lett. **118** (2017) 072001.
 - [116] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.
 - [117] A. Rogozhnikov *et al.*, *New approaches for boosting to uniformity*, JINST **10** (2015) T03002, arXiv:1410.4140.

- [118] P. Ilten, M. Williams, and Y. Yang, *Event generator tuning using Bayesian optimization*, JINST **12** (2017) P04028, [arXiv:1610.08328](#).
- [119] LHCb collaboration, R. Aaij *et al.*, *Observation of the $B^0 \rightarrow \rho^0 \rho^0$ decay from an amplitude analysis of $B^0 \rightarrow (\pi^+ \pi^-)(\pi^+ \pi^-)$ decays*, Phys. Lett. **B747** (2015) 468, [arXiv:1503.07770](#).
- [120] T. D. Evans, *Studies of the decays $D^0 \rightarrow K^\mp \pi^\pm \pi^\pm \pi^\mp$ at CLEO-c and LHCb*, PhD thesis, University of Oxford, 2017, LHCb publication LHCb-PAPER-2017-040 in preparation.
- [121] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. C **38** (2014) 090001.
- [122] T. M. Karbach and M. Schlupp, *Constraints on Yield Parameters in Extended Maximum Likelihood Fits*, [arXiv:1210.7141](#).
- [123] R. Andreassen *et al.*, *GooFit: A library for massively parallelising maximum-likelihood fits*, J. Phys. Conf. Ser. **513** (2014) 052003, [arXiv:1311.1753](#).
- [124] C. Hasse *et al.*, *Amplitude analysis of four-body decays using a massively-parallel fitting framework*, in *22nd International Conference on Computing in High Energy and Nuclear Physics (CHEP 2016) San Francisco, CA, October 14-16, 2016*, 2017. [arXiv:1702.06735](#).
- [125] S. Reichert, *Measurement of the mixing parameters of neutral charm mesons and search for indirect CP violation with $D^0 \rightarrow K_S^0 \pi^+ \pi^-$ decays at LHCb*, PhD thesis, The University of Manchester, 2015.
- [126] M. Gersabeck, private communication.
- [127] ATLAS collaboration, M. Aaboud *et al.*, *Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13\text{TeV}$ with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117** (2016) 182002, [arXiv:1606.02625](#).