

Search for the decay $B^+ \rightarrow D^{*-} K^+ \pi^+$
Thesis submitted for the degree of MSc by research.

Ronnie P. Rera



Acknowledgements

The following thesis could not have been possible were it not for key people. I'd like to thank my advisor Tim Gershon for the opportunity to take on this project as well as always being a source of inspiration and a motivating force. The confidence I had in doing this work was due to having such a knowledgeable advisor that I knew I could trust. I felt that Tim went beyond my expectations of what it means to be a good advisor.

A big thank you is due to Tom Latham. When difficulties were encountered and it seemed unclear how to proceed, Tom was always there to lend a helping hand even long after he had left the office. Thanks to Mark Whitehead as well for your guidance and contributions to the project. I'm very appreciative of Michal Kreps for being a great secondary advisor and providing suggestions on this thesis for which it has surely improved. Each member of the Warwick LHCb team: Charlotte, Abhijit, Tom Blake, Christoph, Anton, David, Daniel, Andrew, Cayo and Chris has contributed to this thesis in some way and I'm very thankful they were always there to lend a helping hand.

For those who inspired me to take on this MSc, thank you to Matt Bellis, John Cummings, Rose Finn, Michele McColgan, John Moustakas, Mark Rosenberry, Graziano Vernizzi, and the rest of the Siena Physics faculty and friends. For igniting my curiosity in physics in the first place, thank you to George Vosburgh.

Contents

1	Introduction	4
2	Theory	5
2.1	The Standard Model	5
2.1.1	Elementary particles	5
2.1.2	Fundamental forces	6
2.1.3	Particle decays	7
2.2	CP violation	8
2.2.1	Neutral kaons	9
2.2.2	The CKM matrix	9
2.2.3	The Unitary Triangle	10
2.2.4	Measurements of γ	11
2.3	Charm spectroscopy	12
2.4	Dalitz plot formalism	13
2.5	Particle accelerators and detectors	14
2.5.1	Interactions of charged particles	15
2.5.2	Tracking and scintillation detectors	16
2.5.3	Cherenkov radiation	17
2.5.4	Calorimeters	17
3	The LHCb detector	19
3.1	VELO	20
3.2	Tracking system	21
3.3	RICH I and II	22
3.4	Spectrometer magnet	24
3.5	Calorimeter system	24
3.6	Muon detection system	25
3.7	Trigger and online system	26
4	Selection	28
4.1	Trigger and stripping	28
4.2	Initial selection	29
4.3	NeuroBayes selection	31
4.4	Particle identification requirements	36
4.5	$B^0 \rightarrow D^{*-} \pi^+$ veto	38
4.6	Yield	38
5	Fitting the B mass candidates	40
5.1	Signal shape	40
5.2	Combinatorial background	41
5.3	Cross-feed background	41
5.4	Partially reconstructed background	42

5.5	Fit to data	43
5.6	Dalitz plots	43
6	Efficiencies	46
6.1	Geometrical efficiencies	46
6.2	Selection efficiencies	47
6.3	PID efficiencies	48
6.4	Total efficiencies	48
7	Branching fraction ratio	49
7.1	Background-subtracted and efficiency-corrected Dalitz plots	50
8	Systematic uncertainties	52
8.1	Fit model	52
8.2	MC modelling of efficiency	52
8.3	Efficiency variation with decay angles	53
8.4	PID efficiency	54
8.4.1	PIDcalib tool	54
8.4.2	Track kinematics	54
8.4.3	Binning	54
8.4.4	Total PID systematic uncertainty	54
8.5	Stripping effect on phase space	54
8.6	Summary	55
8.7	Cross checks	55
9	Final results and summary	57
A	Helicity angle efficiency plots	58
	References	60

1 Introduction

Throughout the history of particle physics, the need to probe smaller length scales has required experimentalists to achieve higher and higher energy collisions in accelerator experiments. The Large Hadron Collider (LHC), built by the European Organization for Nuclear Research (CERN), is currently the largest particle accelerator in operation, achieving center of mass energy collisions of 13 TeV. The four main detector experiments at the LHC are ATLAS, CMS, ALICE, and LHCb.

The LHCb experiment specializes in b -physics, which concerns the interactions of particles containing the bottom quark. Interactions of b -hadrons are of interest because they allow measurements of charge parity symmetry (CP) violating parameters. CP symmetry means that if a particle is switched with its antiparticle and its spatial coordinates are reversed then laws of physics will remain the same. The idea that some particles may violate CP symmetry was proposed as part of the cause of the asymmetry of matter and antimatter in the observable universe [1]. CP violation was first measured in the decays of neutral kaons in 1964 by Val Fitch, Jim Cronin, and collaborators [2] and has since remained an area of great interest in particle physics. Studying b -hadron decays for CP violating parameters has been an area of much experimental and theoretical activity and the LHCb detector is currently the best tool to investigate this physics.

Many decay modes of the B mesons are yet to be discovered. Specifically the $B^+ \rightarrow D^{*-} K^+ \pi^+$ decay mode has not yet been observed. The decay mode under study in this thesis is important as it may provide new ways to measure the CP violation parameter γ as described in section 2.2.4; it could also form a background to other channels used to study γ . In addition, it also provides an excellent way to investigate charm meson spectroscopy, which can help to improve our understanding of QCD which is another of LHCb's goals. The overall goal for this thesis is then to make a measurement of the relative branching fraction ratio with respect to the $B^+ \rightarrow D^{*-} \pi^+ \pi^+$ control mode. Studies of $B^+ \rightarrow D^{*-} K^+ \pi^+$ could improve our understanding of already known particles and potentially lead to the observation of new resonances.

2 Theory

2.1 The Standard Model

The Standard Model of particle physics contains everything known about subatomic particles and their interactions. Physicists desire to classify and understand subatomic particles and their interactions has been ongoing since at least the early 20th century. All of the elementary particles fit into two categories: fermions and bosons. For every particle that exists there is an associated antiparticle which has the same mass but opposite charge. Some particles are their own antiparticles. The interactions of these elementary particles are due to the electromagnetic, weak, or strong forces, which make up all of the fundamental forces in the Standard Model. The development of the Standard Model can be traced back all the way to the first table top electron experiments up to the Higgs boson discovery with the Large Hadron Collider (LHC) at CERN. The LHC is the largest accelerator ever in operation and is still producing new data, supplementing our understanding of, and testing, the Standard Model. For a general reference of the Standard Model, please refer to Ref. [3].

2.1.1 Elementary particles

To our best understanding, there are three types of elementary particles that make up everything in the universe: fermions, gauge bosons, and Higgs bosons. Fermions are half-integer spin particles that can be further categorised as either leptons or quarks. The distinction between leptons and quarks is simply that the former (latter) are not (are) charged under $SU(3)$ and therefore do not (do) interact via the strong interaction. Gauge bosons are the force carrier particles, responsible for the fundamental interactions between elementary particles. Gauge bosons are connected to the gauge symmetries of the Standard Model. The Higgs Boson is a scalar boson (spin 0) and explains why all particles have mass.

Leptons are elementary particles that have half integer spin and integer charges $(-1,0,+1)$. The charged leptons, in increasing mass order, are called the electron, the muon, and the tau. Each charged lepton also has a neutral partner neutrino, namely the electron, muon, and tau neutrinos. Since charged leptons, such as the electron, are easy to detect in laboratory experiments they were discovered well before their neutral neutrino counterparts. Neutrinos were first proposed by Wolfgang Pauli to explain how beta decay could conserve energy, momentum, and spin. Since their discovery, neutrinos have been a popular area of research in particle physics as so little is known about their physical properties. The main interest in neutrinos is that they might be “Majorana fermions”, which means they might be their own antiparticles, and hence different from any other fermion in the Standard Model. Connected with this, CP violation in the lepton sector may then be able to explain the matter-antimatter asymmetry in the Universe [4].

Quarks were first proposed by Murray Gell-Mann and George Zweig as a way to describe some particles like baryons and mesons as bound states of more fundamental particles, baryons being bound states of three quarks while mesons are bound states of a

quark and an anti-quark. Like the leptons, quarks also have half integer spin but are the only known particles which contain charges that are multiples of $\frac{1}{3}$. For each $-\frac{1}{3}$ charged quark there exists a $+\frac{2}{3}$ charged counterpart. The first generation of quarks contains the up ($+\frac{2}{3}$) and down ($-\frac{1}{3}$) quarks, the second generation is made up of the charm ($+\frac{2}{3}$) and strange ($-\frac{1}{3}$) quarks, and the top ($+\frac{2}{3}$) and bottom ($-\frac{1}{3}$) quarks make up the third generation. Quarks bind together to form states of matter such as the meson and baryon as well as other states, such as the pentaquark, states recently discovered by the LHCb collaboration at CERN [5].

The gauge bosons are the gluon, photon, W and Z bosons all of which are spin 1. The only charged boson known is the W which has charge ± 1 . Photons, arguably the first elementary particles known to man, are responsible for interactions via the electromagnetic force. The W and Z bosons carry the weak force, which is responsible for radioactive decay. Holding quarks into bound states are the gluons which carry the strong force.

2.1.2 Fundamental forces

There are three forces with which elementary particles may interact. The electromagnetic, weak, and strong force are the only forces that are relevant at the energy scales of elementary particles. It is not known how gravity can be described in a consistent way with the Standard Model. These three forces fit into a larger mathematical class of theories called Quantum Field Theory or QFT. In a QFT, particles are treated as excitations of an underlying physical field. In gauge theory, field theories are described by Lagrangians that are invariant under symmetry transformation groups. The QFTs that are relevant to the SM are: Quantum Electrodynamics (QED), Quantum Chromodynamics (QCD), and the Electroweak Theory (EWT). QED, QCD, and EWT correspond to U(1), SU(3) and SU(2)×U(1) symmetry groups respectively. The Standard Model is itself a gauge theory and allows a unified framework to describe the quantum field theories. QED combines our knowledge of relativistic quantum mechanics with electromagnetism and describes the interactions of all charged particles by the exchange of photons. Holding the nuclei of atoms together and trapping quarks into bound states is the strong force, which is described mathematically by QCD. Interactions of quarks are described as the exchange of gluons which carry the strong force. Analogous to electric charge, in QCD quarks and gluons have a color charge. EWT is a combination of electromagnetism and the weak force. In the early universe electromagnetism and the weak force were one force called the electroweak force, which broke down into electromagnetism and the weak force as it is observed today. The electroweak symmetry is spontaneously broken by the Higgs mechanism. The weak force is responsible for radioactive decays and is carried by the W and Z bosons. The weak force is interesting for particle experiments as it is the only interaction known to violate *CP* symmetry. A summary of the particles in the Standard Model is shown in Table 1.

Table 1: The Standard Model. Values given by the PDG [6].

Fermions	symbol	electromagnetic	weak	strong	charge	Mass/MeV
electron	e	*	*		+1	$0.510998928 \pm 0.000000011$
neutrino	ν_e		*		0	$< 10^{-6}$
muon	μ	*	*		+1	$105.6583715 \pm 0.0000035$
neutrino	ν_μ		*		0	$< 10^{-6}$
tau	τ	*	*		+1	1776.86 ± 0.12
neutrino	ν_τ		*		0	$< 10^{-6}$
up	u	*	*	*	$+\frac{2}{3}$	$2.3^{+0.7}_{-0.5}$
down	d	*	*	*	$-\frac{1}{3}$	$4.8^{+0.5}_{-0.3}$
charm	c	*	*	*	$+\frac{2}{3}$	1275 ± 20
strange	s	*	*	*	$-\frac{1}{3}$	95 ± 5
top	t	*	*	*	$+\frac{2}{3}$	$173210 \pm 510 \pm 710$
bottom	b	*	*	*	$-\frac{1}{3}$	4180 ± 30
Bosons	symbol	electromagnetic	weak	strong	charge	Mass/MeV
photon	γ	*			0	0
gluon	g			*	0	0
W boson	W	*	*		± 1	80385 ± 10
Z boson	Z		*		0	91187.6 ± 2.1
Higgs boson	H				0	$125090 \pm 210 \pm 110$

2.1.3 Particle decays

Most particles created during the collisions at the LHC are short lived. From Einstein's relativistic kinematics we know that the energy from the pp collision can be converted into mass and that massive particles can decay to make less massive and eventually more stable particles. The particles created, such as the B meson for example, may decay into different final states. Each of these different final states is known as a decay mode. The total decay rate is calculated as the sum of each individual decay rate for each decay mode,

$$\Gamma = \sum_i \Gamma_i. \quad (1)$$

The lifetime of the particle is then the inverse of the total decay rate,

$$\tau = \frac{1}{\Gamma}. \quad (2)$$

The probability for a particle to decay into a particular decay mode is called the branching fraction. The branching fraction of any individual decay mode is calculated as the ratio of the decay rate of the decay mode in question to the total decay rate of the particle:

$$\mathcal{B}(i) = \frac{\Gamma_i}{\Gamma}. \quad (3)$$

To put Eq. 3 into perspective, a branching fraction of 1% means that on average 1 out of 100 decays produces that particular final state. For this project we are interested in measuring a relative branching fraction. This means the branching fraction of the decay mode relative to the branching fraction of a decay mode that is already known. Relative branching fractions are convenient to measure at LHCb and similar experiments, since

Table 2: Properties of particles involved in $B^+ \rightarrow D^{*-} K^+ \pi^+$. Values given by the PDG [6].

Particle	Quark Content	Mass (MeV)	Lifetime (Width)
B^+	$u\bar{b}$	5279.25 ± 0.26	$1.638 \pm 0.004 \times 10^{-12} \text{ s}$
D^0	$c\bar{u}$	1864.80 ± 0.14	$410.1 \pm 1.5 \times 10^{-15} \text{ s}$
D^*	$c\bar{d}$	2010.26 ± 0.07	$(83.4 \pm 1.8 \text{ keV})$
K^+	$u\bar{s}$	493.677 ± 0.013	$1.2379 \pm 0.0021 \times 10^{-8} \text{ s}$
π^+	$u\bar{d}$	139.57018 ± 0.00035	$2.6033 \pm 0.0005 \times 10^{-8} \text{ s}$

they can be determined without precise knowledge of the total number of B mesons produced. An ideal normalisation channel should have similar topology, and be triggered and reconstructed in similar ways to the signal channel. For this reason $B \rightarrow D^{*-} K^+ \pi^+$ is measured relative to $B \rightarrow D^{*-} \pi^+ \pi^+$,

$$\frac{\mathcal{B}(B^+ \rightarrow D^{*-} K^+ \pi^+)}{\mathcal{B}(B^+ \rightarrow D^{*-} \pi^+ \pi^+)}. \quad (4)$$

The properties of the particles that are involved in the $B^+ \rightarrow D^{*-} K^+ \pi^+$ decay are shown in Table 2

2.2 CP violation

The observable universe seems to be dominated by matter with hardly any antimatter. In the early universe, when temperatures were much higher than they are today, there existed an equal number of baryons and antibaryons. The process of baryon and antibaryon annihilation to make two high energy photons that would then recombine to make a baryon and antibaryon happened regularly,

$$p + \bar{p} \leftrightarrow \gamma + \gamma. \quad (5)$$

The universe was rapidly expanding and eventually this reaction was no longer able to occur. The expansion causes the temperature to decrease until the forward reaction ($p + \bar{p} \rightarrow \gamma + \gamma$) is always possible, while the back reaction ($\gamma + \gamma \rightarrow p + \bar{p}$) is not. The number of baryons and antibaryons in the universe then remained fixed. If there had been an equal number of baryons and antibaryons created than we wouldn't have the universe we have today because the baryons and antibaryons would annihilate with each other into photons. Instead, what we observe today is that the number density of baryons greatly exceeds that of antibaryons. This means there must have been a billion and one baryons for every billion antibaryons created in the early universe in order to get the observed asymmetry of matter to antimatter we see today. An explanation for this discrepancy was first proposed by Sakharov in 1967 [1]. Sakharov proposed three conditions: (1) baryon number is not conserved, (2) C and CP (charge and charge parity) symmetry must be violated, (3) the universe must have moved away from thermal equilibrium. The LHCb collaboration is concerned specifically with the second condition. CP symmetry means

that if you exchange a particle for its antiparticle and reverse its spatial coordinates then the laws of physics should remain the same. This symmetry must have been broken or else for every reaction creating a net increase in baryons over antibaryons then there would be an equal reaction producing a net increase in antibaryons over baryons. CP violation was first measured in the mixing of neutral kaon states by Cronin and Fitch. Since neutrinos are now known to have non-zero masses, it is predicted that CP violation can occur in both quarks and leptons however, to date CP violation has only been measured in the quark sector.

2.2.1 Neutral kaons

Neutral kaons K^0 ($d\bar{s}$) and $\bar{K}^0$ ($s\bar{d}$) are allowed to mix, forming states,

$$K_s^0 \approx K_1 = \frac{1}{\sqrt{2}}(K^0 + \bar{K}^0) \text{ and } K_L^0 \approx K_2 = \frac{1}{\sqrt{2}}(K^0 - \bar{K}^0), \quad (6)$$

named “short” and “long” as one is longer lived than the other by about three orders of magnitude. Eq. 6 shows that K^0 and $\bar{K}^0$ can mix into each other, but they form eigenstates (with well-defined masses and lifetimes) that are approximately CP eigenstates. If they mix into each other with equal probabilities, CP is conserved. If the probabilities are unequal, Eq. 6 is not exact, and CP is violated. If CP was conserved, K_1 would be expected to decay to 2π (CP even) and K_2 would decay to 3π (CP odd). This means K_2 will have a longer lifetime than K_1 because the available phase space for a 2π decay is larger than a 3π decay giving K_s^0 a much shorter lifetime. However, the observation of K_L^0 going to $\pi\pi$ established CP symmetry breaking in weak interactions [2].

2.2.2 The CKM matrix

In the Standard Model, CP violation can be accommodated in the Cabibbo Kobayashi Maskawa (CKM) matrix [7]. The CKM matrix describes the coupling of different quarks in flavor changing weak interactions. Essentially it gives the probability for a flavor of quark to turn into any other flavor of quark,

$$\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \times \begin{bmatrix} d \\ s \\ b \end{bmatrix}.$$

The magnitude of each element of the CKM matrix has been measured experimentally with a range of different methods, including leptonic decays, neutrino interactions, B^0 to $\bar{B}^0$ oscillations, as well as other decay modes containing the top quark. The magnitude of each element in the CKM matrix is shown below, with the diagonal showing the most likely quark interactions. These results are from a fit which assumes CKM unitarity:

$$\begin{vmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{vmatrix} = \begin{bmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{bmatrix}.$$

The CKM matrix can be parametrized by three mixing angles θ_{12} , θ_{23} , θ_{13} , and the CP violating phase δ . By expressing these angles as $s_{ij} = \sin \theta_{ij}$ and $c_{ij} = \cos \theta_{ij}$, the CKM matrix may be rewritten as,

$$\begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} = \begin{bmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{bmatrix}.$$

The values of s_{ij} have been determined experimentally to follow $s_{13} \ll s_{23} \ll s_{12} \ll 1$. The magnitudes of the elements appear to suggest a hierarchy of the elements of the CKM matrix. To make this hierarchy more obvious, A convenient way to write the CKM matrix is by introducing the parameter $\lambda = \sin \theta_c = 0.225$. This allows the CKM matrix to be written in terms of the Wolfenstein parametrization. The Wolfenstein parametrization is defined by

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}, \quad s_{23} = A\lambda^2 = \lambda \left| \frac{V_{cb}}{V_{us}} \right|, \quad (7)$$

$$s_{13}e^{i\delta} = V_{ub}^* = A\lambda^3(\rho + i\eta) = \frac{A\lambda^3(\bar{\rho} + i\bar{\eta})\sqrt{1 - A^2\lambda^4}}{\sqrt{1 - \lambda^2}[1 - A^2\lambda^4(\bar{\rho} + i\bar{\eta})]}.$$

This allows the CKM matrix to be written in terms of the Wolfenstein parametrization with elements λ , A , ρ , and η ,

$$\begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} = \begin{bmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{bmatrix} + \mathcal{O}(\lambda^4)$$

2.2.3 The Unitary Triangle

The CKM matrix is a unitary matrix implying that

$$\begin{aligned} |V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 &= 1 \\ |V_{cd}|^2 + |V_{cs}|^2 + |V_{cb}|^2 &= 1 \\ |V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2 &= 1, \end{aligned} \quad (8)$$

and that $\sum_i V_{ij}V_{ik}^* = \delta_{jk}$ and $\sum_j V_{ij}V_{kj}^* = \delta_{ik}$. When $j \neq k$ and $i \neq k$, each of these relations can be represented as a triangle in the complex plane. Each angle of these unitary triangles represents a relative phase between CKM matrix elements. One particularly interesting unitary triangle is shown in Fig. 1. Many experiments measuring CP violating parameters can be used to constrain these angles. Mathematically, the angles of the

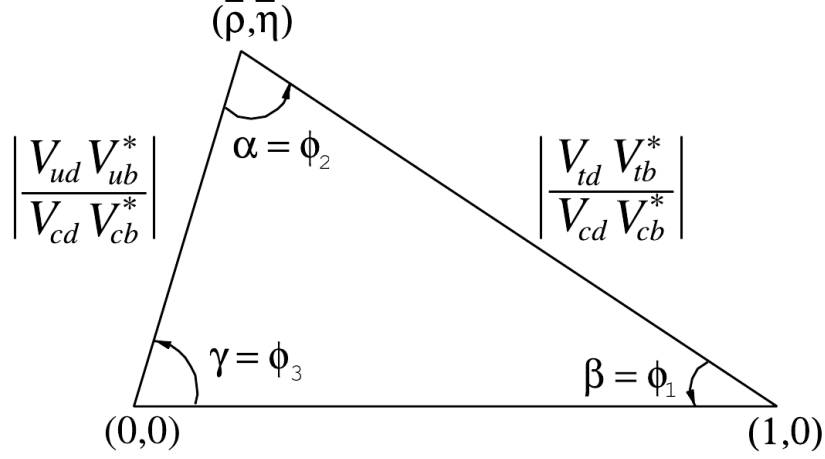


Figure 1: A representation of the unitarity of the CKM matrix. The figure is taken from the PDG [6].

unitary triangle are written as,

$$\begin{aligned}
 \beta = \phi_1 &= \arg \left(- \frac{V_{cd} V_{cb}^*}{V_{td} V_{tb}^*} \right) \\
 \alpha = \phi_2 &= \arg \left(- \frac{V_{td} V_{tb}^*}{V_{ud} V_{ub}^*} \right) \\
 \gamma = \phi_3 &= \arg \left(- \frac{V_{ud} V_{ub}^*}{V_{cd} V_{cb}^*} \right)
 \end{aligned} \tag{9}$$

The reason for both naming conventions is simply historic. The BaBar collaboration used one convention and the Belle group used the other.

2.2.4 Measurements of γ

In the unitary triangle, γ is the least well measured angle. The angle γ can be measured in decays that only receive contributions from tree diagrams, which is not true for the other two angles. This gives it the advantage that its measurement is not very much affected by theoretical uncertainty. Tree processes are ones that do not contain loops in Feynman diagrams. The advantage of tree dominated processes is that they are expected to be dominated by Standard Model contributions. Key decay modes in the measurement of γ are $B \rightarrow Dh$ related decays where h is either a kaon or pion. This is due to the fact that interference between $b \rightarrow u\bar{c}q$ and $b \rightarrow c\bar{u}q$, where q is either s or d , transitions provides sensitivity to γ . The LHCb experiment has been an important contributor to measuring γ , building on work done by previous experiments such as Belle and BaBar. The measurements of γ reported by each experiment are shown in Table 3.

Table 3: Measurements of γ by LHCb [8], BaBar [9], and Belle [10].

Experiment	γ ($^\circ$)
LHCb	$72.2^{+6.8}_{-7.3}$
BaBar	69^{+17}_{-16}
Belle	68^{+15}_{-14}

2.3 Charm spectroscopy

Studying excited states of particles provides a way to improve our current knowledge of properties of already known particles as well as giving a way to discover new resonances. Excited states of hadrons have the same quark content as the ground state but with different quantum numbers and slightly higher masses. Excited states decay quicker as they do so through the electromagnetic and strong forces rather than weakly.

Quantum numbers are values that describe conserved quantities in a quantum system. The quantum numbers of interest in studying excited mesons ($q\bar{q}$) are the angular momentum L and spin projection S . These two quantum numbers are combined to form the total angular momentum J where $J = |L \pm S|$. For a meson, which is comprised of a quark and an antiquark, the spin can either be 1 or 0 meaning the quark spins are aligned parallel or antiparallel. The angular momentum may take positive integer values starting at 0. Each integer value of angular momentum corresponds to a subshell orbital where $S = 0$, $P = 1$, $D = 2$, $F = 3$, etc. Angular momentum tells us about the parity of the system. Parity inversion is a transformation to the mirror image. For a two quark system parity is given by $P = (-1)^{L+1}$. Spin-parity (J^P) in the series $0^+, 1^-, 2^+, 3^- \dots$ is referred to as natural; the others are unnatural. Only natural parity states can decay (by strong or EM interactions, i.e. parity conserving) to two pseudoscalars, so the unnatural states can be harder to study. The natural parity states in Fig. 2, $D_0^*(2400)$ and $D_2^*(2460)$, are well-studied, but there is much less information on the $D_1(2420)$ and $D_1'(2430)$ states. The “*” denotes natural spin parity.

Various excited charm states can contribute as intermediate states in decays of the type $B \rightarrow D^{(*)}hh^{(\prime)}$. In the case of $B \rightarrow D^{(*)}hh^{(\prime)}$ decays, where $h^{(\prime)} = \pi, K$, the constrained initial and final states provide excellent resolution. In addition, these decays have a low background, especially for modes containing the $D^*(2010)$ state. Through these types of decays the quantum numbers of intermediate resonant states can be determined by amplitude analysis through their angular distributions. Decays which are sensitive to resonance states of natural parity have been studied extensively by LHCb [11–13], as well as by collaborations like BaBar [14] and Belle [15, 16]. As for states with unnatural parity, little information is known on such states. The $B \rightarrow D^*hh^{(\prime)}$ decays allow us to study states with unnatural parity. The spectrum of expected $c\bar{u}$ states with masses $< 3.2 \text{ GeV}/c^2$ is shown in Fig. 2. Therefore it is of interest to study $B \rightarrow D^*K\pi$ decays using the 2011 and 2012 LHCb dataset. The immediate goal is a first observation of the three-body decay mode. If successful this could lead on to an amplitude analysis once larger statistics are available.

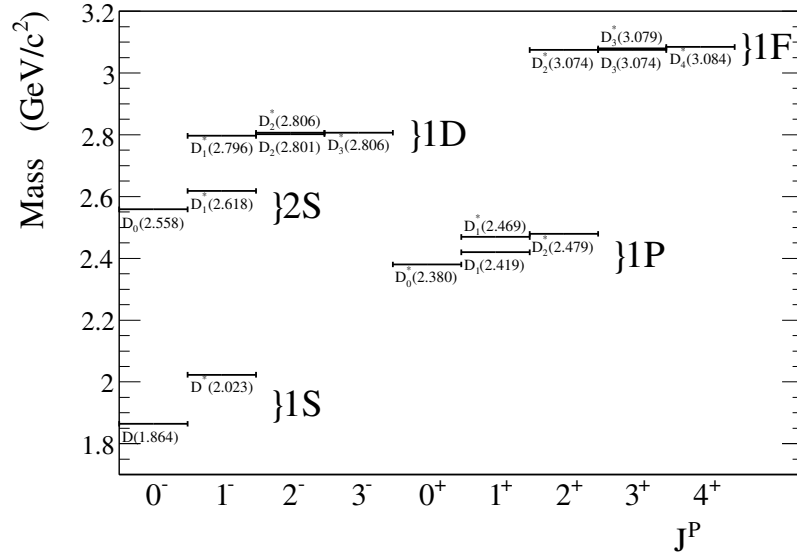


Figure 2: Predicted mass spectrum for $c\bar{u}$ states. From Ref. [17], based on information in Ref. [18]. Note that the mass values given in parentheses are not the latest experimental values.

2.4 Dalitz plot formalism

Particle decays into three body final states can be studied by way of a Dalitz plot. The Dalitz plot allows us to investigate the amplitudes of resonant and nonresonant contributions to the decay. The parent particle may decay weakly, one daughter of which then decays into two more through a strong interaction, creating the three-body decay. This type of process is a resonance and will show up in the Dalitz plot as a band. For

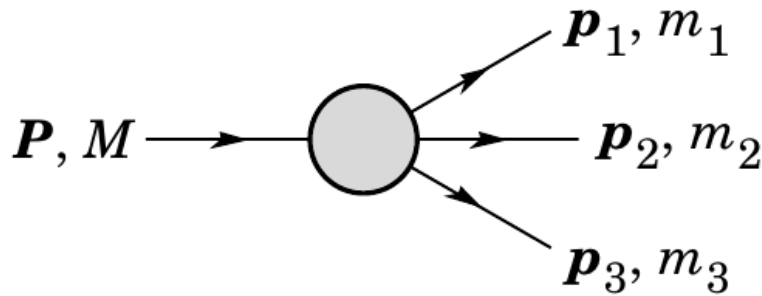


Figure 3: An example of a 3 body decay where M , m_1 , m_2 , and m_3 are the masses of the particles and P , p_1 , p_2 , and p_3 are their momenta. The figure is taken from the PDG [6].

instance, a meson decaying into a three-body final state, in which the parent particle is

spin 0 and all three final-state particles are also spin 0, can be expressed using only two variables, being combinations of the invariant masses of two of the decay products. If the parent particle in Fig. 3 has mass M , then $M^2 = m_1^2 + m_2^2 + m_3^2 = m_{12}^2 + m_{23}^2 + m_{13}^2$. The mass of the third two-body combination is constrained by this relation. Two of the variables may be plotted against each other in a Dalitz plot as shown in Fig. 4. The boundary of the Dalitz plot is constrained by the conservation of energy and momentum, so that for a specific value of m_{12}^2 the minimum and maximum values of m_{23}^2 are defined (as indicated by the dotted vertical line in Fig. 4). The distribution of decays across the Dalitz plot unveils information about the kinematics of the decay, as can be seen in Fig. 5. Bands correspond to resonances, and the number of lobes in each band indicates the spin of the resonance. The boundary of the Dalitz plot is constrained by the conservation of

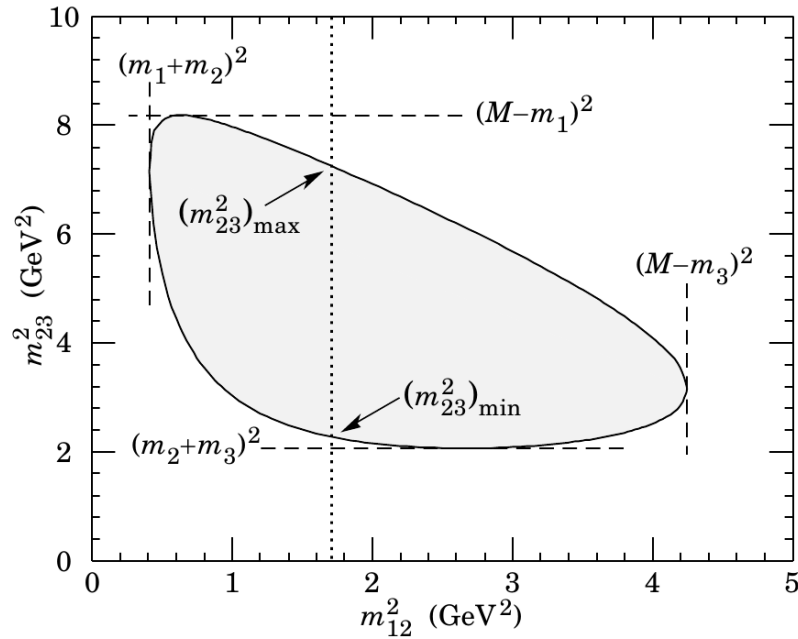


Figure 4: An example of a Dalitz plot for a 3 body decay. The phase space for the decay can be visualised using just two variables. The figure is taken from the PDG [6].

energy and momentum. Bands showing up in the plot correspond to the resonances.

2.5 Particle accelerators and detectors

Accelerators provide an accessible way to study particles and their interactions by creating a controlled environment. The object of detectors is to be able to identify the particles produced in interactions and to measure their kinematic properties. The proton, electron, photon, and neutrino are stable particles, although the neutrino is not detectable. If the average flight distance before decay of a particle is longer than the length of the detector, it is effectively stable. Examples of particles which are effectively stable to the LHCb

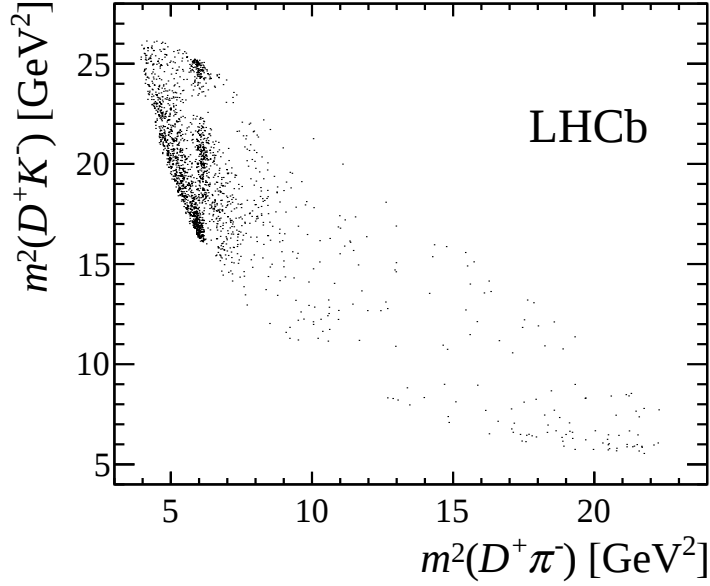


Figure 5: An example of a Dalitz plot from $B^- \rightarrow D^+ K^- \pi^-$ data [12].

detector are neutrons, K_L^0 , and the vast majority of charged pions and kaons. Other produced particles decay either immediately, or after travelling a short distance inside the detector.

2.5.1 Interactions of charged particles

Charged particles passing through the detector interact electromagnetically with atomic electrons and nuclei of the surrounding detectors. These interactions cause particles passing through the detector to lose energy through ionization and hence slow down. The rate at which a single charged particle loses energy as it passes through the detector is given by the Bethe-Bloch equation [3],

$$\frac{dE}{dx} \approx -4\pi\hbar^2 c^2 \alpha^2 \frac{nZ}{m_e v^2} \left(\ln \left[\frac{2\beta^2 \gamma^2 c^2 m_e}{I_e} \right] - \beta^2 \right). \quad (10)$$

In Eq. 10, v is the velocity for a singly charged particle, Z is the atomic number, n is the number density of the medium the particle passes through, I_e is the effective ionisation potential, and α is the fine structure constant. The rate of energy lost due to ionization through different mediums for different particles is shown in Fig. 6. Muons will travel several meters further than most particles in the detector and their energy is mostly lost due to ionization. This feature allows muons to be more easily identifiable as most other particles will lose energy through other interactions.

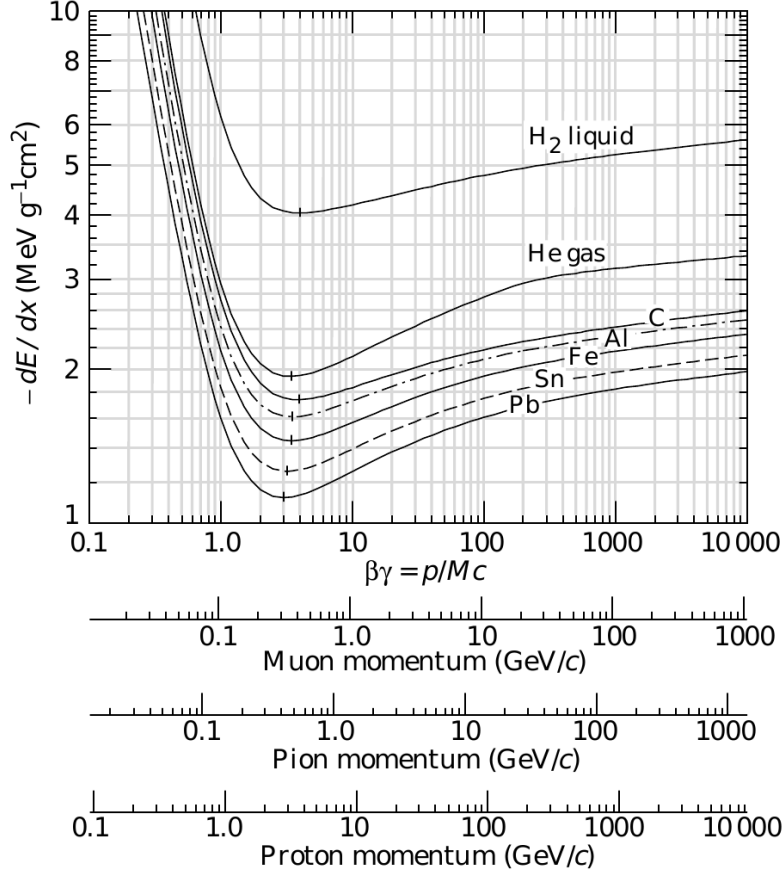


Figure 6: The ionization energy loss for different particles through different mediums. The figure is taken from the PDG [6].

2.5.2 Tracking and scintillation detectors

Particles passing through the detector causing ionization of the surrounding atoms can be tracked. As atoms ionize, an electric potential difference can be present so that the loose electrons from ionization are pulled toward an anode where the signal can be read out. Gases are often used as the medium through which the electrons pass and cause ionisation, however semiconducting materials such as silicon can also be used. A magnetic field forces charged particles to follow a curved path allowing a measurement of the momentum.

Scintillation detectors provide another popular alternative as a way to track particles. The atoms which were ionized by the passing charged particle are left in an excited state. When the excited state decays UV light is emitted. Adding a dye to the scintillating material allows the molecules to absorb the UV light and re-emit the photons as blue light. The blue light may then be detectable by photomultiplier tubes. Typical scintillation materials are liquids and plastics. The down side to using scintillating detectors however is that spatial information can be lost. Scintillation detectors are more popular with neutrino experiments since high spatial resolution is not necessarily required.

2.5.3 Cherenkov radiation

The emission of Cherenkov radiation provides another way charged particles may be identified. As a charged particle passes through a dielectric medium of a refractive index n it causes polarization of the medium. The molecules will then emit photons and return to an unpolarized state. Particles may travel faster than the speed at which light travels in that medium. This causes constructive interference between the emitted photons and as a result, Cherenkov light is emitted as a cone with some angle θ_c as shown in Fig. 7. Charged particles may then be identified by their cone angle θ_c since different particles will generally produce Cherenkov light with different cone angles. If the particle travels at speed $\beta = v/c$ where c is the speed of light through a vacuum then θ_c can be calculated as,

$$\cos \theta_c = \frac{1}{n\beta}. \quad (11)$$

Therefore, knowing θ_c gives you beta. So to get m (and hence particle type) you need p , which is obtained from the tracking system.

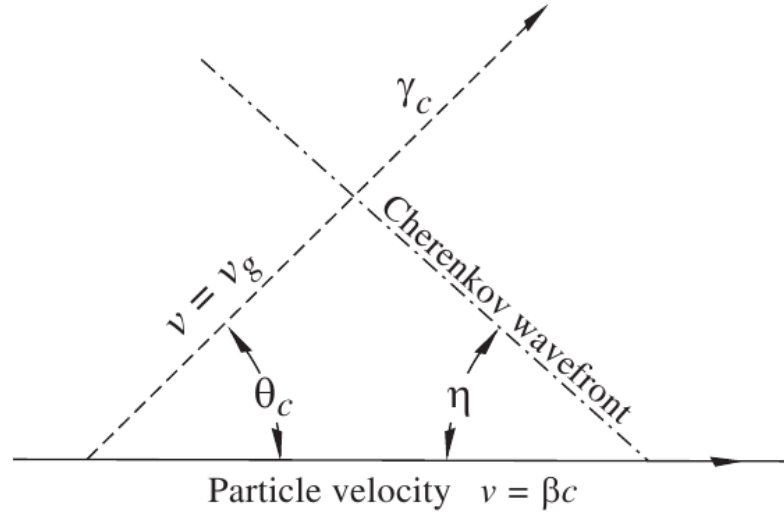


Figure 7: Particles passing through dielectric materials may be identified by the angle θ_c formed from emitting Cherenkov radiation. In the figure η is the cone opening half angle, v_g is the photon group velocity, and γ_c is the Cherenkov Lorentz factor. The figure is taken from the PDG [6].

2.5.4 Calorimeters

Experiments at the LHC use electromagnetic and hadron calorimeters to measure the energies of particles. Electromagnetic calorimeters are used to measure the energies of electrons and photons. The electromagnetic calorimeter is constructed from materials

which cause electromagnetic showers to be produced. The electromagnetic showers are the result of high energy electrons interacting with the surrounding medium and causing bremsstrahlung radiation (photons). The bremsstrahlung radiation will then in turn create e^+e^- pairs. The electromagnetic showers produce scintillation light which can be measured by photon detectors. The amount of scintillation light is proportional to the energy of the electron or photon that caused the original electromagnetic shower.

The HCAL is used mainly in the trigger, where the aim is to catch enough energy to get a positive signal. Hadrons (baryons and mesons) will cause a type of shower similar to the electromagnetic showers. As hadrons pass through the detectors and lose energy they will also interact with atomic nuclei via the strong force. The particles produced in this interaction will continue downstream to interact further and produce a cascade of particles. The hadron shower passes through multiple layers of scintillation detectors and the sum of the energy of the hadron shower is proportional to the energy of the original hadron.

3 The LHCb detector

The LHCb detector is a single arm forward spectrometer covering a polar angle of 10 to 300 millirads horizontally and 250 millirad vertically [19]. The detector is composed of multiple stages. The subsystems are designed in two halves (with the exception of the magnet), one either side of the beam pipe, to allow for assembly and maintenance. The components of the LHCb detector are:

- Spectrometer magnet
- Vertex Locator (VELO)
- Tracking system (TT, IT, OT)
- Ring imaging Cherenkov counters (RICH 1-2)
- Calorimeter system (SPD/PS, ECAL, HCAL)
- Muon detection system
- Trigger
- Online system

A more detailed description of the LHCb detector can be found in Ref. [19,20].

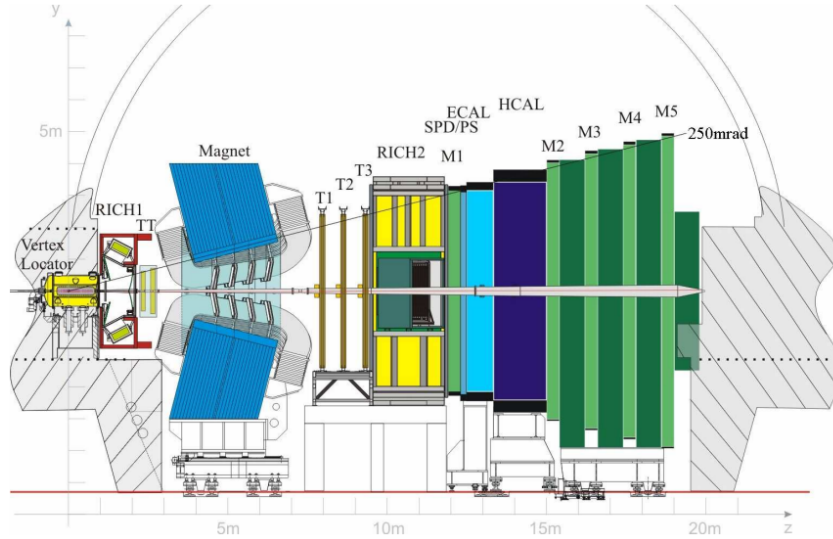


Figure 8: Cross section of the LHCb detector. Figure reproduced from Ref. [19].

3.1 VELO

The first stage of the detector is the Vertex Locator or VELO shown in Fig. 9. The VELO provides precise measurements of track coordinates close to the proton-proton interaction region. The VELO is comprised of two retractable halves with a system of 21 silicon micro-strip tracking modules each. Each module consists of two $300\ \mu\text{m}$ thick sensors allowing for measurements of the r and ϕ coordinates along the direction of the beam. The sensors are placed along the z direction, mounted in a vessel that maintains vacuum by a thin aluminum wall. This is done in order to minimize the material traversed by charged particles before they cross the sensor while protecting the sensors from the RF (radio frequency) waves produced by the beam. The track coordinates are used to reconstruct vertices from beauty and charmed hadrons as well as the primary pp collision vertex to allow for a measurement of decay lifetimes and impact parameter. This information is crucial to discriminate signatures of heavy flavour decays from background. The global performance of the VELO must meet criteria such as having a signal to noise ratio > 14 , efficiency of at least 99% for tracks that traverse at least 4 sensors, and a spatial cluster resolution of $4\ \mu\text{m}$ for 100 mrad tracks.

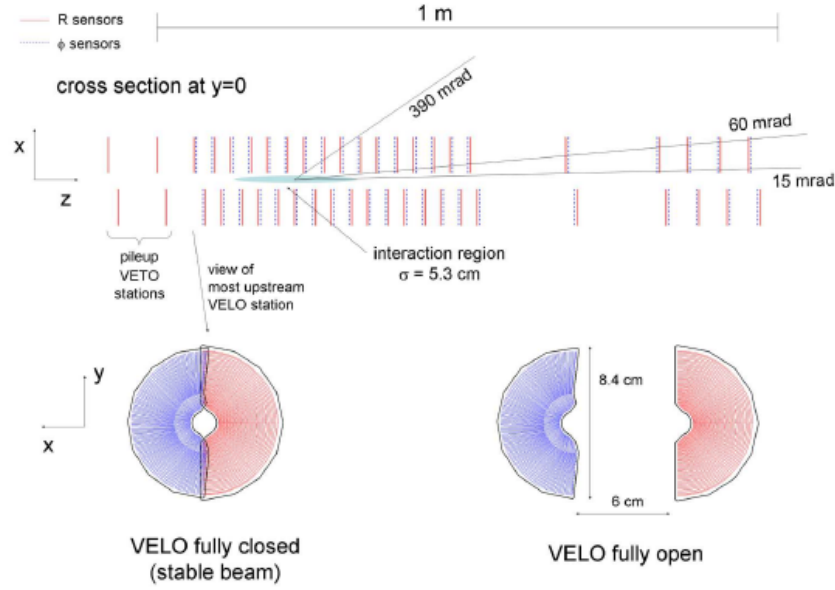


Figure 9: Layout of the VELO detector. Figure reproduced from Ref. [19].

In order to cover the angular acceptance of the downstream detectors, the VELO has the ability to detect particles with pseudorapidity in the range $1.6 < \eta < 4.9$ coming from primary vertices in the range $|z| < 10.6\text{ cm}$. The VELO has an impact parameter resolution of $< 35\ \mu\text{m}$ for particles with transverse momentum $> 1\text{ GeV}/c$ and a primary vertex resolution of $13\ \mu\text{m}$ in the transverse plane and $71\ \mu\text{m}$ along the beam axis for vertices with 25 tracks [21].

3.2 Tracking system

The tracking system measures track momentum. This requires excellent momentum resolution in order to very precisely measure the invariant mass of reconstructed b -hadrons. Shown in Fig. 10, the Tracker Turicensis (TT) is a planar tracking station that is located upstream of the dipole magnet. T-stations are placed after the magnet which have both the Inner Tracker (IT) and Outer Tracker (OT) parts of the tracking system. The TT and IT use the same technology and is therefore referred together as the silicon tracking system (ST). Silicon microstrip sensors are used in the TT and IT each having a strip pitch of $200\ \mu\text{m}$.

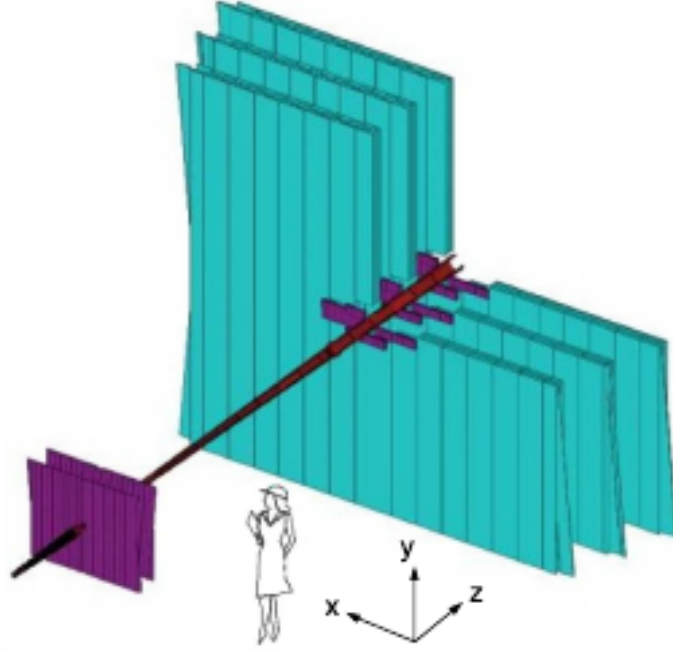


Figure 10: Arrangement of the TT, IT, and OT layer. Figure reproduced from Ref. [19].

The IT covers a 120 cm wide by 40 cm high cross-shaped region in the center of the three tracking stations. There are four ST stations which each have four detection layers with vertical strips in the first and the last layer. The TT, shown in Fig. 11, uses 143,360 readout strips 38 cm long, having an effective area of about $8.4\ \text{m}^2$ while the IT has 12,9024 readout strips that are 11 cm and 22 cm long, giving it an effective area of $4.0\ \text{m}^2$. Parameters such as spatial resolution, hit occupancy, signal shaping time, signal hit efficiency, radiation damage, material budget, and number of readout channels were all taken into consideration for the design of the silicon tracker. The hit efficiency of the TT is greater than 99.7% and for the IT is greater than 99.8% for 2011 data. The hit resolution was found to be $52.6\ \mu\text{m}$ for the TT and $50.3\ \mu\text{m}$ for the IT. Finally the Outer Tracker (OT), made of Kapton/Al straw tubes, is a drift time detector for the measurement of charged particle momentum. The OT has momentum resolution of $\delta p/p \approx 0.4\%$ and an

overall reconstruction efficiency of 80%. The single hit resolution of the OT is $205\ \mu\text{m}$.

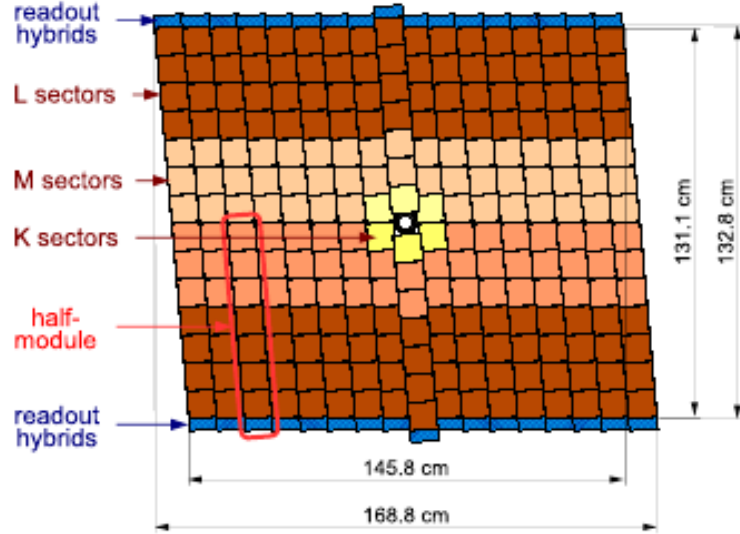


Figure 11: Layout of the TT detector. Figure reproduced from Ref. [19].

3.3 RICH I and II

The Ring Imaging Cherenkov (RICH) detectors provide particle identification information that is particularly important to distinguish between different types of charged hadrons, i.e. pions, kaons and protons. The ability for LHCb to be able to separate pions from kaons in B hadron decays is essential to the experiment. Two RICH detectors are installed in order to cover the full momentum range of charged particles. RICH 1 covers charged particles in the $1\text{--}60\ \text{GeV}/c^2$ range, using aerogel and C_4F_{10} radiators. RICH 2 is able to cover the higher momentum range $15\text{--}100\ \text{GeV}/c^2$, using CF_4 radiators. Charged particles passing through the RICH are identified by their Cherenkov angle as shown in Fig. 12. Focusing of Cherenkov light is achieved by spherical and flat mirrors to reflect the image out of the spectrometer acceptance. Hybrid photon detectors are used to detect Cherenkov photons with $200\text{--}600\ \text{nm}$ wavelength.

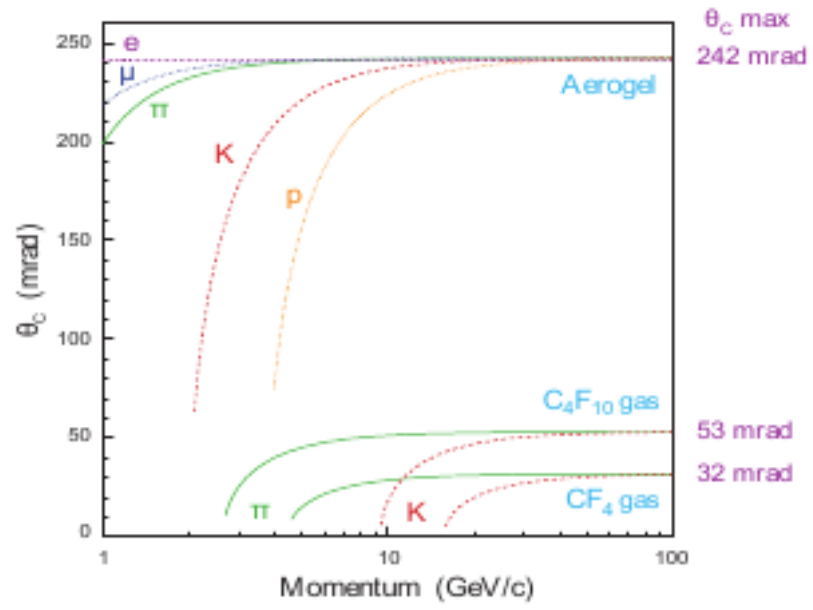


Figure 12: Particles passing through the RICH are identified by their Cherenkov angle. Figure reproduced from Ref. [19].

3.4 Spectrometer magnet

The dipole magnet shown in Fig. 13 provides a 4 Tm magnetic field and is used to measure the momentum of charged particles. The magnet is a warm magnet design held together by saddle shaped coils in a window frame yoke. A magnet control system controls and monitors parameters of the magnet such as the temperature, voltage, water flow, mechanical movement, etc. There is a second independent system which ensures safe operation of the magnet by enforcing a discharge if critical parameters are outside the accepted operating range.

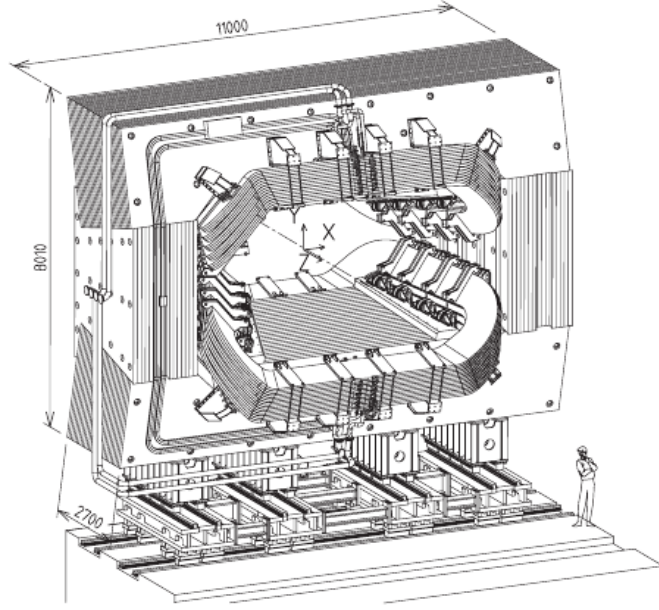


Figure 13: Structure of the magnet. Figure reproduced from Ref. [19].

3.5 Calorimeter system

The calorimeter identifies hadron, electron, and photon candidates for the first trigger level, which must return a decision within $4 \mu\text{sec}$ of the interaction. The calorimeter's ability to reconstruct π^0 and prompt photons is important for studies of certain B meson decays in LHCb. The calorimeter is made up from the electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL) with a scintillator pad detector (SPD) and preshower detector (PS) preceding the ECAL. The SPD/PS system has a sensitive area which is 7.6 m wide and 6.2 m high. The job of the ECAL is to select electrons and photons of high transverse energy while rejecting the high background of charged pions. This is achieved by longitudinal segmentation of the electromagnetic shower. The scintillator pad detector in front of the pre-shower selects charged particles to distinguish electrons from photons.

The HCAL is a sampling device made from iron and scintillating tiles which run parallel to the beam axis. The HCAL provides a measurement of the transverse energy of hadrons. Scintillation light is read out by Photo Multiplier Tubes (PMT) using wavelength-

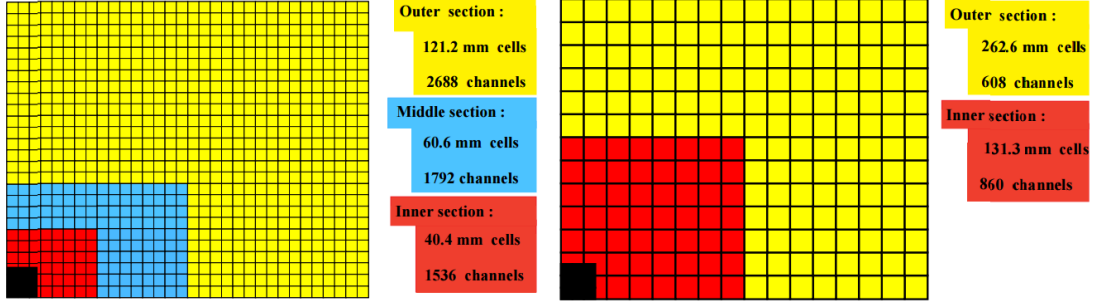


Figure 14: The SPD/PS and ECAL (left) and HCAL (right) seen from a lateral layout. Figure reproduced from Ref. [19].

shifting fibers (WLS). The SPD/PS fibers are then read out by a 64 channel multianode photomultiplier tubes (MAPMT) and individual phototubes for the ECAL and HCAL are set in proportion to the gain and distance to the beampipe. This has the advantage that it allows for a constant E_T scale. The HCAL tubes operate at a higher gain than those for the ECAL due to the light yield at the HCAL being smaller by a factor of 30. Shown in Figure 14, the calorimeter system (SPD/PS, ECAL, and HCAL) has a lateral segmentation. This is due to a varying hit density by two orders of magnitude over the surface of the calorimeter. The ECAL was chosen to split in to three sections while the HCAL is segmented into two zones due to the size of hadronic showers.

3.6 Muon detection system

Muon identification is crucial to LHCb because of the muons which are present in final states of several interesting B decays. LHCb's muon system provides information for the high transverse momentum muon trigger at level 0 and muon identification for the high level trigger and offline analysis. The muon system is composed of five parts (M1-M5) placed along the beam axis. The M2-M5 stages placed downstream of the calorimeter are interleaved with iron absorbers in order to select penetrating muons. The M1 stage, placed in front of the calorimeters, improves the transverse momentum measurement in the trigger. The muon trigger relies on muon track reconstruction and requires a hit in all five stations. Stations M1-M3 have high spatial resolution in the x coordinate which is used to define track direction and calculate transverse momentum of muon candidates to a resolution of 20%. The main purpose of stations M4-M5 is to identify penetrating particles.

3.7 Trigger and online system

The LHCb trigger is used to lower the data output rate from the beam crossing rate of 10 MHz down to 5 kHz. The reduction is achieved through level-0 (L0) and the high level trigger (HLT). The L0 and HLT triggers are hardware and software based respectively. The trigger is optimized to achieve the highest efficiency for events that are selected in the offline analysis while strongly rejecting background events. The L0 trigger attempts to reconstruct:

- High transverse energy hadron, electron, and photon cluster in the calorimeter
- One or two high transverse momentum muons in the muon chamber

At the HLT, the beam crossing rate is further reduced from 1 MHz down to 5 kHz. The HLT has the ability to access all data in one event but given the 1 MHz output rate of the L0 trigger and limitations on computing power, only part of the full event data is used to reject the majority of uninteresting events.

The HLT is split into HLT1 and HLT2 stages. The HLT1 reconstructs particles in the VELO and identifies tracks with large impact parameter with respect to the primary vertices. The HLT2 then implements inclusive trigger algorithms with the B decay being only partially reconstructed, and exclusive trigger algorithms to fully reconstruct B -hadron final states. An important aspect of the HLT1 is that it reduces the rate to 30 kHz to save CPU power and allows for full pattern recognition of the remaining events. Before the final selection, tracks are selected with loose cuts placed on their momentum and impact parameter. Cuts are placed on the invariant mass or on pointing of the B -momentum to the primary vertex. As a result, the selection further reduces the rate sufficiently to allow for data to be written to storage for future analysis.

The Online system ensures that the data is transferred from the front end electronics into permanent storage as well as ensuring all detector clocks are synchronized with the LHC clock. The online system is made of three parts:

- Data acquisition system DAQ
- Timing and Fast control TFC
- Experiment control system ECS

The transportation from the electronics to storage of data belonging to a bunch crossing and identified by the trigger is handled by the DAQ. The TFC drives all storage of data from the front end electronics and online processing. It does this by distributing the beam-synchronous clock, the L0 trigger, synchronous resets, and fast control commands. The operational state of the LHCb detector is controlled and monitored by the ECS. The operational state of the detector encompasses parameters such as voltage, temperature, gas flow, and pressures. The ECS is also responsible for monitoring the trigger, TFC, and DAQ systems.

After the trigger stage, in an important step called “stripping”, streams of selected events are produced to be used in individual analyses. Information from reduced raw data is used to determine the four-momentum of measured particles, locate the primary and secondary vertices, and reconstuct the composite particles. A pre-selection algorithm for each channel of interest allows for a variety of physics anaylses to be done. Events passing the selection requirements are fully reconstructed and contain the full information on each event.

4 Selection

In this section the steps in the selection process and the cuts applied to the data sample are discussed. The discussion here is explicitly for the $B^- \rightarrow D^{*-} K^+ \pi^+$ signal channel; appropriate corresponding requirements are applied for the $B^- \rightarrow D^{*-} \pi^+ \pi^+$ control channel.

Requirements on the trigger and stripping lines allow events coming from the relevant decay modes to be used for the analysis that follows. The initial selection helps to reduce the data to a manageable quantity and clean up the background through having well-defined requirements, that we can evaluate efficiencies for. The main discrimination between signal and background consists of training a neural network on signal events using Monte Carlo and background on data. This allows a fit to the signal events to be performed as described in Sec. 5. Particle identification requirements are then optimized and placed on final state tracks to reduce the contribution from mis-identified background decays.

4.1 Trigger and stripping

Similar to other analyses of similar final states in LHCb, it is required that the Level 0 trigger is either `L0Hadron_T0S` or `L0Global_TIS`. In the case of `L0Hadron_T0S` (Triggered On Signal), this means a particle from the signal decay was measured with a large E_T in the calorimeters. As for `L0Global_TIS` (Triggered Independent of Signal), this means that the event contained a high p_T muon or one of the calorimeters measured a large E_T object which is not consistent with coming from the signal decay. As described in Sec. 3.7, for hadronic final states, significant yields are obtained when the decay products leave deposits in the calorimeter, and hence are `L0Hadron_T0S`, but also when something else in the event, typically from the decay of the other b -hadron produced in the $pp \rightarrow b\bar{b} X$ collision, causes a trigger `L0Global_TIS`. During offline analysis, we can associate the trigger decisions with reconstructed particles, so we can know if a candidate is in one of these two categories or not. The two distinct Level 0 trigger paths provide an opportunity to subdivide the data and check for systematic effects, as discussed in Sec. 8. Both are retained throughout the analysis.

The HLT requirements are that at least one trigger line from the following three has fired: `Hlt2Topo2BodyBBDT_T0S`, `Hlt2Topo3BodyBBDT_T0S` and `Hlt2Topo4BodyBBDT_T0S`. This requirement is explicitly made in the offline selection. The `Hlt2Topo` refers to the HLT2 topological trigger line which is designed to efficiently trigger on B decays with at least two charged daughters. The "BBDT" in the line names refer to the use of a boosted decision tree algorithm in the trigger [22]. To save on CPU time, $p_T > 500 \text{ MeV}/c$ and $p > 5 \text{ GeV}/c$ are required on one D daughter and one bachelor particle in HTL.

The cuts from the `B2D0KPiPiD2HHPIDBeauty2CharmLine` stripping line are summarised in Table 4; those for the `B2D0PiPiPiD2HHPIDBeauty2CharmLine` line used to select the normalisation channel are identical except for the mass hypothesis of one of the bachelor tracks. The variables used in the stripping lines select charged B meson candidates

decaying to a neutral $\bar{D}^0$ meson with three other charged tracks known as bachelor particles. In the table, ‘One D daughter and one bachelor’ means that one of the D daughters and one of the bachelor tracks must satisfy the criteria, whilst ‘One track’ means that any of the 5 final state tracks must satisfy the criteria. The variables shown are: the reconstructed mass ($m_{B^+}^{\text{reco}}$); the reconstructed decay time ($\tau_{\text{reconstructed}}$); the minimum χ_{IP}^2 between the reconstructed B candidate and any primary vertex in the event (here χ_{IP}^2 is the difference between the χ^2 of the primary vertex reconstruction with and without the considered particle); the quality of the fit performed to the B production vertex ($(\chi^2/\text{ndf})_{\text{vertex}}$); the cosine of the angle between the B candidate’s reconstructed momentum and the vector pointing from the event primary vertex to the B decay vertex ($\cos(\text{Dir. angle w.r.t. own PV})$); the total p_{T} of all final state tracks ($\Sigma_{\text{all}} p_{\text{T}}$); and the output variable of a BDT which is used to identify the secondary vertex of a B decay (BBDT output).

The stripping lines also include requirements on the $\bar{D}^0$ candidate with two combined bachelor tracks which are: the absolute difference of the reconstructed mass of the $\bar{D}^0$ and the $\bar{D}^0$ as given by the PDG ($|m_{\bar{D}^0}^{\text{reco}} - m_{\bar{D}^0}^{\text{PDG}}|$); the square of the distance between the reconstructed decay vertex and the most likely primary vertex divided by its uncertainty ($\chi_{\text{Distance to best PV}}^2$); the maximum distance of closest approach (DOCA) between any of the daughter tracks ($\max_{\text{daughters}}(\text{D.O.C.A.})$). Additional requirements are placed on the track p , p_{T} , probability of being a false track (Ghost probability), and fit quality ($(\chi^2/\text{ndf})_{\text{track}}$), and the total number of long tracks in the event ($N_{\text{long tracks}}$).

Signal channel data are selected with the `B2D0KPiPiD2HHPIDBeauty2CharmLine` stripping line. It should be noted that this line was designed to select the four body decay $B^+ \rightarrow \bar{D}^0 K^+ \pi^+ \pi^-$, as studied previously in LHCb [23, 24]. As such, it lets through a much larger rate than is necessary for the part of phase-space corresponding to $D^{*-} \rightarrow \bar{D}^0 \pi^-$. To reduce the rate, a requirement of $m(K^+ \pi^+ \pi^-) < 3.0 \text{ GeV}/c^2$ is imposed in the stripping. This results in a small inefficiency for $B^- \rightarrow D^{*-} K^+ \pi^+$ decays, which must be corrected for, with associated systematic uncertainty assigned, later in the analysis.

4.2 Initial selection

Initial selection requirements to remove obvious background events are applied. The initial selection allows a reduction in the size of the output from the stripping. As discussed in Sec. 4.3, the main signal/combinatorial background discrimination is achieved with a multivariate classifier. Before that stage however, the initial selection requirements applied will help improve the training.

These initial requirements are summarised in Table 5; they comprise tighter B^+ candidate mass cuts compared to the stripping, together with D^{*-} candidate mass cuts and loose particle identification requirements. To improve B mass resolution, a D mass constraint is applied to the vertex fit.

One important requirement, applied at this stage, is on the difference between the D^{*-} and $\bar{D}^0$ candidate masses. The small difference between the sum of the $\bar{D}^0$ and π^- masses and the D^{*-} mass, known as the Q value, results in a strong correlation between these

Table 4: Cuts applied to the data sample in the B2D0KPiPiD2HHPIDBeauty2CharmLine stripping selection.

Particle	Parameter	Cut value
B^+	$m_{B^+}^{\text{reco}}$	$< 7.0 \text{ GeV}/c^2$
	$m_{B^+}^{\text{reco}}$	$> 4.75 \text{ GeV}/c^2$
	$\tau_{\text{reconstructed}}$	$> 0.2 \text{ ps}$
	$\min_{\text{PVs}} \chi_{\text{IP}}^2$	< 25
	$(\chi^2/\text{ndf})_{\text{vertex}}$	< 10
	$\cos(\text{Dir. angle w.r.t. own PV})$	> 0.999
	$\Sigma_{\text{all}} p_{\text{T}}$	$> 5.0 \text{ GeV}/c$
	BBDT output	> 0.05
$\bar{D}^0$	$ \mathbf{m}_{\bar{D}^0}^{\text{reco}} - \mathbf{m}_{\bar{D}^0}^{\text{PDG}} $	$< 100 \text{ MeV}/c^2$
	$(\chi^2/\text{ndf})_{\text{vertex}}$	< 10
	$\chi_{\text{Distance to best PV}}^2$	> 36
	$\cos(\text{Dir. angle w.r.t. own PV})$	> 0.0
	$\max_{\text{daughters}}(\text{D.O.C.A.})$	$< 0.5 \text{ mm}$
	$\Sigma_{\text{daughters}} p_{\text{T}}$	$> 1.8 \text{ GeV}/c$
	m^{reco}	$< 3.0 \text{ GeV}/c^2$
	$(\chi^2/\text{ndf})_{\text{vertex}}$	< 8
Recombined bachelors	$\chi_{\text{Distance to best PV}}^2$	> 16
	$\cos(\text{Dir. angle w.r.t. own PV})$	> 0.98
	$\max_{\text{bachelors}}(\text{D.O.C.A.})$	$< 0.4 \text{ mm}$
	$\Sigma_{\text{bachelors}} p_{\text{T}}$	$> 1.25 \text{ GeV}/c$
	p_{T}	$> 100 \text{ MeV}/c$
	$p_{\bar{D}^0 \text{ daughter}}$	$> 1.0 \text{ GeV}/c$
Charged tracks	p_{bachelor}	$> 2.0 \text{ GeV}/c$
	$\min_{\text{PVs}} \chi_{\text{IP}}^2$	> 4
	$(\chi^2/\text{ndf})_{\text{track}}$	< 3
	Ghost probability	< 0.3
	p_{T}	$> 500 \text{ MeV}/c$
	p	$> 5.0 \text{ GeV}/c$
One D daughter and one bachelor	$(\chi^2/\text{ndf})_{\text{track}}$	< 2.5
	p_{T}	$> 1.7 \text{ GeV}/c$
	p	$> 10.0 \text{ GeV}/c$
One track	$(\chi^2/\text{ndf})_{\text{track}}$	< 2.5
	$\min_{\text{PVs}} \chi_{\text{IP}}^2$	> 16
	$\min_{\text{PVs}} \text{IP}$	$> 0.1 \text{ mm}$
	$N_{\text{long tracks}}$	< 250
Global Event Cut	$N_{\text{long tracks}}$	< 250

two values as shown in Fig. 15.

A requirement on the difference between the candidate masses, denoted $\Delta m \equiv$

Table 5: Requirements applied to both signal and control data samples in the initial selection. Note that a D^0 meson mass constraint is applied when calculating $m_{D^{*-}}^{\text{reco}}$.

Particle	Variable	Requirement
B^+	$m_{B^+}^{\text{reco}}$	> 5000 and $< 6000 \text{ GeV}/c^2$
	χ_{IP}^2	< 20
$D^{*-} - \bar{D}^0$	$ \Delta m - 145.421 \text{ MeV}/c^2 $	$< 5 \text{ MeV}/c^2$

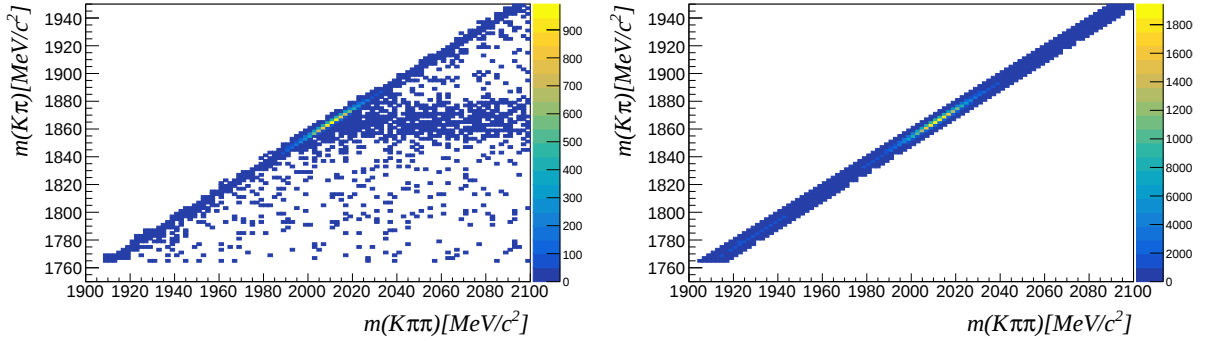


Figure 15: Correlations between the reconstructed masses of the D^* and D candidates from (left) signal MC and (right) data. Note that the data already has a requirement on the $D^* - D$ mass difference applied. The kinematic threshold $m(K^+\pi^-\pi^-) > m(K^+\pi^-) + m_\pi$ ensures that the top left of the plot is empty.

$m(K^+\pi^-\pi^-) - m(K^+\pi^-)$ is therefore the best way to select signal decays, that contain true D^{*-} decays, while suppressing background. The Δm distribution is shown in Fig. 16. The selection is so clean that after the Δm requirement essentially all B^+ candidates contain real D^{*-} decays. This greatly simplifies the set of possible background decays that need to be taken into account: for example there is no need to consider background from charmless B decays.

4.3 NeuroBayes selection

For this analysis the multivariate analyser NeuroBayes is used to suppress the combinatorial background. The signal sample in the neural network training is taken from $B^+ \rightarrow D^{*-} K^+ \pi^+$ Monte Carlo simulation (MC) in the B candidate mass range $5200 \text{ MeV}/c^2$ to $5400 \text{ MeV}/c^2$, while data in the B candidate mass range $> 5600 \text{ MeV}/c^2$ is used as a background sample. Due to the number of events in the data compared to the MC only a quarter of the events from the data are used while training the neural network.

The variables used to train the neural network are listed in Table 6. In order to ensure the neural network is not training signal and background with any artificial effects due to differences between data and MC caused by inaccurate modelling, and to ensure that identical initial selection cuts are made in both signal and background samples, the impact

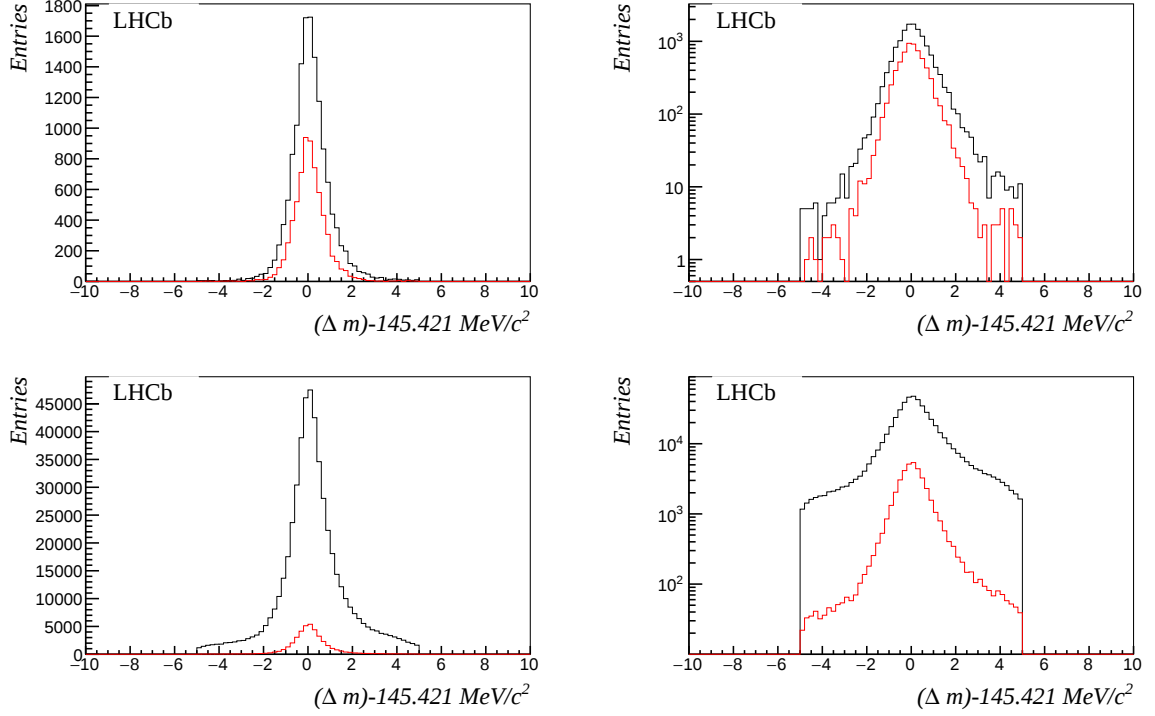


Figure 16: The difference Δm between the D^* and D candidate masses, after subtracting the PDG expected value of $145.421 \text{ MeV}/c^2$ for (top) MC and (bottom) data with (left) linear and (right) logarithmic y -axis scales. The black and red lines are, respectively, before and after the application of the neural network requirement discussed in the next section.

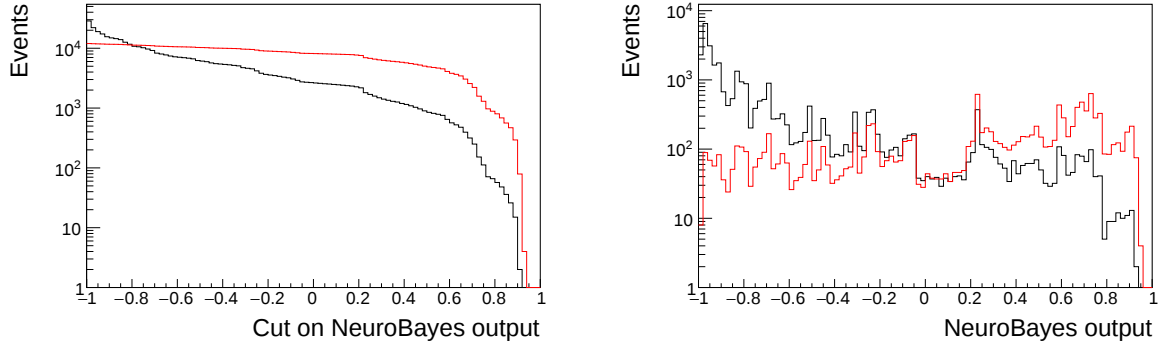


Figure 17: Result of the NeuroBayes training applied to $D^{*-}K^+\pi^+$ with (red) signal and (black) background. (Left) relative yields as a function of requirement on the NeuroBayes output. (Right) distributions in the NeuroBayes output variable (logarithmic scale).

parameter with respect to the primary vertex of the B is required to be < 20 prior to training. Cuts are also made to the decay times of the B and D candidates, and the χ^2 of the decay time of the D to eliminate extreme values. Candidates with bachelor pion or

Table 6: Variables trained in the NeuroBayes selection and their ranking given by the neural network output for $D^*K\pi$. Variables labelled † are calculated after a D mass constraint.

Particle	Variables	NeuroBayes ranking	
B^+	$^\dagger\chi_{\text{end vertex}}^2$	1	(best)
	$^\dagger\chi_{\text{flight distance}}^2$	2	
	D.O.C.A.	4	
	$\chi_{\text{impact parameter w.r.t. own PV}}^2$	5	
	p_z	9	
	$^\dagger\chi_{\text{min IP}}^2$	12	(worst)
	τ	14	
	p_T	16	
	χ_{tau}^2	3	
	$^\dagger\chi_{\text{flight distance}}^2$	6	
$\bar{D}^0$	$^\dagger\text{DIRA ORIVX}$	7	
	Dira	10	
	τ	11	
	FD OWNPV	13	
	$^\dagger\chi_{\text{min IP}}^2$	15	
	$\chi_{\text{original vertex}}^2$	8	
K^+			

kaon momentum $> 100 \text{ GeV}/c$, or with $p_T > 20 \text{ GeV}/c$, are also removed. The results of the training as well as signal and background yields are shown in Fig. 17.

The variables which are used in the training are checked for any correlation with the Dalitz plot variables or with the B mass. Correlations with the B mass are shown in Table 7 and with the Dalitz plot variables in Table 8. All correlations are found to be $\mathcal{O}(10\%)$, in magnitude, or smaller and are considered to be acceptable.

Once the training is completed a Punzi Figure of Merit (FoM) plot is generated in order to find the optimal cut on the neural network output. The Punzi FoM is defined as $\epsilon_{\text{sig}} / \left(\frac{a}{2} + \sqrt{B} \right)$ [25] where a is the required significance level which is taken to be 5, ϵ_{sig} is the signal efficiency determined from MC and B is the estimated background. For each value of the requirement, the value of B is determined by fitting the surviving data in the B candidate mass sideband region with a quadratic function, and extrapolating the fit results to the signal region in order to estimate the number of background events in that region. The Punzi FoM plot is shown in Fig. 18. It can be seen that the FoM plot peaks at about 0.4, which is the value chosen as the minimum requirement to be applied to the data.

Table 7: Correlations between the variables used in the input of NeuroBayes and the B mass for $D^*K\pi$ in the signal region ($5200 < m(B) < 5400 \text{ MeV}/c^2$) with simulation and the background region $m(B) > 5600 \text{ MeV}/c^2$ with data. The TH2::GetCorrelationFactor method was used in generating the correlations. Variables labelled † are calculated after a D mass constraint.

Variable		Correlation in %	
		Data	MC
B^+	p_T	1.30	-7.02
	$^\dagger\chi^2_{\text{end vertex}}$	-2.47	7.61
	$^\dagger\chi^2_{\text{flight distance}}$	-2.56	-5.25
	D.O.C.A.	0.40	7.01
	$\chi^2_{\text{impact parameter w.r.t. own PV}}$	0.19	8.81
	$^\dagger\chi^2_{\text{min IP}}$	-0.11	4.92
	τ	-4.50	-4.77
	χ^2_{tau}	-2.65	1.07
$\bar{D}^0$	$^\dagger\chi^2_{\text{flight distance}}$	-1.58	-5.97
	$^\dagger\text{Dira}$	1.45	-6.75
	Dira ORIVX	1.74	-2.22
	τ	0.70	1.65
	FD OWNPV	0.34	-6.60
	$^\dagger\chi^2_{\text{min IP}}$	-4.26	-4.19
K^+	$\chi^2_{\text{original vertex}}$	1.99	6.83

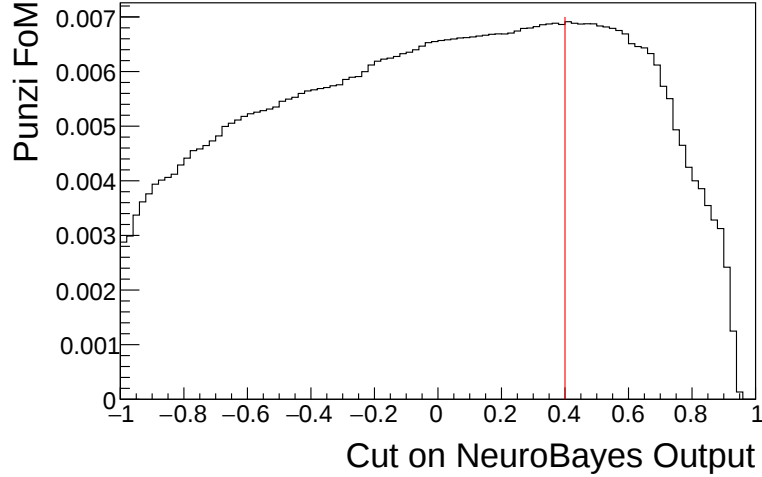


Figure 18: Punzi figure of merit of the cut on the NeuroBayes output.

Table 8: Correlations between the variables used in the input of NeuroBayes and the B mass for $D^*K\pi$ in the signal region ($5200 < m(B) < 5400 \text{ MeV}/c^2$) with MC and the background region ($m(B) > 5600 \text{ MeV}/c^2$) with data. The TH2::GetCorrelationFactor method was used in generating the correlations. Variables labelled † are calculated after a D mass constraint.

Variable		Correlation (%)					
		$m^2(D^*K)$		$m^2(D^*\pi)$		$m^2(K\pi)$	
		Data	MC	Data	MC	Data	MC
B^+	p_T	8.36	3.47	-19.1	-4.05	14.5	2.31
	$^\dagger\chi^2_{\text{end vertex}}$	5.13	-1.55	-10.5	1.96	6.83	-3.55
	$^\dagger\chi^2_{\text{flight distance}}$	0.56	5.77	11.4	-2.78	-9.24	-0.53
	D.O.C.A.	8.68	-8.11	0.47	-1.48	-2.21	4.74
	$\chi^2_{\text{impact parameter w.r.t. own PV}}$	1.70	-2.33	-6.61	2.25	3.75	0.63
	$^\dagger\chi^2_{\text{min IP}}$	3.96	-0.96	-6.01	1.24	2.89	0.26
	τ	3.11	5.99	16.1	-0.58	-15.7	-2.19
$\bar{D}^0$	χ^2_{tau}	7.89	6.35	-0.24	2.58	-5.56	-7.36
	$^\dagger\chi^2_{\text{flight distance}}$	2.87	4.36	5.77	-1.60	-5.54	0.85
	$^\dagger\text{Dira}$	-4.11	6.97	-9.57	-1.88	10.6	-0.60
	Dira ORIVX	-1.85	-1.16	-3.84	-1.84	4.85	4.06
	τ	0.18	1.39	-2.84	0.65	2.98	0.26
	FD OWNPV	-2.20	5.07	5.80	-1.29	12.1	1.73
	$^\dagger\chi^2_{\text{min IP}}$	8.64	1.29	19.1	-1.44	-18.5	0.93
K^+	$\chi^2_{\text{original vertex}}$	1.40	-1.51	-3.82	0.85	3.33	-2.58

4.4 Particle identification requirements

Due to the inability of MC to accurately generate PID efficiencies, the `PIDCalib` package was used to determine the PID efficiencies using `ProbNN` variables. The `PIDCalib` software package uses a data-driven method to estimate the efficiencies of PID requirements applied to simulated samples. The `ProbNN` variables are in the range 0 to 1. A PID cut of the form $\text{ProbNNh} \times (1 - \text{ProbNNh}') > N$ where N is between 0 and 1 is chosen. Here if $\mathbf{h} = \mathbf{K}$ then $\mathbf{h}' = \mathbf{\pi}$ and vice versa. The bachelor kaon cut $\text{ProbNNK} \times (1 - \text{ProbNNpi}) > N$ on the bachelor kaon is optimised using a Punzi FoM plot. Similarly the bachelor pion cut is optimised by considering the Punzi FoM evaluated as a function of N in $\text{ProbNNpi} \times (1 - \text{ProbNNK}) > N$.

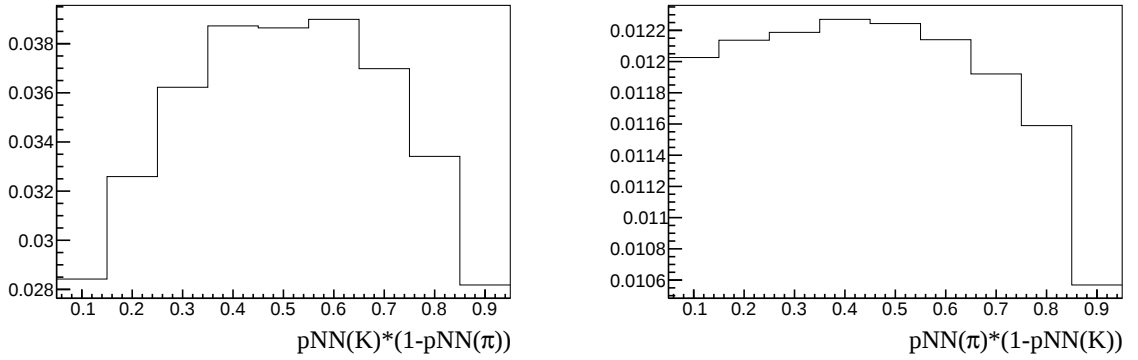


Figure 19: Punzi figure of merit for the PID requirements on the bachelor kaon and pion.

To determine the Punzi FoM, the efficiencies are obtained using MC with corrections obtained using `PIDCalib`, while the background level is estimated from the B mass sideband, as discussed above. From Fig. 19 it is shown that the optimum `ProbNN` cut for the bachelor kaon is at 0.6 while for the bachelor pion the optimum `ProbNN` cut is at 0.4. A requirement, optimised in a similar way, is also made on the pion coming from the decay of the D meson. The Punzi FoM plot for the daughter pion shown in Fig. 20 suggests that a cut at 0.3 was the optimum value. No requirement is found to be necessary on the kaon from the D decay (the FoM peaks at a cut of zero). The particle identification requirements are summarised in Table 9.

Table 9: Particle identification requirements.

Particle	Variable	Requirement
K^+	$\text{ProbNNK} \times (1 - \text{ProbNNpi})$	> 0.6
π^+	$\text{ProbNNpi} \times (1 - \text{ProbNNK})$	> 0.4
D daughter π^-	$\text{ProbNNpi} \times (1 - \text{ProbNNK})$	> 0.3

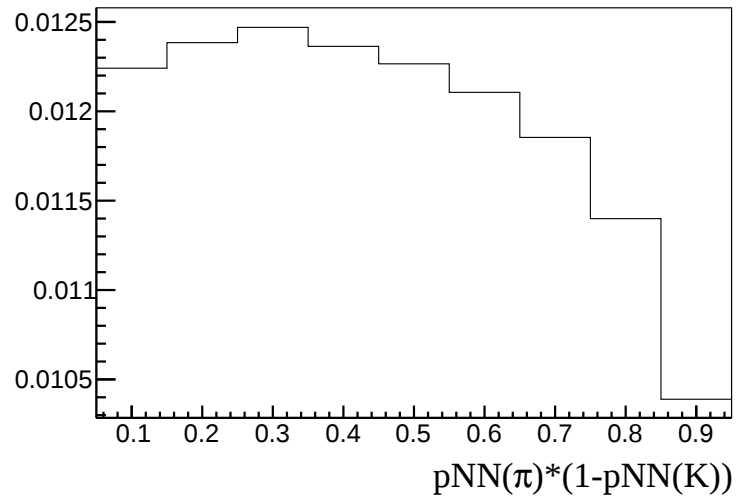


Figure 20: Punzi figure of merit for the PID requirement on the D daughter pion.

4.5 $B^0 \rightarrow D^{*-}\pi^+$ veto

After the selection cuts discussed above are applied, a peaking background from $B^0 \rightarrow D^{*-}\pi^+$ decays can be seen in the $D^{*-}\pi^+\pi^+$ mass distribution around $5460 \text{ MeV}/c^2$. Peaking background occurs when an alternative decay mode is incorrectly reconstructed as signal. To remove this background, the $D^{*-}\pi^+$ invariant mass is calculated (two combinations per candidate; no B mass constraint applied unlike in the calculations of Dalitz plot variables) and those candidates under the $B^0 \rightarrow D^{*-}\pi^+$ peak, $5200 < m(D^{*-}\pi^+) < 5400 \text{ MeV}/c^2$, are vetoed. This requirement effectively removes the $D^*\pi$ background with negligible loss of signal, as shown in Fig. 21.

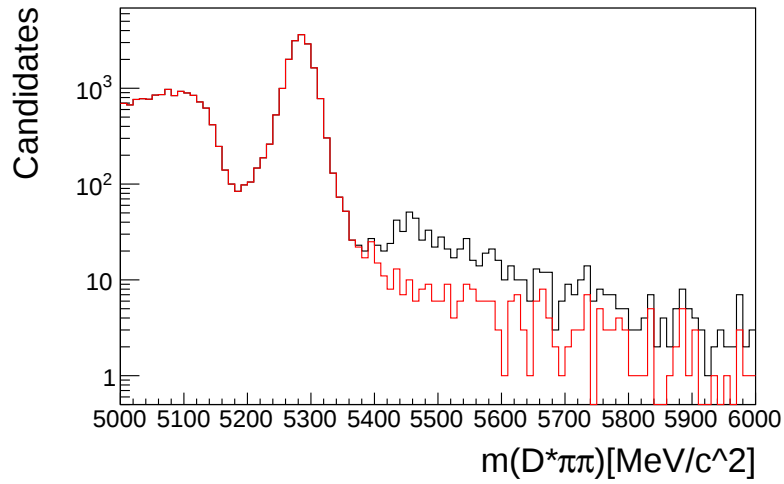


Figure 21: Invariant mass distribution for $D^{*-}\pi^+\pi^+$ candidates (black) before and (red) after the $B^0 \rightarrow D^{*-}\pi^+$ veto.

4.6 Yield

The selection requirements discussed above are applied to the $D^{*-}K^+\pi^+$ data set resulting in 1 408 candidates. Of those candidates, 609 are in the B mass window, $m_{B^+}^{\text{PDG}} \pm 25 \text{ MeV}/c^2$. For the $D^{*-}\pi^+\pi^+$ data set 30 210 candidates are selected, while 13 490 of those candidates fall within the B mass window $m_{B^+}^{\text{PDG}} \pm 25 \text{ MeV}/c^2$. The mass distributions of $D^{*-}K^+\pi^+$ and $D^{*-}\pi^+\pi^+$ candidates are shown in Fig. 22

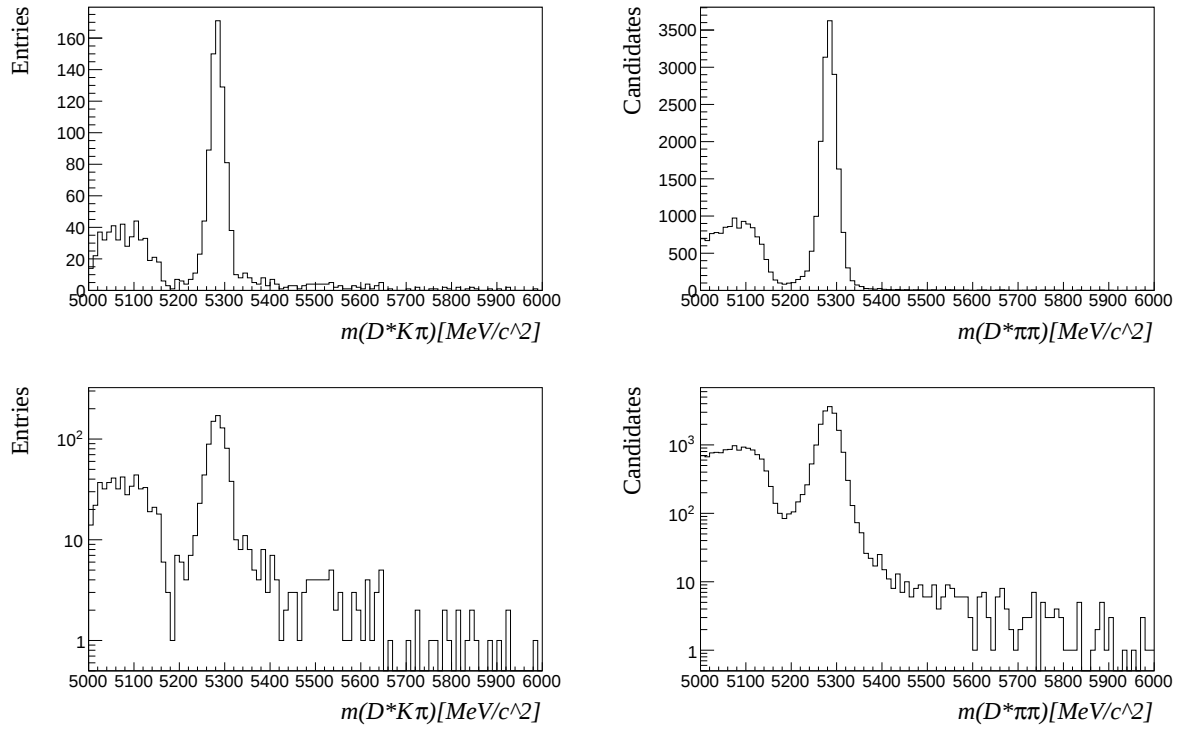


Figure 22: Mass distribution of (left) $D^{*-}K^+\pi^+$ and (right) $D^{*-}\pi^+\pi^+$ after initial selection cuts applied to data. Below each figure the same plot is shown in log scale.

5 Fitting the B mass candidates

One-dimensional extended likelihood fits to the $B^+ \rightarrow D^{*-} K^+ \pi^+$ and $B^+ \rightarrow D^{*-} \pi^+ \pi^+$ candidate mass distributions are performed in order to determine the yields. The fits are performed using **RooFit**, a toolkit for data modelling that allows for modelling probability distributions, and include signal and background contributions as follows:

- $B^+ \rightarrow D^{*-} K^+ \pi^+$
 - signal
 - combinatorial background
 - cross feed from $B^+ \rightarrow D^{*-} \pi^+ \pi^+$
 - partially reconstructed background
- $B^+ \rightarrow D^{*-} \pi^+ \pi^+$
 - signal
 - combinatorial background
 - partially reconstructed background

Combinatorial background occurs when a real or fake D meson is reconstructed with two random tracks. Partially reconstructed backgrounds come from the decay of a B meson to a D meson with two charged tracks and unreconstructed particles. The data for both $D^{*-} K^+ \pi^+$ and the $D^{*-} \pi^+ \pi^+$ control channel are fitted in the range $5000 \text{ MeV}/c^2$ to $6000 \text{ MeV}/c^2$. The fit models are simpler than in many other Dhh' final states since the very clean signature of the $D^{*-} \rightarrow \bar{D}^0 \pi^-$ decay greatly reduces the range of possible backgrounds.

5.1 Signal shape

The signal probability distribution functions (PDFs) for both $D^{*-} K^+ \pi^+$ and $D^{*-} \pi^+ \pi^+$ consist of a double Crystal Ball function [26]. The Crystal Ball function has the form,

$$\text{CB}(x; \mu, \sigma, \alpha, n) = \frac{\left(\frac{n}{|\alpha|}\right)^n \exp\left(-\frac{1}{2}\alpha^2\right)}{\left(\frac{n}{|\alpha|} - |\alpha| - \frac{x-\mu}{\sigma}\right)^n} \bigg|_{x < \mu - |\alpha|\sigma}, \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \bigg|_{x > \mu - |\alpha|\sigma}, \quad (12)$$

where the parameters are the mean μ , width σ , and tail parameters n and α . The double Crystal Ball function is then,

$$\mathcal{P}_{\text{sig}}(m) = f \text{CB}(m; \mu, \sigma, \alpha_1, n_1) + (1 - f) \text{CB}(m; \mu, \sigma, \alpha_2, n_2), \quad (13)$$

where the tails of the two components are on opposite sides of the peak.¹

¹ This is implemented by switching the ranges in which the two parts of Eq. (12) are valid to $x > \mu + |\alpha|\sigma$ and $x < \mu + |\alpha|\sigma$. The two cases are indicated by positive and negative signs for α_1 and α_2 , respectively.

Signal MC samples are fitted in order to obtain the values for the tail parameters which are subsequently fixed in the fit to the data. In order to stabilise the fit, and reduce the large correlations which otherwise occur, $\alpha_2 = -\alpha_1$ is imposed. In fits to the MC samples where α_1 and α_2 are treated as independent parameters they are found to have consistent values. The MC samples have been processed with the full selection criteria except the PID cuts. The effects of the PID criteria are implemented by weighting each candidate according to the corresponding PID efficiencies given by `PIDCalib` (see Sec. 4.4). No truth matching is applied, since it is not possible to determine whether a candidate is signal or not in an unambiguous yet perfect way.² Consequently, a small combinatorial background contribution may exist in the MC samples; this is modelled in the fits with a component described by an exponential function. The results of the fits to the MC samples are shown in Fig. 23, with a summary of the fitted parameters in Table 10.

Table 10: Signal PDF parameters obtained from fits to MC samples. Units of MeV/c^2 are implied for μ and σ .

Parameter	$D^{*-}K^+\pi^+$	$D^{*-}\pi^+\pi^+$
μ	5280.16 ± 0.24	5280.0 ± 0.27
σ	15.11 ± 0.23	15.70 ± 0.30
α_1	1.68 ± 0.067	1.44 ± 0.076
α_2	$-\alpha_1$	$-\alpha_1$
n_1	1.63 ± 0.15	2.14 ± 0.21
n_2	0.95 ± 0.096	2.68 ± 0.34
f	0.81 ± 0.031	0.74 ± 0.042

5.2 Combinatorial background

The combinatorial background PDF for both $B^+ \rightarrow D^{*-}K^+\pi^+$ and $B^+ \rightarrow D^{*-}\pi^+\pi^+$ channels is described with an exponential function. The exponent and yield are allowed to vary during the fit to data.

5.3 Cross-feed background

Cross-feed background is expected from $B^+ \rightarrow D^{*-}\pi^+\pi^+$ decays misidentified as $B^+ \rightarrow D^{*-}K^+\pi^+$ candidates. Indeed, this contribution can be seen in the data (Fig. 22) as a shoulder at $\sim 5350 \text{ MeV}/c^2$ in the $D^{*-}K^+\pi^+$ mass distribution. The shape for this contribution was obtained from data by reconstructing $B^+ \rightarrow D^{*-}\pi^+\pi^+$ decays as $D^{*-}K^+\pi^+$ and using `PIDCalib` to weight the events according to their misidentification probability. The shape is modelled by a single Crystal Ball function. The shape is subsequently fixed in the fit to the $B^+ \rightarrow D^{*-}K^+\pi^+$ candidates, and the yield left free to vary.

² See Appendix E of Ref. [12].

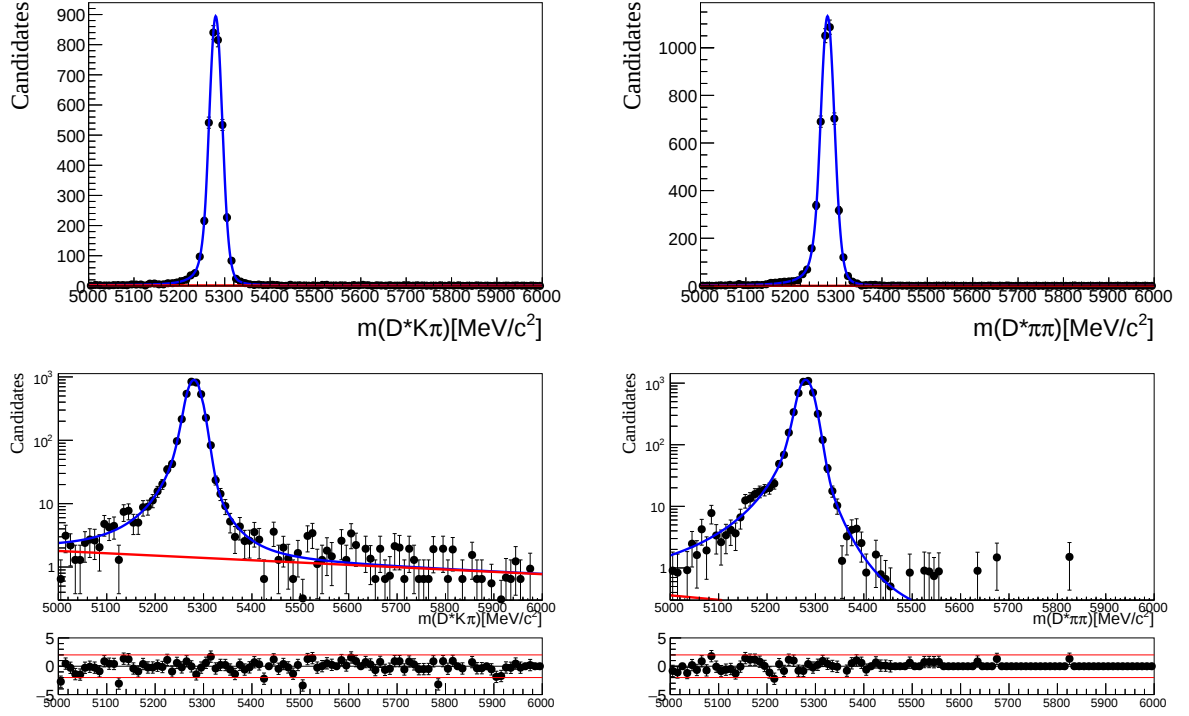


Figure 23: Fit to B mass candidates in (left) $D^{*-}K^+\pi^+$ and (right) $D^{*-}\pi^+\pi^+$ MC samples. The (blue) signal shapes are fit with a double Crystal Ball function while the (red) combinatorial background (invisible) is described with an exponential model. The top (bottom) plots are identical but with linear (logarithmic) y -axis scales.

In principle there could also be cross-feed in the other direction, *i.e.* from $B^+ \rightarrow D^{*-}K^+\pi^+$ decays misidentified as $B^+ \rightarrow D^{*-}\pi^+\pi^+$ candidates. However, the relative branching fractions and (mis)identification probabilities render the contribution negligible, therefore it is not included in the fit.

5.4 Partially reconstructed background

Partially reconstructed backgrounds are expected and observed to contribute, in both $D^{*-}\pi^+\pi^+$ and $D^{*-}K^+\pi^+$ channels, with a broad peaking structure below around $5200 \text{ MeV}/c^2$. For the signal channel, this component includes possible contributions from $D^{*-}K^+\pi^+\pi$ where the extra charged or neutral pion is “missing”, *i.e.* is not included in the reconstruction of the candidate. Examples of decays that would give such a final state are $D^{**}K\pi$ where $D^{**} \rightarrow D^*\pi$, and $B^0 \rightarrow D^{*-}K^{*0}\pi^+$ with $K^{*0} \rightarrow K^+\pi^-$.

The partially reconstructed background is described in the same way as in some other analyses, for example in Ref. [27], namely the convolution of an ARGUS function and

Gaussian function. The ARGUS function [28] is defined by,

$$f(m) = m \left(1 - \left(\frac{m}{m_0} \right)^2 \right)^{\frac{1}{2}} \times \exp \left(c \left(1 - \left(\frac{m}{m_0} \right)^2 \right) \right). \quad (14)$$

This is convolved with a Gaussian function with mean set to zero and a width that is shared with the signal Crystal Ball function. In the ARGUS function, Eq. (14), m is the variable corresponding to the B candidate invariant mass, m_0 is the cutoff energy, and c is the shape parameter. Both m_0 and c are floated during the fit. The parameter m_0 is expected to take a value close to $m_B - m_\pi$, but is floated in the fit since this is found to be possible without any reduction in fit stability.

5.5 Fit to data

The total PDFs for the fits to the $D^{*-}K^+\pi^+$ and $D^{*-}\pi^+\pi^+$ samples include the signal shape, combinatorial background, partially reconstructed background, and ($B^+ \rightarrow D^{*-}\pi^+\pi^+ \rightarrow B^+ \rightarrow D^{*-}K^+\pi^+$) cross-feed. The yields for each component are allowed to vary in the fit, as are the peak position (μ) and width (σ) of the signal PDF, the exponential (τ) for the combinatorial background shape, and the m_0 and c parameters of the partially reconstructed background PDF. The fits are shown in Fig. 24. The results of the fit to the $D^{*-}K^+\pi^+$ sample are given in Table 11, with those from the fit to the $D^{*-}\pi^+\pi^+$ sample in Table 12.

The significance of the $B^+ \rightarrow D^{*-}K^+\pi^+$ signal is clearly overwhelming, but is nonetheless estimated from the difference in negative log-likelihood between the nominal fit and an alternative fit in which the signal yield is fixed to zero. In the alternative fit, the signal μ and σ parameters are fixed to nominal values. It is found to be 31.5σ considering only statistical uncertainty.

The obtained yield of $B^+ \rightarrow D^{*-}\pi^+\pi^+ \rightarrow B^+ \rightarrow D^{*-}K^+\pi^+$ is cross-checked to make sure the cross-feed is consistent with the expected rate from the known particle (mis)identification probabilities and the correctly reconstructed $B^+ \rightarrow D^{*-}\pi^+\pi^+$ yield. The ratio of (mis)identification probabilities is 0.00261, while the ratio of the cross-feed to correctly reconstructed yields is 0.00305 ± 0.00086 .

5.6 Dalitz plots

The data in the B mass signal region in Fig. 24 can be used to show the Dalitz plot distributions. These distributions, shown in Fig. 25, contain a small amount of background and are not efficiency corrected. The B mass signal regions are taken as $5200 < m(D^{*-}K^+\pi^+) < 5360 \text{ MeV}/c^2$ and similarly for $D^{*-}\pi^+\pi^+$. Their projections onto $m(D^{*-}\pi^+)$ or $m(D^{*-}\pi^+)_{\min}$ are shown in Fig. 26. A contribution from the $D_1(2420)$ resonance is immediately clear.

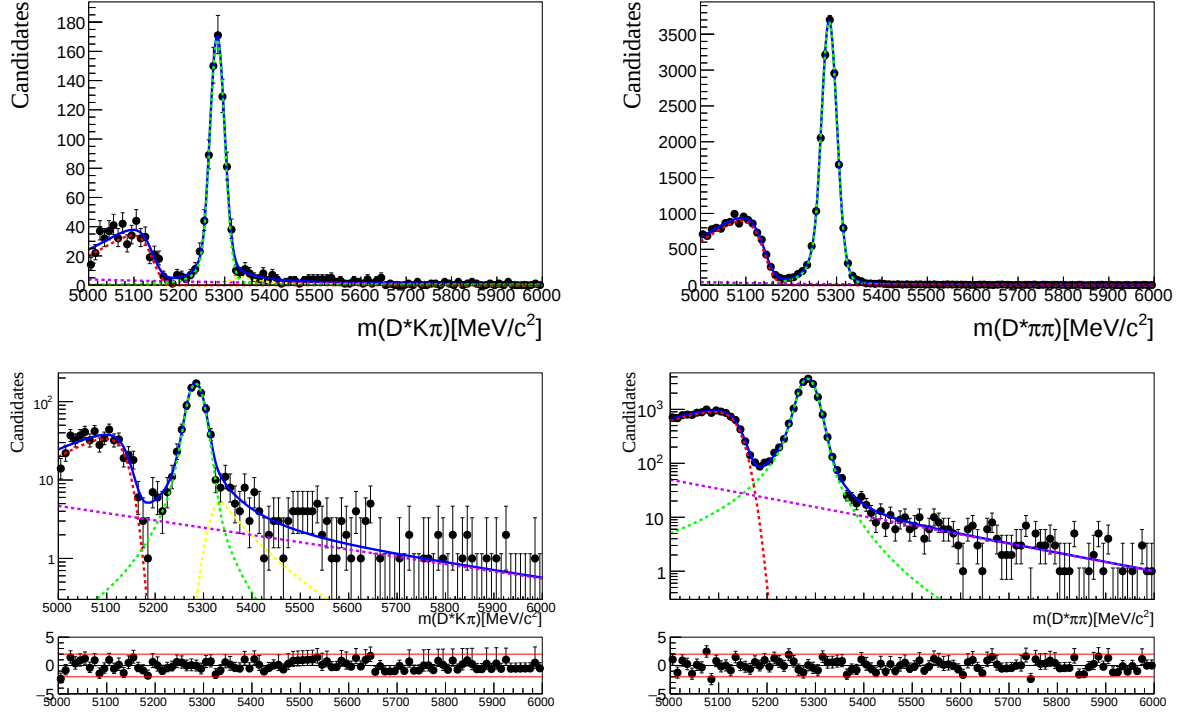


Figure 24: Fit to B candidate mass distributions for (left) $D^{*-}K^+\pi^+$ and (right) $D^{*-}\pi^+\pi^+$ samples with (top) linear and (bottom) logarithmic y -axis scales. Included under the logarithmic plots are the pull distributions where the red line is 2 sigma. The individual components are (blue) total PDF, (green) signal shape, (violet) combinatorial background, (red) partially reconstructed background, and (yellow) $D^{*-}\pi^+\pi^+ \rightarrow D^{*-}K^+\pi^+$ cross-feed.

Table 11: Parameters obtained from fit to $D^{*-}K^+\pi^+$ data. Units of MeV/c^2 (c^2/MeV) are implied for μ , σ and m_0 (τ).

Parameter	$D^{*-}K^+\pi^+$
N_{signal}	770.7 ± 30.4
$N_{\text{comb. bkgd.}}$	157.3 ± 39.8
$N_{B^+ \rightarrow D^{*-}\pi^+\pi^+}$	55.1 ± 15.6
N_{PRB}	424.9 ± 30.3
μ	5283.57 ± 0.72
σ	16.57 ± 0.61
τ	-0.0022 ± 0.0006
m_0	5148.6 ± 3.3
c	-28.8 ± 4.6

Table 12: Parameters obtained from fit to $D^{*-}\pi^+\pi^+$ data. Units of MeV/c^2 (c^2/MeV) are implied for μ , σ and m_0 (τ).

Parameter	$D^{*-}\pi^+\pi^+$
N_{signal}	17593.5 ± 142.7
$N_{\text{comb. bkgd.}}$	11520.6 ± 158.1
N_{PRB}	1096.0 ± 163.4
μ	5283.6 ± 0.16
σ	17.5 ± 0.14
τ	-0.0039 ± 0.0004
m_0	5148.6 ± 0.9
c	-24.7 ± 1.0

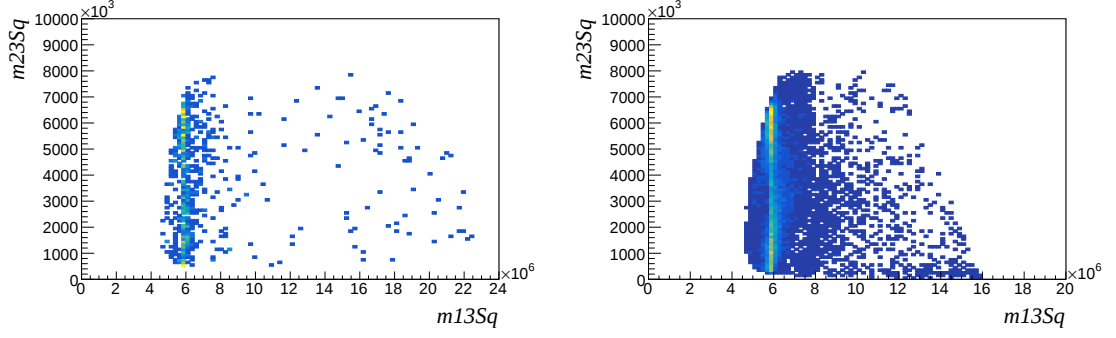


Figure 25: Dalitz plot distributions for candidates in the signal regions of (left) $B^+ \rightarrow D^{*-}K^+\pi^+$ and (right) $B^+ \rightarrow D^{*-}\pi^+\pi^+$ decays.

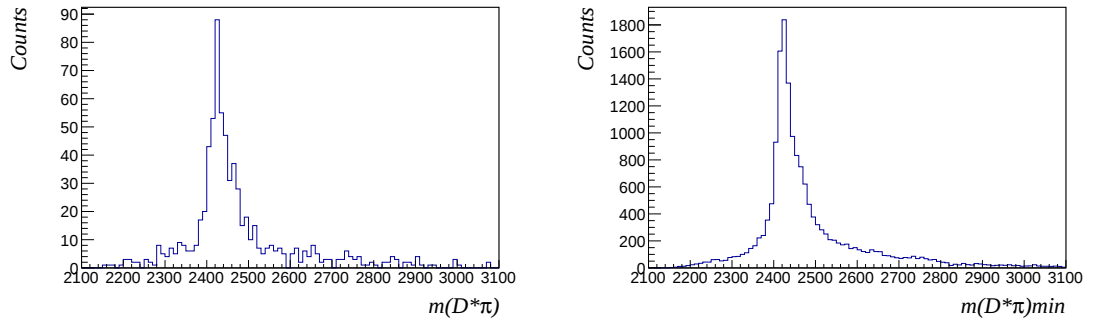


Figure 26: One dimensional distributions for candidates in the signal regions of (left) $B^+ \rightarrow D^{*-}K^+\pi^+$ and (right) $B^+ \rightarrow D^{*-}\pi^+\pi^+$ decays.

6 Efficiencies

A selection which selects decays in any region of the phase space without any bias is ideal. However, the efficiencies vary as a function of the Dalitz plot and must therefore be accounted for. The different contributions to the efficiencies are studied individually in order to understand each separately.

Each of the efficiencies is evaluated using MC apart from that due to the PID requirements, which is taken from data control samples. Geometrical efficiencies account for whether or not the daughter particles are within the LHCb acceptance. The geometrical efficiency is denoted by ϵ^{geom} . The selection efficiency includes the selection requirements from Sec. 4 including offline selection requirements as well as trigger and stripping, with the exception of the PID requirements. The selection efficiency is denoted by $\epsilon^{\text{sel|geom}}$, with the notation implying that this is evaluated for candidates that are within the geometrical acceptance only. The PID efficiencies are determined using the `PIDCalib` package to obtain the appropriate efficiencies for tracks with kinematics as given by the signal MC samples. The PID efficiency is denoted by $\epsilon^{\text{PID|sel\&geom}}$. The total efficiency is calculated as

$$\epsilon^{\text{tot}} = \epsilon^{\text{geom}} \times \epsilon^{\text{sel|geom}} \times \epsilon^{\text{PID|sel\&geom}}. \quad (15)$$

The phase-space of the $B^+ \rightarrow D^{*-} K^+ \pi^+$ decay has four degrees of freedom, having a spin-0 particle in the initial state and one spin-1 plus two spin-0 particles in the final state. These degrees of freedom can be chosen to be the Dalitz plot variables $m^2(D^{*-}\pi^+)$ and $m^2(D^{*-}K^+)$ together with the D^{*-} helicity angle, $\theta_{\text{hel}}^{D^{*-}}$, and the angle χ between the decay planes. As shown in Appendix A, the efficiency function can be assumed to factorise as

$$\epsilon\left(m^2(D^{*-}\pi^+), m^2(D^{*-}K^+), \theta_{\text{hel}}^{D^{*-}}, \chi\right) = \epsilon\left(m^2(D^{*-}\pi^+), m^2(D^{*-}K^+)\right) \times \epsilon\left(\theta_{\text{hel}}^{D^{*-}}\right) \times \epsilon(\chi). \quad (16)$$

The variation of the efficiency with respect to $\theta_{\text{hel}}^{D^{*-}}$ and χ is found to be small, as shown in Appendix A. Thus, it is neglected in the baseline results, and is treated as a source of systematic uncertainty. The variation of the efficiency with respect to the Dalitz plot variables, however, is significant and is accounted for as described in the remainder of this section.

6.1 Geometrical efficiencies

The geometrical efficiencies for $D^{*-}K^+\pi^+$ and $D^{*-}\pi^+\pi^+$ are generated from 2011 and 2012 MC. The 2011 and 2012 geometrical efficiencies are then averaged together, accounting for the different integrated luminosities in the two years. It is required that $0.01 < \theta < 0.40$ for all the daughter particles, which is the standard definition of the LHCb acceptance. This ensures that the tracks of the daughter particles are within the LHCb acceptance. Fig. 27 shows the variation of the geometrical efficiencies over the Dalitz plot. The plots show the ratio of the Dalitz plot with the geometrical cuts applied to the Dalitz plot with no cuts applied. The Dalitz plot for $B^+ \rightarrow D^{*-}K^+\pi^+$ is shown in terms of

$m^2(K^+\pi^+)$ vs $m^2(D^{*-}\pi^+)$, while that for $B^+ \rightarrow D^{*-}\pi^+\pi^+$ is shown in terms of $m^2(\pi^+\pi^+)$ vs $m^2(D^{*-}\pi^+)_{\min}$. Here $m^2(D^{*-}\pi^+)_{\min}$ is the smaller of the two values of $m^2(D^{*-}\pi^+)$ that can be calculated in a $B^+ \rightarrow D^{*-}\pi^+\pi^+$ decay; hence the $D^{*-}\pi^+\pi^+$ Dalitz plot is “folded”. This results in different, but similar, shapes for the Dalitz plots for the two decays. The geometrical efficiency tends to be slightly lower around the borders of the Dalitz plot.

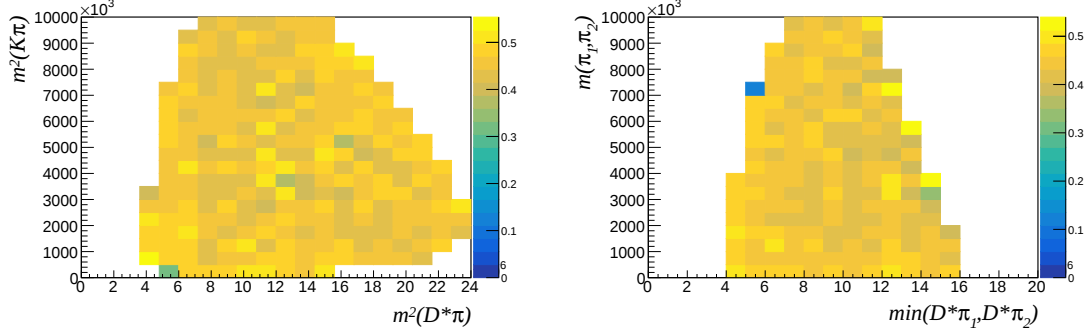


Figure 27: Geometrical efficiency for (left) $D^{*-}K^+\pi^+$ and (right) $D^{*-}\pi^+\pi^+$ over the Dalitz plot. The units of both x and y axes are MeV^2/c^4 , with the “ $\times 10^6$ ” label on the x -axis obscured.

6.2 Selection efficiencies

The variation of the selection efficiencies over the Dalitz plot is studied by taking the ratio of the yields of selected B mass candidates after all selection requirements are applied, to those produced in standard LHCb MC production, with the standard acceptance requirement applied. The selection efficiencies are shown in Fig. 28. Similarly to the geometrical efficiency, the selection efficiency tends to be lower near to the borders of the Dalitz plot.

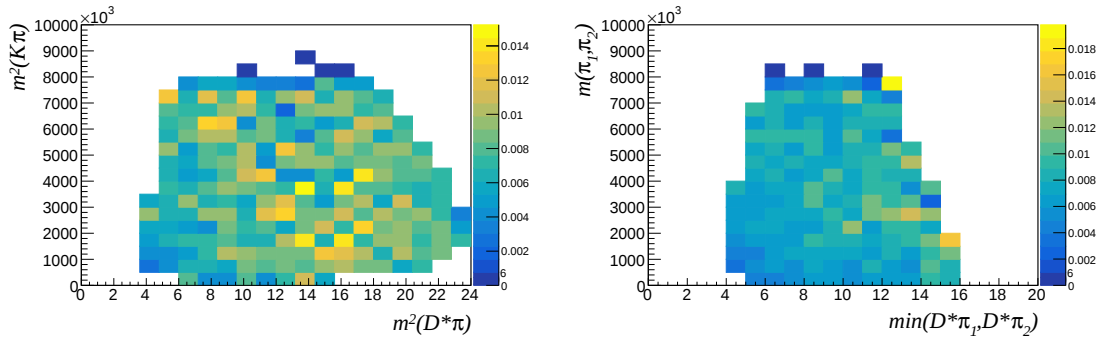


Figure 28: Selection efficiency for (left) $D^{*-}K^+\pi^+$ and (right) $D^{*-}\pi^+\pi^+$ over the Dalitz plot. The units of both x and y axes are MeV^2/c^4 , with the “ $\times 10^6$ ” label on the x -axis obscured.

6.3 PID efficiencies

The variation of the PID efficiencies over the Dalitz plot is studied using PIDCalib. For both decay modes, relatively little variation is seen over the Dalitz plot, however the $D^{*-}\pi^+\pi^+$ PID efficiencies are much closer to 1 compared to those for $D^{*-}K^+\pi^+$. This is due to having comparatively looser PID requirements for $D^{*-}\pi^+\pi^+$.

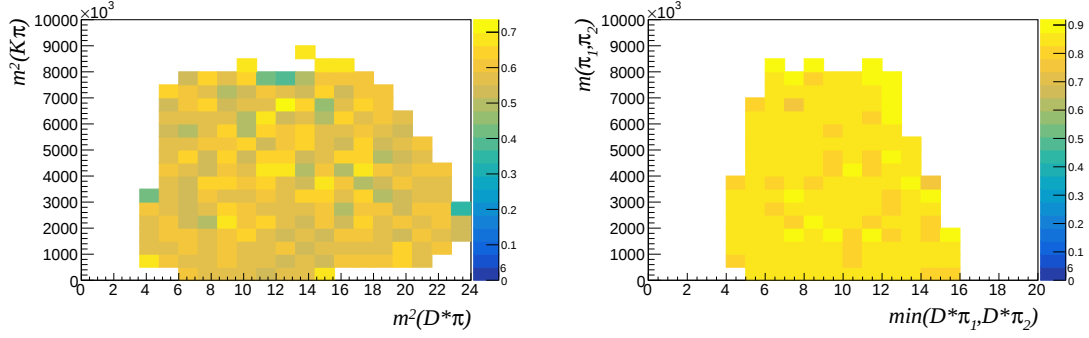


Figure 29: PID efficiency for (left) $D^{*-}K^+\pi^+$ and (right) $D^{*-}\pi^+\pi^+$ over the Dalitz plot. The units of both x and y axes are MeV^2/c^4 , with the “ $\times 10^6$ ” label on the x -axis obscured.

6.4 Total efficiencies

The efficiencies discussed above, ϵ^{geom} , $\epsilon^{\text{sel|geom}}$, and $\epsilon^{\text{PID|sel\&geom}}$, are multiplied together to make the total efficiency as in Eq. (15). These total efficiency histograms, shown in Fig. 30, are used to calculate the relative branching fraction ratio discussed later in Sec. 7.

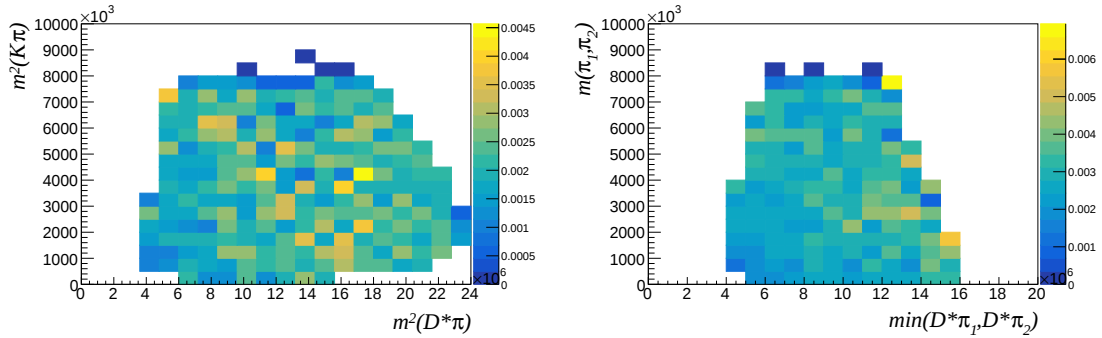


Figure 30: Total efficiency for (left) $D^{*-}K^+\pi^+$ and (right) $D^{*-}\pi^+\pi^+$ over the Dalitz plot. The units of both x and y axes are MeV^2/c^4 , with the “ $\times 10^6$ ” label on the x -axis obscured.

7 Branching fraction ratio

The branching fraction ratio is calculated by

$$\frac{\mathcal{B}(B^+ \rightarrow D^{*-} K^+ \pi^+)}{\mathcal{B}(B^+ \rightarrow D^{*-} \pi^+ \pi^+)} = \frac{N^{\text{corr}}(D^* K \pi)}{N^{\text{corr}}(D^* \pi \pi)}. \quad (17)$$

Here

$$N^{\text{corr}}(D^* h h) = \sum_i \frac{W_i}{\epsilon_i^{\text{tot}}} = \sum_i \frac{W_i}{\epsilon^{\text{geom}} \times \epsilon^{\text{sel|geom}} \times \epsilon^{\text{PID|sel\&geom}}}. \quad (18)$$

In Eq. (18), W_i are the sWeights for event i , ϵ_i^{tot} are the total efficiencies discussed in Sec. 6, and i runs over all candidates in the B mass distribution used in the fit. The sWeights are obtained from the mass fits presented in Sec. 5.5, modified so that only the yields are floated, as required to obtain correct sWeights [29]. The corrected yields of candidates for $D^{*-} K^+ \pi^+$ and $D^{*-} \pi^+ \pi^+$ are shown in Table 13. The uncertainty on N^{corr} is calculated using the equation

$$\sigma(N^{\text{corr}}) = \sqrt{\sum_i \left(\frac{W_i}{\epsilon_i^{\text{tot}}} \right)^2}. \quad (19)$$

Table 13: Yields and corrected yields for $B^+ \rightarrow D^{*-} K^+ \pi^+$ and $B^+ \rightarrow D^{*-} \pi^+ \pi^+$ decays. The yields N are taken from the results of the fits, and the corrected yields N^{corr} are calculated using Eq. (18).

	$D^{*-} K^+ \pi^+$	$D^{*-} \pi^+ \pi^+$
N	771 ± 30	$17\,600 \pm 143$
N^{corr}	$441\,000 \pm 17\,400$	$6\,800\,000 \pm 54\,700$

However, it is possible to have the uncertainty on the yield from the **sWeights** to be less than the uncertainty returned by the nominal fit since only the yields are floating during the sWeighting. To correct for this effect, our uncertainty equation must be modified and properly scaled [30]. This is calculated by

$$\sigma^{\text{corr}}(N^{\text{corr}}) = \sqrt{\sigma(N^{\text{corr}})^2 + \left(\frac{N^{\text{corr}}}{N} \sigma^{\text{shape}}(N) \right)^2}, \quad (20)$$

where $\sigma^{\text{shape}} = \sqrt{\sigma^{\text{fit}}(N)^2 - \sigma^{\text{yields only}}(N)^2}$. The values of the uncertainties are listed in Tab. 14.

Using Eq. (17), The ratio of the branching fractions is found to be

$$\frac{\mathcal{B}(B^+ \rightarrow D^{*-} K^+ \pi^+)}{\mathcal{B}(B^+ \rightarrow D^{*-} \pi^+ \pi^+)} = \frac{N^{\text{corr}}(D^* K \pi)}{N^{\text{corr}}(D^* \pi \pi)} = 0.0649 \pm 0.0026, \quad (21)$$

where the uncertainty is statistical only.

Table 14: Various uncertainties calculated for $D^*K\pi$ and $D^*\pi\pi$

	$D^*K\pi$	$D^*\pi\pi$
σ^{fit}	30.4	142.7
$\sigma^{\text{yields only}}$	29.7	137.1
σ^{shape}	6.4	39.4
σ^{corr}	17 400	54 700

7.1 Background-subtracted and efficiency-corrected Dalitz plots

The determination of the signal weights used to evaluate the branching fraction also allows background-subtracted and efficiency-corrected Dalitz plot distributions to be obtained. Figure 31 shows the distribution obtained by weighting each candidate in the whole mass fit range by W_i – this therefore gives the background-subtracted Dalitz plot distribution [29]. Their projections onto $m(D^{*-}\pi^+)$ or $m(D^{*-}\pi^+)_{\text{min}}$ are shown in Fig. 32. Figure 33 shows the distribution obtained by weighting each candidate in the whole mass fit range by $W_i/\epsilon_i^{\text{tot}}$ – this therefore gives the background-subtracted and efficiency-corrected Dalitz plot distribution. Their projections onto $m(D^{*-}\pi^+)$ or $m(D^{*-}\pi^+)_{\text{min}}$ are shown in Fig. 34. As before, the $D_1(2420)$ resonance contribution is very clear; the background subtraction and efficiency correction do not change the picture qualitatively.

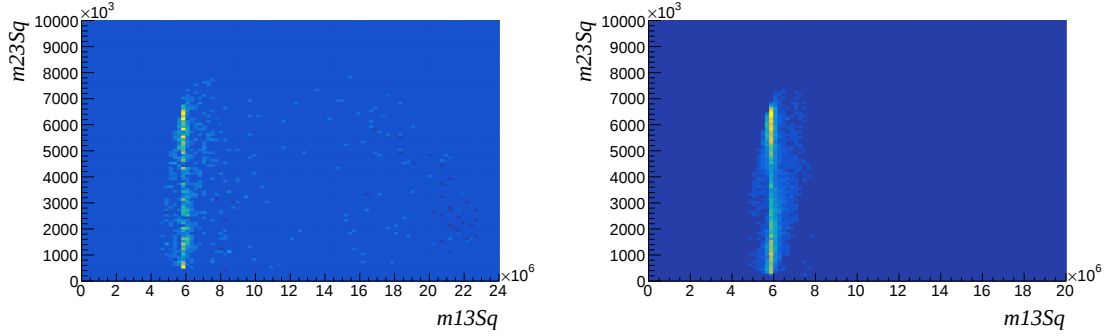


Figure 31: Background-subtracted Dalitz plot distributions, obtained by weighting each candidate in the whole mass fit range by W_i , of (left) $B^+ \rightarrow D^{*-}K^+\pi^+$ and (right) $B^+ \rightarrow D^{*-}\pi^+\pi^+$ decays.

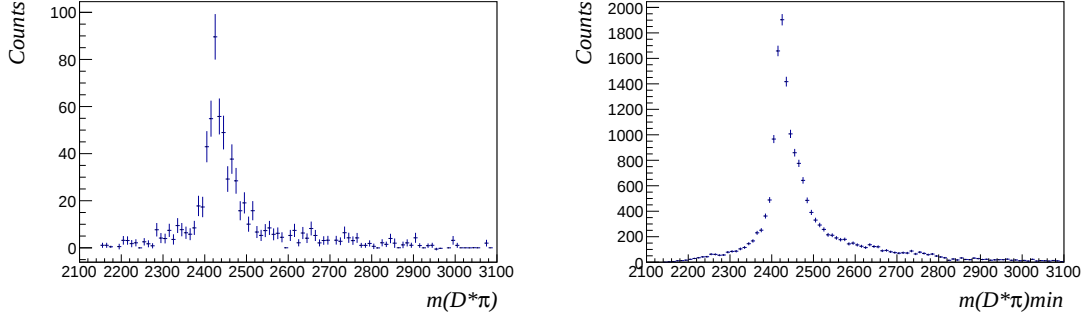


Figure 32: Background-subtracted one dimensional plot distributions, obtained by weighting each candidate in the whole mass fit range by W_i , of (left) $B^+ \rightarrow D^{*-}K^+\pi^+$ and (right) $B^+ \rightarrow D^{*-}\pi^+\pi^+$ decays.

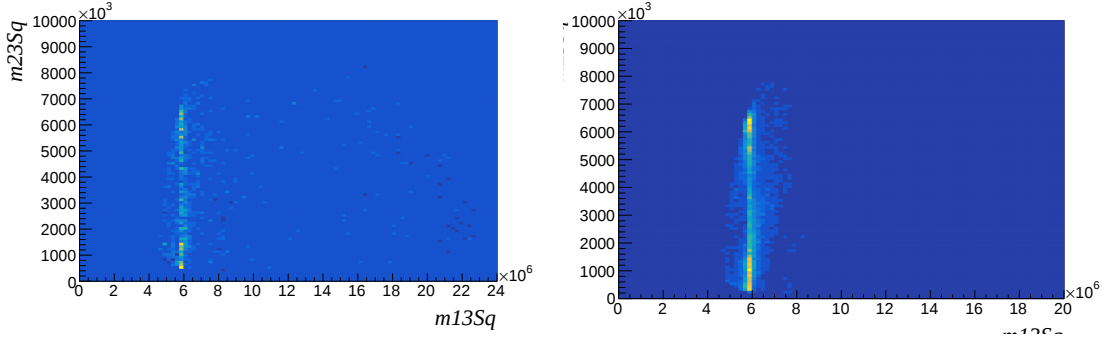


Figure 33: Background-subtracted and efficiency-corrected Dalitz plot distributions, obtained by weighting each candidate in the whole mass fit range by $W_i/\epsilon_i^{\text{tot}}$, of (left) $B^+ \rightarrow D^{*-}K^+\pi^+$ and (right) $B^+ \rightarrow D^{*-}\pi^+\pi^+$ decays.

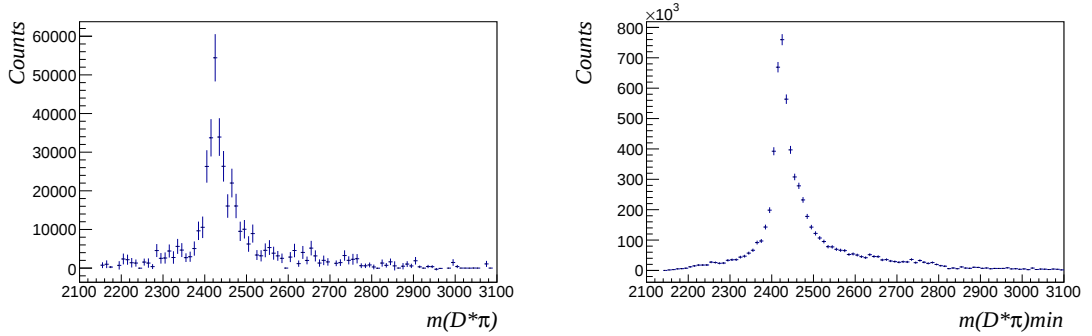


Figure 34: Background-subtracted and efficiency-corrected one dimensional plot distributions, obtained by weighting each candidate in the whole mass fit range by $W_i/\epsilon_i^{\text{tot}}$, of (left) $B^+ \rightarrow D^{*-}K^+\pi^+$ and (right) $B^+ \rightarrow D^{*-}\pi^+\pi^+$ decays.

8 Systematic uncertainties

Some analyses might be troubled with uncertainties coming from the luminosity, $b\bar{b}$ production cross-section, and fragmentation fractions; however these effects cancel exactly in the ratio, described in Sec. 7, of the branching fraction using the methods employed in this analysis. The analysis is also largely robust against uncertainties in the geometrical acceptance and selection requirements. There are however uncertainties caused by differences in the selection for the signal and normalisation modes, specifically for particle identification requirements and for the requirement imposed in the stripping line that removes part of the phase space (see Sec. 4.1). The sources of systematic uncertainty that are evaluated are therefore those due to the fit model, the MC modelling of efficiency variation across the Dalitz plot, possible efficiency variation with the decay angles (see Sec. 6) that is neglected in the baseline result, the determination of PID efficiencies using the `PIDCalib` package, and the possible effect of the stripping requirements. In all cases methods to determine the uncertainties that have been used in previous analyses (*e.g.* Ref. [12]) are exploited. All uncertainties are quoted relative to the central value of the result.

8.1 Fit model

The following variations in the fit model are considered. In each case, the relative difference in the signal yield obtained with the alternative model to that obtained with the baseline fit is assigned as the corresponding uncertainty.

- variation of fixed parameters in the signal shape;
- replacing baseline signal model (double Crystal Ball function) with alternative shape (Crystal Ball + Gaussian);
- fix the threshold value (m_0) to its expected value of $m_B - m_\pi$ instead of allowing it to float;
- change the functional form that describes combinatorial background to a second-order polynomial function instead of exponential;
- in the $D^{*-}\pi^+\pi^+$ fit, include a component to describe possible cross-feed from $D^{*-}K^+\pi^+$ decays, which is neglected in the baseline fit.

The systematic uncertainties due to each of these variations are shown in Table 15. The total is obtained by combining them all in quadrature.

8.2 MC modelling of efficiency

The efficiency is modelled using MC as described in Sec. 6 (the PID efficiency, which is determined in a data-driven way, is discussed separately below). Each contribution to the efficiency, ϵ^{geom} and $\epsilon^{\text{sel|geom}}$, is modelled by binning the Dalitz plot. The systematic

Table 15: Fit model systematic uncertainties.

Source	Uncertainty
Fixed parameters	1.0%
Signal model	2.7%
PR background threshold	1.2%
Combinatorial shape	1.1%
Cross-feed to $D^{*-}\pi^+\pi^+$	0.1%
Total	3.3%

uncertainty due to finite simulation statistics can therefore be evaluated by fluctuating independently the value of the efficiency in each Dalitz plot bin. This is repeated many times, and the spread (RMS) of results obtained in the ensemble, shown in Fig. 35, is taken as an estimate of the associated systematic uncertainty. It is found to be 5.4%.

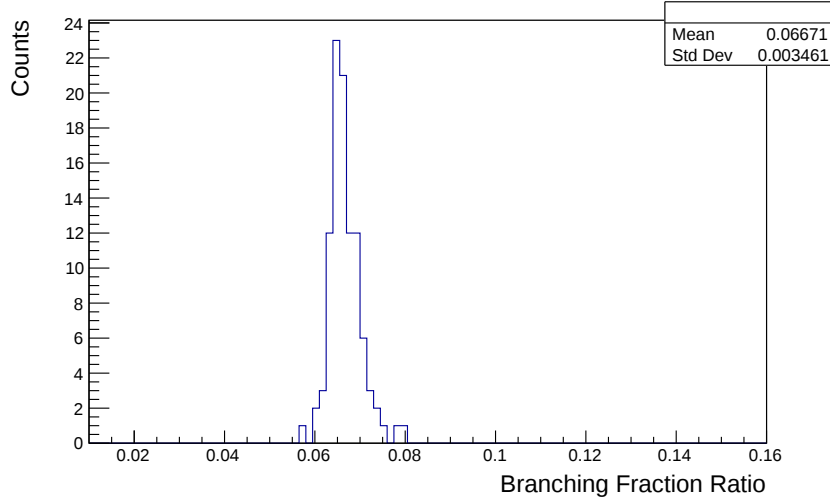


Figure 35: Results for the branching fraction ratio from an ensemble in which the value of the efficiency in each Dalitz plot bin is fluctuated.

8.3 Efficiency variation with decay angles

In App. A, the variation of efficiency with the decay angles χ and θ_{D^*} is shown to be a small effect, and it is therefore neglected in the baseline result. To evaluate the systematic uncertainty due to neglecting the small possible variation, the branching fraction ratio is determined using Eqs. (17) and (18) where in the latter of the efficiencies the dependence on each of the decay angles is included, assuming that they factorise (see Fig. 37). The change in the branching fraction ratio compared to the baseline result is assigned as the uncertainty. This leads to an uncertainty of 0.34%.

8.4 PID efficiency

8.4.1 PIDcalib tool

For each track which differs in the two decays, uncertainties of 0.5 % per track due to the kaon identification requirements and 0.5 % per track due to the pion identification requirements are assigned. To account for possible correlations in these uncertainties, they are added linearly to obtain a total uncertainty of 1.0 %. Note that the effects of the PID requirements placed on the D daughters cancel in the ratio of branching fractions that is measured.

8.4.2 Track kinematics

Our nominal PID efficiency is obtained using a data calibration sample to reweight events according to the kinematics of the signal MC samples. There is therefore a systematic uncertainty due to possible differences between the kinematics of the tracks in data and in MC. Although such differences are known to be small, an uncertainty of 2.0 % based on values obtained in similar analyses [12] is assigned.

8.4.3 Binning

A systematic uncertainty is assigned for the effect of using finer binning in the `PIDcalib` weighting histograms. A binning scheme is chosen for track p_T to keep relative errors roughly constant across the range. Calibration histograms were remade with twice the number of bins in p_T but p binning unchanged. Additionally, the binning in `nTracks` was changed from 3 bins to 4 bins while keeping that in p and p_T the same. This leads to a systematic uncertainty of 1.3%

8.4.4 Total PID systematic uncertainty

The uncertainties from these effects are added in quadrature to give an overall PID systematic uncertainty of 2.7 %.

8.5 Stripping effect on phase space

As discussed in Sec. 4.1, the stripping lines contain a requirement of $m(K^+\pi^+\pi^-) < 3.0 \text{ GeV}/c^2$ (and similarly for $m(\pi^+\pi^+\pi^-)$), included to reduce the rate. This removes part of the phase space for $B^+ \rightarrow D^{*-}K^+\pi^+$ (and $B^+ \rightarrow D^{*-}\pi^+\pi^+$) decays, and therefore introduces a systematic uncertainty if the fraction of signal lost is not the same in the two modes. In order to evaluate this effect, the cut is tightened to $m(K^+\pi^+\pi^-) < 2.7 \text{ GeV}/c^2$ (and similarly for $m(\pi^+\pi^+\pi^-)$), the result re-evaluated, and the shift assigned as the corresponding systematic uncertainty. This leads to a systematic uncertainty of 3.2%

8.6 Summary

The contributions from different sources of systematic uncertainty to the branching fraction ratio of Eq. (17) are summarised in Table 16. The total systematic uncertainty is obtained by combining all contributions in quadrature.

Table 16: Systematic uncertainties on the branching fraction ratio.

Source	Uncertainty
Fit model	3.3%
MC statistics	5.4%
Efficiency variation with decay angles	0.3%
PID efficiency	2.7%
Stripping effect on phase space	3.2%
Total	7.6%

The largest source of systematic uncertainty is due to MC statistics. Changes to the DP binning scheme were investigated to see if this could be reduced, but it was found not to be possible. Production of a larger signal MC sample would obviously reduce this source of uncertainty, although this would be desirable it would cause a significant delay in the analysis and would also have a non-negligible cost in terms of CPU and storage required to generate, process and study the larger samples. Moreover, the uncertainty on the absolute branching fraction due to the normalisation channel branching fraction (shown in Sec. 9) is a factor of two larger than the total of the other sources of systematic uncertainty. We therefore have not attempted to reduce the uncertainty due to MC statistics.

Since the result is systematically limited, the effect of the systematic uncertainties on the significance has been estimated. As the outcome will obviously be far in excess of 5σ , the approximate formula is used

$$s_{\text{tot}} = \frac{s_{\text{stat}}}{\sqrt{1 + \left(\frac{\sigma^{\text{syst}}}{\sigma^{\text{stat}}}\right)^2}}. \quad (22)$$

Here, s_{stat} is the statistical significance (31.5σ) obtained from the change in $-2\ln\mathcal{L}$, and σ^{syst} includes all systematics that effect the yield, which in our case is only the fit model (3.3%). The outcome is a significance of 24.5σ .

8.7 Cross checks

The stability of the result is cross-checked by subdividing the data according to year of data-taking or hardware trigger decision. The results are presented in Table 17 and do not indicate any problems.

Table 17: Cross-checks of the stability of the result when subdividing the data.

	$N^{\text{corr}}(D^*K\pi)$	$N^{\text{corr}}(D^*\pi\pi)$	BF ratio
2011	139 600	2 293 400	0.06085
2012	306 100	5 016 400	0.06102
LOHadron_T0S	301 600	5 018 500	0.06009
LOGlobal_TIS	144 100	2 296 000	0.06277

9 Final results and summary

In this thesis, the methods and procedures outlined in the preceding sections have resulted in the first search for the decay $B^+ \rightarrow D^{*-} K^+ \pi^+$. The analysis is based on the 3 fb^{-1} pp collision data sample collected in 2011 and 2012 with the LHCb detector. The $B^+ \rightarrow D^{*-} K^+ \pi^+$ decay is observed with overwhelming significance and its branching fraction is measured relative to that for $B^+ \rightarrow D^{*-} \pi^+ \pi^+$ to be

$$\frac{\mathcal{B}(B^+ \rightarrow D^{*-} K^+ \pi^+)}{\mathcal{B}(B^+ \rightarrow D^{*-} \pi^+ \pi^+)} = 0.0649 \pm 0.0026 (\text{stat}) \pm 0.0049 (\text{syst}).$$

Using the known value, $\mathcal{B}(B^+ \rightarrow D^{*-} \pi^+ \pi^+) = (1.35 \pm 0.22) \times 10^{-3}$ [6, 16], results in

$$\mathcal{B}(B^+ \rightarrow D^{*-} K^+ \pi^+) = (8.8 \pm 0.4 (\text{stat}) \pm 0.7 (\text{syst}) \pm 1.4 (\text{norm})) \times 10^{-5},$$

where the third uncertainty is due to the precision of the knowledge of the normalisation channel branching fraction.

Appendices

A Helicity angle efficiency plots

The variations of each of the contributions to the efficiency, and the total efficiency, with respect to χ and $\theta_{\text{hel}}^{D^{*-}}$ are presented in Fig. 36. The efficiencies are here integrated over the Dalitz plot variables. No significant effect is seen.

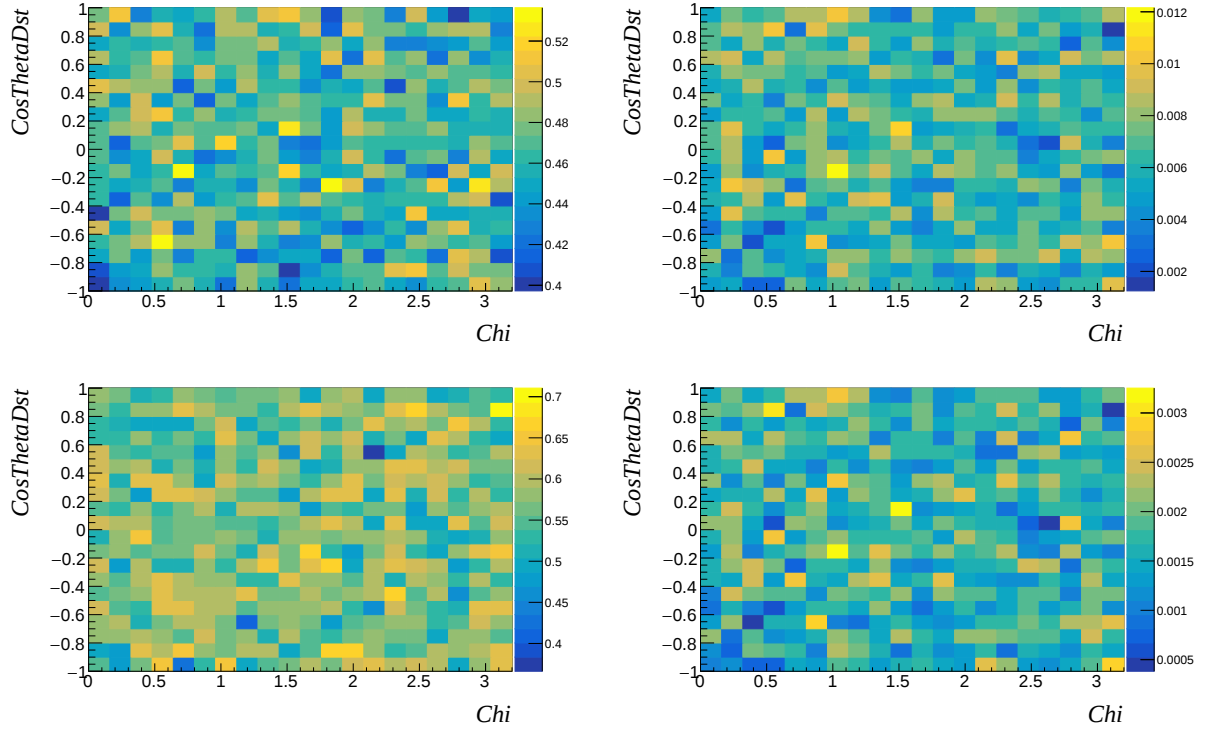


Figure 36: Variation in the efficiency as a function of the helicity angles for $D^{*-}K^+\pi^+$.

The one-dimensional projections of the efficiencies onto χ and $\theta_{\text{hel}}^{D^{*-}}$ are also investigated, as shown in Fig. 37. These are consistent with being flat. The effects of possible small dependencies of the efficiencies on these variables will be accounted for in the systematic uncertainties.

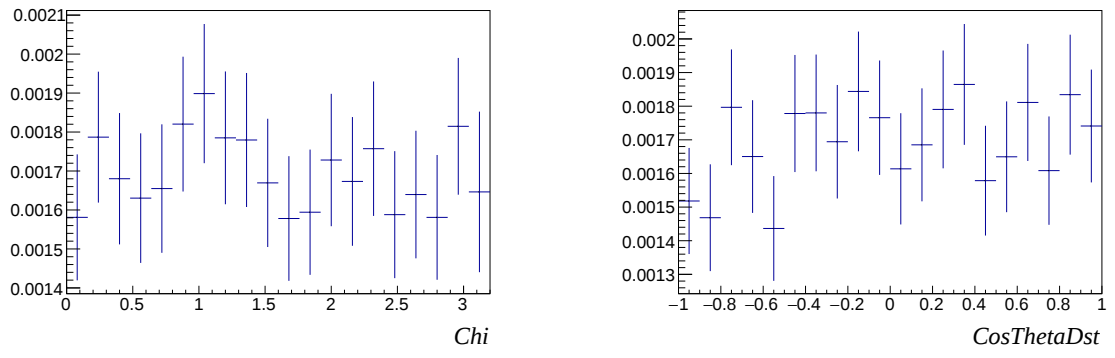


Figure 37: Variation in the one dimensional efficiency as a function of the χ and θ_{D^*} angles.

References

- [1] A. D. Sakharov, *Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32, [Usp. Fiz. Nauk161,61(1991)].
- [2] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π decay of the K_2^0 meson*, Phys. Rev. Lett. **13** (1964) 138.
- [3] M. Thomson, *Modern Particle Physics*, Modern Particle Physics, Cambridge University Press, 2013.
- [4] P. Di Bari, *An introduction to leptogenesis and neutrino properties*, Contemp. Phys. **53** (2012), no. 4 315, [arXiv:1206.3168](#).
- [5] 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](#).
- [6] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001, and 2015 update.
- [7] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973), no. 2 652, [arXiv:http://ptp.oxfordjournals.org/content/49/2/652.full.pdf+html](#).
- [8] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CKM angle γ from a combination of LHCb results*, JHEP **12** (2016) 087, [arXiv:1611.03076](#).
- [9] BaBar collaboration, J. P. Lees *et al.*, *Observation of direct CP violation in the measurement of the Cabibbo-Kobayashi-Maskawa angle gamma with $B^\pm \rightarrow D^{(*)} K^{(*)\pm}$ decays*, Phys. Rev. **D87** (2013), no. 5 052015, [arXiv:1301.1029](#).
- [10] Belle collaboration, K. Trabelsi, *Study of direct CP in charmed B decays and measurement of the CKM angle gamma at Belle*, in *7th International Workshop on the CKM Unitarity Triangle (CKM 2012) Cincinnati, Ohio, USA, September 28-October 2, 2012*, 2013. [arXiv:1301.2033](#).
- [11] LHCb collaboration, R. Aaij *et al.*, *Measurements of the branching fractions of the decays $B_s^0 \rightarrow \bar{D}^0 K^- \pi^+$ and $B^0 \rightarrow \bar{D}^0 K^+ \pi^-$* , Phys. Rev. **D87** (2013) 112009, [arXiv:1304.6317](#).
- [12] LHCb collaboration, R. Aaij *et al.*, *First observation and amplitude analysis of the $B^- \rightarrow D^+ K^- \pi^-$ decay*, Phys. Rev. **D91** (2015) 092002, [arXiv:1503.02995](#).
- [13] LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis of $B^- \rightarrow D^+ \pi^+ \pi^-$ decays*, Phys. Rev. **D94** (2016) 072001, [arXiv:1608.01289](#).

- [14] BaBar collaboration, P. del Amo Sanchez *et al.*, *Dalitz-plot Analysis of $B^0 \rightarrow \bar{D}^0 \pi^+ \pi^-$* , [arXiv:1007.4464](#).
- [15] Belle collaboration, A. Kuzmin *et al.*, *Study of $\bar{B}^0 \rightarrow D^0 \pi^+ \pi^-$ decays*, Phys. Rev. **D76** (2007) 012006, [arXiv:hep-ex/0611054](#).
- [16] Belle collaboration, K. Abe *et al.*, *Study of $B^- \rightarrow D^{*0} \pi^-$ ($D^{*0} \rightarrow D^{(*)+} \pi^-$) decays*, Phys. Rev. **D69** (2004) 112002, [arXiv:hep-ex/0307021](#).
- [17] LHCb collaboration, R. Aaij *et al.*, *Study of D_J meson decays to $D^+ \pi^-$, $D^0 \pi^+$ and $D^{*+} \pi^-$ final states in pp collisions*, JHEP **09** (2013) 145, [arXiv:1307.4556](#).
- [18] S. Godfrey and N. Isgur, *Mesons in a relativized quark model with chromodynamics*, Phys. Rev. **D32** (1985) 189.
- [19] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.
- [20] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, [arXiv:1412.6352](#).
- [21] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, [arXiv:1405.7808](#).
- [22] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, [arXiv:1210.6861](#).
- [23] LHCb collaboration, R. Aaij *et al.*, *First observation of the decays $\bar{B}^0 \rightarrow D^+ K^- \pi^+ \pi^-$ and $B^- \rightarrow D^0 K^- \pi^+ \pi^-$* , Phys. Rev. Lett. **108** (2012) 161801, [arXiv:1201.4402](#).
- [24] LHCb collaboration, R. Aaij *et al.*, *Study of $B^- \rightarrow DK^- \pi^+ \pi^-$ and $B^- \rightarrow D \pi^- \pi^+ \pi^-$ decays and determination of the CKM angle γ* , Phys. Rev. **D92** (2015) 112005, [arXiv:1505.07044](#).
- [25] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79, 2003. [arXiv:physics/0308063](#).
- [26] 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.
- [27] LHCb collaboration, R. Aaij *et al.*, *Observation of the annihilation decay mode $B^0 \rightarrow K^+ K^-$* , Phys. Rev. Lett. **118** (2017) 081801, [arXiv:1610.08288](#).
- [28] ARGUS collaboration, H. Albrecht *et al.*, *Search for hadronic $b \rightarrow u$ decays*, Phys. Lett. **B241** (1990) 278.

- [29] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, [arXiv:physics/0402083](#).
- [30] LHCb collaboration, R. Aaij *et al.*, *Observation of $B^0 \rightarrow \overline{D}^0 K^+ K^-$ and evidence for $B_s^0 \rightarrow \overline{D}^0 K^+ K^-$* , Phys. Rev. Lett. **109** (2012) 131801, [arXiv:1207.5991](#).