Rare and Challenging Charm Decays at LHCb

Ed Greening
St. John's College

A thesis submitted to the University of Oxford for the degree of
Doctor of Philosophy in Particle Physics
July 2014

Abstract

This thesis presents two studies of charged charm meson decays using proton-proton collision data, corresponding to an integrated luminosity of 3 fb^{-1}, recorded by the LHCb experiment.

A search for $D^+_{(s)} \to \pi^+\mu^+\mu^-$ decays is performed. Rare decays of charged charm mesons have not previously been observed unlike their counterpart B meson and kaon decays. Such decays are $c \to u$ quark transitions and take place via loop diagrams within the Standard Model. Measurements of their branching fractions are an indirect test of New Physics, whose virtual contributions may enhance such quantities. No signals are observed and the 90% (95%) CL limits on the branching fractions, the most stringent to date, are found to be,

$$\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-) < 2.9\,(3.3) \times 10^{-8},$$
$$\mathcal{B}(D^+_s \to \pi^+\mu^+\mu^-) < 1.76\,(1.91) \times 10^{-7}.$$

The charge asymmetry measurement of $D^+ \to \pi^+\pi^0$ using $\pi^0 \to e^+e^-\gamma$ decays is also described. Charge asymmetry in the Standard Model is expected to be small for charm meson decays. Any measured deviation from zero would be evidence of New Physics. This decay has not previously been observed at a hadron collider and the measurement acts as a 'proof of principle' study allowing a future analysis to examine the decay mode. No evidence of CP violation is found.

Acknowledgements

I wish to thank the University of Oxford's Particle Physics Department - in particular the LHCb group - and St John's College for giving me the opportunity to perform this research. I would like to thank the Science and Technology Funding Council for giving particle physics students the opportunity to work at the site of their experiment. The time that I spent at CERN was both valuable and enjoyable. I also thank Liverpool University's Physics Department for giving undergraduate students the chance to do summer projects on LHCb, inspiring stuff.

I would like to thank Malcolm John for his excellent supervision and guidance. For allowing me to complete 180 hours of French classes, *je te remercie*, as Google Translate informs me they say. Benoit Viaud and Brian Meadows were similarly invaluable when it came to this research (and both wonderful people!). Help received from Andrea, Chris, Dan, Denis, Hamish, Nick, Mat and Phil was vital in helping me to navigate the tools of particle physics. Thank you all for your time.

Finally, to all of the brilliant friends that I have made in Oxford, in particular: Gaëlle, Madu, Richard, Rhorry, and the entire volleyball team, I'm truly sad to be leaving. I thank all of the friends in Geneva who kept me there for an extra year, especially: Andrew, Cecile, Espen, Jon and the members of our declining futsal team, Le Box, and our twice champion relay team, VELOcity. My family too, thanks for repeatedly laughing at my thesis title. And following your warning, Mother, you'll be glad to know that I successfully avoided burns from all of the corrosive chemicals and Bunsen burners used at CERN.

Declaration

All work presented in this thesis is my own, except where explicitly stated. The thesis has not been submitted for another qualification to this or any other university. Chapter 3 describes an analysis that updates an analysis that was published by the LHCb Collaboration in 2013. The published work used $1\,\text{fb}^{-1}$ of LHCb data and was prepared in close collaboration with Benoit Viaud and Olga Kochebina (LAL, Orsay). This thesis describes a new analysis of 3 fb^{-1} from 2011 and 2012 and was completed solely by the author for this document.

Contents

Glossary

BDT boosted decision tree. 58

CERN European Organisation for Nuclear Research. 14

CKM Cabibbo-Kobayashi-Maskawa. 6

CL confidence level. 4

CP charge-parity. 3

CPV CP violation. 7

DOCA distance of closest approach. 54

DT decision tree. 56

ECAL electromagnetic calorimeter. 34

EM electromagnetic. 16

FCNC flavour-changing neutral current. 4

FOI field of interest. 39

GIM Glashow-Iliopoulos-Maiani. 4

HCAL hadronic calorimeter. 34

HLT higher level trigger. 40

IP impact parameter. 29

IT Inner Tracker. 25

L0 Level-0. 33

LD long-distance. 46

LHC Large Hadron Collider. 14

Linac2 linear accelerator 2. 14

NP New Physics. 8

OT Outer Tracker. 25

PID particle identification. 18

PS pre-shower. 34

PS Proton Synchrotron. 15

PSB Proton Synchrotron Booster. 14

PV primary vertex. 21

RF radio frequency. 16

RICH Ring Imaging Cherenkov. 18

SD short-distance. 46

SM Standard Model of particle physics. 1

SPD scintillator pad detector. 34

SPS Super Proton Synchrotron. 15

SUSY supersymmetry. 12

SV secondary vertex. 21

TIS trigger independently of signal. 97

TOS trigger on signal. 97

TT Tracker Turicensis. 19

VELO Vertex Locator. 18

Chapter 1

Introduction

The mathematics underpinning the most basic physical processes are described. Elementary particles and the fundamental interactions are discussed. The phenomena and history of charm physics is also given to contextualise later parts of the thesis.

1.1 Building blocks of Nature

The mathematical framework which describes the fundamental interactions and particles is called the Standard Model of particle physics (SM). The model is built around twelve fermions and five gauge bosons. The dynamics of the fundamental particles are described by the SM Lagrangian (density),

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EW} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}, \tag{1.1}$$

which is invariant under local gauge transformations of the symmetry group,

$$SU(3)_C \otimes SU(2)_I \otimes U(1)_Y. \tag{1.2}$$

The symmetry is guaranteed by the introduction of gauge bosons (spin-1) corresponding to each symmetry group. For instance, $SU(3)_C$ introduces gluons and represents the symmetry group for the strong interaction, resulting in the conservation of C (colour charge). $SU(2)_I \otimes U(1)_Y$ introduces the $W^{\pm}$, the Z^0 and the

γ and is the symmetry group for the electroweak interaction, conserving I (weak isospin) and Y (weak hypercharge). Hence, electric charge, the sum of weak hypercharge and the third component of weak isospin, $Q = \frac{Y}{2} + I_z$, is also conserved. These are the three fundamental interactions of the SM and act on the fermions with varying strengths. The SM gauge bosons are detailed in Table 1.1. The final particle within the SM is the Higgs boson, H^0. This is a scalar (spin-0) particle resulting from electroweak symmetry breaking, in which the $W^{\pm}$ and the Z^0 gain mass due to the nonzero vacuum expectation value of the Higgs field. Couplings of the top quark, the bottom quarks and the tau to the Higgs boson suggest that, as predicted, fermions (with the exception of neutrinos, whose mass generating mechanism is unconfirmed) also gain their masses via their interactions with the Higgs field. The properties of the Higgs boson are also shown in Table 1.1.

Interaction	**Boson**	**Charge**	**Mass**
Strong	Gluon (g)	0	0
EM	Photon (γ)	0	0
Weak	$W^{\pm}$	± 1	80.4 GeV/c^2
	Z^0	0	91.2 GeV/c^2
Higgs	H^0	0	125.7 ± 0.4 GeV/c^2

Table 1.1: Properties of the SM bosons. Mass values taken from [1].

Fermions (spin-$\frac{1}{2}$) are sub-divided into quarks and leptons, depending upon how they interact. Quarks contain all three of the quantum numbers conserved within the SM Lagrangian, so interact via all three fundamental interactions. Quarks have one of six possible flavour quantum numbers. The six flavours of quark each couple with a different strength to the Higgs field, hence the quarks have different masses. Leptons only have weak isospin and weak hypercharge conserved quantum numbers, so do not interact with gluons. Leptons also come in six distinct flavours. Each fermion has an associated antiparticle with the same mass but opposite lepton/baryon number. If the fermion is electromagnetically charged, its antiparticle has opposite electromagnetic charge. The SM fermions

are detailed in Table 1.2.

Generation	**Quarks**		**Leptons**	
1	Up (u)	Down (d)	Electron (e)	Electron neutrino (ν_e)
2	Charm (c)	Strange (s)	Muon (μ)	Muon neutrino (ν_μ)
3	Top (t)	Bottom (b)	Tau (τ)	Tau neutrino (ν_τ)
Charge	$+\frac{2}{3}$	$-\frac{1}{3}$	-1	0

Table 1.2: Properties of the SM fermions.

Although the SM provides a robust framework for studying particle physics, it has a number of deficiencies. Some of its biggest problems are:

- **Gravity**, the fourth fundamental interaction is not incorporated into the SM as it cannot be described at the quantum level.
- **Neutrino masses**, lack of observation of right-handed neutrinos means that neutrinos may not gain mass in the same way as the other fermions.
- **Baryogenesis**, at the level observed, the abundance of matter over anti-matter cannot be explained by the charge-parity (CP) violating effects in the quark sector of the SM.
- **Hierarchy**, unknown nature of the mechanism which cancels quantum corrections to the observed Higgs boson mass.
- **Dark matter and dark energy**, which make up the majority (24% and 71%, respectively) of mass within the Universe. There are no candidate quanta for either in the SM.

These puzzles are being investigated by different experiments. Quantum gravitational effects are extremely weak, so are difficult to test. Collaborations such as BICEP2 and Planck probe the polarisation of primordial photons caused by gravitational waves. These observations may provide insight into inflationary models

of the Universe. The Majorana nature of neutrinos [2] is easily tested by observation of neutrinoless double-beta decay. In this process, Majorana neutrinos, being their own antiparticles, annihilate. The EXO-200 collaboration set a lower limit on the half-life of neutrinoless double-beta decay of ^{136}Xe nuclei, $t_{1/2} > 1.6 \times 10^{25}$ years at 90% confidence level (CL). Lepton number violating decays are also used to impose limits upon the branching fractions of decays mediated by Majorana neutrinos. LHCb has set the world's most stringent limits on two of these decays, with $\mathcal{B}(B^+ \to \pi^-\mu^+\mu^+) < 4.0\times10^{-9}$ [3] and $\mathcal{B}(D^+ \to \pi^-\mu^+\mu^+) < 2.5\times10^{-8}$ [4] at 95% CL. The amount of CP violation within the SM is measured by experiments which explore quark mixing. These searches are discussed in more detail later in this chapter. Supersymmetric models offer a solution to the hierarchy problem and offer a natural dark matter candidate. Observations of supersymmetric particles are needed to confirm these models. As of yet, the general purpose detectors at the LHC have only been able to set mass limits upon potential supersymmetric particles [5, 6].

1.2 Historical development of charm physics

In 1963, when only the three lightest quarks were known, Cabibbo [7] suggested a mechanism to describe the observed rates of semi-leptonic decays. Factors of $\cos\theta_C$ and $\sin\theta_C$ were introduced to the $d \to u + W^-$ and $s \to u + W^-$ vertices, respectively. By comparing decay rates with that of $\mu^- \to \nu_\mu + W^-$ decays the value of the Cabibbo angle was determined, $\theta_C \simeq 13^\circ$.

1.2.1 GIM mechanism

The Glashow-Iliopoulos-Maiani (GIM) mechanism was first predicted in 1970 [8] to explain the non-existence of the flavour-changing neutral current (FCNC) [9,

10], whilst also accounting for observations of neutral kaon mixing [11, 12]. The GIM mechanism added an additional fourth generation quark to provide FCNC suppression at the one-loop level. The new quark was required to have an electric charge of +2/3 and to compete with the up quark in FCNC decays. Vertex couplings of $-\sin\theta_C$ and $\cos\theta_C$ in $d \to c + W^-$ and $s \to c + W^-$ transitions, respectively, lead to an almost exact cancellation in the amplitude (a small rate remains due to the mass difference of the quarks).

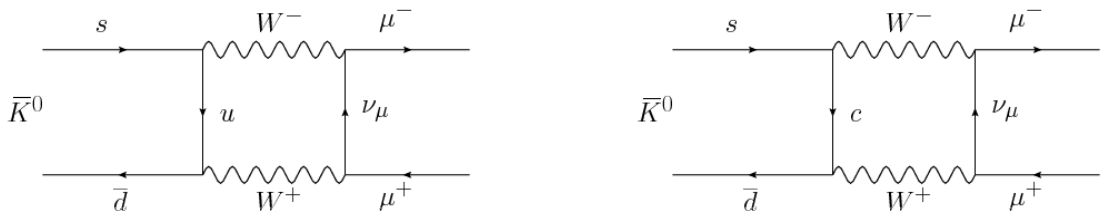

Figure 1.1: Feynman diagrams contributing to the FCNC decay $\bar{K}^0 \to \mu^+\mu^-$. The amplitude of each diagram is proportional to the product of the coupling at the vertices, (left) $\cos\theta_C \cdot \sin\theta_C$ and (right) $\cos\theta_C \cdot -\sin\theta_C$. The total amplitude is found by summing these contributions.

1.2.2 Discovery of charmonium

The discovery of a charmonium $c\bar{c}$ state subsequently lead to a period of rapid advancement in particle physics, known as the November Revolution. Two papers [13, 14] published in 1974 on consecutive pages of the same journal both claimed observation of a heavy narrow resonance with a invariant mass of 3.1 GeV/c^2, see Fig. 1.2. The groups responsible for the discovery named the particle J and ψ, respectively. An amalgamation of these was taken as the particle's name.

This lead to further searches for other charm hadrons. The first evidence of charm baryons ($\Sigma_c^{++}(2455)$ and Λ_c^+) appeared in 1975 [15], followed a year later by observations of charm mesons (D^0 and D^+) [16, 17]. These discoveries confirmed the interpretation of J/ψ as $c\bar{c}$ and, more importantly, validated the quark model.

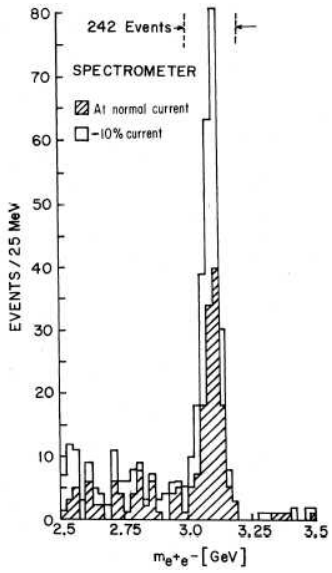

Figure 1.2: Invariant mass spectrum of e^+e^- events showing the discovery of a heavy resonance in $p^+ + Be \rightarrow (J/\psi \rightarrow e^+e^-)X$ decays. Data are plotted for two different spectrometer currents showing that the peak is independent of this. Taken from [13].

1.3 Quarks in the Standard Model

1.3.1 Yukawa couplings

Quarks, like all fermions within the SM, gain their observed masses from interactions with the Higgs field. The strength of the coupling between the Higgs field and a fermion is proportional to the observed mass of the fermion, and is an example of a Yukawa coupling. Coupling strengths for each fermion are not predicted by the SM. Instead they are calculated from the observed masses of the fermions.

1.3.2 The CKM matrix

The Cabibbo-Kobayashi-Maskawa (CKM) matrix [18] describes the relationship between the strong and weak quark eigenstates. As it is not perfectly diagonal, it allows mixing between three quark generations. It is an extension of the GIM

mechanism for an additional generation of quarks. The extension increases the number of free parameters from one (θ_C) to four: three mixing angles and one complex phase. The introduction of a complex phase into the quark mixing matrix accounts for CP violation (CPV). Complex phases (there is also one in the neutrino mixing matrix) are the only source of CPV in the SM. The most general form of the matrix is,

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.3}$$

Where d', s' and b' are the weak eigenstates, and d, s and b are the mass eigenstates. Each matrix element, V_{ij}, represents the coupling strength of the transition from flavour i to j. Weak universality requires that,

$$\sum_k |V_{ik}|^2 = \sum_k |V_{kj}|^2 = 1. \tag{1.4}$$

This stems from the fact that all quarks couple with the same strength to the gauge bosons of the electroweak interaction. Note that it is only due to convention that the down-type quarks are rotated, rather than the up-type quarks.

1.3.3 CKM parameterisation

One of the many CKM parameterisations is the Wolfenstein parameterisation [19],

$$V_{CKM} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda[1 + \frac{1}{2}A^2\lambda^4(2\rho - 1) + iA^2\lambda^4\eta] & 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}(4A^2 + 1)\lambda^4 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2[1 + \frac{1}{2}\lambda^2(2\rho - 1) + i\lambda^2\eta] & 1 - \frac{1}{2}A^2\lambda^4 \end{pmatrix} \tag{1.5}$$

where $\lambda = sin\theta_C \simeq 0.23$. Notice how the magnitudes of the diagonal terms are close to unity and those of the off-diagonal terms are small. CPV is represented by complex elements. The least suppressed, $\mathcal{O}(\lambda^3)$, complex terms appear in matrix

elements V_{ub} and V_{td}. These matrix elements are accessible in B meson decays. Kaon decays can also access V_{td}, where the large mass of the top quark helps to weaken the suppression. CPV is weaker in charm meson decays as they can only access complex terms which are suppressed by at least $\mathcal{O}(\lambda^5)$. For example, $c \to b \to u$ and $c \to d \to u$ transitions access complex elements V_{ub} and V_{cd}, respectively.

1.4 Some charm meson phenomenology

Off-diagonal terms in the CKM matrix are responsible for a number of interesting charm meson phenomena. These can be studied to allow the magnitude of the elements of the CKM matrix to be calculated. Exploration of the SM and the hunt for New Physics (NP) is conducted via a range of analysis techniques, probing interesting phenomena. Three principle effects are described here, two of which are exploited in the analyses presented in this thesis.

1.4.1 Mixing

Neutral mesons, such as the D^0, can oscillate to their antiparticles via a second-order weak process. As the mass eigenstates, D_1 and D_2, are not the flavour eigenstates, D^0 and $\bar{D}^0$, they can be written in terms of one another,

$$|D_1\rangle = p\left|D^0\right\rangle + q\left|\bar{D}^0\right\rangle, \tag{1.6}$$

$$|D_2\rangle = p\left|D^0\right\rangle - q\left|\bar{D}^0\right\rangle. \tag{1.7}$$

Where p and q can be complex and satisfy $|p|^2 + |q|^2 = 1$. Oscillations have been observed in neutral strange, bottom and charm mesons. The rate of mixing is smallest in the charm system due to the effective GIM cancellation of contributions from internal down and strange quarks. The relatively small size of V_{cb} in the

CKM matrix also limits the contribution from internal bottom quarks. This level of suppression is not present for neutral kaons and bottom mesons due to the high top quark mass. In fact, $B_s^0 - \bar{B}_s^0$ oscillations occur with a high frequency of about 3×10^{12} Hz, or about 9 times during their 1.5 ps lifetime [20], see Fig. 1.3. This is a result of the magnitude of V_{cs} and V_{tb} being near-unity. $D^0 - \bar{D}^0$ oscillations are therefore very slow, with a period one hundred times the average lifetime of the D^0. Nevertheless, with the large samples collected by LHCb the beginning of the oscillation has been observed, as shown in Fig. 1.4.

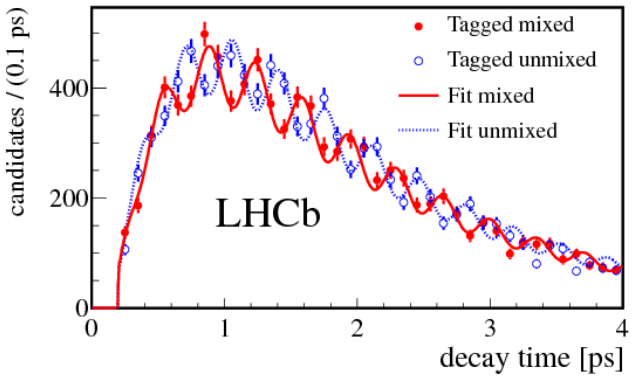

Figure 1.3: Oscillations of B_s^0 measured by LHCb. Shown are mixed (different flavour at decay and production) and unmixed (same flavour at decay and production) samples. Taken from [20].

1.4.2 Charge-parity violation

The complex phase present in the CKM matrix allows decays mediated by the weak interaction to violate CP symmetry, in contrast to strong and electromagnetic decays. There are three types of CPV:

1. **Direct CPV**, occurs when the amplitude of a decay and its CP conjugate have different magnitudes.

2. **Indirect CPV**, occurs when $p \neq q$ (Eqs. 1.6 and 1.7), meaning that the probability of $D^0 \to \bar{D}^0$ is different from $\bar{D}^0 \to D^0$.

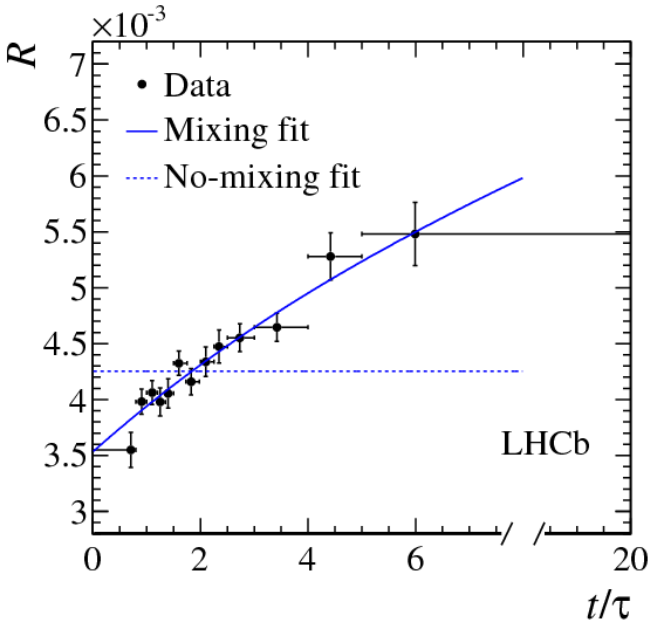

Figure 1.4: Ratio, R, of reconstructed $D^0 \to K^+\pi^-$ decays from D^0 and $\bar{D}^0$ mesons as a function of their lifetime where the flavour of the D meson is tagged upon its production. If oscillation did not occur, the ratio would be constant. Taken from [21].

3. **Interference**, between a direct decay, $D^0 \to f$, and a decay involving mixing, $D^0 \to \bar{D}^0 \to f$.

CPV is an interference effect and as such, two competing amplitudes are needed for it to occur. In the SM, CP violating CKM elements do not occur at tree level, but may appear in loop diagrams. For example, loop diagrams featuring $c \to b \to u$ transitions involve V_{ub}, which has a complex component. This transition is suppressed by the CKM matrix element's magnitude and the relatively low mass of the bottom quark. Therefore CP violating effects in charm are expected to be small, with SM predictions of the order of 0.1% [22]. This should be contrasted with B meson decays which exhibit effects of the order of 10% [23]. Recent improvements in the sensitivity of charm decay measurements have disfavoured CP asymmetry (A_{CP}) down to 1%, see Table 1.3. With uncertainties approaching 0.1%, CPV searches start to probe SM expectations. An A_{CP} measurement is

presented in Chapter 4.

Decay	**Data** (fb^{-1})	A_{CP} (%)	**Ref.**
$D^{\pm} \to \phi\pi^{\pm}$	1.0	$-0.04 \pm 0.14 \pm 0.14$	[24]
$D^{\pm}_{(s)} \to K^0_S\pi^{\pm}$	3.0	$+0.38 \pm 0.46 \pm 0.17$	[25]
$D^{\pm}_{(s)} \to K^0_S K^{\pm}$	3.0	$+0.03 \pm 0.17 \pm 0.14$	[25]
ΔA_{CP}	1.0	$-0.34 \pm 0.15 \pm 0.10$	[26]
ΔA_{CP}	3.0	$+0.14 \pm 0.16 \pm 0.08$	[27]
A_D	3.0	-0.7 ± 1.9	[28]

Table 1.3: Values of A_{CP} observed in charm decays at LHCb, where the first uncertainty is statistical and the second systematic. The size of the data sample used is also shown. ΔA_{CP} corresponds to the difference between A_{CP} ($D^0 \to K^+K^-$) and A_{CP} ($D^0 \to \pi^+\pi^-$). A_D is a measure of the $D^0 - \bar{D}^0$ mixing asymmetry and is found by comparing the ratio of $D^0 \to K^+\pi^-$ to $D^0 \to K^-\pi^+$ rates with the corresponding ratio for the charge-conjugate processes.

1.4.3 Flavour changing neutral current decays

FCNC processes are rare within the SM as they cannot occur at tree level. At loop level, as with mixing in the charm system, loop contributions from down and strange quarks, which have amplitudes of $\mathcal{O}(\lambda)$ and $\mathcal{O}(-\lambda)$, respectively, all but cancel. There is a only a very small loop contribution from bottom quarks because $V_{cb}V_{ub}$ is $\mathcal{O}(\lambda^5)$. Recent LHCb searches for the rare charm decays $D^0 \to \mu^+\mu^-$ [29] and $D^0 \to \pi^+\pi^-\mu^+\mu^-$ [30] did not observe any signal, see Fig. 1.5. Limits on the branching fractions of 6.2×10^{-9} and 5.5×10^{-7} at 90% CL were set, respectively. The search for $D^+_{(s)} \to \pi^+\mu^+\mu^-$ is presented in Chapter 3.

FCNC decays are nevertheless well established in processes that involve K and B mesons. For example, the decay $K^+ \to \pi^+\mu^+\mu^-$ was observed by the NA48/2 collaboration (amongst others) with $\mathcal{B} = (9.62 \pm 0.25) \times 10^{-8}$ [31]. Here, although the amplitudes from up and charm loops also have equal and opposing amplitudes, the mass difference between the quarks eases the suppression, $(m_c - m_u) \sim 1\,\text{GeV}/c^2$, *c.f.* $(m_s - m_d) \sim 0.1\,\text{GeV}/c^2$ for D system. The decays $B^+ \to \pi^+\mu^+\mu^-$ and $B^+ \to K^+\mu^+\mu^-$ has also been observed by LHCb, $\mathcal{B} = (2.4 \pm 0.6) \times$

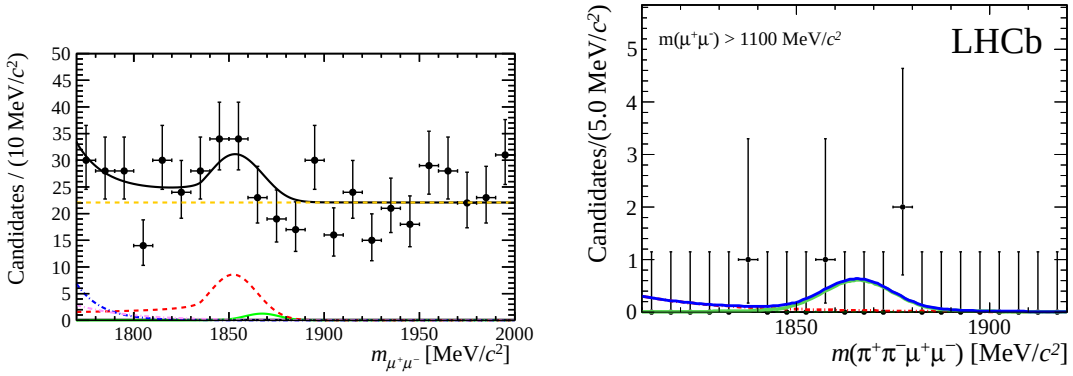

Figure 1.5: Invariant mass distributions of (left) $m(\mu^+\mu^-)$ and (right) $m(\pi^+\pi^-\mu^+\mu^-)$ (in the high-$m(\mu^+\mu^-)$ region, away from any resonant structures). Neither plot shows significant signals, described by the solid green lines. Taken from [29] and [30], respectively.

10^{-8} [32] and $\mathcal{B} = (4.36 \pm 0.23) \times 10^{-7}$ [33], respectively. In the B meson system, the near-unity value of $|V_{tb}|$ and very high mass of the top quark in the loop weaken the suppression, resulting in amplitudes of $\mathcal{O}(\lambda^3)$. Rare decays with final state muons are often targeted as muons are more easily differentiated from large pionic backgrounds produced by hadron colliders.

1.5 Effect of New Physics

The incompleteness of the SM motivates searches for NP. One of the most attractive NP scenarios is provided by supersymmetry (SUSY) which, amongst several theoretical benefits, has the potential to solve the hierarchy problem and simultaneously provide a natural dark matter candidate. SUSY introduces supersymmetric partners for each SM particle. For every fermion in the SM there is a SUSY boson, called a "sfermion", and for every boson there is fermion whose name is made by adding the suffix "-ino". The couplings between these particles and the Higgs boson almost exactly cancel those between the SM particles and the Higgs boson and hence avoid ultraviolet divergences in calculating the Higgs boson's mass. Without this SUSY cancellation, the expected mass of the Higgs boson

naturally rises to the Grand Unified Theory scale, which is the energy at which the three fundamental forces become one.

There are many different SUSY models, some of which conserve a property known as R-parity. R-parity is a binary symmetry that distinguishes SUSY particles from non-SUSY particles. It is required for conservation of baryon and lepton numbers as it stops the lightest SUSY particle decaying into (lighter) SM particles. With any further decay forbidden, the lightest SUSY particle is a dark matter candidate.

Given that there have been no experimental observations of SUSY particles, it suggests that their masses are much greater than their SM particle equivalents. Even if these masses are unobtainable at LHC collisions, SUSY particles can appear as virtual particles in loop diagrams. Physical quantitites, such as branching fractions and angular asymmetries (forward-backward), could receive an enhancement as a result of their presence.

Measured properties of FCNC decays of B and K mesons, which proceed via virtual up-type quarks, are consistent with SM predictions. There is no reason, *a priori*, that requires NP to affect up- and down-type quarks in the same way, however. Given that FCNC decays of charm mesons, which proceed via virtual down-type quarks, have not been observed, the possibility that they are enhanced by NP is not ruled out. As discussed in Sec. 1.4.3, previous searches for charm meson FCNC decays have not shown any significant enhancement to expected SM branching fractions.

Chapter 2

The LHCb experiment

LHCb is a dedicated forward-arm spectrometer located at the Large Hadron Collider (LHC). The detector is designed to study heavy flavour physics by making CP measurements and searching for rare decays of bottom and charm hadrons. The accelerator complex, the beam conditions and the LHCb experiment are described. The data-taking process is also outlined.

2.1 The LHC complex

The LHC is a particle collider located at the European Organisation for Nuclear Research (CERN) [34]. The collider accelerates two stored beams of hadrons, either protons or lead ions, in opposite directions around 27 km of tunnel. Designed to collide protons at a centre-of-mass energy of up to $\sqrt{s} = 14$ TeV, the collider is yet to reach this limit having run at lower energies until Long Shutdown 1 (LS1), $2010 - 2013$. The analyses detailed in this thesis use all data from this period.

The LHC accelerates protons provided by CERN's accelerator complex, as illustrated in Fig. 2.1. The proton beam originates from a hydrogen gas source. The gas is placed in a metal cylinder before it is ionised by applying an electric field. The protons are first accelerated to an energy of 50 MeV by linear accelerator 2 (Linac2). Pulses of protons from Linac2 are injected into the Proton

Synchrotron Booster (PSB). The PSB is composed of four parallel accelerator rings that increase the energy of the protons to 1.4 GeV. Upon injection from the PSB into the Proton Synchrotron (PS) the protons are divided into bunches and accelerated to 26 GeV. Once they reach this energy the 6.5 km-circumference Super Proton Synchrotron (SPS) accelerates them to 450 GeV before they are injected into the LHC. At injection into the LHC the proton bunches are divided into beams travelling in clockwise and anti-clockwise directions.

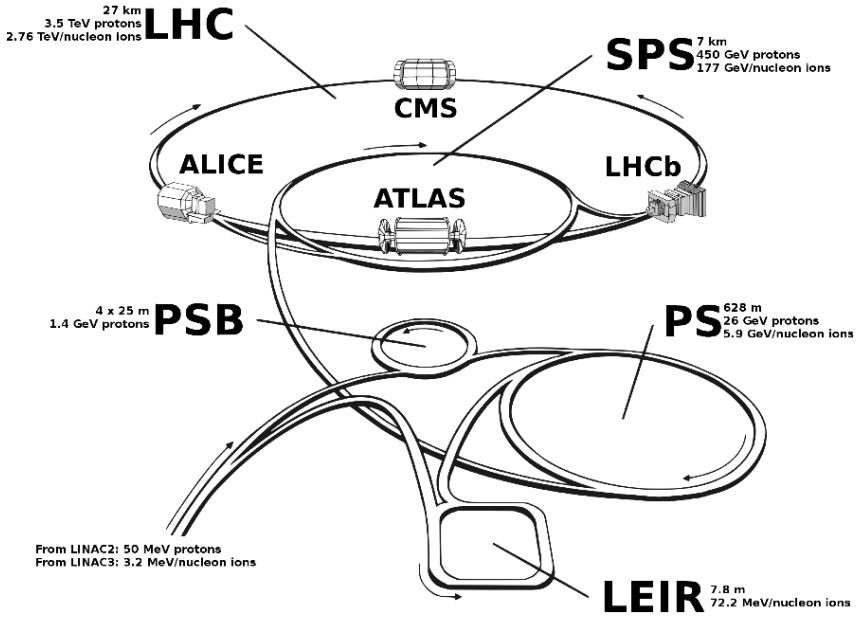

Figure 2.1: Schematic of the CERN accelerator complex. The four experiments on the LHC are also shown. Taken from [35].

The LHC uses 1,232 dipole magnets to keep the protons on their approximately circular path. The strength of the magnetic field provided by the dipoles is increased to up to 8 T as the energy of the protons in each beam increases to 7 TeV. 392 quadrupole magnets focus the bunches of protons. The bunches are ellipsoidal in shape ($\sim 16\,\mu\text{m} \times 16\,\mu\text{m} \times 7\,\text{cm}$) and separated by approximately 7.5 m. Superfluid ^{4}He is used to cool the copper-clad niobium-titanium magnets to a super-conducting temperature of 1.9 K.

Sixteen radio frequency (RF) cavities are used to accelerate the protons. An RF cavity is a metallic chamber containing an electromagnetic (EM) field. The shape of the chamber is designed so that EM waves resonate, with a frequency of 400 MHz, inside the cavity. The protons feel a force as a result of the EM field which pushes them in the direction of acceleration. Protons arriving too early (late) are given a smaller (larger) energy boost by the EM field. The RF cavities are housed in cryomodules which keep them at 4.5 K, to avoid losing energy to electrical resistance. High-power klystrons (tubes containing electron beams which amplify high-frequency signals) drive each RF cavity using a waveguide to couple energy into the cavity.

2.2 Beam conditions

LHC data taking can be divided into three eras based on centre-of-mass energy and instantaneous luminosity. The 2010 run recorded $\mathcal{L} = 0.04$ fb^{-1} of data at a centre-of-mass energy of $\sqrt{s} = 7$ TeV. In 2011, the instantaneous luminosity was doubled and $\mathcal{L} = 1.10$ fb^{-1} of data was recorded. Finally, in 2012, the LHC collided protons at a centre-of-mass energy of $\sqrt{s} = 8$ TeV, recording $\mathcal{L} = 2.08$ fb^{-1} of data. The absolute luminosity recorded in each year of running is shown in Fig 2.2. The nominal and actual running conditions in each year are shown in Table 2.1.

Condition	**2010**	**2011**	**2012**	**Nominal**
$\sqrt{s}$ [TeV]	7	7	8	14
Instantaneous $\mathcal{L}$ [cm^{-2}s^{-1}]	1×10^{32}	4×10^{32}	4×10^{32}	2×10^{32}
Bunches	344	1380	1380	2808
Intensity per bunch [10^{11}]	1.15	1.70	1.70	0.65
μ	$0.5 - 2.5$	1.5	1.6	0.4

Table 2.1: Running conditions for the LHC in 2010, 2011, 2012 and the nominal values. The intensity per bunch refers to the number of protons in each bunch. The pileup, μ, is the average number of proton-proton collisions per visible bunch crossing.

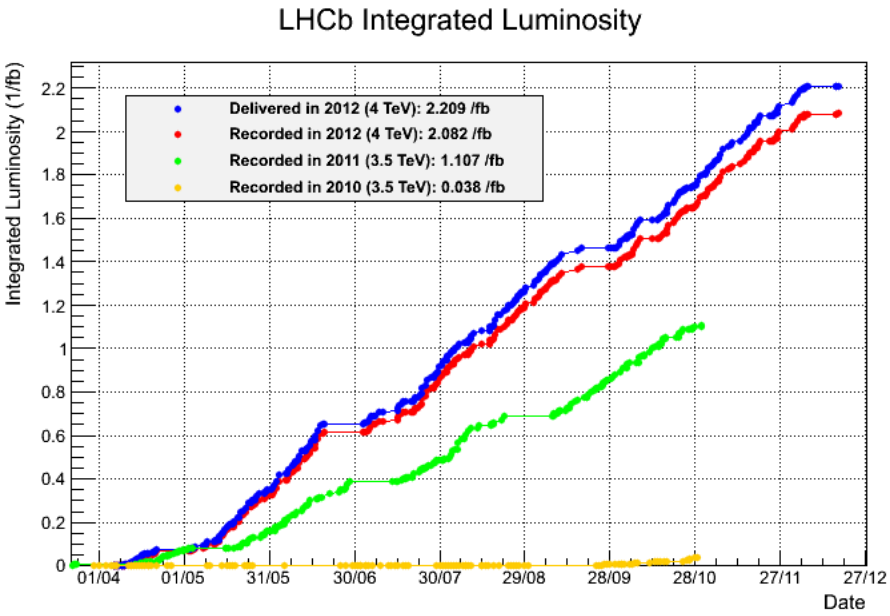

Figure 2.2: Recorded luminosity at LHCb for 2010, 2011 and 2012. The amount of luminosity delivered in 2012 is also shown. Taken from [36].

In the design of LHCb, the average number of visible interactions per bunch crossing, μ, was chosen to be low, 0.4, to allow a clean environment to simplify event reconstruction. However, μ was significantly higher during running as it was found that the trigger and reconstruction software could cope with such an environment because the bunches were spaced 50 ns apart rather than the planned 25 ns. The data taking efficiency, i.e. the ratio of recorded to delivered luminosity, at LHCb has been at approximately 90% during all running periods.

2.2.1 Charm production

The $c\bar{c}$ cross section at the LHC has been measured to be $\sigma(c\bar{c}) = 1419 \pm 134\,\mu$b [37] within LHCb's acceptance. This is approximately 20 times that of $b\bar{b}$ [38]. This amounts to 6×10^{11} $c\bar{c}$ events produced within LHCb's acceptance for $2010-2012$ running. The number of charm and beauty events on disk is comparable however, due to the reduced trigger efficiencies and bandwidth available to charm events.

The trigger efficiency is lower for charm due to, on average, lower momentum tracks and a smaller displacement between the charm meson's production and decay vertex.

2.3 Overview of the detector design

Due to their relatively low mass compared to the parton collision, beauty and charm hadrons created at the LHC are produced highly boosted and at angles close to the beam axis, as shown in Fig. 2.3. The LHCb detector is built covering a forward angle from 10 mrad to 250 (300) mrad in the vertical (horizontal) plane to exploit this fact. This covers a range of pseudorapidity, $1.8 < \eta < 4.9$, where η is defined as,

$$\eta = -\log\left[\tan\frac{\theta}{2}\right], \tag{2.1}$$

and θ is the angle between a particle's momentum and the beam axis. The acceptance is slightly larger in the horizontal plane to allow for the bending of charged particle trajectories due to the dipole magnet. By comparison, the tracking systems within ATLAS and CMS cover $-2.5 < \eta < 2.5$.

An illustration of the LHCb detector is shown in Fig 2.4. The detector consists of tracking stations, a dipole magnet, calorimeters and systems for particle identification (PID). In LHCb's coordinate system, the z-axis is defined by the ideal orbit of a proton inside the LHC beam pipe. The main component of the dipole's magnetic field is in the y-axis (vertical). The resultant Lorentz force felt by charged particles bends their trajectories in the x-axis (horizontal), allowing their momentum to be accurately measured by finely-spaced vertical tracking elements. LHC collisions occur inside the Vertex Locator (VELO). This is the first of the tracking sub-detectors comprising the tracking system, which is spread over 10 m. Two Ring Imaging Cherenkov (RICH) systems provide PID information

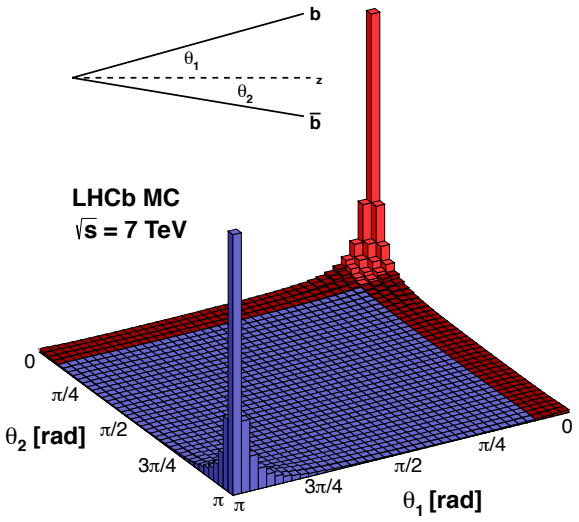

Figure 2.3: Angular distribution of $b\bar{b}$ production at 7 TeV LHC collisions. The red area shows LHCb's acceptance. Taken from [39].

critical for separating pion, kaon and proton candidates. Two types of calorimeters measure the energy deposited by electrons, photons and hadrons. Furthest from the collision point are the muon stations, used for identifying and measuring the energy of muons.

2.4 The particle tracking system

The tracking system is made of the VELO, the dipole magnet and four tracking stations; Tracker Turicensis (TT) and T1 to T3. The VELO and TT are located upstream of the magnet and composed entirely out of silicon strips. T1 − T3 are downstream of the magnet and composed of a mixture of silicon and straw tubes.

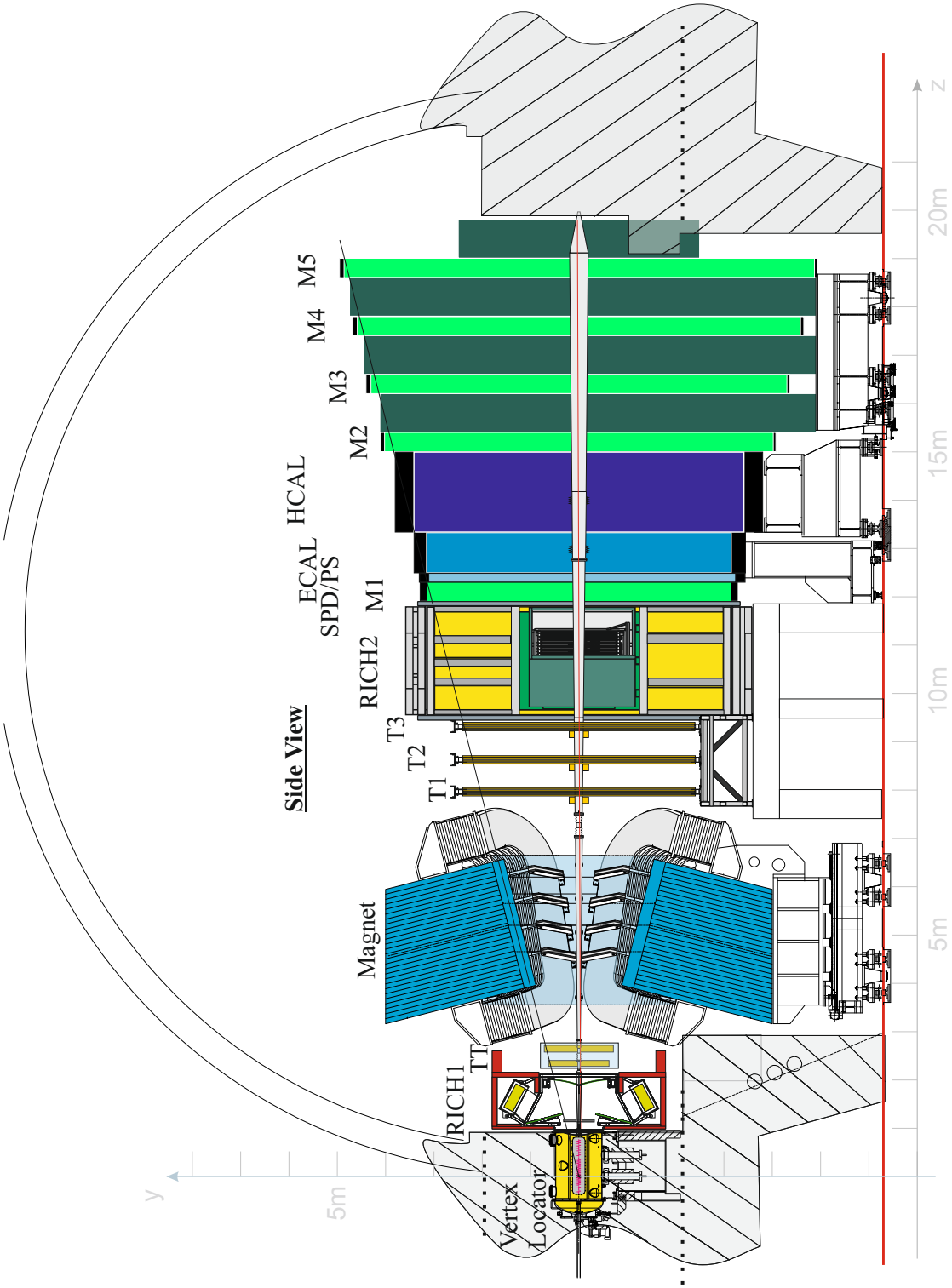

Figure 2.4: Schematic side view of the LHCb detector. Taken from [40].

2.4.1 Vertex locator

The VELO is a silicon micro-strip detector deployed close to the interaction region to provide precise pattern recognition of the collision. Hit coordinates identify the displacement of secondary vertex (SV) from primary vertex (PV) that are characteristic of charm and beauty hadron decays. This information is used in filtering algorithms and is a powerful discriminator for separating charm and beauty decays. A further function of the VELO is to determine the luminosity delivered by the LHC during van der Meer scans.

The VELO was created as two almost identical detector halves. Each half contains 21 modules with an R- and a ϕ- sensor facing parallel to the beam-axis. During data-taking, the sensors are just 8.2 mm from the beam. They are mounted on 'Roman Pots' and are retractable, by 29 mm, to avoid damage to the silicon from the beam during beam injection. The positioning of the modules in the x- and y-directions is adjustable to an accuracy of 10 μm. Real-time tracking algorithms provide images of the beams and help determine the motion steps and final detector position. The two configurations are shown in Fig. 2.5. There is an overlap between the two halves of the detector to ensure complete coverage at high pseudorapidity. One half of the detector is displaced by 150 mm in the z-direction to allow this. This is a possible source of systematic asymmetry in the detector. To correct for this and other sources of detector geometry systematics, the polarity of the dipole magnet is switched regularly throughout data-taking.

Both sensor types have 2048 strips within a semi-circular shape of diameter 84 mm and thickness 300 μm. Strips on the R-sensor are orientated concentrically from the centre with gradually increasing inter-strip (strip pitch) spacing. Strips on the ϕ-sensor are orientated approximately radially and divided into two regions. Those in the inner region are shorter, and the two regions are skewed with

respect to one another. This angular-offset aids pattern recognition by removing ambiguity. The strip pitch on these sensors varies from $40 - 100$ μm. The layout of both types of sensor are shown in Fig. 2.6.

The best hit position resolution of the VELO is 4 μm. This is achievable for tracks which cross more of the sensor ($7 - 11^{\circ}$ projected angle) at the point of the thinest silicon strips (40 μm), see Fig. 2.7. This allows for a lifetime resolution of 50 fs [41], an order of magnitude smaller than the lifetimes of beauty and charm mesons.

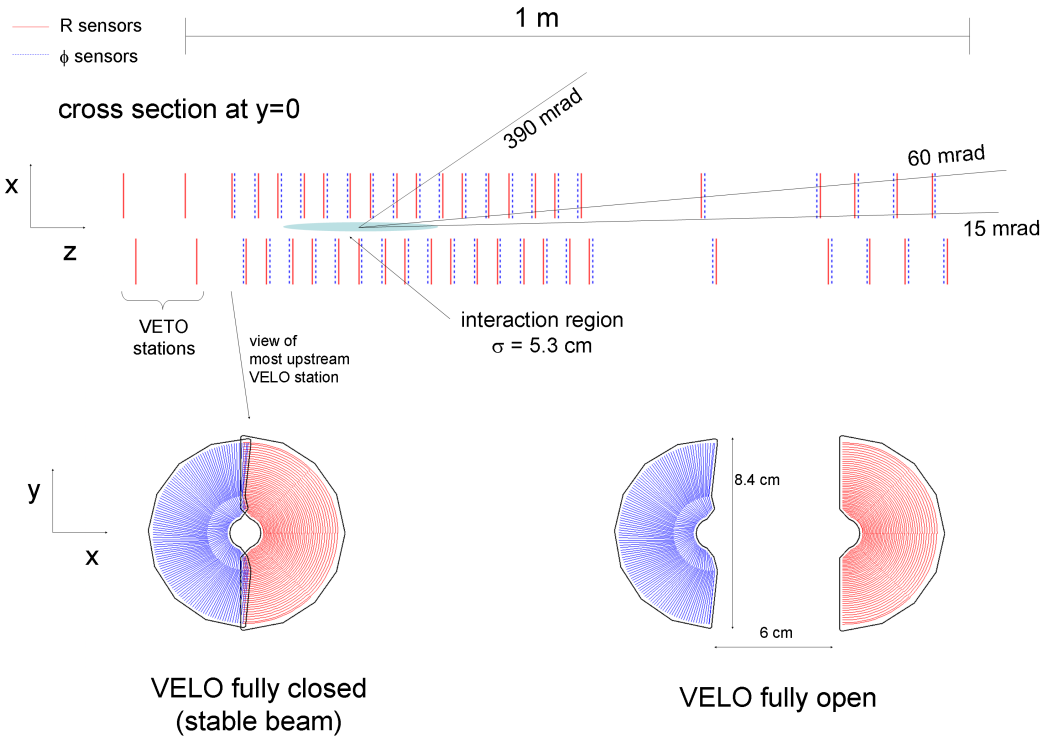

Figure 2.5: Schematic showing the cross section in the (x,z) plane of the VELO sensors at y=0. Both configurations of pairs of sensors from opposite halves of the detector are also shown. Taken from [40].

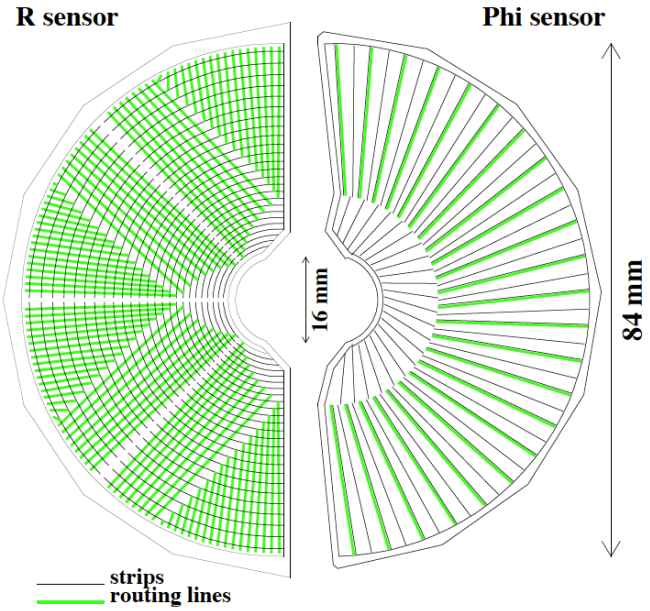

Figure 2.6: Layout of the strips and the routing lines in (left) R- and (right) ϕ-sensors. Taken from [42].

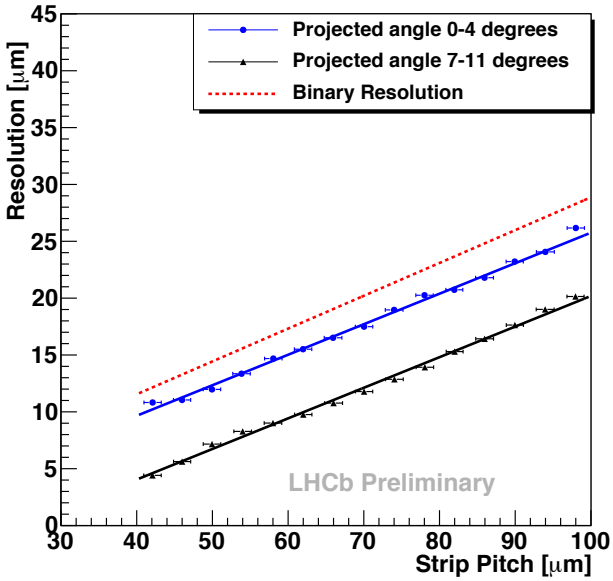

Figure 2.7: Hit position resolution versus strip pitch for 2 bins of the projected angles, evaluated for a R sensor using long tracks for 2011 data. The red line shows the binary resolution (resolution for a digital detector behaviour), the VELO performs considerably better due to the analogue pulse height readout. The best resolution is 4 μm for 40 μm strip pitch, this is the best vertex detector resolution at the LHC. Taken from [43].

2.4.2 Tracker turicensis

The TT, like the VELO, is a silicon micro-strip detector. It is sandwiched in between RICH1 and the dipole magnet, close enough to the magnet to be within the magnet's fringe field. This allows the VELO−TT combination to make momentum measurements with an accuracy of approximately 20%. Low momentum tracks which are bent out of the acceptance of the other tracking stations by the magnet have their momentum measured exclusively by the TT and can be used in a limited number of analyses. Information from the TT for all other tracks is used as part of the full tracking system momentum measurement.

The TT is formed of four layers in a $xuvx$ arrangement, as illustrated in Fig. 2.8. The u and v layers are rotated by $\pm 5°$ with respect to the x layers to obtain a stereo view of the particle trajectory. Layers vx are displaced from xu by 270 mm along the beam-axis.

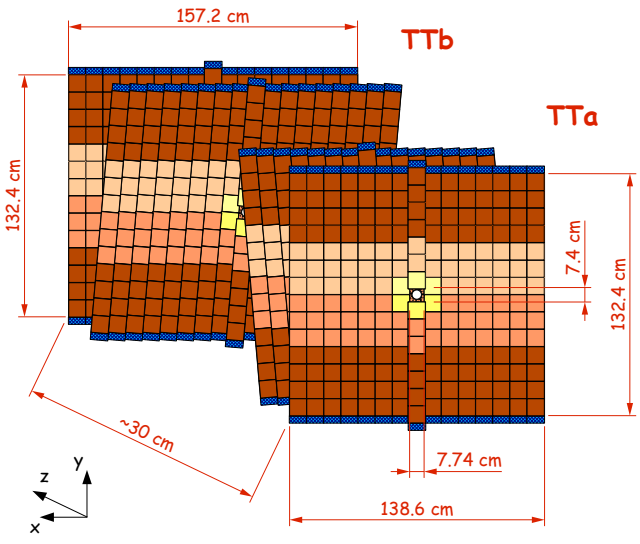

Figure 2.8: Arrangement of the silicon layers in the TT. TTa is the pairing of xu and TTb is of vx. The dimensions of the layers as well as the rotation of the u and v layers in the beam-axis is shown. Taken from [44].

2.4.3 Tracking stations

The tracking stations, T1-T3 can be classed into two distinct parts: the Inner Tracker (IT) and the Outer Tracker (OT). The IT is similar in material make up and design to the TT in that it is a silicon micro-strip detector composed of four layers in the same *xuvx* configuration. It is composed of four detector boxes arranged into a cross shape which surround the beam pipe, as illustrated in Fig. 2.9. Silicon technology is expensive but offers a fine spatial resolution. Hence the IT is in the centre of each tracking stations, where the occupancy is highest.

The dimensions of the IT are chosen to keep the occupancy of the OT at a manageable rate (<10%). The OT covers the much larger remaining area of the tracking stations. It is a drift-time detector based on straw-tubes. The inner surface of each tube acts as a cathode (negatively charged) while a wire in the centre acts as an anode (positively charged). Tubes are filled with a gas mixture that is ionised when traversed by a charged particle. A potential difference across the tubes initiates the liberated electrons to cause an avalanche effect as they move towards the wire, resulting in an electrical current. Tubes are filled with a mixture of argon (70%) and carbon dioxide (30%). This mixture is used for a fast drift-time (50 ns) whilst maintaining sufficient drift coordinate resolution (200 μm). Modules are created out of two staggered rows of straw-tubes, as shown in Fig. 2.10(a). Each tracking station is composed of four layers of these modules, arranged in the same *xuvx* formation, as shown in Fig. 2.10(a)

2.4.4 Magnet

A 5 m long (warm) dipole magnet is used by the tracking system to allow momentum measurements. The design of the magnet with an integrated 4 Tm magnetic field had to allow for minimal field protrusion into the RICHs, whilst allowing

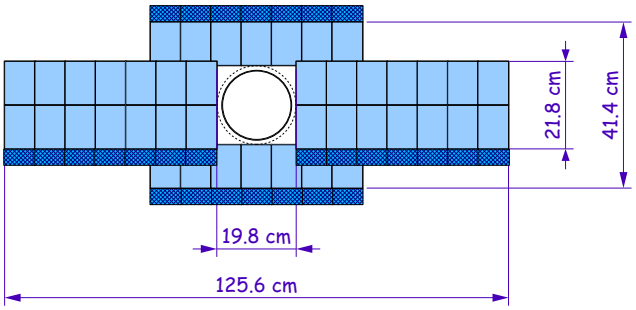

Figure 2.9: Arrangement of the detector boxes in the IT. The dimensions of the boxes are shown. Taken from [40].

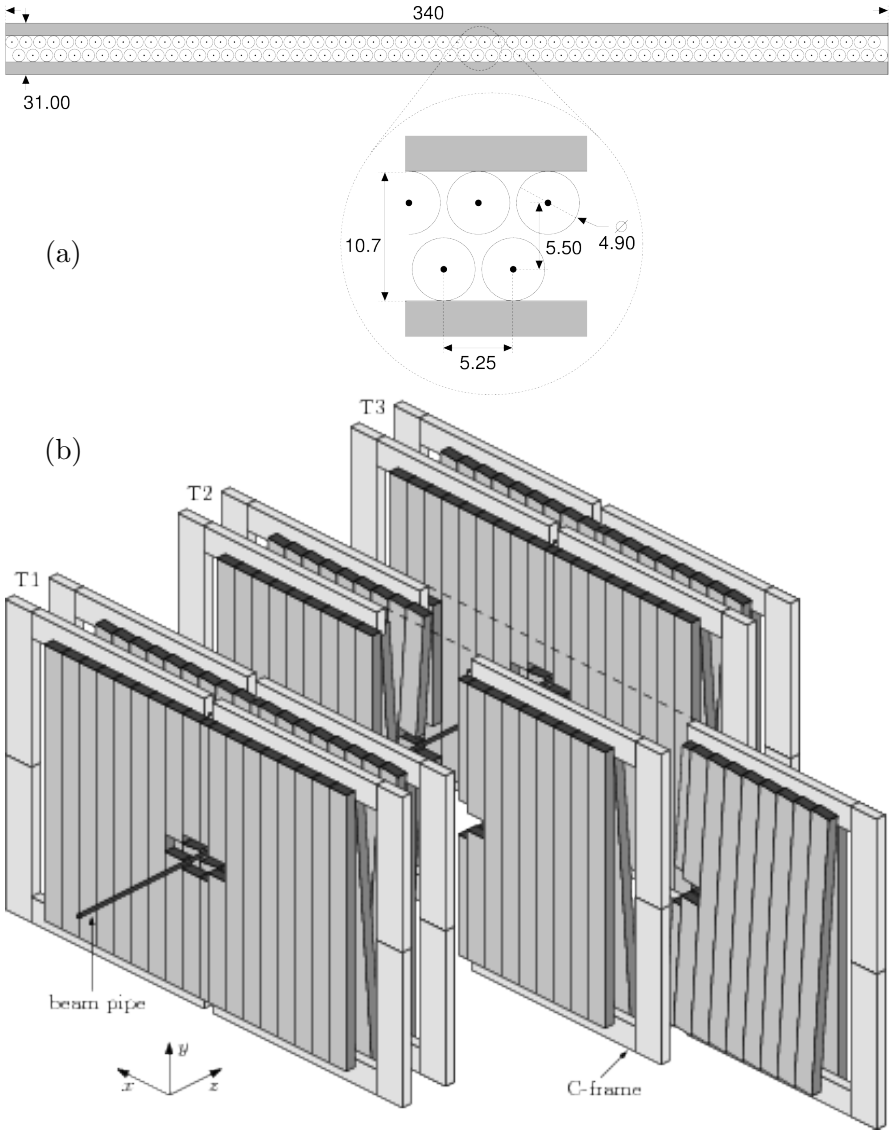

Figure 2.10: (a) Module cross section and (b) arrangement of OT straw-tube modules in layers and stations. Taken from [45].

a field between the VELO and TT. The magnetic field's main component lies in the y-axis and its polarity is switched regularly to allow the investigation of systematics in charge asymmetry measurements.

2.4.5 Track reconstruction

Tracking algorithms combine hits in the VELO, TT and tracking stations to determine the trajectory of charged tracks from the VELO to the calorimeters. Tracks are classed as one of five types, depending on the position of the associated hits. The schematic shown in Fig. 2.11 illustrates this. The track types are the following:

- **Long tracks**, traversing all of the tracking sub-detectors. These tracks have the most precise momentum resolution.

- **Upstream tracks**, traversing only the VELO and TT stations. These are tracks with lower momentum, hence after they pass RICH1, where they may emit Cherenkov photons if their velocity is above threshold, they are bent out of the detector's acceptance by the magnetic field.

- **Downstream tracks**, traversing only the TT and tracking stations. They are the result of charged particles produced in secondary interactions, typically in K^0_S and Λ decays.

- **T tracks**, traversing only the tracking stations. Generally not used but can be useful to seed tracks back through the magnet.

- **VELO tracks**, traversing only the VELO. This occurs when a track is produced at small or negative values of η. These tracks are useful for primary vertex reconstruction.

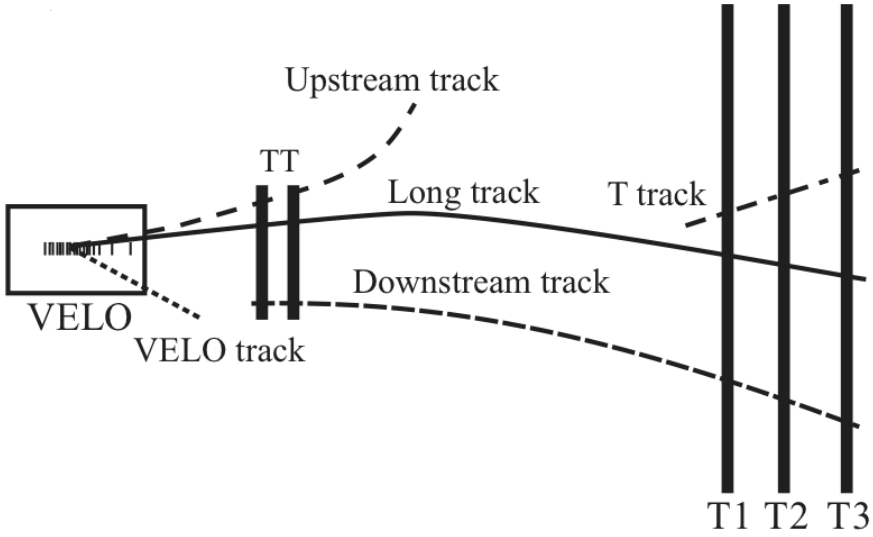

Figure 2.11: Schematic illustration of the track types. The Lorentz force felt by charged particle is perpendicular to the particle's velocity and the main B-field component, B_y (the y-axis comes out of the page). Taken from [40].

Track reconstruction searches for straight line track segments within the VELO and tracking stations. An exact description of the magnetic field strength is not needed as linear extrapolation of two segments describing the same trajectory will intersect in the magnetic field at an almost constant z-position. The track segments are matched by comparing their track parameters in x, y (position) and dy/dz (slope). From these a χ^2 may be calculated, with the combinations with the lowest χ^2 being matched. This idea is illustrated in Fig. 2.12. The track trajectories are then refitted using a Kalman filter algorithm [46] which takes account of multiple scattering and energy loss within the detector.

2.4.6 Tracking performance

The performance of the tracking system is evaluated using simulated $B^0 \to J/\psi K^0_S$ events [40]. The average number of fully reconstructed tracks in fully simulated $b\bar{b}$ events is about 72. Approximately one third of these are long tracks. The efficiency of finding as a long track the trajectory of a particle with momentum

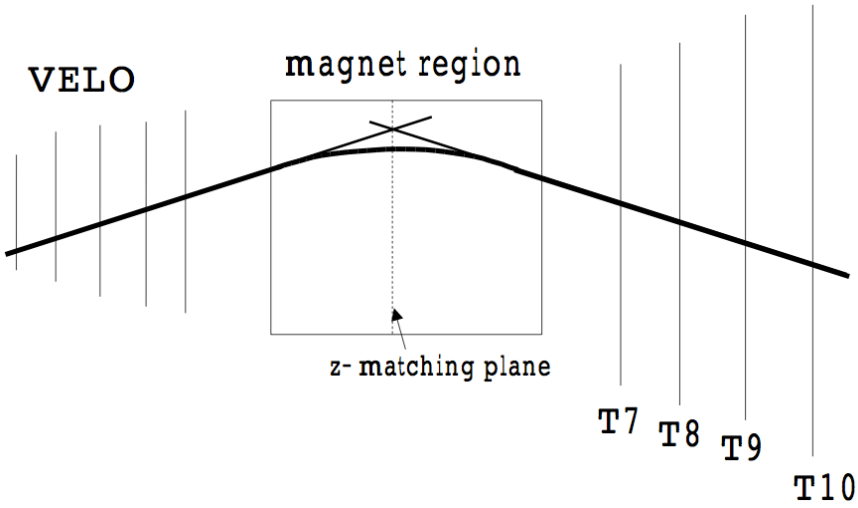

Figure 2.12: Principle of matching two straight line segments from VELO and the tracking stations, drawn in the bending plane of the magnet. The matching can be done without detailed knowledge of the inhomogeneous magnetic field. Taken from [47].

larger than 10 GeV/c is on average 94%. Ghost tracks are defined as those that are not associated with a simulated particle. The corresponding average ghost fraction is about 9%, but most ghost tracks have a low reconstructed p_T. The momentum resolution increases as a function of p_T, $0.35 < \delta p/p < 0.55\%$ at a momentum scale of $3 < p < 100$ GeV/c. The impact parameter (IP) resolution also carries a dependence on p_T that can be parametrized as $\sigma_{IP} = 14\,\mu\mathrm{m} + 35\mu\,\mathrm{m/p_T}$ with p_T in GeV/c. The PV resolution along the beam axis is shown in Fig. 2.13.

2.5 Ring imaging Cherenkov detectors

The RICH system provides LHCb with charged PID over a large momentum range, $2 < p < 100\,\mathrm{GeV}/c$. This unique resource at the LHC is vital to suppress physics backgrounds which are topologically identical to signal. For example, the invariant mass fits shown in Fig. 2.14 highlight the improvement in signal significance after requiring PID.

RICH detectors collect Cherenkov light produced by charged particles as they

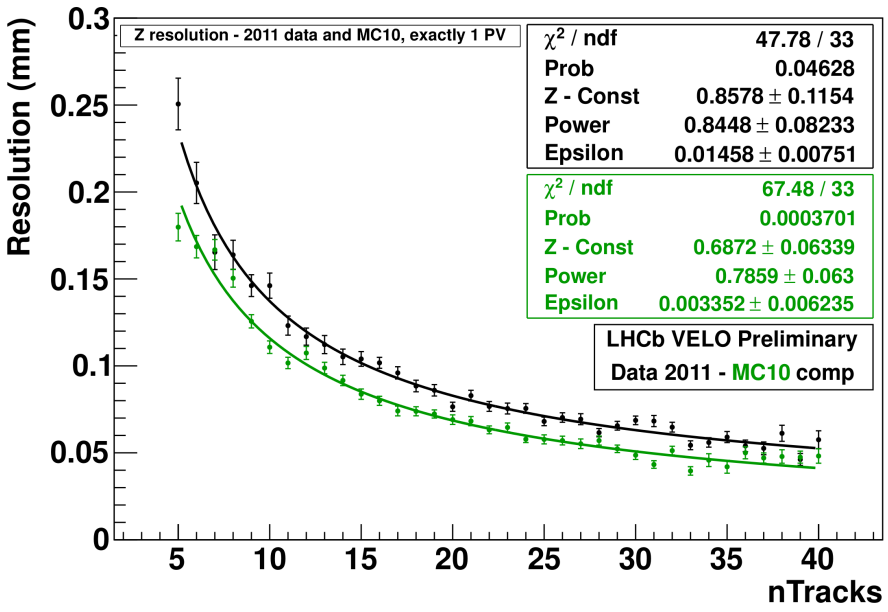

Figure 2.13: Vertex resolution along the beam axis as function of the number of tracks. The black (green) data points and line describe the 2011 data (simulation). Taken from [43].

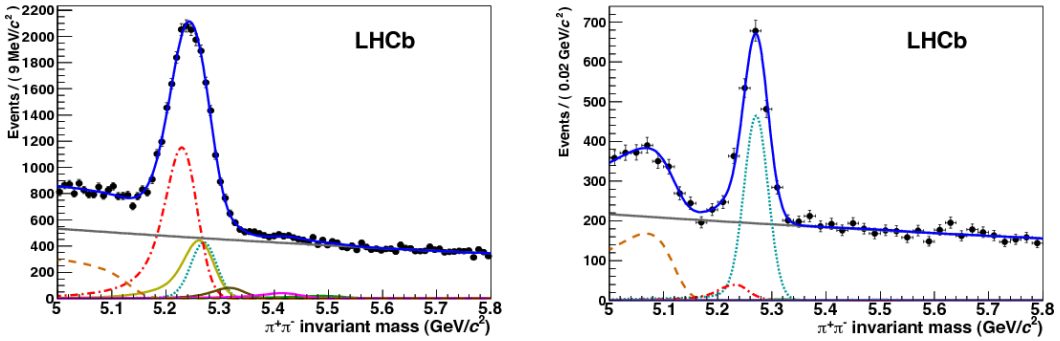

Figure 2.14: Invariant mass distribution of $B \to h^+h^-$ decays [48] in the LHCb data before the use of the RICH information (left), and after applying RICH particle identification (right). The signal under study is the decay $B^0 \to \pi^+\pi^-$, represented by the turquoise dotted line. The contributions from different b-hadron decay modes ($B^0 \to K\pi$ red dashed-dotted line, $B^0 \to$ 3-body orange dashed-dashed line, $B_s \to KK$ yellow line, $B_s \to K\pi$ brown line, $\Lambda_b \to pK$ purple line, $\Lambda_b \to p\pi$ green line), are eliminated by positive identification of pions, kaons and protons and only the signal and two background contributions remain visible in the plot on the right. The grey solid line is the combinatorial background.

exceed the speed of light in a medium. Light is produced throughout a circular cone around the direction of travel. The opening angle of the cone, θ_c, is related to the velocity, β, of the particle,

$$\cos\theta_c = \frac{1}{\beta n}, \tag{2.2}$$

where n is the refractive index of the medium. The likelihood of the particle under various mass hypotheses is compared by combining β with momentum measurements from the tracking system.

The RICH system is composed of two detectors: RICH 1 and RICH 2. RICH 1 covers particles with low to medium momenta, $2 < p < 40\,\mathrm{GeV}/c$, over LHCb's full acceptance range. It is located as close as possible to the interaction region to reduce its overall volume. RICH 2 identifies high momenta particles, $15 < p < 100\,\mathrm{GeV}/c$, and is placed downstream of the magnet, covering a smaller acceptance range, $15 - 120$ mrad. This is allowed because high momentum particles are bent less by the magnetic field so tend to traverse the detector at smaller angles to the beam line.

Fluorocarbon gases at room temperature and pressure are used as Cherenkov radiators. RICH 1 uses C_4F_{10}, with the addition of a silica aerogel layer to identify low momentum kaons, while RICH 2 contains CF_4. Mirrors focus emitted photons into ring images on photon detector planes outside of the acceptance, as shown in Fig. 2.15. Corrections to the images are made to account for distortion by the fringe field of the magnet.

2.5.1 Charged hadron identification

The identification of hadrons is performed by combining Cherenkov angle information, provided by the RICH systems, with the track momentum measured by

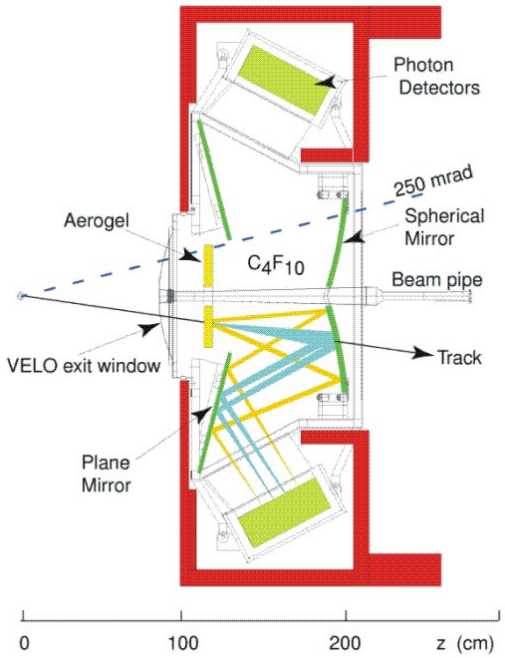

Figure 2.15: Schematic side view of RICH 1. The charged particle shown emits light as it traverses both the aerogel and gas radiators. Mirrors focus the light upon photon detectors which turn it into an electrical signal. Taken from [40].

the tracking system. As the RICH systems operate in a high occupancy environment, Cherenkov cones predominately overlap. An overall event likelihood, where all tracks in an event and in both RICH detectors are considered simultaneously, is used for optimal performance.

Starting from the 'all tracks are pions' hypothesis (due to the abundance of pions created during collisions), for each track in turn the likelihood is recomputed changing the mass hypothesis to e, μ, π, K and p^+, whilst leaving all other track hypotheses unchanged. The mass hypothesis for a track that gives the largest increase in the event likelihood identifies that track. This procedure is then repeated until all tracks have been set to their optimal hypothesis, and no further improvement in the event likelihood is found. For isolated tracks (tracks associated with a non-overlapping Cherenkov ring and occurring in $\sim 2\%$ of cases) the reconstructed Cherenkov angle forms a distinct band structure, shown in Fig. 2.16.

Calibration samples are used to evaluate the performance of discriminating variables under different mass hypotheses. These samples are isolated using non-biasing kinematic cuts. The efficiency with which kaons are separated from pions as a function of momentum is measured in data and shown in Fig. 2.17. Requiring that the kaon hypothesis is more likely than the pion hypothesis, $\mathrm{DLL}_{\mathrm{K}-\pi} > 0$, and averaging over the momentum range $2 < p < 100$ GeV/c, the kaon efficiency is $\epsilon(K \to K) \sim 95\%$ for a pion misidentification efficiency of $\varrho(\pi \to K) \sim 10\%$ [49].

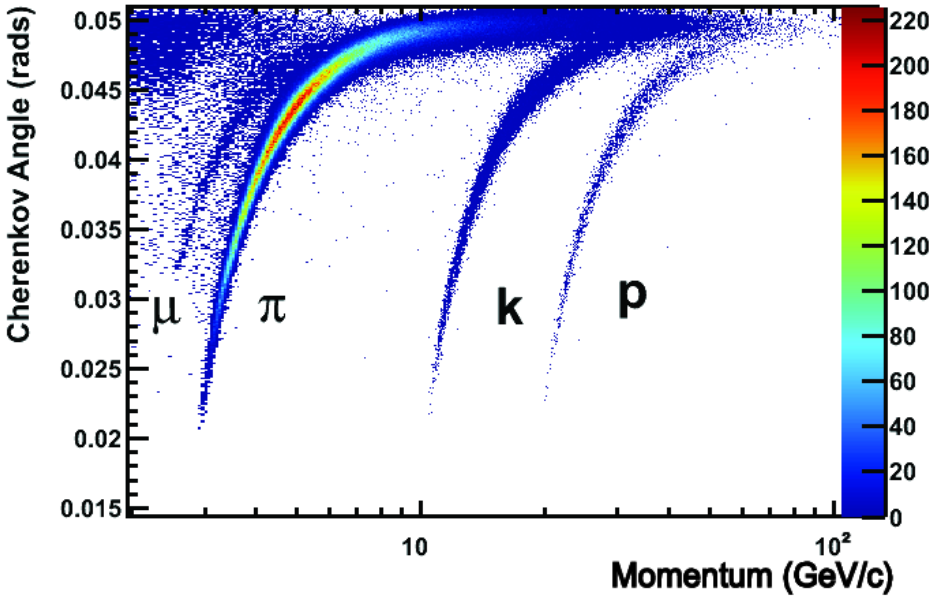

Figure 2.16: Reconstructed Cherenkov angle measured in RICH1 as a function of track momentum in the C_4F_{10} radiator for isolated tracks. Taken from [49].

2.6 Calorimeters

Calorimeters are used to help identify and to measure the energy and position of both charged and neutral particles inside LHCb. Information from the calorimeters is used to select high transverse energy $h^{\pm}$, $e^{\pm}$ and γ candidates in the Level-0 (L0) trigger. When certain particles traverse a material they initiate a particle shower by producing a cascade of particles. The total energy of all secondary

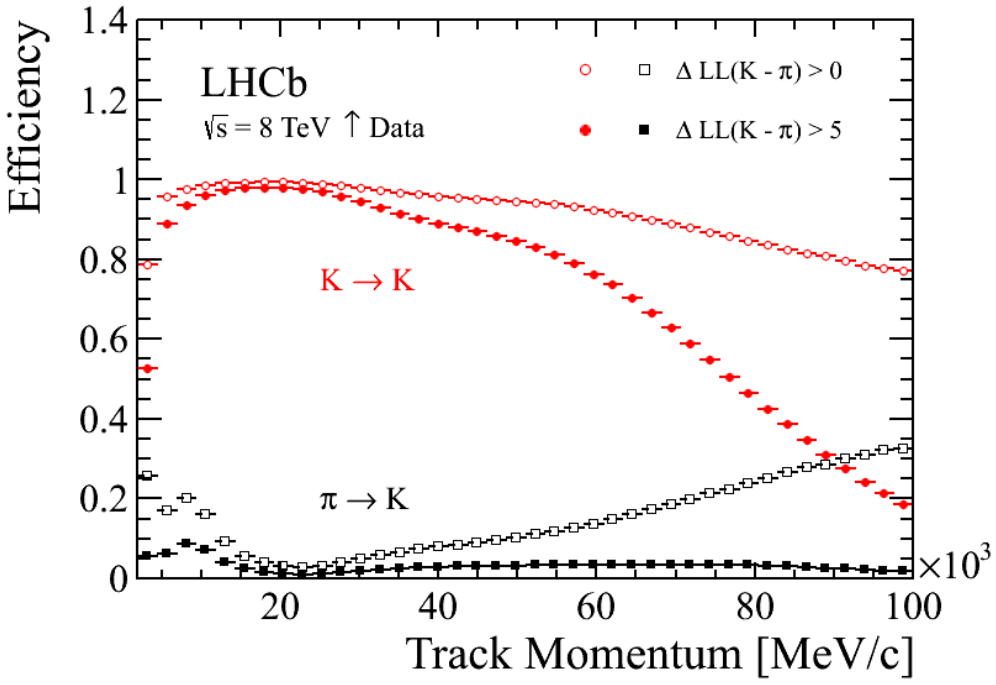

Figure 2.17: Kaon identification efficiency and pion misidentification rate measured on data as a function of track momentum for two different $\mathrm{DLL}_{\mathrm{K}-\pi}$ cuts. Taken from [49].

particles is collected to give a measure of the traversing particle's energy.

Particle showers come in two types: electromagnetic, measured by the electromagnetic calorimeter (ECAL), and hadronic, measured by the hadronic calorimeter (HCAL). Electromagnetic showers are started by particles which interact primarily or exclusively through ionisation, bremsstrahlung or pair-production (for low energy $e^{\pm}$, high energy $e^{\pm}$, and γ, respectively). The scale of a shower is the radiation length, X_0, and is specific to that material. This is also the length after which a particle's energy is reduced to a fraction (dependent upon the type of interaction) of its original size. Hadronic showers stem from strong interactions, with a scale which is the nuclear absorption length, λ_l. This is typically greater than X_0 which means that HCALs are normally larger and denser than their EM counterparts.

The scintillator pad detector (SPD) and pre-shower (PS) detector are needed to

identify electrons and photons. A particle's electromagnetic charge is determined by the SPD as only charged particles cause the emission of light. This allows for separation of $e^{\pm}$ and γ. 15 mm of lead absorber, corresponding to $2.5X_0$, is placed after the SPD but before the PS. By segmenting the electromagnetic cascades in the PS and ECAL, the shower profile is used to differentiate $\pi^{\pm}$ from $e^{\pm}$.

The ECAL and HCAL are sampling calorimeters. This means that they are composed of alternating layers of absorber and scintillater. Scintillation light is transmitted to photon multiplier tubes via waveguides before being turned into an electrical signal. In the ECAL, optimal energy resolution requires full shower containment from high energy photons. To capture the shower, 66 layers of 2 mm thick lead and 4 mm thick scintillating material, corresponding to $25X_0$, provide an energy resolution of $\frac{8\%}{\sqrt{E(\mathrm{GeV})}} \oplus 0.8\%$. This compares favourably to BaBar's ECAL at higher energies, $\frac{2\%}{\sqrt{E(\mathrm{GeV})}} \oplus 1.4\%$ [50]. There is no such requirement in the HCAL, so it is made up of $5.6\lambda_l$ of alternating iron and scintillating material layers, resulting in an energy resolution of $\frac{67\%}{\sqrt{E(\mathrm{GeV})}} \oplus 9\%$. At lower energies, this compares favourably with the HCAL in the CMS experiment, $\frac{100\%}{\sqrt{E(\mathrm{GeV})}} \oplus 5\%$ [51].

2.6.1 Electron reconstruction

The ECAL and tracking system play an important role in electron identification. Identification is mainly based on the balance between track momentum and the energy of the charged cluster in the ECAL, see Fig. 2.18 [40], and the goodness of the cluster-track matching which is classified by the variable χ^2_{track}. Improvement in electron identification is obtained by discriminating based on the energy deposited in the preshower detector and along the extrapolated particle trajectory in the HCAL.

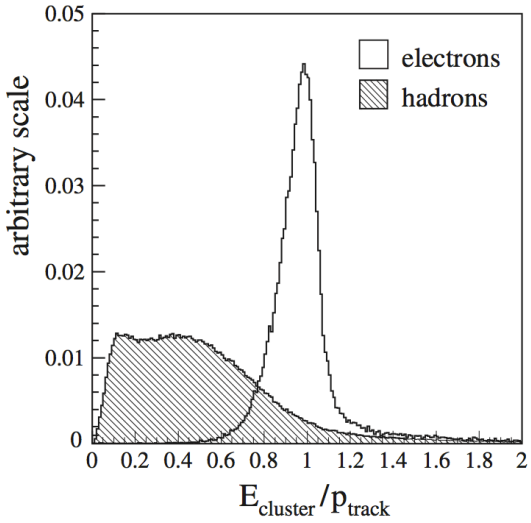

Figure 2.18: The ratio of uncorrected energy of the charged cluster in ECAL to the momentum of reconstructed tracks for electrons (open histogram) and hadrons (shaded histogram). Taken from [40].

Reconstructing an electron means that energy emitted in the form of bremsstrahlung photons needs to be taken into account. An electron may release bremsstrahlung photons as it travels through LHCb. If this occurs after the magnet, the photon will hit the same calorimeter cells as the electron, and the energy is recuperated. Emission before the magnet means that the trajectory of the electron will differ from its original path which is still that of the photon. The energy of the two are then deposited in different parts of the calorimeter, as seen in Fig. 2.19 [40]. Should the energy of the bremsstrahlung not be recoverable, the invariant mass resolution will be degraded.

2.6.2 Photon reconstruction

Neutral clusters may be identified using matching techniques with reconstructed tracks. Reconstructed tracks are extrapolated to the calorimeter reference plane. The geometrical distance between the cluster position and the extrapolated charged tracks is used to determine χ^2_{track}. Photon candidates are required to have χ^2_{track}

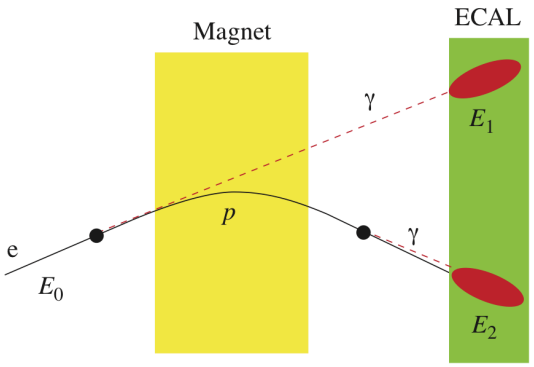

Figure 2.19: Schematic illustration of bremsstrahlung photon recovery. An electron may radiate photons when passing through material before or after the magnet. In the first case, a well defined cluster corresponding to the photon is seen in the ECAL, with energy E1, thus the energy of the electron at the origin, E0 = E1 + E2. In the second case the bremsstrahlung energy forms part of the electron cluster with energy E = p, the momentum measured in the spectrometer. Taken from [40].

exceeding 4. This cut suppresses the clusters due to charged tracks such as electrons and charged hadrons.

The energy and direction of the photon are evaluated and corrected for based on energy deposits in 3×3 ECAL cells around the cluster seed. The corrections applied depend on the ECAL region and take into account the eventual conversion of the photon into an electron-positron pair. The presence of a hit in the SPD determines whether a photon converted. Simulation studies [52] show clusters associated with hits in the SPD are converted photons at $(79.1 \pm 0.1)\%$. Those without SPD hits are 90% (negligible statistical error) unconverted photons. Averaging across p_T, 44% of photons from the interaction point convert by the time they reach the calorimeter.

Conversion of photons can only occur within material and, of the 44% of photons that convert, half convert before the magnet. Conversion before the magnet means that the electrons are likely to be resolvable inside the ECAL, as they will have been separated by the magnetic field. In this case however, the χ^2_{track} cut on charged tracks may remove such candidates or the magnet may sweep

the electrons out of the calorimeter's acceptance. Hence, the efficiency is lower for this sample than for unconverted photons. Those that convert after the magnet are likely to hit a single cell within the ECAL. The efficiency for this sample is similar to unconverted photons. Only unconverted photons are used in the charge asymmetry analysis presented in this thesis.

2.7 Muon stations

Rare decay analyses often search for final state muons because of their clean signature. As well as a combinatorial background, muonic decays suffer from topologically identical hadronic backgrounds with large branching fraction. Muon stations must therefore be highly efficient at identifying muons whilst maintaining a low hadronic misidentification rate.

The muon system is composed of five muon stations, M1 to M5, and attempts to exploit the penetration power of muons. The geometry of the stations is projective, meaning that the stations traverse dimensions scale with their distance from the interaction region. This provides a coverage of 16 (20) mrad to 258 (306) mrad in the vertical (horizontal) planes, essentially LHCb's full acceptance. M1 is located upstream of the calorimeters so that it can help improve p_T measurements in the trigger. Stations M2 to M5 are located downstream of the calorimeters. Particles produced in the interaction region must travel 6.6 nuclear absorption lengths before arriving at M2. M2 to M5 are interleaved with 80 cm thick iron absorbers to filter out hadrons and electrons. At M5, particles produced in the interaction region have travelled approximately 20 nuclear absorption lengths.

Using similar technologies to the drift chambers used in the Tracking Stations, each station provides spatial hit measurements that allow the track trajectories to be plotted. This information is obtained by partitioning the system into rectan-

gular logical pads whose dimensions determine the (x,y) resolution. Each station is divided into four regions of approximately equal occupancy. The area of each region is double the size of the previous region, so the resolution increase by a factor of eight for the outermost region.

2.7.1 Muon identification

Muons are identified by a binary selection (`IsMuon`) requiring tracks to have associated hits in the muon stations within a field of interest (FOI) extrapolated from the tracking stations. The exact number of stations requiring hits, see Table 2.2, and the size of the field of interest are dependent upon track momenta.

Momentum range [GeV/c]	**Muon stations**
$3 < p < 6$	M2 and M3
$6 < p < 10$	M2 and M3 and (M4 or M5)
$p > 10$	M2 and M3 and M4 and M5

Table 2.2: Muon stations required to trigger the binary selection as a function of track momentum. Muons with $p < 3$ GeV/c do not reach the muon stations.

Using a sample of $J/\psi \to \mu^+\mu^-$ events [53] the muon identification efficiency is measured, $\epsilon(\mu \to \mu) = (98.13 \pm 0.04)\%$, with a corresponding pion misidentification, $\varrho(\pi \to \mu) = (1.025 \pm 0.003)\%$. Further discrimination between muons and non-muons is based on the test-statistic used for the RICH, $\mathrm{DLL}_{\mu-\pi}$. The efficiency, $\epsilon(\mu \to \mu)$, and misidentification rate, $\varrho(\pi \to \mu)$, for $\mathrm{DLL}_{\mu-\pi}$ are shown in Figs. 2.20. The $\mathrm{DLL}_{\mu-\pi}$ cut chosen for the rare decay analysis is $\mathrm{DLL}_{\mu-\pi} > 3$. The momentum distribution for muon candidates from the rare decay analysis is shown in Fig. 2.21. Crucially, the majority of muon candidates have a momentum greater than 10 GeV/c. Candidates with a lower momentum than this have an efficiency of zero with the PID cut, $\mathrm{DLL}_{\mu-\pi} > 3$.

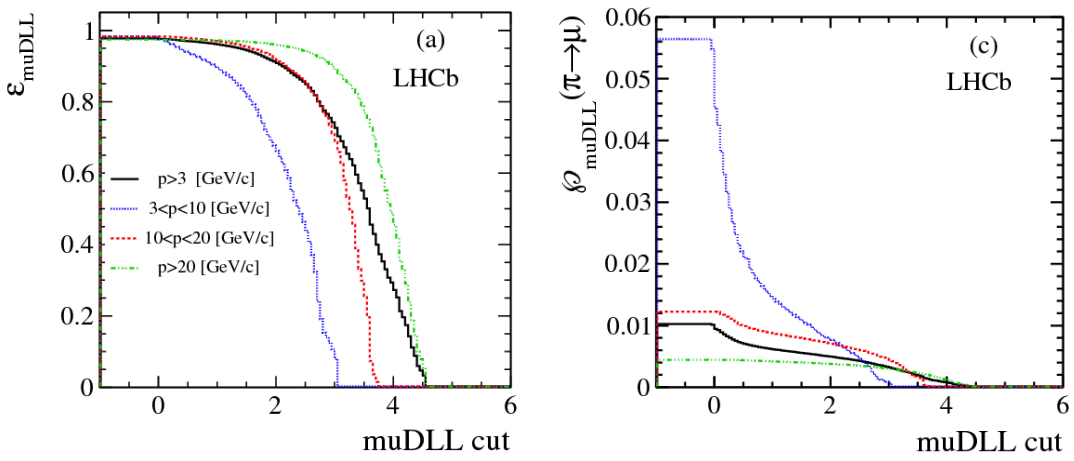

Figure 2.20: (Left) Efficiency, $\epsilon(\mu \to \mu)$, and (right) misidentification rate, $\varrho(\pi \to \mu)$, as a function of $\mathrm{DLL}_{\mu-\pi}$ for different momentum bins. The curves intercept the y-axis at the point corresponding to the isMuon requirement. The cut used in the rare decay analysis is $\mathrm{DLL}_{\mu-\pi} > 3$ and is seen to be approximately 75% efficient, with a pion misidentification rate of approximately 0.3%. Taken from [53].

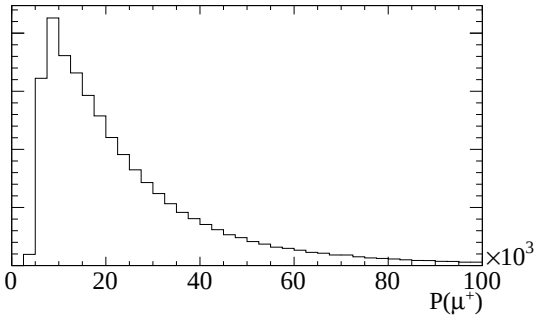

Figure 2.21: Momentum distribution of muon candidates from the rare decay analysis, before an offline selection is applied.

2.8 Triggering and event filtering

LHCb's trigger uses information from all sub-detectors. The trigger is divided into L0 and a higher level trigger (HLT). L0 is hardware-based and uses input from the calorimeters and muon system. This trigger reduces the rate of considered events from the rate of beam crossings with at least one inelastic pp interaction, 15 MHz, to 1 MHz, the rate at which the whole detector can be readout. Events

passing L0 are passed to the HLT which further reduces the rate to 5 kHz. All events passing the HLT are stored on disk, where full track reconstruction occurs. This is known as offline reconstruction, as opposed to the online reconstruction performed by the HLT.

2.8.1 Level 0

The L0 trigger is fully synchronous with 40 MHz bunch crossing signal from the LHC. Decisions are made by custom electronics within 4 μs of this signal's arrival. L0 selects events of physical interest through two independent triggers: the L0-Calorimeter trigger and the L0-Muon trigger.

The L0-Calorimeter trigger uses information from the SPD, PS, ECAL and HCAL. An event is accepted by this part of the L0 if at least one candidate in the event passes an E_T threshold. A candidate's E_T is the energy deposited by a track in a cluster of 2×2 adjacent cells of the calorimeters.

- $h^{\pm}$ **candidate (L0Hadron)**, highest E_T HCAL cluster. If the highest E_T ECAL cluster is in front of this cluster, combine the measurements.
- γ **candidate (L0Photon)**, highest E_T ECAL cluster with hits in front of the cluster in the PS but not in the SPD.
- $e^{\pm}$ **candidate (L0Electron)**, highest E_T ECAL cluster with hits in front of the cluster in the PS and SPD.

In the muon stations, for each logical pad hit in M3, straight line extrapolation and FOIs are used to find hits in the other stations. It is assumed that the tracks come from the interaction region and receive a kick from the magnet. Stations are divided into towers (inter-station) and quadrants (intra-station). A single quadrant of the muon system is shown in Fig. 2.22. Each system quadrant identifies

its two highest p_T tracks and so eight tracks are passed to a decision unit. There is no exchange of information between quadrants so tracks traversing quadrant boundaries are not reconstructed. An event is accepted by the L0 trigger if one of the eight tracks has a p_T above threshold (L0Muon), or $p_T^{largest} \times p_T^{2ndlargest}$ is above threshold (L0Dimuon).

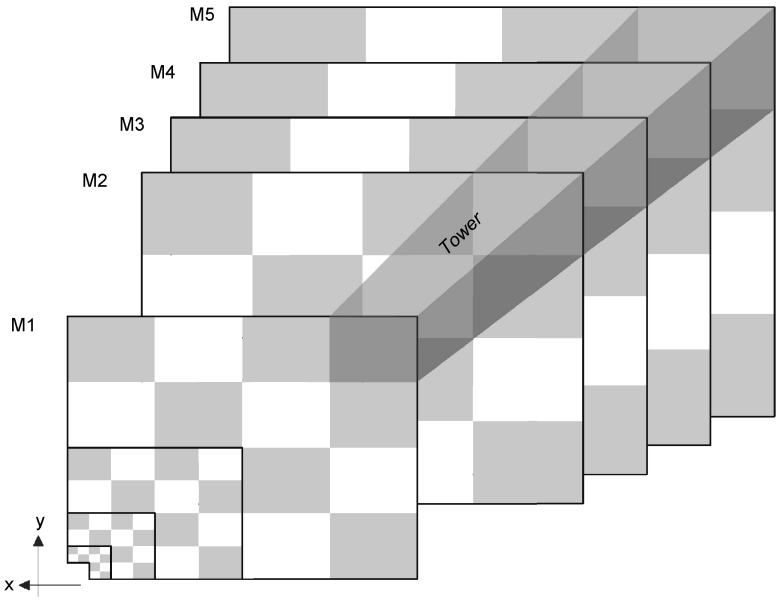

Figure 2.22: A quadrant of the muon detector with its 48 towers pointing towards the interaction point. Taken from [54].

2.8.2 High level trigger

The HLT runs on an event filter farm some distance from the actual detector and is a composition of trigger decisions implemented in software. The HLT is conceptually divided into two stages, HLT1 and HLT2. During 2011 running, the HLT contained 38 HLT1 and 131 HLT2 lines. Slightly more were added for 2012 running. An event passes each stage if it is accepted by any line.

HLT1 processes the full L0 rate and partially reconstructs events to reduce the rate to 43 kHz. Its purpose is to reconstruct the L0 tracks with hits in the VELO and Tracking Stations, or, for γ or π^0, to confirm the absence of an associated

charged track. VELO hits are first used by HLT1 to identify the position of the pp interaction region. This is measured to be stable to within a few micrometers. The VELO then reconstructs vertices with at least five tracks. Vertices within a 300 μm radius cylinder around the interaction region are defined as PVs. For non-muon lines tracks are selected based on the IP and the quality of the track (relating number of hits to expected hits). For muon lines, tracks are extrapolated from the VELO to the M3. They are accepted if, upon further extrapolation, at least another hit is found in another muon station and the track is of reasonable quality. Momentum and p_T cuts remove lower momentum tracks using information gained from adding hits in the tracking stations (forward tracking). Finally a Kalman filter track fit calculates the χ^2 and covariance matrix of each track.

At this rate, on all high momentum tracks, HLT2 can perform full reconstruction of exclusive events and partial reconstruction of inclusive B-meson decays. Electron and muon candidates are identified with information from the ECAL and Muon System, respectively. Two-prong SVs are identified. Selection cuts are applied on such quantities as the invariant mass, vertex quality and the sum of the momenta. Lower momentum tracks are added to the vertices to form exclusively reconstructed candidates. Selection cuts on decay topology and RICH PID information may also be applied. The total rate that is saved to disk is 5 kHz, this is broken down in Fig. 2.23. The performance of two leading HLT2 charm lines are shown in Fig. 2.24.

2.8.3 Event stripping

Each analysis is based on data selected by a stripping line. Each stripping line runs LHCb's full offline reconstruction, using information from all sub-detectors, to find events containing specific decays. Events which successfully pass are stored in easily-handled data files, typically a few to tens of gigabytes in size.

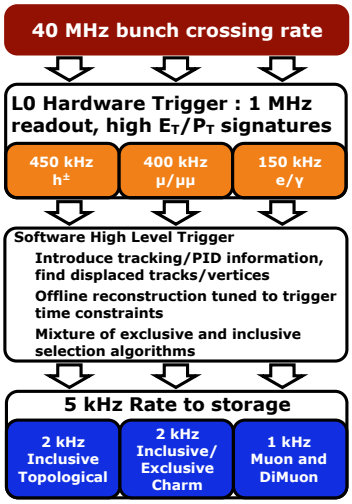

Figure 2.23: LHCb's trigger scheme. Read-out rates and the information available at each stage are shown. Taken from [55].

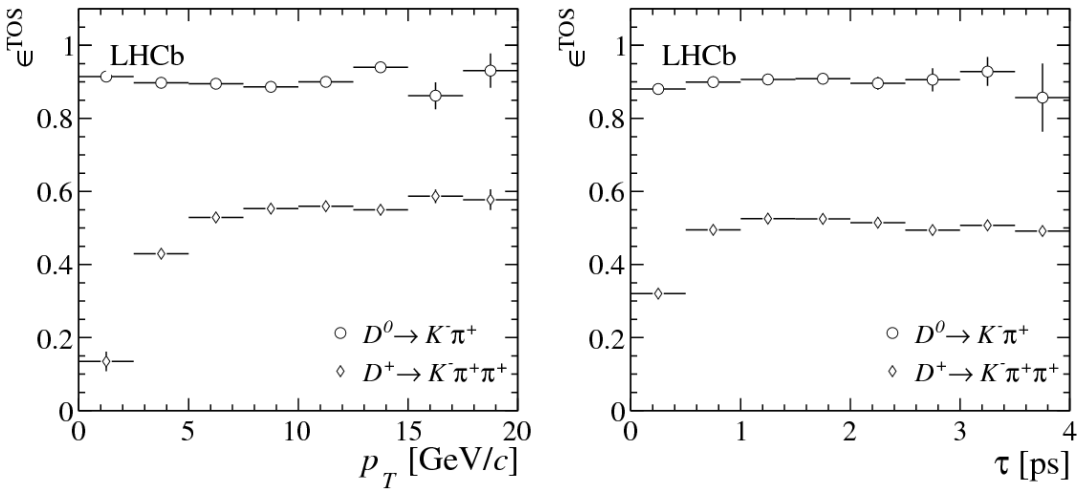

Figure 2.24: Efficiency of the two- and three-body charm lines as a function of (left) p_T and (right) lifetime of the D meson. These lines triggered on the signal (TOS). The efficiency is much higher for the two-body mode as the cuts applied here are very similar to those applied in the offline selection. Taken from [56].

Chapter 3

A comprehensive study of the $\pi^+\mu^+\mu^-$ final state from charged charm mesons

A search for non-resonant $D^+_{(s)} \to \pi^+\mu^+\mu^-$ decays is performed using proton-proton collision data, corresponding to an integrated luminosity of 3.0 fb^{-1}, recorded by the LHCb experiment. No signals are observed and the 90% (95%) CL limits on the branching fractions, the most stringent to date, are found to be,

$$\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-) < 2.9\,(3.3) \times 10^{-8},$$
$$\mathcal{B}(D^+_s \to \pi^+\mu^+\mu^-) < 1.76\,(1.91) \times 10^{-7}.$$

3.1 Introduction

Decays that feature loop diagrams may be affected by the couplings of non-SM particles. Measuring the branching fractions and angular asymmetries of FCNC processes therefore indirectly tests NP. Unlike in the B and K systems, FCNC decays have never been observed in charm mesons. The copious amount of charm mesons produced at LHCb places the experiment in a position to search for such decays. The analysis presented in this thesis is an update to my LHCb publication [4]. That paper used 1 fb^{-1} of data from 2011. Here the 2012 data is also

studied.

3.1.1 Phenomenological motivation

Flavour-changing neutral current (FCNC) processes are rare within the Standard Model (SM) as they cannot occur at tree level. At loop level, they are suppressed by both the Glashow-Iliopoulos-Maiani (GIM) [8] and Cabibbo-Kobayashi-Maskawa (CKM) [18] mechanisms but are nevertheless well established in processes in K and B mesons. In contrast to the B meson system, where the near-unity value of $|V_{tb}|$ and high mass of the top quark in the loop weakens GIM suppression, that cancellation is almost exact in D meson decays leading to lower SM branching fractions ($\mathcal{B}$). Feynman diagrams for the decay $D^+ \to \pi^+\mu^+\mu^-$ are shown in Fig. 3.1(a, b).

The decay $D_s^+ \to \pi^+\mu^+\mu^-$ proceeds via the weak annihilation diagram shown in Fig. 3.1(c). In the event that a signal is seen this can be used to check the nature of a potential $D^+ \to \pi^+\mu^+\mu^-$ signal where an analogous weak annihilation diagram proceeds, albeit suppressed by a factor $|V_{cd}|^2$. This check is needed in order to distinguish between FCNC and weak annihilation contributions. Note that throughout, inclusion of conjugate processes is implied.

The diagrams in Fig. 3.1 make up the non-resonant, or short-distance (SD), contribution to the decays. Accompanying this is the long-distance (LD) contribution proceeding via intermediate resonances: $D^+_{(s)} \to \pi^+(X \to \mu^+\mu^-)$, where X can either be the pseudoscalar meson, η, or one of the vector mesons, $\rho^0(770)$, $\phi(1020)$ or, for the D_s^+ decay, $\omega(782)$. The decays $D^+_{(s)} \to \pi^+(\phi \to \mu^+\mu^-)$, see Fig. 3.2, have the largest of the LD branching fractions: $\mathcal{B}(D^+ \to \pi^+ \ (\phi \to \mu^+\mu^-)) = (1.56 \pm 0.12) \times 10^{-6}$ and $\mathcal{B}(D_s^+ \to \pi^+ \ (\phi \to \mu^+\mu^-)) = (1.29 \pm 0.14) \times 10^{-5}$ [1]. As an analysis strategy the search for the SD component is normalised by this almost-identical LD decay. To separate the SD and different LD com-

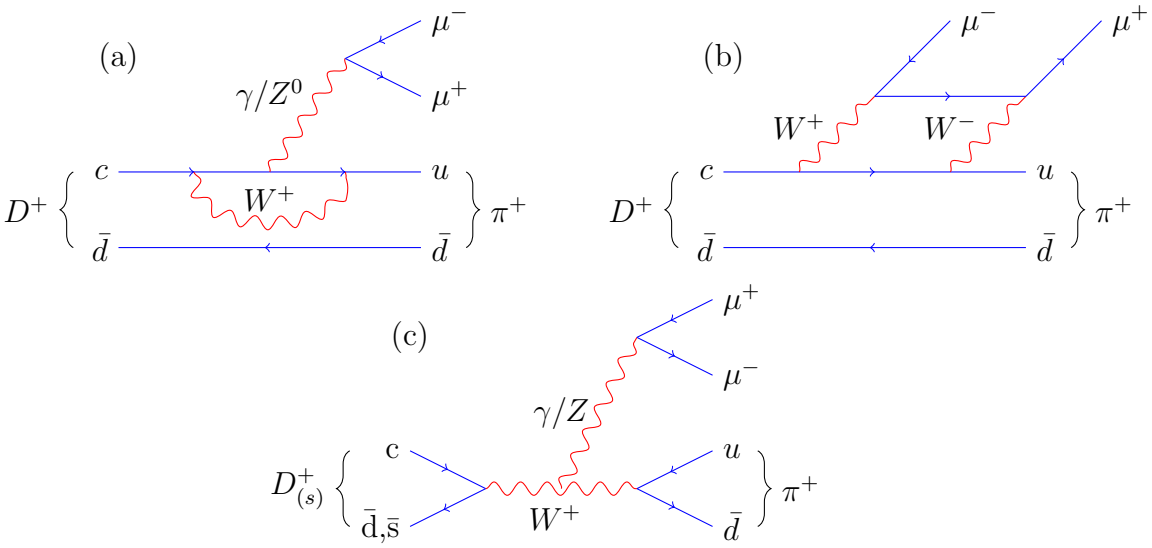

Figure 3.1: Feynman diagrams for (a, b) the FCNC decay $D^+ \to \pi^+\mu^+\mu^-$, and (c) the weak annihilation of a $D^+_{(s)}$ meson.

ponents, candidates are binned in dimuon invariant mass, $m(\mu^+\mu^-)$. Bins are chosen to fully enclose each resonance present in $m(\mu^+\mu^-)$ (although the ρ^0 and ω contributions are kept together), protecting the two regions sensitive to the SD contribution. The distribution of $m(\mu^+\mu^-)$ versus $m(\pi^+\mu^+\mu^-)$ for preselected candidates and the $m(\mu^+\mu^-)$ bin definitions are shown in Fig. 3.3. The exact bin ranges are detailed in Sec. 3.5.

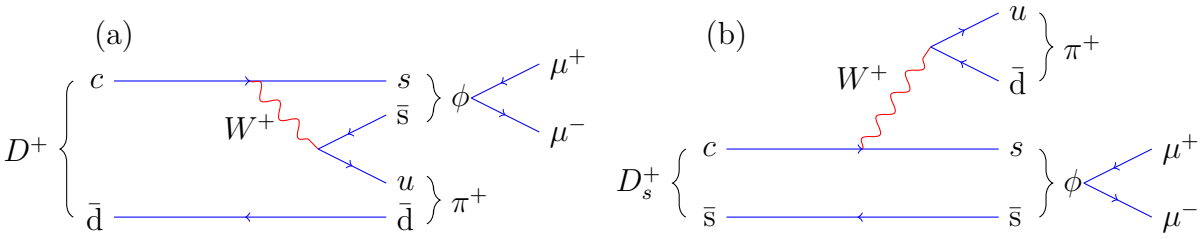

Figure 3.2: Feynman diagrams for the normalisation decays (a) $D^+ \to \pi^+(\phi \to \mu^+\mu^-)$ and (b) $D^+_s \to \pi^+(\phi \to \mu^+\mu^-)$. The D^+ decay is colour-suppressed with respect to the D^+_s decay.

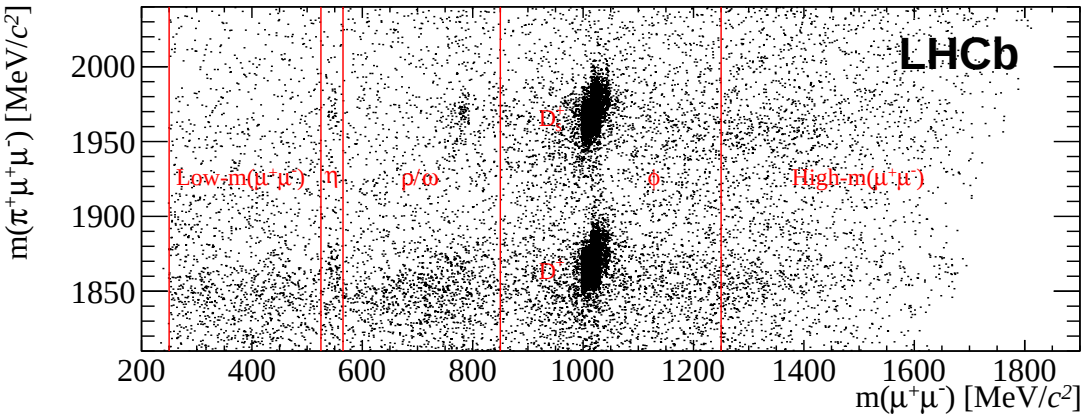

Figure 3.3: Distribution of $m(\mu^+\mu^-)$ versus $m(\pi^+\mu^+\mu^-)$ for preselected events. The dimuon invariant mass bins and their boundaries are shown. The ϕ resonances from $D^+_{(s)}$ candidates (labelled) are already clearly visible.

3.1.2 Theoretical predictions

The SM SD contribution is expected to have a branching fraction of the order of $10^{-8} - 10^{-9}$, see Table 3.1. Suppression of this contribution provides an opportunity to probe the effects of NP on the coupling of up-type quarks in electroweak processes. This is because NP models may introduce additional diagrams that, *a priori*, need not be suppressed in the same manner as the SM contributions. Extensions of the SM, such as Supersymmetric models with R-parity violation, may increase the branching fraction up to 10^{-6} in regions of $m(\mu^+\mu^-)$ away from the LD contribution [57], as shown in Fig 3.4.

Year	$\mathcal{B}(\times 10^{-9})$	**Ref.**
2001	0.2	[58]
2001	9.4	[59]
2011	3.7	[60]

Table 3.1: Predicted SM SD branching fractions for $D^+ \to \pi^+\mu^+\mu^-$ decays.

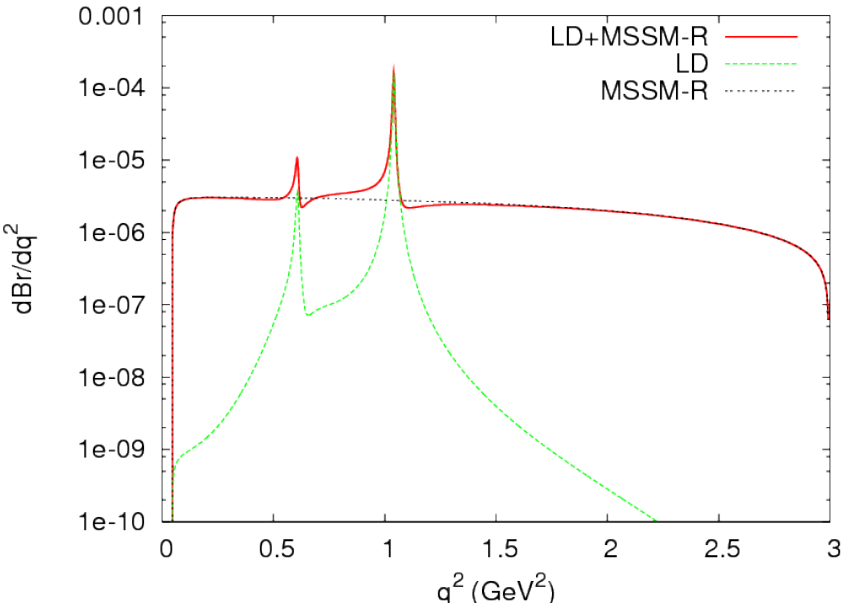

Figure 3.4: Distributions of the LD, SUSY and combined branching fractions for $D^+ \to \pi^+\mu^+\mu^-$ decays across dimuon invariant mass, $m(\mu^+\mu^-)$. Taken from [57].

3.1.3 Previous results

The most stringent limit published prior to my work was $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-) < 3.9 \times 10^{-6}$ (90% CL) by the D0 collaboration [61]. The FOCUS collaboration places the most stringent limit on the D_s^+ weak annihilation decay with $\mathcal{B}(D_s^+ \to \pi^+\mu^+\mu^-) < 2.6 \times 10^{-5}$ [62]. The limits are summarised in Table 3.2 and the invariant mass plots used to create them are shown in Fig. 3.5.

Decay	$\mathcal{B}$**(90% CL)**	**Collaboration**
$D^+ \to \pi^+\mu^+\mu^-$	$< 3.9 \times 10^{-6}$	D0
$D_s^+ \to \pi^+\mu^+\mu^-$	$< 2.6 \times 10^{-5}$	FOCUS

Table 3.2: Previous limits on the branching fractions of $D^+_{(s)} \to \pi^+\mu^+\mu^-$ decays.

3.2 Event selection

The analysis is performed using proton-proton collision data recorded during three different data-taking periods, shown in Table 3.3. Data recorded in 2010 is used

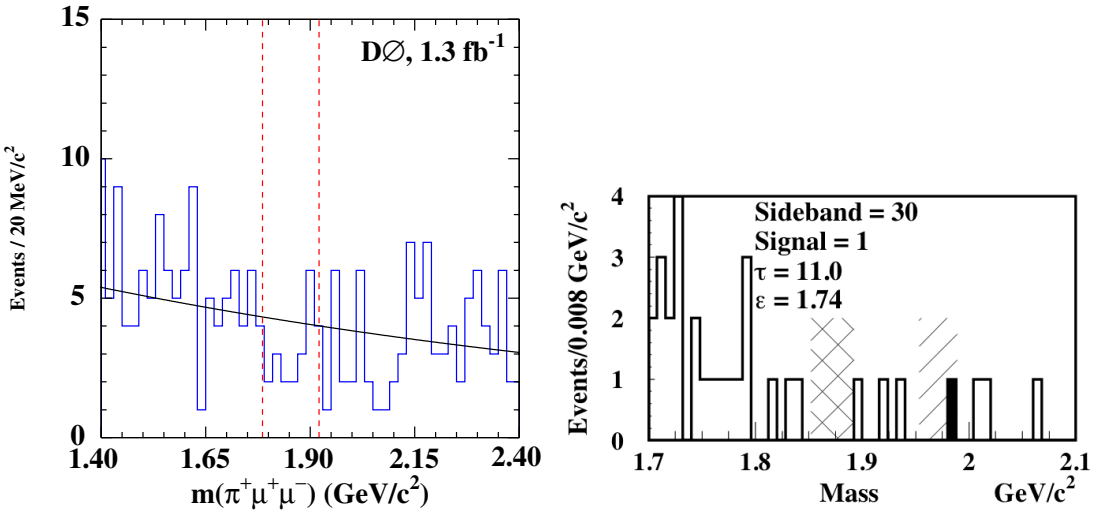

Figure 3.5: Invariant mass fits of $\pi^+\mu^+\mu^-$ candidates by the (left) D0 and (right) FOCUS collaborations. Taken from [61] and [62], respectively.

to train a selection. Improvements are not found when using part of the 2011 or 2012 dataset to train a selection. Two data samples are used in the analysis:

- $D^+_{(s)} \to \pi^+\mu^+\mu^-$ combinations, hereafter referred to as 'the signal'.

- $D^+_{(s)} \to \pi^+\pi^+\pi^-$ combinations, used to understand the pionic background peaking close to the nominal $D^+_{(s)}$ masses.

The final data samples contain candidates which pass a number of selection algorithms. First they must be accepted by LHCb's trigger, then $D^+_{(s)}$ candidates are constructed in the required final state. The subsequent offline selection is more sophisticated and involves a multivariate analysis and particle identification to separate the signal from both combinatorial and physics backgrounds. Simulated signal data was generated in order to guide the analysis. The details of the simulation are given in Table 3.4. The efficiency of reconstruction/stripping is relatively high so approximately 220k simulated events are available to the analysis.

Beam conditions	Luminosity [fb^{-1}]	Used
2010	0.036	Train selection
2011	1.03	Signal extraction
2012	2.05	Signal extraction

Table 3.3: Data used in the analysis.

Beam conditions	Events per polarity
2011	1.0 M
2012	2.0 M

Table 3.4: Simulated $D^+ \to \pi^+\mu^+\mu^-$ decays produced for this analysis. Both magnetic polarities are modelled. The simulated signal was created using Pythia 8.175 [63] and EvtGen 13.6 [64].

3.2.1 Trigger selection

The trigger criteria which $\pi^+\mu^+\mu^-$ candidates are required to pass depends upon the candidate's dimuon invariant mass, $m(\mu^+\mu^-)$. This is because the HLT dimuon lines feature a $m(\mu^+\mu^-)$ cut at 1.0 GeV/c^2, close to the ϕ resonance. As the number of candidates in the ϕ peak helps determine the signal's branching fraction in the low- and high-$m(\mu^+\mu^-)$ bins, a sharp turn-on in the $m(\mu^+\mu^-)$ spectrum must be avoided. The two trigger criteria are shown in Table 3.5.

	L0		HLT1		HLT2	
$m(\mu^+\mu^-)$ bin	Muon	Dimuon	Muon	Dimuon	$\pi^+\mu^+\mu^-$	Dimuon
Low, η, ρ^0/ω, ϕ_{low}	✓	✓	✓	×	✓	×
ϕ_{high}, high	✓	✓	✓	✓	✓	✓

Table 3.5: Trigger criteria for candidates in different $m(\mu^+\mu^-)$ bins. Candidates must pass at least one trigger line at each stage of the trigger. The ϕ_{high} bin contains all of the candidates in the ϕ_{low} bin plus extra ones which pass the HLT1/2 dimuon lines (as well as one of the required L0 lines).

The L0 muon and dimuon lines search for high momentum tracks in the muon stations, see Table 3.6. Information from the VELO and other tracking stations is available for decisions made by the HLT1 trigger lines, see Table 3.7. Finally, an analysis specific HLT2 line builds $D^+_{(s)}$ candidates. Candidates in the high-

$m(\mu^+\mu^-)$ and ϕ_{high} bin are also accepted if they pass a generic HLT2 dimuon line. Table 3.8 shows the cuts in both HLT2 lines. The ϕ_{high} bin contains all of the candidates in the ϕ_{low} bin plus extra ones which pass the HLT1/2 dimuon lines (as well as one of the required L0 lines). $\pi^+\pi^+\pi^-$ candidates may pass any combination of LHCb's trigger lines, in order to maximise statistics.

Line	**Requirement**	**Proportion (%)**
Muon	$p_T > 1.48$ GeV/c	98
Dimuon	$p_{T,\mu_1} \cdot p_{T,\mu_2} > (1.30 \text{ GeV}/c)^2$	63

Table 3.6: Criteria of the L0 trigger lines. Candidates must pass at least one of the lines. The proportion of all final selection candidates passing these lines is shown.

Line	**Requirement**	**Proportion (%)**
Muon	$p > 8.0$ GeV/c $p_T > 1.0$ GeV/c IP > 100 μm IP $\chi^2 > 16$	96
Dimuon	$p_{T,\mu_1} > 0.5$ GeV/c $p_{T,\mu_2} > 0.5$ GeV/c $p_{\mu_1} > 6.0$ GeV/c $p_{\mu_2} > 6.0$ GeV/c IP $\chi^2 > 3$ Track $\chi^2_{ndf} < 4$ $m(\mu^+\mu^-) > 1.0$ GeV/c^2	81

Table 3.7: Criteria of the HLT1 trigger lines. Candidates from the low-$m(\mu^+\mu^-)$, η, ρ^0/ω and ϕ_{low} bins must pass the muon line. The proportion of high-$m(\mu^+\mu^-)$ and ϕ_{high} final selection candidates passing these lines is shown.

3.2.2 Stripping selection

$D^+_{(s)}$ candidates are built in an offline processing. The selection applied at this stage is shown in Table 3.9, with the retention rate of the stripping shown in Table 3.10. Both $\pi^+\mu^+\mu^-$ and $\pi^+\pi^+\pi^-$ candidates are required to pass the same selection. Having both data samples pass the same stripping selection is vital as it allows the pionic peaking backgrounds' shapes to be used in the final signal fit.

Line	Requirement	Proportion (%)
Dimuon	$p_{T,\mu_1} \cdot p_{T,\mu_2} > (1.5\ \text{GeV}/c)^2$ Track $\chi^2_{ndf} < 5$ $m(\mu^+\mu^-) > 1.0\ \text{GeV}/c^2$	88
$\pi^+\mu^+\mu^-$	$\sum p_{T,daughter} > 2.0\ \text{GeV}/c$ $\max(p_{T,daughter}) > 1.0\ \text{GeV}/c$ DOCA $< 200\ \mu$m $m(\pi^+\mu^+\mu^-) > 1.8\ \text{GeV}/c^2$ $m(\pi^+\mu^+\mu^-) < 2.1\ \text{GeV}/c^2$	85

Table 3.8: Criteria of the HLT2 trigger lines. Candidates from the low-$m(\mu^+\mu^-)$, η, ρ^0/ω and ϕ_{low} bins must pass the $\pi^+\mu^+\mu^-$ line. The proportion of high-$m(\mu^+\mu^-)$ and ϕ_{high} final selection candidates passing these lines is shown. The variable, DOCA, is defined in the next section.

The algorithm selecting $\pi^+\pi^+\pi^-$ candidates is pre-scaled by a factor one-hundred due to the channel's large branching fraction, shown in Table 3.14.

Object	Requirement
All tracks	Track $\chi^2_{ndf} < 8$ IP $\chi^2 > 4$ DOCA < 0.15 mm $\mathcal{P}(\text{ghost}) < 0.5$
π^+	$p > 2000\ \text{MeV}/c$ $p_T > 300\ \text{MeV}/c$
$\mu^\pm$	$p > 3000\ \text{MeV}/c$ $p_T > 500\ \text{MeV}/c$
$\mu^+\mu^-$	$m > 250\ \text{MeV}/c^2$
$\pi^+\mu^+\mu^-$	SV $\chi^2_{ndf} < 5$
D^+	$\cos(\theta_D) > 0.9999$ IP $\chi^2 < 25$ $m > 1763\ \text{MeV}/c^2$ $\lvert\Delta m\rvert < 200\ \text{MeV}/c^2$

Table 3.9: Cuts used in the stripping line, StrippingDXMuMuSS. $|\Delta m|$ is the difference between the candidate's invariant mass and the nominal mass of the D^+ [1].

The variables used in the stripping selection are defined below. The pointing angle (θ_D) is defined as the angular separation of the momentum and displacement vectors of the candidate, where the displacement vector is the displacement of the SV from the associated PV. The PV associated with a given candidate is that with

Line	Retention rate %
$D^+_{(s)} \to \pi^+\mu^+\mu^-$	0.051
$D^+_{(s)} \to \pi^+\pi^+\pi^-$	0.034
Total	0.109

Table 3.10: Retention rate for lines used in the analysis as tested on 10M events. The retention rate is the percentage of events that are triggered on by LHCb which pass the stripping line. The rate of $\pi^+\mu^+\mu^-$ and $\pi^+\pi^+\pi^-$ candidates is similar despite the pre-scales applied to the $\pi^+\pi^+\pi^-$ data sample.

respect to which the candidate has the smallest IP. Only $D^+_{(s)}$ candidates satisfying a maximum requirement on the IP χ^2 are accepted. The three daughter particles must make a good quality SV, so a maximum requirement is placed on the SV χ^2.

To remove ghost and clone tracks minimum requirements are placed on the track quality χ^2_{ndf}. Tracks originating from a PV are removed by a minimum requirement on the IP χ^2 of the tracks. To ensure that tracks come from a common origin a cut is placed on the distance of closest approach (DOCA). This is the maximum value of the minimum separation distance between any two of the charged tracks.

3.2.3 Final selection

A final selection, shown in Table 3.11, prepares the data for a signal extraction fit by carefully maximising signals and reducing contributions from combinatorial and physics backgrounds. The most powerful discriminating variable is a multivariate classification, a BDT, and is described in detail in the next section. Cutting upon this variable strongly reduces the combinatorial background component present in data.

Applying a tight muon particle identification requirement to the $\pi^+\mu^+\mu^-$ sample is found be important. The optimal working-point for the BDT and muon

$\mathrm{DLL}_{\mu-\pi}$ cuts are found by an algorithm that examines the expected upper limit on the signal branching fraction as a function of the cuts. A looser requirement, $\mathrm{DLL}_{\mu-\pi} > -2$, is applied to the 'muons' in the $\pi^+\pi^+\pi^-$ data sample to retain statistics. Applying different $\mathrm{DLL}_{\mu-\pi}$ cuts is not found to significantly alter the invariant mass distribution of candidates in the $\pi^+\pi^+\pi^-$ sample when reconstructed as $\pi^+\mu^+\mu^-$, see Fig. 3.6. However, statistics are low for $\mathrm{DLL}_{\mu-\pi} > 3$, so a systematic uncertainty is later added to account for any possible change.

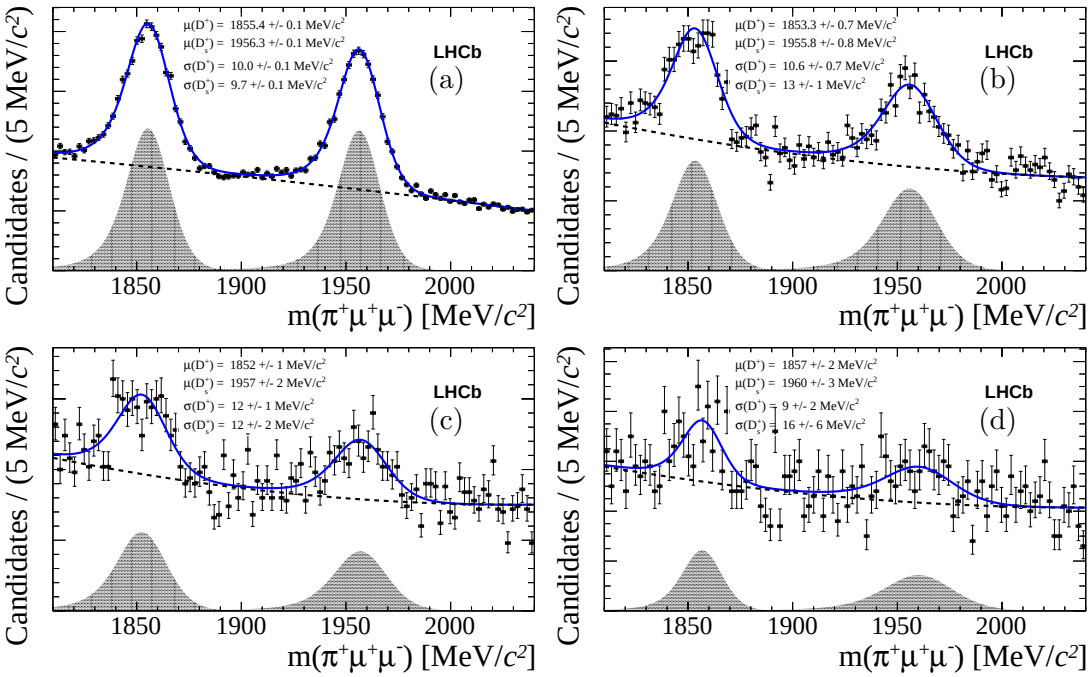

Figure 3.6: Invariant mass of preselected $\pi^+\pi^+\pi^-$ candidates reconstructed as $\pi^+\mu^+\mu^-$ with PID cuts on the pions designated muons: (a) $\mathrm{DLL}_{\mu-\pi} > -2$, (b) $\mathrm{DLL}_{\mu-\pi} > 1$, (c) $\mathrm{DLL}_{\mu-\pi} > 2$ and (d) $\mathrm{DLL}_{\mu-\pi} > 3$. $\mathrm{DLL}_{\mu-\pi} > -2$ is the cut applied in the final selection to $\pi^+\pi^+\pi^-$ candidates while $\mathrm{DLL}_{\mu-\pi} > 3$ is applied to $\pi^+\mu^+\mu^-$ candidates.

For events containing multiple $D^+_{(s)}$ candidates, the best candidate is chosen based on SV χ^2. This removes candidates which share tracks. Multiple candidates occur in approximately $\sim 1/10^3$ events, thus giving negligible biases for the 37,632 total candidates in the fit. The muons are forbidden from sharing hits with each other or any other tracks in the muon stations by requiring each to satisfy

$N_{shared} = 0$. The pion candidate is required to have $\mathrm{DLL}_{\mu-\pi} < 0$ and $\mathrm{DLL}_{\mathrm{K}-\pi} < 0$, the latter removing kaonic backgrounds, which are Cabibbo favoured D_s^+ decays. Distributions of the particle identification variables used in the final selection and N_{shared} for $\pi^+\mu^+\mu^-$ candidates are shown in Fig. 3.7.

Particle/Event	**Requirement**
Event	BDT>0.98
	$N_{candidates} = 1$
π^+	$\mathrm{DLL}_{\mu-\pi} < 0$
	$\mathrm{DLL}_{\mathrm{K}-\pi} < 0$
$\mu^\pm$	$N_{shared} = 0$
	$\mathrm{DLL}_{\mu-\pi} > 3$

Table 3.11: Criteria of the final selection for $\pi^+\mu^+\mu^-$ candidates. The same criteria apply for $\pi^+\pi^+\pi^-$ candidates, except for the $\mathrm{DLL}_{\mu-\pi}$ cut on the pions designated muons, which is loosened ($\mathrm{DLL}_{\mu-\pi} > -2$).

3.2.4 Multivariate analysis

Critical background suppression is achieved with a variable resulting from multivariate analysis. Multivariate analysis based on machine learning techniques is a powerful tool for separating signals within data. In high energy physics, it is extremely useful for separating physics signals from combinatorial background. Machine learning algorithms use experience gained on labelled examples, such as simulated signal decays, to generalise and respond to unseen inputs based on their similarity to the examples.

Decision tree (DT) learning is one such machine learning algorithm. The way a tree works is easily visualised, see Fig. 3.8. A series of binary decisions is undertaken splitting the data at each step (node). Events passing the condition go one-way, while those failing go the other. This continues until a stop-criteria, such as total number of splittings or the number of events remaining, is met.

There are two types of DT: classification and regression. Classification DTs

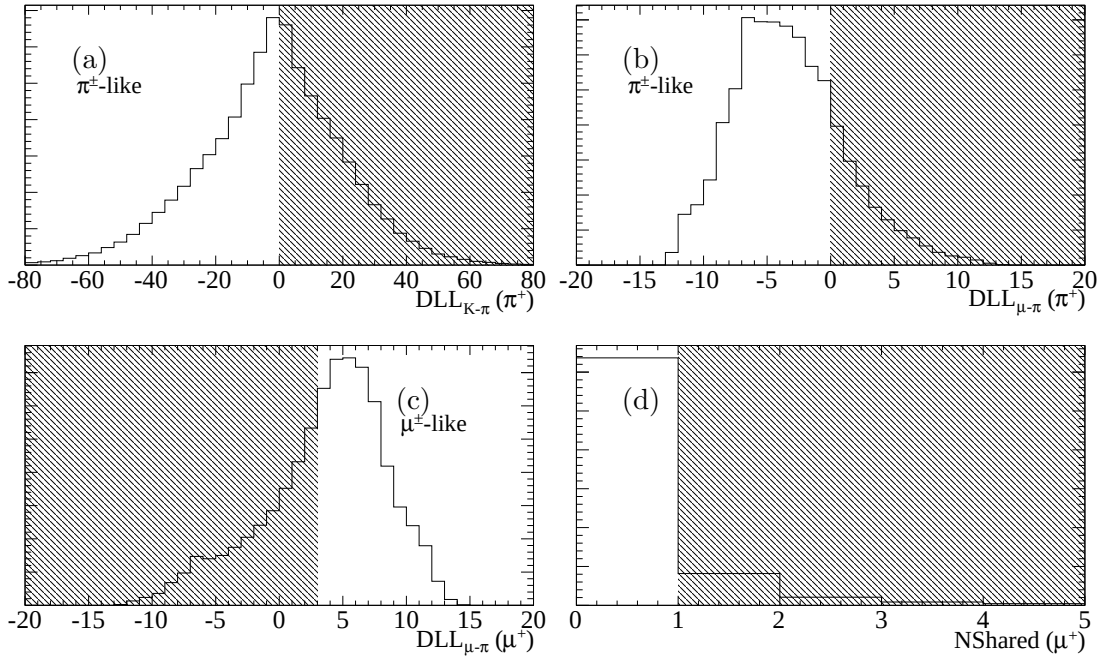

Figure 3.7: Data distributions of the PID variables and N_{shared} with all cuts applied except for PID and N_{shared} cuts. The distributions shown are (a) $\mathrm{DLL}_{\mathrm{K}-\pi}$ (π^+), (b) $\mathrm{DLL}_{\mu-\pi}$ (π^+) (c) $\mathrm{DLL}_{\mu-\pi}$ (μ^+) and (d) N_{shared}. The cuts placed on these distributions are shown by the shading (excluded regions) and in Table 3.11.

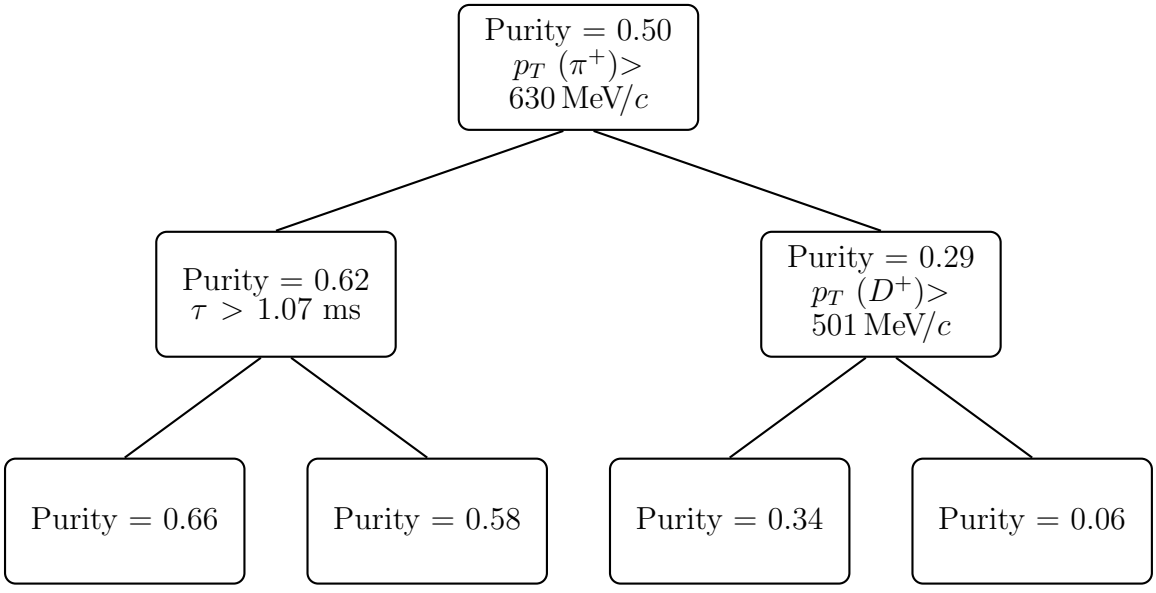

Figure 3.8: The 78th decision tree of the BDT. At each node the binary split with the smallest average squared error is applied to the data. The response at each node is related to the gradient of a function comparing the model response with the true values. Purity is defined as the number of signal events at the node to the total number of events at that node.

assign a class to events at the terminal node, where the class chosen is that of the majority of events at the node. For these DTs, the variable and cut used to split the data at each node is chosen based on the purity at the daughter nodes. Unlike classification DTs, regression DTs give a specific value to events at the terminal node. For regression DTs, the variable which produces the smallest average squared error is chosen to split the data at each node. The average squared error is defined by,

$$f(x) = \frac{1}{N}\sum_{i}^{N}(x_i - \hat{x})^2, \tag{3.1}$$

where x_i is the tree response for i-th event, $\hat{x}$ is the mean tree response over all events in the node, and N is the total number of events in the data sample at that node. For both types of DT, the same variable may be used for splitting at several nodes, while others might not be used at all.

A shortcoming of DTs is their instability with respect to statistical fluctuations in the training sample from which the tree structure is derived. Boosting is a technique that stabilises this response by extending the concept of a DT from a single tree to a forest. A number of boosting algorithms exist, of which the gradient boost (GradBoost) algorithm is one. The GradBoost algorithm uses regression DTs internally to minimise the loss function [65, 66]. The boosted decision tree (BDT) response is then based on the combined response of all trees within the forest.

A GradBoost BDT from the TMVA software package [67] is trained for this analysis with simulated $D^+ \to \pi^+\mu^+\mu^-$ decays to characterise the signal, and sidebands around the $D^+_{(s)}$ peaks in 2010 signal data to characterise the combinatorial background. A simulated signal sample was created in a 1:2 ratio under the 2011 (7 TeV) and 2012 (8 TeV) beam conditions to emulate the ratio in data, as previously shown in Table 3.4. Two categories of BDT are trained and used to

classify events with different categories of the variable, A_{p_T}.

3.2.5 BDT categories ($A_{p_T} \neq 1$, $A_{p_T} = 1$)

Information from the rest of the event is employed as a BDT input via an isolation variable, A_{p_T}, that considers the imbalance in the p_T of nearby tracks compared to that of the $D^+_{(s)}$ candidate,

$$A_{p_T} = \frac{p_T(D^+_{(s)}) - (\sum \vec{p})_{\mathrm{T}}}{p_T(D^+_{(s)}) + (\sum \vec{p})_{\mathrm{T}}}, \tag{3.2}$$

where $p_T(D^+_{(s)})$ is the p_T of the $D^+_{(s)}$ meson and $(\sum \vec{p})_{\mathrm{T}}$ is the transverse component of the vector sum momenta of all charged particles within a cone around the candidate, excluding the three signal tracks. The cone is defined by a circle of radius 1.5 in the plane of pseudorapidity and azimuthal angle, measured in radians around the $D^+_{(s)}$ candidate direction. Signal $D^+_{(s)}$ decays tends to be more isolated with a greater p_T asymmetry than combinatorial background. Two categories of BDT are trained. For events for which,

- $A_{p_T} \neq 1$, and,
- $A_{p_T} = 1$.

This is done because when background tracks do not appear in the cone ($A_{p_T} = 1$), the variable offers no discrimination. The distributions of this variable for simulation and data are shown in Fig. 3.9.

3.2.6 BDT performance

The configuration options for the BDT are recorded in Table 3.12. Half of the data available to the BDT is used for training, while the other half is used to validate its

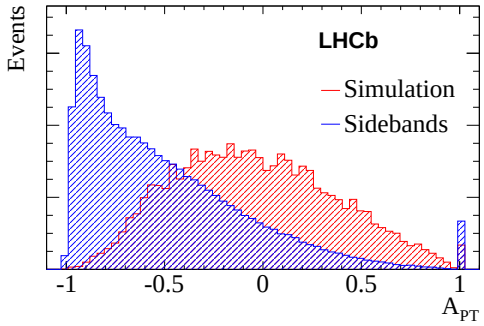

Figure 3.9: Distributions of the transverse momentum asymmetry of the event for simulated signal (red) and signal sideband (blue) candidates.

separation power and to test performance and for overtraining. The output of the BDT is shown in Fig. 3.10. The variables are ranked in order of their importance, see Tab. 3.13. The top-ranked variables are consistent with the idea that the main form of background is low momentum particles originating from the PV. Distributions of the input variables for simulated signal under the two different beam conditions are observed to very similar, see Figs. 1 to 3 in the Appendix. The differences in the distributions of these variables for the simulated signal and 2010 data sidebands can be seen in Figs. 4 to 6 in the Appendix and enable the BDT to discriminate between the signal and the combinatorial background. The flight distance χ^2 is the significance of the flight distance (magnitude of the displacement vector) in units of χ^2.

Option	**Value**	**Description**
Boost type	GradBoost	Boosting algorithm
Number of trees	1000	Number of trees in the forest
Shrinkage	0.1	Weight of response from individual trees
Number of cuts	20	Number of steps during cut optimisation
Maximum nodes	5	Maximum number of nodes per tree

Table 3.12: Configuration options chosen for the trained BDT. The default TMVA GradBoost configuration is used. This algorithm was found to produce the separation between signal and background. By applying this shrinkage value the BDT response is stabilised because each tree contributes less and more trees are grown.

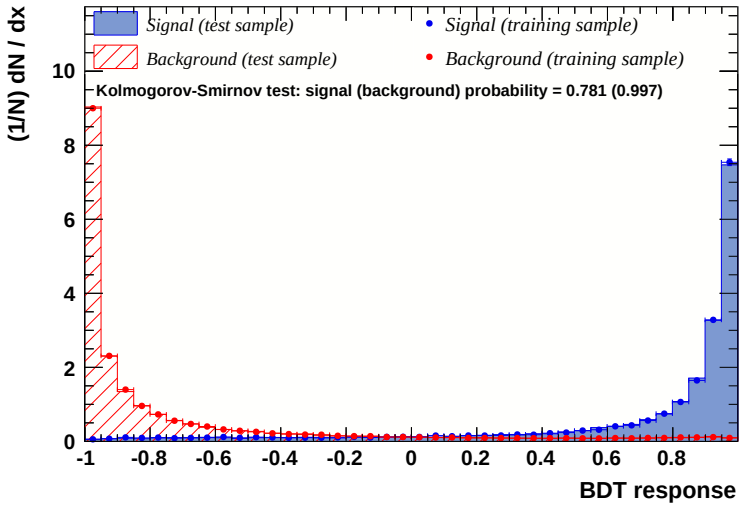

Figure 3.10: BDT response evaluated on the training (data points) and test (histograms) samples. Combinatorial background candidates have values peaked at -1, while simulated signal candidates peak at +1. The high-level of agreement visible, and confirmed by Kolmogorov-Smirnov tests (a measure of agreement between two distributions), indicate that the BDT has not been overtrained.

Events with $A_{p_T} \neq 1$		**Events with $A_{p_T} = 1$**	
Variable	**Importance** $[10^{-2}]$	**Variable**	**Importance** $[10^{-2}]$
$\cos(\theta_D)$	9.3	IP $\chi^2(\pi^+)$	10.2
IP $\chi^2(D^+)$	9.0	$p_T(D^+)$	9.7
EV $\chi^2(D^+)$	9.0	$\tau(D^+)$	9.3
A_{p_T}	8.3	FD $\chi^2(D^+)$	8.0
$p_T(\pi^+)$	7.3	EV $\chi^2(D^+)$	6.9
DOCA	7.0	IP $\chi^2(D^+)$	6.7
IP $\chi^2(\pi^+)$	6.4	min(IP $\chi^2(\mu^\pm)$)	6.7
$p(\pi^+)$	5.2	max(IP $\chi^2(\mu^\pm)$)	6.0
max($p(\mu^\pm)$)	4.7	min($p(\mu^\pm)$)	5.2
min($p(\mu^\pm)$)	4.5	DOCA	5.0
max($p_T(\mu^\pm)$)	4.3	$p_T(\pi^+)$	5.0
$\tau(D^+)$	4.0	max($p_T(\mu^\pm)$)	4.5
min($p_T(\mu^\pm)$)	3.9	max($p(\mu^\pm)$)	4.5
min(IP $\chi^2(\mu^\pm)$)	3.9	$p(D^+)$	4.0
max(IP $\chi^2(\mu^\pm)$)	3.7	min($p_T(\mu^\pm)$)	3.2
$p_T(D^+)$	3.6	$\cos(\theta_D)$	2.9
FD $\chi^2(D^+)$	3.1	$p(\pi^+)$	2.1
$p(D^+)$	2.8		

Table 3.13: BDT input variables in rank of importance for events with (left) $A_{p_T} \neq 1$ and (right) $A_{p_T} = 1$. The importance is equal to the number of splittings made on each variable weighted by the improvement in separation at the splitting.

3.3 Characterisation of the signal

Illustrative distributions of $D^+_{(s)}$ candidates can be plotted using the *sPlot* technique [68]. The technique is used to unfold data composed of a variety of sources, typically signal and background. The variables characterising the data are assumed to fall into one of two categories: discriminating and control variables. Discriminating variables are defined by the fact that their distributions are known *a priori* for all sources of events within the data. Alternatively, the distributions of control variables are unknown, or otherwise treated as such, for all sources of events within the data. Following a maximum likelihood fit to a discriminating variable of the data, x, a weighting is applied on an event-by-event basis. The weighting, known as an *sWeight*, has the form,

$$w = \frac{\sum_{j=1}^{N_s} V_{nj} f_j(x)}{\sum_{k=1}^{N_s} N_k f_k(x)}, \tag{3.3}$$

where N_k is the number of events of the species k, $f_i(x)$ is the PDF of the discriminating variable for the i-th species, n is the species to which the desired control variables belong and V_{nj} is the covariance matrix of the maximum likelihood fit to the data. The number of species in the data is N_s. With this *sWeight*, it can be shown that the distributions of control variables reproduce, on average, the true distributions [68].

For this analysis, the invariant mass, $m(\pi^+\mu^+\mu^-)$, is used as the discriminating variable to which the maximum likelihood fit is applied, see Fig. 3.11. Fig. 3.12 shows the distributions of the control variable $m(\mu^+\mu^-)$ for $D^+_{(s)}$ candidates. The LD contribution is clearly observed in the form of resonance structures at the expected meson masses. For D^+_s candidates, the proximity of the ρ^0 and ω meson masses and large width of the ρ^0 meson mass means that their individual contributions are not identifiable. Other decay variables, together with their simulated

signal distributions, are shown in Figs. 3.13 to 3.14. The signal and simulated signal distributions are in good agreement, validating the signal selection used in the analysis.

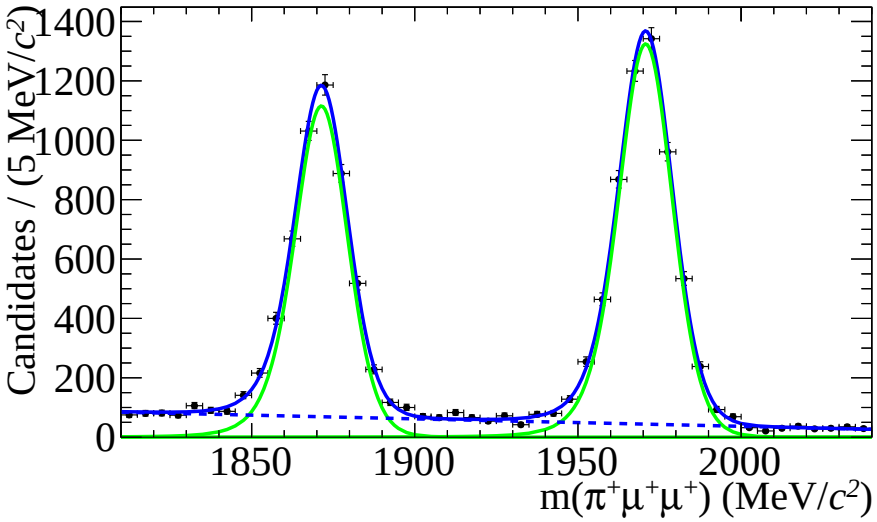

Figure 3.11: Invariant mass distribution for $D^+_{(s)} \to \pi^+\mu^+\mu^-$ candidates across the entire dimuon mass spectrum. The fitted model describing the signal (green line) and combinatorial background (dotted blue line), discussed in Sec. 3.5, are used to derive the *sWeight* for each candidate.

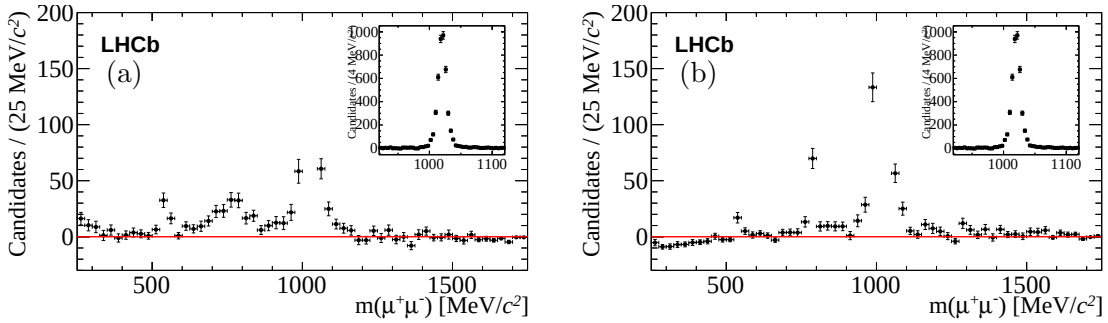

Figure 3.12: Background-subtracted $m(\mu^+\mu^-)$ spectrum of (a) $D^+ \to \pi^+\mu^+\mu^-$ and (b) $D^+_s \to \pi^+\mu^+\mu^-$ candidates that pass the final selection. The inset shows the ϕ contribution, and the main figure shows the η (550 MeV/c^2), ρ^0 (770 MeV/c^2) and, for the D^+_s, ω (782 MeV/c^2) contributions.

3.4 Peaking backgrounds

As well as the LD contribution to the signal decay, there are two peaking backgrounds which affect the analysis. These are residual misidentified $D^+_{(s)} \to \pi^+\pi^+\pi^-$

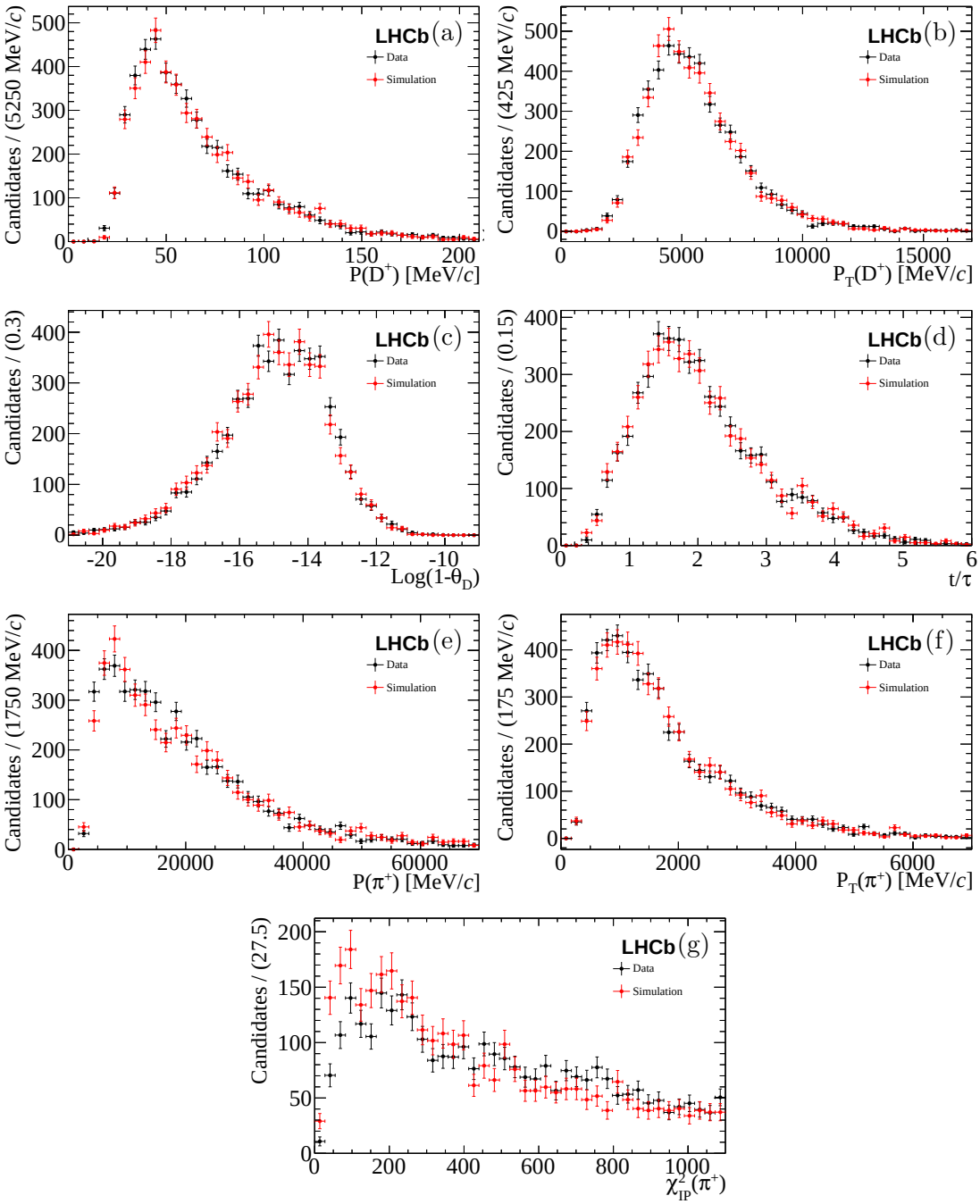

Figure 3.13: Distributions of simulated signal (red) and background-subtracted data (black) after the final selection. The distributions shown are (a) $p(D^+)$, (b) $p_T(D^+)$, (c) Log$(1-\theta_D)$, (d) time of flight divided by mean lifetime, (e) $p(\pi^+)$, (f) $p_T(\pi^+)$ and (g) π^+ impact parameter significance.

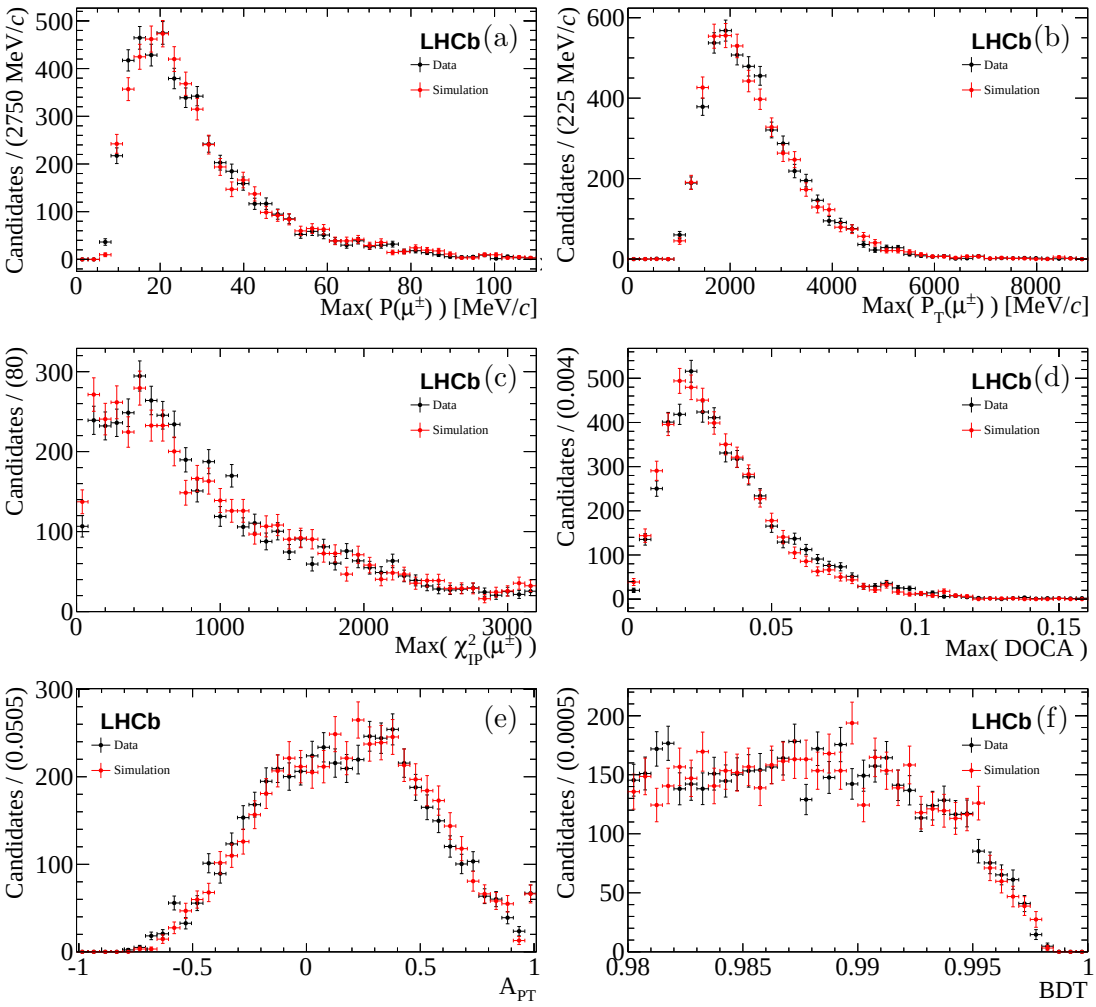

Figure 3.14: Distributions of simulated signal (red) and background-subtracted data (black) after the final selection. The distributions shown are (a) maximum $p(\mu^{\pm})$, (b) maximum $p_T(\mu^{\pm})$ and (c) maximum $\mu^{\pm}$ impact parameter significance, (d) the maximum distance of closest approach between any two of the final state particles, (e) the transverse momentum asymmetry of the event and (f) the BDT classifier response.

and partially-reconstructed $D_s^+ \to \pi^+(\eta \to \mu^+\mu^-\gamma)$ decays. Table 3.14 shows their branching fraction. For comparison, the branching fraction of the LD contributions are shown in Table 3.15 and Table 3.16. Other potential sources of background are investigated and found not to permeate the fitted range of the invariant mass distribution, see Fig. 3.15. Their contributions are therefore ignored. The contribution from $D_s^+ \to K^+\pi^+\pi^-$ decays, although partially filtering into the fit region, is not accommodated in the fit. The contribution from this decay is negligible because its branching fraction, $\mathcal{B} = (6.5 \pm 0.4) \times 10^{-3}$ [1], is half that of $D_s^+ \to \pi^+\pi^+\pi^-$ decays and it only partially falls within the fit region. The decay is additionally suppressed (by approximately a factor of ten) by the $\mathrm{DLL}_{\mathrm{K}-\pi}$ cut applied to π^+ candidates.

Decay	$\mathcal{B}$
$D^+ \to \pi^+\pi^+\pi^-$	$(3.18 \pm 0.18) \times 10^{-3}$
$D_s^+ \to \pi^+\pi^+\pi^-$	$(1.09 \pm 0.05) \times 10^{-2}$
$D_s^+ \to \pi^+(\eta \to \mu^+\mu^-\gamma)$	$(5.61 \pm 0.87) \times 10^{-6}$

Table 3.14: Branching fraction of decays appearing as peaking backgrounds in the $m(\pi^+\mu^+\mu^-)$ invariant mass distribution. [1]

X	$\mathcal{B}(D^+ \to \pi^+ X)$	$\mathcal{B}(X \to \mu^+\mu^-)$	$\mathcal{B}(D^+ \to \pi^+ (X \to \mu^+\mu^-))$
η	$(3.53 \pm 0.21) \times 10^{-3}$	$(5.8 \pm 0.8) \times 10^{-6}$	$(2.05 \pm 0.31) \times 10^{-8}$
ρ^0	$(8.1 \pm 1.5) \times 10^{-4}$	$(4.55 \pm 0.28) \times 10^{-5}$	$(3.69 \pm 0.72) \times 10^{-8}$
ω	$< 3.4 \times 10^{-4}$	$(9.0 \pm 3.1) \times 10^{-5}$	$< 3.1 \times 10^{-8}$
ϕ	$(5.42 \pm 0.19) \times 10^{-3}$	$(2.87 \pm 0.19) \times 10^{-4}$	$(1.56 \pm 0.12) \times 10^{-6}$

Table 3.15: Branching fraction of LD contributions to $D^+ \to \pi^+\mu^+\mu^-$ decays. [1]

X	$\mathcal{B}(D_s^+ \to \pi^+ X)$	$\mathcal{B}(X \to \mu^+\mu^-)$	$\mathcal{B}(D_s^+ \to \pi^+ (X \to \mu^+\mu^-))$
η	$(1.83 \pm 0.15) \times 10^{-2}$	$(5.8 \pm 0.8) \times 10^{-6}$	$(1.06 \pm 0.17) \times 10^{-7}$
ρ^0	$(2.0 \pm 1.2) \times 10^{-4}$	$(4.55 \pm 0.28) \times 10^{-5}$	$(9.10 \pm 5.49) \times 10^{-9}$
ω	$(2.5 \pm 0.7) \times 10^{-3}$	$(9.0 \pm 3.1) \times 10^{-5}$	$(2.25 \pm 1.00) \times 10^{-7}$
ϕ	$(4.5 \pm 0.4) \times 10^{-2}$	$(2.87 \pm 0.19) \times 10^{-4}$	$(1.29 \pm 0.14) \times 10^{-5}$

Table 3.16: Branching fraction of LD contributions to $D_s^+ \to \pi^+\mu^+\mu^-$ decays. [1]

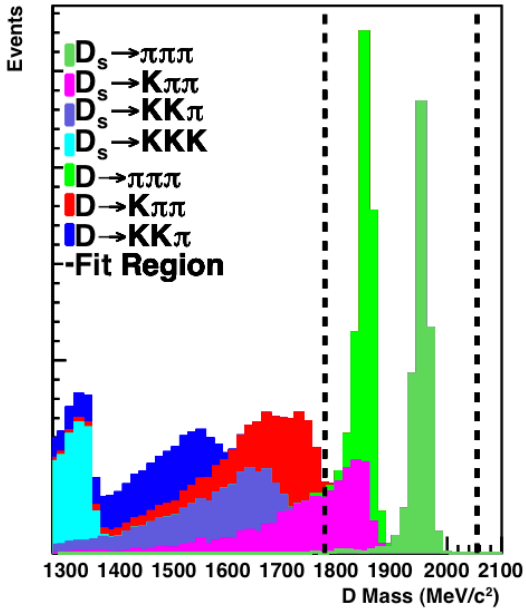

Figure 3.15: Hadronic $D^+_{(s)}$ decays reconstructed as $\pi^+\mu^+\mu^-$. The fit region is also shown. All histograms are normalised to unity. The actual size of the histogram describing $D^+_s \to K^+\pi^+\pi^-$ decays is reduced by a factor of twenty compared to that describing $D^+_s \to \pi^+\pi^+\pi^-$ decays because of a smaller branching fraction and additional $\mathrm{DLL}_{\mathrm{K}-\pi}$ suppression.

3.4.1 $D^+_{(s)} \to \pi^+\pi^+\pi^-$

$D^+_{(s)} \to \pi^+\pi^+\pi^-$ decays have branching fractions orders of magnitude larger than the LD contribution to the signal decays. In the invariant mass fit to signal data, PID cuts are placed on both muon candidates to suppress this background. This reduces the misidentification rate by a factor of approximately 10^{-4}, consistent with the reference percent-level misidentification rate per candidate given in Sec. 2.7.1. The shape of the pionic decays when reconstructed under the signal mass hypothesis, $m(\pi^+\mu^+\mu^-)$, are taken from the $\pi^+\pi^+\pi^-$ data sample described in Sec. 3.2. These shapes are displaced enough to allow the yield to be floated in the fit. The invariant mass distribution of $\pi^+\pi^+\pi^-$ candidates when reconstructed as $\pi^+\pi^+\pi^-$ and $\pi^+\mu^+\mu^-$ are shown in Fig. 3.16.

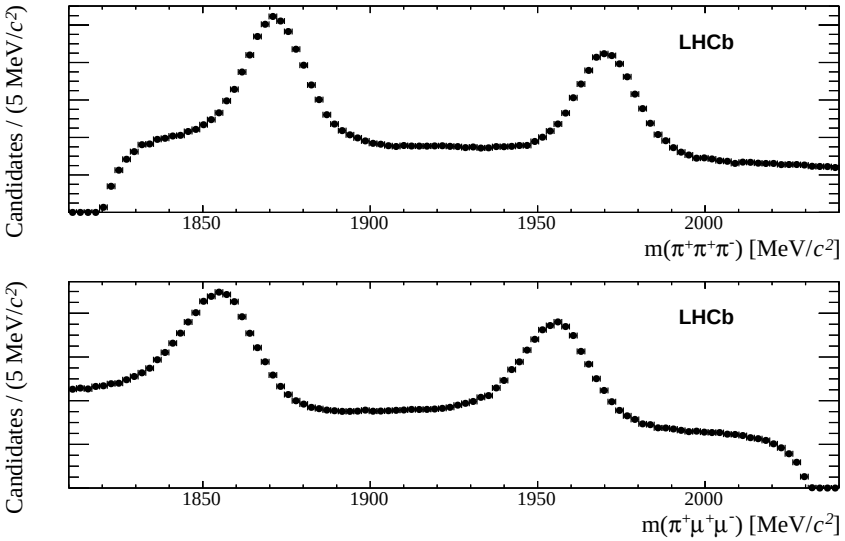

Figure 3.16: Invariant mass of $\pi^+\pi^+\pi^-$ candidates reconstructed as (top) $\pi^+\pi^+\pi^-$ and (bottom) $\pi^+\mu^+\mu^-$. The mean of the $D^+_{(s)}$ peaks are displaced by approximately 20 MeV/c^2 upon reassignment of the mass hypotheses.

3.4.2 $D^+_s \to \pi^+(\eta \to \mu^+\mu^-\gamma)$

$\eta \to \mu^+\mu^-\gamma$ decays have a branching fraction approximately 50 greater than that of $\eta \to \mu^+\mu^-$ decays, driven by the lifting of the helicity suppression present in the dimuon decay of this spin-0 meson. Though the majority of the background falls below the lower limit of the fit range, the contribution is still prominent enough to require a fit component in the low-$m(\mu^+\mu^-)$ bin. The small fraction that resides in the η bin ($\sim 1\%$), see Fig. 3.17, is negligible. A suitable PDF is taken from generator-level simulation, see Fig. 3.18, parameterised by the addition of a bifurcated Gaussian and an ARGUS function [69]. This fixed shape is then convoluted with the $D^+_s \to \pi^+\mu^+\mu^-$ mass peak PDF to account for detector resolution effects and the emission of final state radiation. The yield of this component floats in the fit.

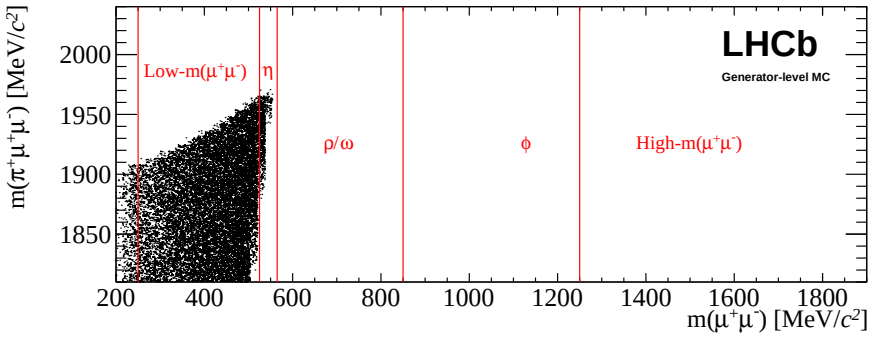

Figure 3.17: Distribution of $m(\mu^+\mu^-)$ versus $m(\pi^+\mu^+\mu^-)$ for partially-reconstructed generator $D_s^+ \to \pi^+(\eta \to \mu^+\mu^-\gamma)$ events. The dimuon invariant mass bins and their boundaries are shown in red.

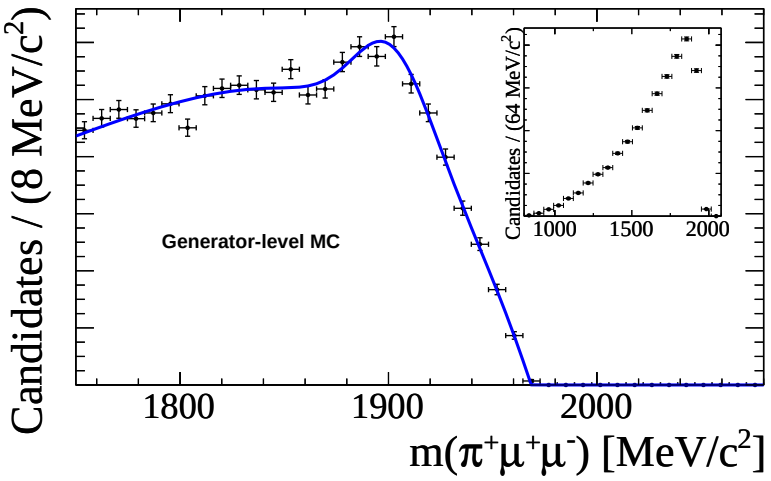

Figure 3.18: Partially-reconstructed $D_s^+ \to \pi^+(\eta \to \mu^+\mu^-\gamma)$ in arbitrary units. The main plot shows the distribution in a truncated $m(\pi^+\mu^+\mu^-)$ invariant mass range (slightly wider than that used in the fits to the signal data) and the insert plot shows the full distribution. The total PDF is shown by the solid blue line.

3.5 Fits to $D^+_{(s)} \to \pi^+\mu^+\mu^-$

The upper limits are extracted with a RooFit [70] binned maximum likelihood fit to the invariant mass distributions of $\pi^+\mu^+\mu^-$ candidates in the range $1810 - 2040$ MeV/c^2. This range is chosen to fully contain the PDFs of the correctly identified D^+ and D^+_s candidates as well as those of the kinematically similar $D^+_{(s)} \to \pi^+\pi^+\pi^-$ background.

To separate the SD and different LD contributions to $D^+_{(s)} \to \pi^+\mu^+\mu^-$ decays, the data are split into bins of dimuon invariant mass, $m(\mu^+\mu^-)$. The boundaries of each of these bins is given in Table 3.17.

Bin description	$m(\mu^+\mu^-)$ **range** [MeV/c^2]
low-$m(\mu^+\mu^-)$	250 – 525
η	525 – 565
ρ^0/ω	565 – 850
ϕ	850 – 1250
high-$m(\mu^+\mu^-)$	1250 – 2000

Table 3.17: Dimuon invariant mass bin boundaries.

The result is extracted from two independent fits due to the requirement of some of the software triggers that $m(\mu^+\mu^-)$ exceeds 1.0 GeV/c^2, as discussed in Sec. 3.2.1. A fit to signal candidates excluding these trigger lines is simultaneously performed in the low-$m(\mu^+\mu^-)$, η, ρ^0/ω and ϕ bins. Another fit, including these trigger lines, is performed in the high-$m(\mu^+\mu^-)$ and ϕ bins. The ϕ bin is included in both fits as it helps define the signal PDF and normalises the signal yield.

3.5.1 The maximum likelihood technique

Given a dataset, the maximum likelihood technique provides estimates for the parameters of a given model. The data are assumed to be independent and identically distributed. The data points, x_i of n entries, and the parameters of the

model, θ, are related via the probability density function (PDF),

$$f(x_1, x_2, ..., x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta). \tag{3.4}$$

This may be rewritten as a likelihood function by treating the data as fixed parameters of the model and allowing the parameters to vary freely,

$$L(\theta; x_i) = \prod_{i=1}^{n} f(x_i; \theta). \tag{3.5}$$

The maximum likelihood technique can be extended to include the PDF yield as a parameter. The extended likelihood convolutes the PDF with a poisson distribution, allowing the normalisation of the PDF to vary freely. The extended likelihood is,

$$L(\theta, N; x_i) = \frac{e^{-N}}{n!} N^n \prod_{i=1}^{n} f(x_i; \theta) = \frac{e^{-N}}{n!} \prod_{i=1}^{n} N f(x_i; \theta), \tag{3.6}$$

where N is the mean of the Poisson distribution describing the yield.

The values of the parameters which maximise the (extended) likelihood or, equivalently, minimise the negative (extended) log likelihood,

$$NLL = -\ln(L(\theta, N; x_i)) = N + \ln(n!) - \sum_{i=1}^{n} \ln(N f(x_i; \theta)), \tag{3.7}$$

provide estimates for the PDF's parameters. In some cases, such as for the yield, the estimators converge upon the true values,

$$\frac{d(NLL)}{dN} = 1 - \sum_{i=1}^{n} \frac{1}{N} = 1 - \frac{n}{N} = 0 \Leftrightarrow n = N. \tag{3.8}$$

In other cases, the estimators are at least consistent (bias disappears as n tends to infinity). The estimator for the parameter variance is the inverse of the second

derivative of the log-likelihood. Alternatively, this can be thought of as the change in the parameter that alters NLL by ± 0.5. In situations where more than one PDF is used to describe the data, the formula becomes,

$$NLL = -\ln(L(\theta, N_s; x_i)) = \sum_{s=1}^{n_s} N_s - \sum_{i=1}^{n} \ln(\sum_{s=1}^{n_s} N_s f_s(x_i; \theta)), \tag{3.9}$$

where n_s is the total number of PDF species and s is the index running over the species. The $\ln(n!)$ term has been dropped as it is constant and independent of fit parameters.

It becomes nontrivial to analytically find solutions to problems where the model contains many parameters (multiple PDFs) and the fit is performed over multiple dimensions (not just x_i). In such situations, numerical methods must be used. In this analysis, MINUIT [71] is used and two algorithms are employed:

- MIGRAD, minimises the NLL .
- HESSE, calculates the covariance matrix from the second derivative of the NLL .

A maximum likelihood fit is chosen rather than an alternative, such as a χ^2 fit, due to its robustness when fitting to low statistics. Its robustness is a consequence of the fact that, unlike the χ^2 fit, the maximum likelihood fit doesn't make incorrect Gaussian uncertainty assumptions for low statistics bins. The uncertainty in such bins is best described by a Poisson distribution. A binned fit is chosen to save CPU time. The only additional disadvantage of a binned fit is that it may miss features smaller than the bin width. Hence, a bin width of 5 MeV/c^2 is carefully chosen to avoid this. A disadvantage of all discussed methods is their bias, which is proportional to 1/N.

3.5.2 The simultaneous probability density function

The $D^+_{(s)} \to \pi^+\mu^+\mu^-$ signals are described by the function

$$f(x_i;\theta)_{sig} \propto \exp\left(\frac{-(x_i-\mu)^2}{2\sigma^2+(x_i-\mu)^2\alpha_{L,R}}\right), \tag{3.10}$$

which is a Gaussian-like peak, known as a Cruijff [72], of mean μ, width σ and where $\alpha_L(x < \mu)$ and $\alpha_R(x > \mu)$ parameterise the tails. The parameters of these shapes are determined simultaneously across all $m(\mu^+\mu^-)$ bins of a given fit. Effectively this means the signal peaks are defined by the ϕ bin.

The $D^+_{(s)} \to \pi^+\pi^+\pi^-$ peaking background data are also split into the predefined $m(\mu^+\mu^-)$ bins and fitted with Eq. 3.10. This provides a high-statistics, well-defined shape for this prominent background, which is imported into the corresponding $m(\mu^+\mu^-)$ bin of the signal fit. The misidentification rate (the ratio of the yield in the $\pi^+\mu^+\mu^-$ data sample to that in the $\pi^+\pi^+\pi^-$ sample) is allowed to vary but is assumed to be constant across all $m(\mu^+\mu^-)$ bins in the fit. This assumption is flawed so a systematic uncertainty is later added.

A component describing the partially-reconstructed $D^+_s \to \pi^+(\eta \to \mu^+\mu^-\gamma)$ peaking background is added to the the low-$m(\mu^+\mu^-)$ bin. The shape of this component is derived from simulation, see Sec. 3.4.2.

A Chebyshev series of order two is made of the first three Chebyshev polynomials. It has the form,

$$f(x_i;\theta)_{comb} = c_1(2x_i^2-1) + c_0 x_i + 1 \tag{3.11}$$

and is used to describe the PDF of all other combinatorial or partially reconstructed backgrounds that vary smoothly across the fit range. The coefficients of the series are permitted to vary independently in each $m(\mu^+\mu^-)$ bin. Chebyshev polynomials are chosen as the basis of the polynomial rather than a power series

basis as the coefficients of Chebyshev polynomials are not correlated. This makes the maximum likelihood fit more robust.

3.5.3 Fit results

Invariant mass spectra together with the fits overlaid are shown in Fig. 3.19 and Fig. 3.20 for the two independent fits. The total number of candidates are shown in Table 3.18. The fitted signal yields for the two fits are shown in Tables 3.19 and 3.20. All other fit parameters are shown in Tables 1 to 3 in the Appendix. Residuals for these fits are shown in Fig. 7 and Fig. 8 in the Appendix. The simultaneous fits to $D^+_{(s)} \to \pi^+\pi^+\pi^-$ mis-reconstructed as $D^+_{(s)} \to \pi^+\mu^+\mu^-$ are shown in Fig. 3.21 and Fig. 3.22.

The statistical significance of a signal peak can be calculated using Wilks' theorem [73],

$$\alpha = -2\ln\left(\frac{L_B}{L_{S+B}}\right) = 2\left(NLL_B - NLL_{S+B}\right). \tag{3.12}$$

Significances of 8.8 σ and 7.1 σ are found for $D^+ \to \pi^+(\eta \to \mu^+\mu^-)$ and $D^+_s \to \pi^+(\eta \to \mu^+\mu^-)$ decays, respectively. This is the first observation of these decay chains. Normalising to the branching fraction of $D^+_{(s)} \to \pi^+(\phi \to \mu^+\mu^-)$, it is calculated that,

$$\mathcal{B}(D^+ \to \pi^+(\eta \to \mu^+\mu^-)) = (2.3 \pm 0.5) \times 10^{-8}, \tag{3.13}$$

$$\mathcal{B}(D^+_s \to \pi^+(\eta \to \mu^+\mu^-)) = (6.7 \pm 1.8) \times 10^{-8}. \tag{3.14}$$

These branching fractions match those expected based on the $D^+_{(s)} \to \eta\pi^+$ and $\eta \to \mu^+\mu^-$ branching fractions [1]. No significant excess of candidates is seen in the low-$m(\mu^+\mu^-)$ or high-$m(\mu^+\mu^-)$ bins, hence there is no evidence for the SD contribution.

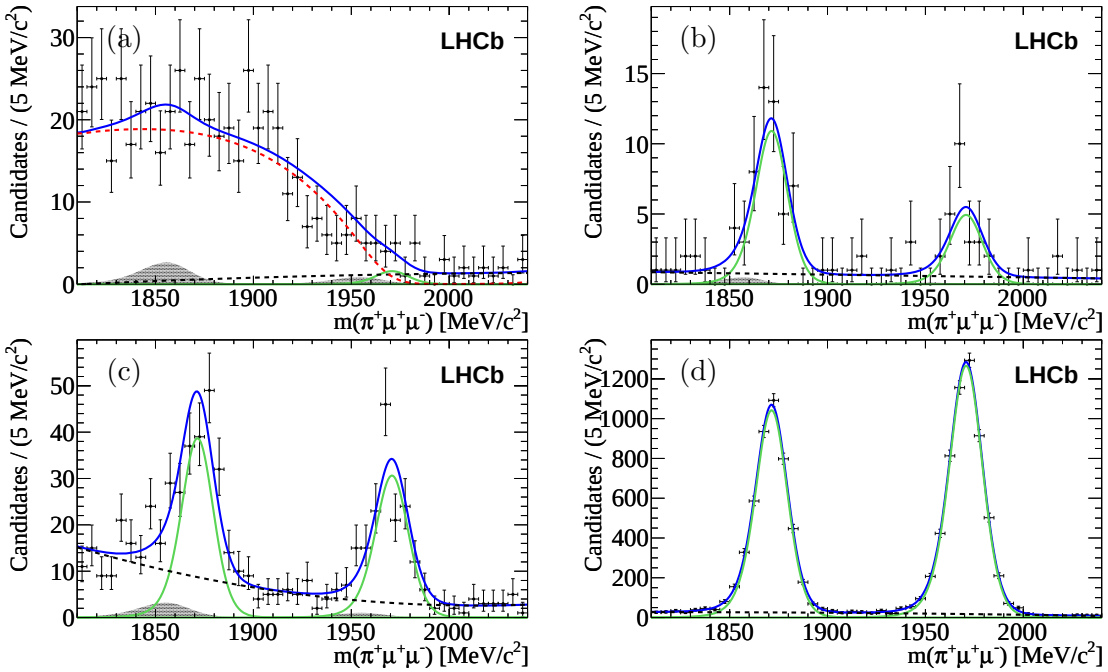

Figure 3.19: Invariant mass distributions for $\pi^+\mu^+\mu^-$ candidates in the lower mass fit (excluding trigger lines with $m(\mu^+\mu^-) > 1.0$ GeV/c^2). Shown are the (a) low-$m(\mu^+\mu^-)$, (b) η, (c) ρ^0/ω and (d) ϕ bins. The data are shown as points (black) and the total PDF (dark blue line) is overlaid. The components of the fit are also shown: the signal (solid green line), the non-peaking background (dashed black line) and the mis-reconstructed $D^+_{(s)} \to \pi^+\pi^+\pi^-$ peaking background (solid area), and, in the low-$m(\mu^+\mu^-)$ bin, the partially-reconstructed $D^+_s \to \pi^+(\eta \to \mu^+\mu^-\gamma)$ peaking background (dashed red line).

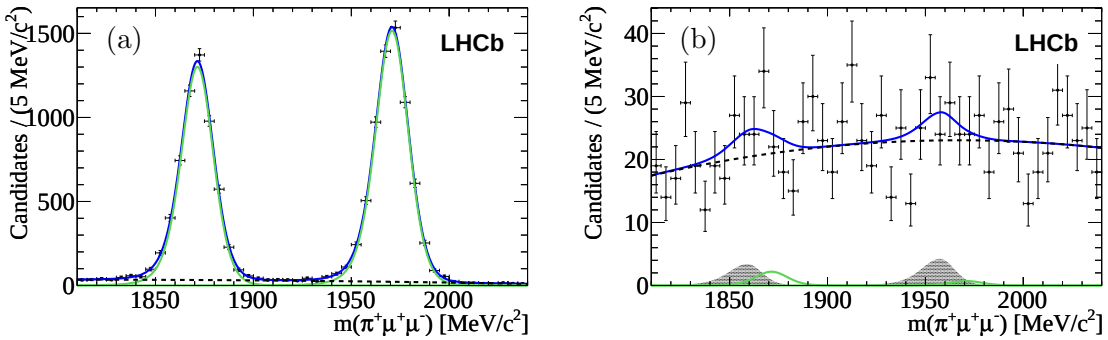

Figure 3.20: Invariant mass distributions for $\pi^+\mu^+\mu^-$ candidates in the higher mass fit (including trigger lines with $m(\mu^+\mu^-) > 1.0$ GeV/c^2). Shown are the (a) ϕ and (b) high-$m(\mu^+\mu^-)$ bins. The fit components are described in the caption of Fig. 3.19.

Fit	**Candidates**	**low-$m(\mu^+\mu^-)$**	η	ρ^0/ω	ϕ	**high-$m(\mu^+\mu^-)$**	**Total**
lower mass	$\pi^+\mu^+\mu^-$	537	101	621	11063	–	12322
	$\pi^+\pi^+\pi^-$	2372	402	2896	4476	–	10146
higher mass	$\pi^+\mu^+\mu^-$	–	–	–	13522	1044	14566
	$\pi^+\pi^+\pi^-$	–	–	–	4476	3918	8394

Table 3.18: The number of $\pi^+\mu^+\mu^-$ and $\pi^+\pi^+\pi^-$ candidates in the $m(\mu^+\mu^-)$ bins of the independent invariant mass fits.

Bin description	$m(\mu^+\mu^-)$ **range** [MeV/c^2]	D^+ **yield**	D^+_s **yield**
low-$m(\mu^+\mu^-)$	250 – 525	7 ± 6	6 ± 6
η	525 – 565	48 ± 10	22 ± 5
ρ^0/ω	565 – 850	166 ± 17	135 ± 14
ϕ	850 – 1250	4474 ± 84	5639 ± 96

Table 3.19: Signal yields for the lower mass invariant mass fit (excluding trigger lines with $m(\mu^+\mu^-) > 1.0$ GeV/c^2).

Bin description	$m(\mu^+\mu^-)$ **range** [MeV/c^2]	D^+ **yield**	D^+_s **yield**
ϕ	850 – 1250	5578 ± 60	6702 ± 58
high-$m(\mu^+\mu^-)$	1250 – 2000	9 ± 7	3 ± 3

Table 3.20: Signal yields for the higher mass invariant mass fit (including trigger lines with $m(\mu^+\mu^-) > 1.0$ GeV/c^2).

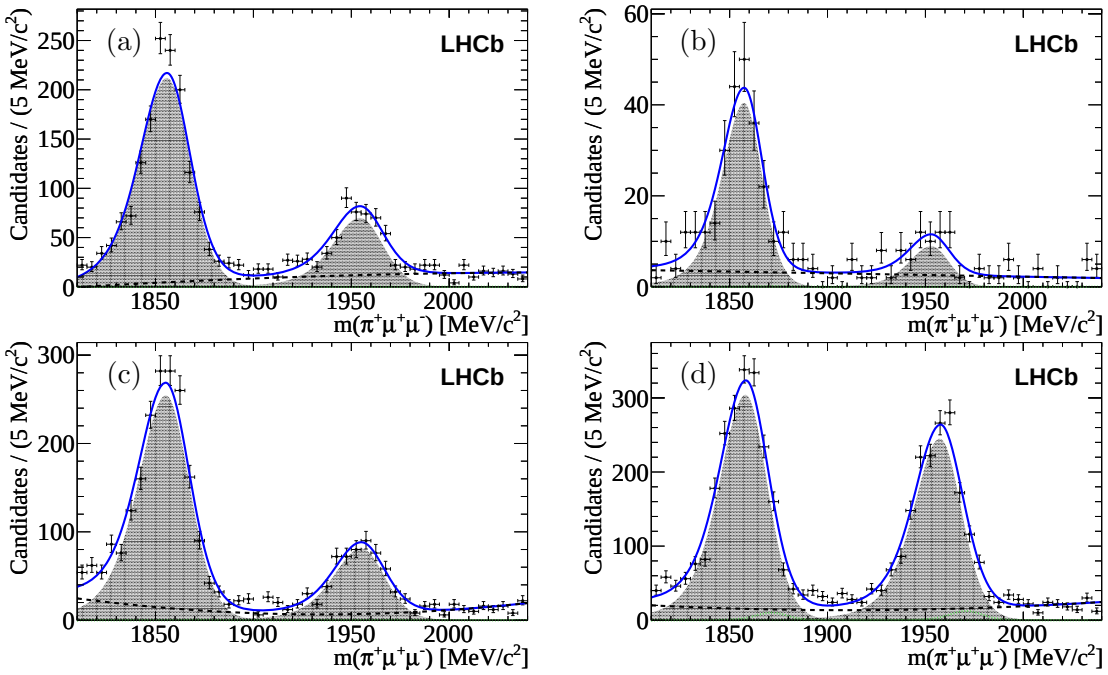

Figure 3.21: Invariant mass distributions for $\pi^+\pi^+\pi^-$ candidates reconstructed as $D^+_{(s)} \to \pi^+\mu^+\mu^-$ in the lower mass fit (excluding trigger lines with $m(\mu^+\mu^-) > 1.0$ GeV/c^2). Shown are the (a) low-$m(\mu^+\mu^-)$, (b) η, (c) ρ^0/ω and (d) ϕ bins. The fit components are described in the caption of Fig. 3.19.

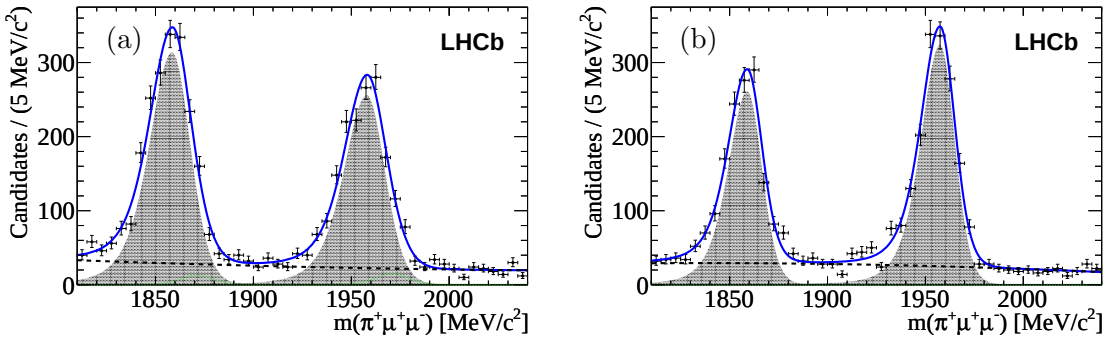

Figure 3.22: Invariant mass distributions for $\pi^+\pi^+\pi^-$ candidates reconstructed as $D^+_{(s)} \to \pi^+\mu^+\mu^-$ in the higher mass fit (including trigger lines with $m(\mu^+\mu^-) > 1.0$ GeV/c^2). Shown are the (a) ϕ and (b) high-$m(\mu^+\mu^-)$ bins. The fit components are described in the caption of Fig. 3.19.

3.6 Efficiency corrections

Branching fractions in the signal regions are calculated using the following relation,

$$\mathcal{B}(D^+_{(s)} \to \pi^+\mu^+\mu^-) = \frac{N_{signal}}{N_\phi} \cdot \frac{\mathcal{B}(D^+_{(s)} \to \pi^+(\phi \to \mu^+\mu^-))}{\mathcal{R}}, \tag{3.15}$$

where N_{signal} and N_ϕ are the yields in the signal and ϕ bins, respectively, and $\mathcal{R} = \frac{\epsilon_{signal}}{\epsilon_\phi}$. Branching fractions are calculated in both the low-$m(\mu^+\mu^-)$ and high-$m(\mu^+\mu^-)$ bins.

As the kinematics of the final state particles vary across the $m(\mu^+\mu^-)$ bins, cuts on variables at various stages of the analysis are found to affect $\mathcal{R}$. Simulated signal is used to quantify this at the following stages:

- **Acceptance**, the efficiency with which the final state particles lie within LHCb. This is calculated for each $m(\mu^+\mu^-)$ bin by generating events without any requirements for the final state particles to traverse LHCb. The ratio of these events to those traversing LHCb is found for each $m(\mu^+\mu^-)$ bin.

- **Reconstruction/Stripping**, the efficiency with which the decay is reconstructed and passes the stripping. Simulated signal is used to calculate the ratio of events generated within LHCb to those which are reconstructed and pass the stripping for each $m(\mu^+\mu^-)$ bin.

- **BDT and trigger cuts**, the efficiency with which the decay passes the BDT and trigger cuts. For each $m(\mu^+\mu^-)$ bin, the ratio of generated events before and after these cuts determines the efficiency.

Cuts on PID variables are not found to significantly alter $\mathcal{R}$ so are ignored.

The efficiency in each bin is shown in Table 4 and Table 5 in the Appendix for the 2011 and 2012 data, respectively. Efficiency ratios in each dimuon mass bin

with respect to the ϕ bin for the 2011 data, the 2012 data and their weighted mean are shown in Table 3.21. The weighted mean is illustrated in Fig. 3.23. Although the efficiency ratios are calculated using simulated $D^+ \to \pi^+\mu^+\mu^-$ decays, they are also used to calculate D_s^+ branching fractions, given the similar kinematics of the two decays. A systematic uncertainty is added to account for this.

Parameter	**Low-**$m(\mu^+\mu^-)$	η	ρ^0	**High-**$m(\mu^+\mu^-)$
2011	0.793 ± 0.047	0.835 ± 0.100	0.775 ± 0.038	1.140 ± 0.049
2012	0.792 ± 0.041	0.706 ± 0.075	0.812 ± 0.034	1.030 ± 0.036
Total	0.792 ± 0.032	0.749 ± 0.060	0.800 ± 0.026	1.067 ± 0.029

Table 3.21: Efficiency ratios with respect to the ϕ bin for 2011 and 2012 signal candidates. The total (weighted mean) is calculated by weighted the 2011 and 2012 values by the ratio of their absolute luminosities.

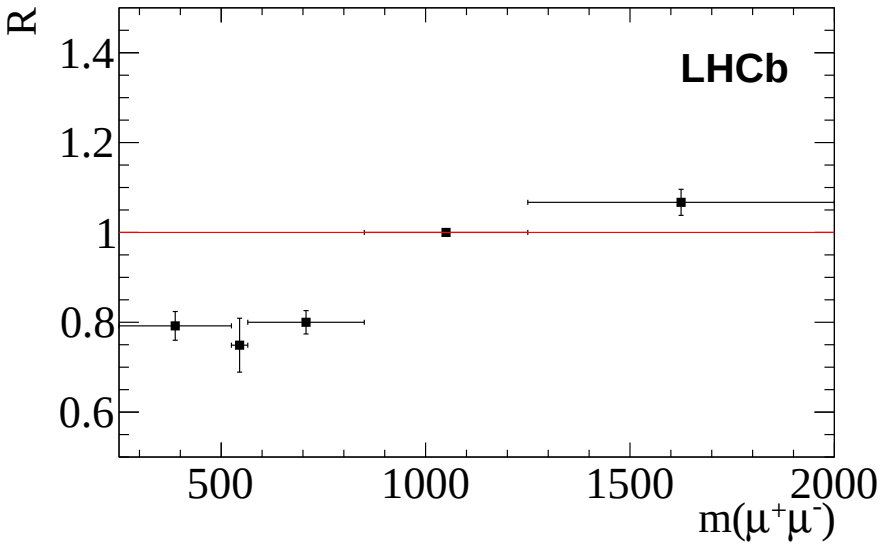

Figure 3.23: Weighted mean efficiency ratios with respect to the ϕ bin for 2011 and 2012 signal candidates.

3.7 Systematic uncertainty

Systematic uncertainties in the calculation of the signal branching fractions arise due to imperfect knowledge of the normalisation mode branching fraction, the

efficiency ratio, and the yield ratio. The systematic uncertainties on the normalisation mode branching fraction and the efficiency ratio are evaluated. The systematic uncertainties on the yield ratio are integrated directly into the limit setting process.

The greatest source of systematic uncertainty is from the normalisation mode branching fractions, $\mathcal{B}(D^+ \to \pi^+(\phi \to \mu^+\mu^-)) = (1.56\pm0.12)\times10^{-6}$ and $\mathcal{B}(D_s^+ \to \pi^+(\phi \to \mu^+\mu^-)) = (1.29\pm0.14)\times10^{-5}$. These translate to systematic uncertainties of 8.1% and 10.9% for the D^+ and D_s^+, respectively.

Systematic uncertainties affecting the efficiency ratio are introduced due to the limited size of the simulated sample and differences between the distributions of variables in data and in simulation. The efficiency ratio due to LHCb's geometric acceptance differs from unity due to the kinematics of the decay. The simulated signal is generated assuming a phase-space $m(\mu^+\mu^-)$ distribution. A flat $m(\mu^+\mu^-)$ distribution results in a efficiency ratio with a 1% difference [74]. This figure is taken as the systematic uncertainty on the acceptance. Variables used in the stripping, such as IP χ^2, are not perfectly reproduced by simulation, as seen in Fig. 3.13(g). Track properties are smeared to reproduce the distributions observed in $D_{(s)}^+ \to \pi^+(\phi \to \mu^+\mu^-)$ decays in data. The corresponding variation in the efficiency ratio indicates an uncertainty of 4% [74]. Comparing the efficiency of the optimal BDT cut on data and simulated signal, see Fig. 3.14(f), gives a conservative estimate of 5% for the resulting systematic uncertainty. The systematic uncertainty on the trigger criteria is evaluated by loosening the criteria. This produces a 3% systematic uncertainty. The systematic uncertainties discussed so far are summarised in Table 3.22. A statistical component due to the limited size of the simulated sample sizes must also be taken into account. As the size of the simulated sample differs in each region, the statistical uncertainty also differs, see Table 3.21, but is generally around 4%. This component

is added in quadrature to the total systematic uncertainty in Table 3.22 and shown in Table 3.23. An additional 5% uncertainty, the ratio of the invariant masses, is assigned for the D_s^+ as the efficiency ratio was calculated using simulated $D^+ \to \pi^+\mu^+\mu^-$ signal. The total systematic uncertainty on the normalisation factor, $\mathcal{B}\left(D^+_{(s)} \to \pi^+(\phi \to \mu^+\mu^-)\right)/\mathcal{R}$, is folded into the limit setting process described in the next section.

Source	**Uncertainty (%)**
$\mathcal{B}\left(D^+_{(s)} \to \pi^+(\phi \to \mu^+\mu^-)\right)$	8.1 (10.9)
Geometric Acceptance	1.0
Reco/Stripping	4.0
BDT	5.0
Trigger	3.0
TOTAL	10.8 ($13.0 \oplus 5$)

Table 3.22: Systematic uncertainties on the efficiency ratio. Values in parentheses are for the $D^+_{(s)}$.

Bin	$D^+ \to \pi^+\mu^+\mu^-$ (%)	$D_s^+ \to \pi^+\mu^+\mu^-$ (%)
Low-$m(\mu^+\mu^-)$	11.3	14.3
η	12.3	15.1
ρ^0	11.1	14.2
High-$m(\mu^+\mu^-)$	11.2	14.2

Table 3.23: Total systematic uncertainty on the normalisation factor in each region including the statistical component.

The final class of systematic uncertainties result from various assumptions made in the yield extraction. These systematic uncertainties do not affect the normalisation factor, only the measured yields of the control and signal modes. They are automatically taken in account in the limit setting procedure directly, as described in the next section.

3.8 Probabilistic interpretation

Signals observed in the low and high $m(\mu^+\mu^-)$ bins of the invariant mass fits are compatible with the null (background-only) hypothesis so limits are set on their branching fractions. A test statistic is defined to quantify evidence against the null hypothesis,

$$Q = \frac{L(\theta, B; x_i)}{L(\theta, S+B; x_i)}, \tag{3.16}$$

where L has previously been defined in Eq. 3.6. S and B are the fitted signal and background yields, respectively. A large value of the observed test-statistic Q_{obs} is evidence for rejecting the null hypothesis.

PDFs describing Q for the null and alternative (signal-plus-background) hypotheses are created using pseudo-experiments. A pseudo-experiment is data, generated based upon a fit to the observed data. As the null hypothesis is being tested against, the fit to the observed data is assumed to follow the null hypothesis, i.e. there is no signal component present. Sets of pseudo-experiments are produced for the null (zero signal candidates generated) and various alternative (predefined signal branching fraction) hypotheses. For each pseudo-experiment, the full maximum likelihood fit is run twice, once for a background-only model, and again for a model with a floating signal contribution. PDFs describing Q are then created for each hypothesis.

The confidence in a hypothesis, H, is given by $P(Q \leq Q_{obs}|H)$ and obtained by integrating the corresponding PDF. The probability that the null hypothesis could produce a value of Q equal to that observed, or less, is,

$$CL_B = P(Q \leq Q_{obs}|H_0) = \int_{-\infty}^{Q_{obs}} Q_B dQ_B. \tag{3.17}$$

Similarly, for an alternative hypothesis,

$$CL_{S+B} = P(Q \leq Q_{obs}|H_1) = \int_{-\infty}^{Q_{obs}} Q_{S+B} dQ_{S+B}. \tag{3.18}$$

These two values allow a decision regarding whether to reject the null hypothesis and to accept the alternative hypothesis. Individually though, these values make no allowance for experimental sensitivity. This becomes important when hypotheses are experimentally indistinguishable. If the observed data is (in)compatible with both hypotheses, this can lead to statements that are not meaningful. The CL_S method [75, 76] corrects for this by penalising the confidence in the alternative hypothesis by that in the null hypothesis,

$$CL_S = \frac{CL_{S+B}}{CL_B}. \tag{3.19}$$

The CL_S ratio is used to exclude signal branching fractions, as illustrated in Fig. 3.24(left). CL_S is also used to set an upper limit on the signal branching fraction. The upper limit is defined as the branching fraction which is excluded (1-CL_S) at 90% (also 95%). To find this, sets of pseudo-experiments are generated for multiple signal branching fractions. For each signal branching fraction, a PDF describing Q results, illustrated in Fig. 3.24(right). CL_S is calculated for each signal branching fraction by integrating the respective PDF from infinity to $-2\ln Q_{obs}$.

Plots of signal branching fraction versus CL_S are used to find excluded branching fractions, as illustrated in Fig. 3.25. The expected CL_S values are derived by calculating CL_S for each entry of the null hypothesis's PDF. This means integrating the null and alternative hypothesis's PDFs, starting from infinity and until each entry in the null hypothesis's PDF. These values of CL_S are then ordered in terms of size, and the median is the expected value for that signal branching fraction. The CL_S values $\pm 1/2\sigma$ from the median are also plotted as bands.

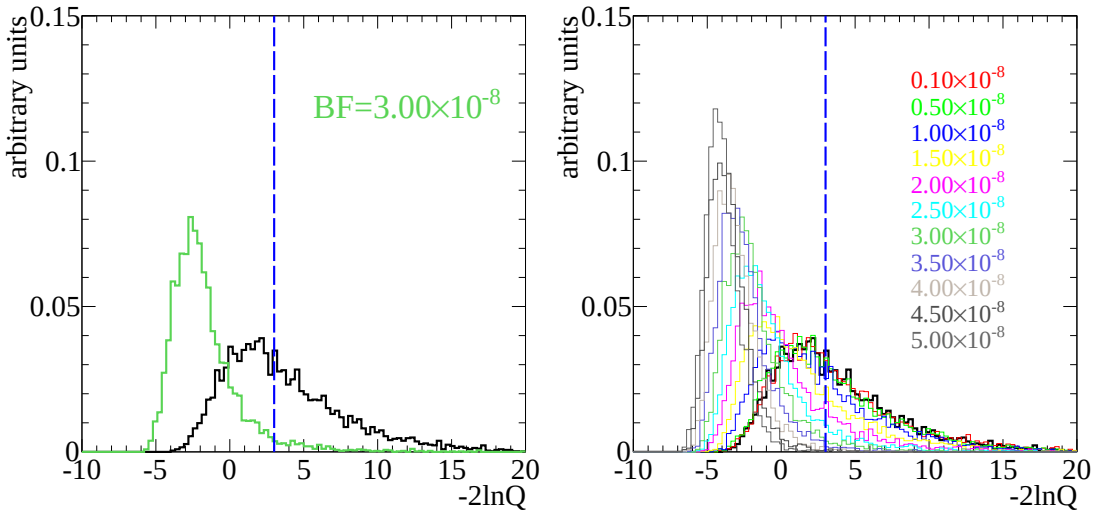

Figure 3.24: Distribution of $-2\ln Q$ for pseudo-experiments generated under the null (black) and (left) a single and (right) multiple alternative hypotheses. The dotted line indicates the value of $-2\ln Q_{obs}$. The ratio of the areas to the right of the line is CL_S for a given signal branching fraction, indicated by the legend. (Left) The green PDF has a CL_S value below 10%, so is excluded at 90% CL. The asymmetry of the distributions is a result of the asymmetry of the Poisson distribution from which the number of generated candidates are drawn.

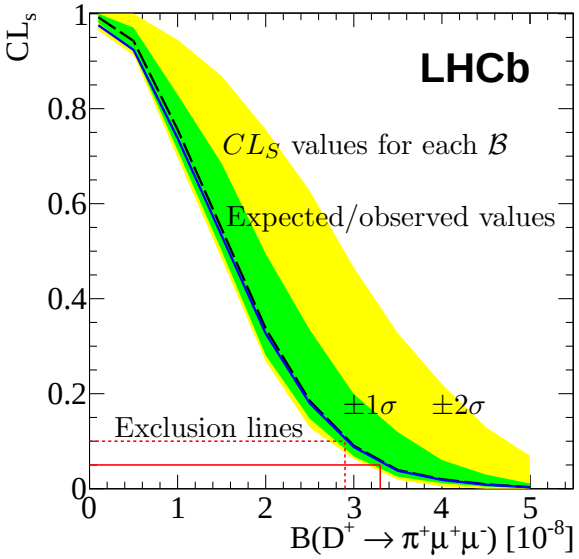

Figure 3.25: CL_S values as a function of signal branching fraction. The key features are indicated.

Upper limits are set on $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ and $\mathcal{B}(D_s^+ \to \pi^+\mu^+\mu^-)$ in the lower and upper $m(\mu^+\mu^-)$ bins individually. These are only partial branching fractions since they only cover a proportion of the total phase-space. Upper limits on the total SD branching fraction are calculated using the upper $m(\mu^+\mu^-)$ bin and assuming a flat distribution across three-body phase-space.

Systematic uncertainties are folded into the CL_S method by fluctuating the value of the parameters in the generation of pseudo-experiments. Parameters of all background and signal PDFs are varied by an amount dependent upon their uncertainty in the fit to the observed data. The normalisation factors are varied according to the figures in Table 3.23. Various assumptions upon which the fit is based also require a systematic uncertainty:

- **Signal PDF the same in all $m(\mu^+\mu^-)$ bins**, so the width of the generating PDF is varied.
- **Pion-to-muon misidentification rate the same in all $m(\mu^+\mu^-)$ bins**, so the generating rate is varied.
- $D_{(s)}^+ \to \pi^+\pi^+\pi^-$ **shape not altered by** $\mathrm{DLL}_{\mu-\pi}$ **cut**, so the width and mean of the generating PDF is varied.
- **Imperfect simulation-based** $D_s^+ \to \pi^+(\eta \to \mu^+\mu^-\gamma)$ **PDF**, so slope and power law of generating ARGUS function is varied.
- **Use of constants in the fit to increase stability of pseudo-experiments**, so generating constants varied.

Observed distributions of CL_S as a function of the total branching fraction for $D_{(s)}^+ \to \pi^+\mu^+\mu^-$ decays are shown in Fig. 3.26. Upper limits at 90% and 95% CL $(1 - CL_S)$ and p-values $(1 - CL_B)$ for the background-only hypothesis are shown in Table 3.24. The p-value represents the probability that the background

alone could produce the observed number of events, or more. The total branching fractions at the 90% (95%) confidence level (CL) are,

$$\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-) < 2.9\,(3.3) \times 10^{-8}, \tag{3.20}$$

$$\mathcal{B}(D_s^+ \to \pi^+\mu^+\mu^-) < 1.76\,(1.91) \times 10^{-7}. \tag{3.21}$$

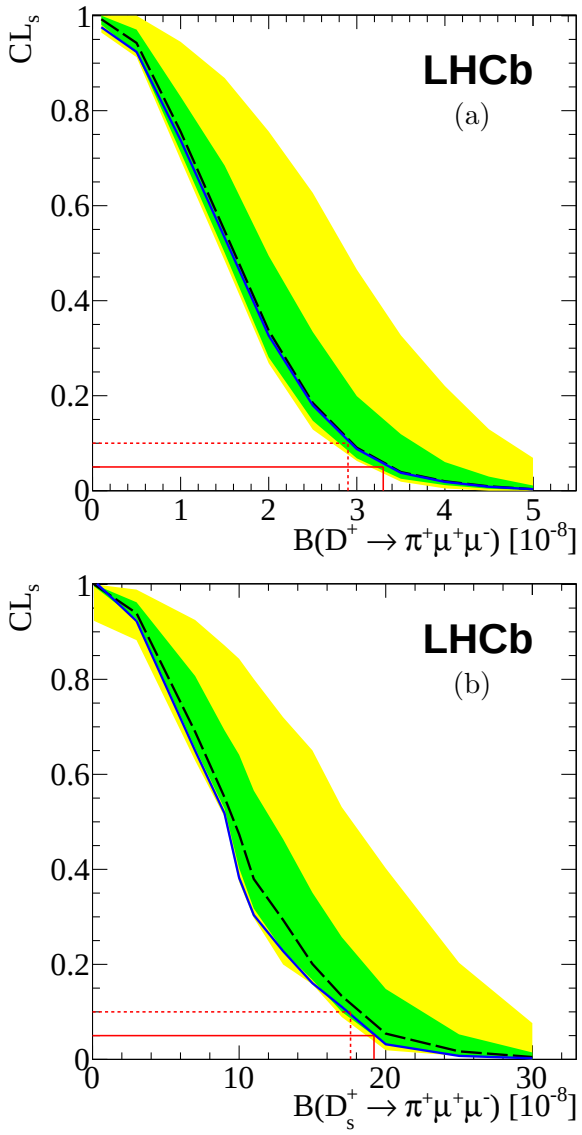

Figure 3.26: Observed (solid curve) and expected (dashed curve) CL_S values as a function of total (a) $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ and (b) $\mathcal{B}(D_s^+ \to \pi^+\mu^+\mu^-)$. The green (yellow) shaded area contains the $\pm1\sigma$ ($\pm2\sigma$) interval of possible results compatible with the expected value if only background is observed. The upper limits at the 90% (95%) CL are indicated by the dashed (solid) line.

Decay	Bin	90% [$\times 10^{-8}$]	95% [$\times 10^{-8}$]	p-value
$D^+ \to \pi^+\mu^+\mu^-$	low-$m(\mu^+\mu^-)$	1.5	1.7	0.57
	high-$m(\mu^+\mu^-)$	2.0	2.2	0.53
	Total	2.9	3.3	0.54
$D_s^+ \to \pi^+\mu^+\mu^-$	low-$m(\mu^+\mu^-)$	5.9	6.8	0.30
	high-$m(\mu^+\mu^-)$	8.3	9.0	0.86
	Total	17.6	19.1	0.80

Table 3.24: Upper limits in each $m(\mu^+\mu^-)$ bin and total branching fractions at the 90% and 95% CL and p-values for the background-only hypothesis.

3.9 Discussion

The limits presented here are the world's most stringent. They represent an improvement compared to those set by LHCb using 1 fb^{-1} of data, see Table 3.25, by approximately a factor of 2.5, similar to the expectation for a 3-fold increase in data. Given the upper limit on $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ presented here, it seems reasonable that when combined with the Run 2 data sample (expecting a further 5 fb^{-1} of data recorded at $\sqrt{s} = 13$ TeV) and smaller systematic uncertainties (such as the branching fraction of the control mode), the upper limit could probe below 10^{-8}. This remains significantly above the predicted SM branching fraction of a few $\times 10^{-9}$, leaving space for the discovery of possible NP enhancements.

Decay	$\mathcal{L}$ [fb^{-1}]	90% [$\times 10^{-8}$]	95% [$\times 10^{-8}$]
$D^+ \to \pi^+\mu^+\mu^-$	1	7.3	8.3
	3	2.9	3.3
$D_s^+ \to \pi^+\mu^+\mu^-$	1	41.0	47.7
	3	17.6	19.1

Table 3.25: Comparison of LHCb upper limits on the total branching fractions for SD contributions to $D^+_{(s)} \to \pi^+\mu^+\mu^-$ decays at the 90% and 95% CL.

3.10 Extending to rare kaonic decays

The strange quark spectates in the FCNC decay, $D_s^+ \to K^+\mu^+\mu^-$. This decay proceeds via analogous Feynman diagrams to those for $D^+ \to \pi^+\mu^+\mu^-$. $D_{(s)}^+ \to K^+\mu^+\mu^-$ decays can proceed by electroweak annihilation, but are Cabibbo suppressed in comparison to their pionic analogs. The most stringent limits on these decays at 90% CL are $\mathcal{B}(D_s^+ \to K^+\mu^+\mu^-) < 2.1 \times 10^{-5}$ and $\mathcal{B}(D^+ \to K^+\mu^+\mu^-) < 4.3 \times 10^{-6}$, set by the BaBar collaboration [77].

As well as SD contributions, there are LD contributions to the decay of the form $D_{(s)}^+ \to K^+(X \to \mu^+\mu^-)$, which are suppressed by at least a factor of ten compared to their pionic counterparts. The branching fraction of the LD contributions are shown in Table 3.26 and Table 3.27. None of the LD contributions to this final state have previously been observed from a $D_{(s)}^+$ decay.

X	$\mathcal{B}(D^+ \to K^+ X)$	$\mathcal{B}(X \to \mu^+\mu^-)$	$\mathcal{B}(D^+ \to K^+ (X \to \mu^+\mu^-))$
η	$(1.08 \pm 0.17) \times 10^{-3}$	$(5.8 \pm 0.8) \times 10^{-6}$	$(6.26 \pm 1.31) \times 10^{-9}$
ρ^0	$(2.0 \pm 0.5) \times 10^{-4}$	$(4.55 \pm 0.28) \times 10^{-5}$	$(9.10 \pm 2.34) \times 10^{-9}$

Table 3.26: Branching fraction of LD contributions to $D^+ \to K^+\mu^+\mu^-$ decays.

X	$\mathcal{B}(D_s^+ \to K^+ X)$	$\mathcal{B}(X \to \mu^+\mu^-)$	$\mathcal{B}(D_s^+ \to K^+ (X \to \mu^+\mu^-))$
η	$(1.75 \pm 0.35) \times 10^{-3}$	$(5.8 \pm 0.8) \times 10^{-6}$	$(1.02 \pm 0.25) \times 10^{-8}$
ρ^0	$(2.7 \pm 0.5) \times 10^{-3}$	$(4.55 \pm 0.28) \times 10^{-5}$	$(1.23 \pm 0.24) \times 10^{-7}$
ω	$< 2.4 \times 10^{-3}$	$(9.0 \pm 3.1) \times 10^{-5}$	$< 2.16 \times 10^{-7}$
ϕ	$(1.84 \pm 0.43) \times 10^{-4}$	$(2.87 \pm 0.19) \times 10^{-4}$	$(5.28 \pm 0.36) \times 10^{-8}$

Table 3.27: Branching fraction of LD contributions to $D_s^+ \to K^+\mu^+\mu^-$ decays.

Lepton number violating (LNV) processes such as $D_{(s)}^+ \to K^-\mu^+\mu^+$ are forbidden in the SM, because they may only occur through lepton mixing facilitated by a non-SM particle such as a Majorana neutrino. The most stringent limits on these decays are also set by the BaBar collaboration. At 90% CL, these limits are $\mathcal{B}(D_s^+ \to K^-\mu^+\mu^+) < 1.3 \times 10^{-5}$ and $\mathcal{B}(D^+ \to K^-\mu^+\mu^+) < 1.0 \times 10^{-5}$ [77].

An identical analysis selection (including the BDT) is performed on $D_{(s)}^+ \to$

$K^+\mu^+\mu^-$ and $D^+_{(s)} \to K^-\mu^+\mu^+$ candidates, albeit with a hadron PID criteria favouring charged kaons rather than pions ($\mathrm{DLL}_{\mathrm{K}-\pi} > 0$). The invariant mass spectrum for $K^-\mu^+\mu^+$ candidates is plotted in Fig. 3.27. A peak from the misidentified, Cabibbo-favoured, $D^+ \to K^-\pi^+\pi^+$ decay is clearly visible. There are no signs of any other signal. A two-dimensional plot showing $m(K^-\mu^+\mu^+)$ against $m(K^-\mu^+)$ is also shown. $D^+ \to K^-\pi^+\pi^+$ decays are seen to proceed via the intermediate resonance, $\bar{K}^*(892)^0$.

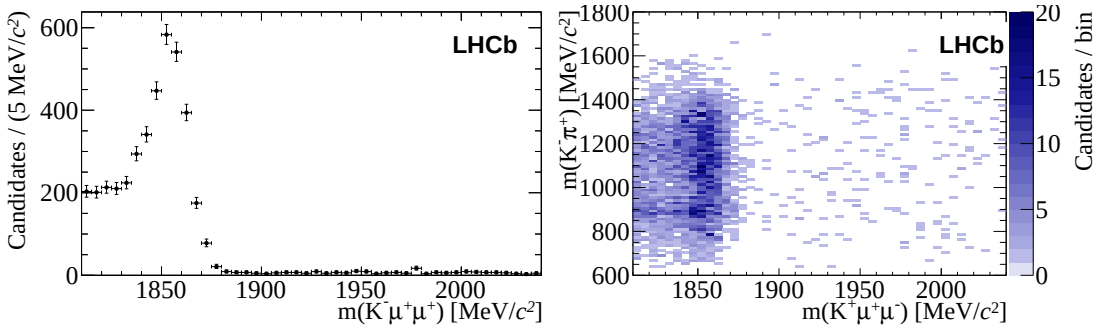

Figure 3.27: Invariant mass distributions for $D^+_{(s)} \to K^-\mu^+\mu^+$ candidates. The distributions are (left) m($K^-\mu^+\mu^+$) and (right) m($K^-\mu^+\mu^+$) versus m($K^-\mu^+$).

The invariant mass spectrum for $K^+\mu^+\mu^-$ candidates is plotted in Fig. 3.28. A two-dimensional plot showing $m(K^+\mu^+\mu^-)$ against $m(\mu^+\mu^-)$ is also shown. The LD contribution to $D^+_s \to K^+\mu^+\mu^-$ decays of the ρ^0 and ϕ resonances are visible. The ϕ resonance population located above the D^+_s mass (~ 2030 MeV/c^2) is due to misidentified $D^+ \to \pi^+\mu^+\mu^-$ decays.

Maximum likelihood fits to the invariant mass spectrum for $K^+\mu^+\mu^-$ candidates in $m(\mu^+\mu^-)$ bins around the ρ^0 and ϕ resonances are shown in Fig. 3.29. The PDFs describing $D^+_{(s)} \to \pi^+\mu^+\mu^-$ decays are reused and fixed to describe $D^+_{(s)} \to K^+\mu^+\mu^-$ decays. The PDFs describing $D^+_{(s)} \to \pi^+\pi^+\pi^-$ decays are reused and fixed to describe $D^+_{(s)} \to K^+\pi^+\pi^-$ decays in this quick study. Parameters describing the combinatorial background shape and all yields float in the fit. The fitted yields and their significances are shown in Table 3.28.

Bin description	$m(\mu^+\mu^-)$ **range** [MeV/c^2]	D^+ **yield**	D_s^+ **yield**	$\alpha(D^+)$	$\alpha(D_s^+)$
ρ^0/ω	565 − 850	49 ± 20	59 ± 21	2.6 σ	3.1 σ
ϕ	850 − 1200	13 ± 21	58 ± 26	0.2 σ	2.3 σ

Table 3.28: Signal yields and significances (α) for $D_s^+ \to K^+\mu^+\mu^-$ decays.

The significance of the $D_s^+ \to K^+(\rho^0 \to \mu^+\mu^-)$ peak is 3.1 σ, making this observation the first evidence for the decay chain. From Table 3.27, the branching fraction of this decay is approximately a factor of two larger than that of $D_s^+ \to K^+(\phi \to \mu^+\mu^-)$ decays. Although the observed yields for the two components are approximately equal, the large statistical uncertainty on both yields means that their ratio is consistent with this.

Given that the branching fraction of $D \to K^+(\rho^0 \to \mu^+\mu^-)$ is approximately a factor of ten smaller than its D_s^+ counterpart, see Tables 3.26 and 3.27, its large yield is surprising. Table 3.24 suggests that D^+ sensitivity is a factor six better than for D_s^+, rectifying this discrepancy to a large degree.

The observed branching fraction of $D_s^+ \to K^+(\rho^0 \to \mu^+\mu^-)$ is calculated using,

$$\mathcal{B}(D_s^+ \to K^+(\rho^0 \to \mu^+\mu^-)) = \frac{\mathcal{N}_\rho}{\mathcal{N}_\omega} \cdot \frac{\mathcal{B}(D_s^+ \to \pi^+(\omega \to \mu^+\mu^-))}{\mathcal{R}}, \tag{3.22}$$

where N_ρ and N_ω are the observed yields for $D_s^+ \to K^+(\rho^0 \to \mu^+\mu^-)$ and $D_s^+ \to \pi^+(\omega \to \mu^+\mu^-))$ decays, respectively. This normalisation mode is chosen as it shares the same $m(\mu^+\mu^-)$ bin, so $\mathcal{R}$ will be similar. These are shown in Tables 3.28 and 3.19, respectively. The normalising branching fraction is, $\mathcal{B}(D_s^+ \to \pi^+(\omega \to \mu^+\mu^-)) = (2.25 \pm 1.00) \times 10^{-7}$ and, $\mathcal{R}$, the ratio of efficiencies, simplifies to,

$$\mathcal{R} = \frac{\epsilon(\pi \to \pi, \mathrm{DLL_{K-\pi}} < 0)}{\epsilon(K \to K, \mathrm{DLL_{K-\pi}} > 0)} = \frac{0.90 \pm 0.05}{0.95 \pm 0.05}. \tag{3.23}$$

These values are given in Sec. 2.5.1 and an uncertainty of 5 % is assigned. The observed branching fraction is,

$$\mathcal{B}(D_s^+ \to K^+(\rho^0 \to \mu^+\mu^-)) = (1.04 \pm 0.61) \times 10^{-7}, \tag{3.24}$$

consistent with the expectation shown in Table 3.27.

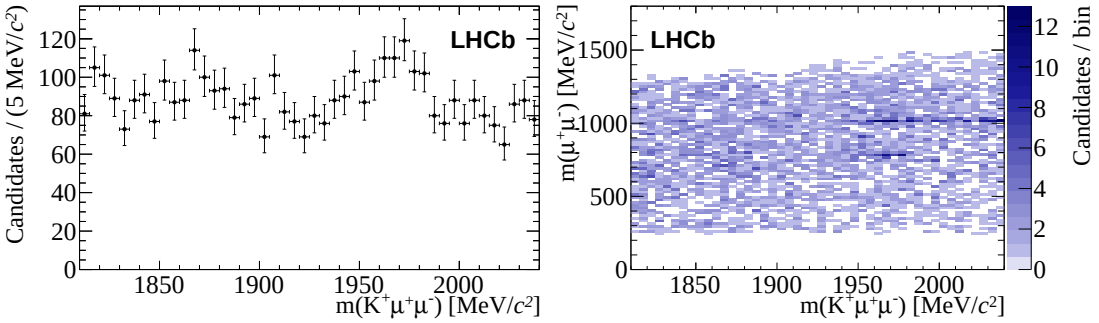

Figure 3.28: Invariant mass distributions for $D^+_{(s)} \to K^+\mu^+\mu^-$ candidates. The distributions are (left) m($K^+\mu^+\mu^-$) and (right) m($K^+\mu^+\mu^-$) versus m($\mu^+\mu^-$).

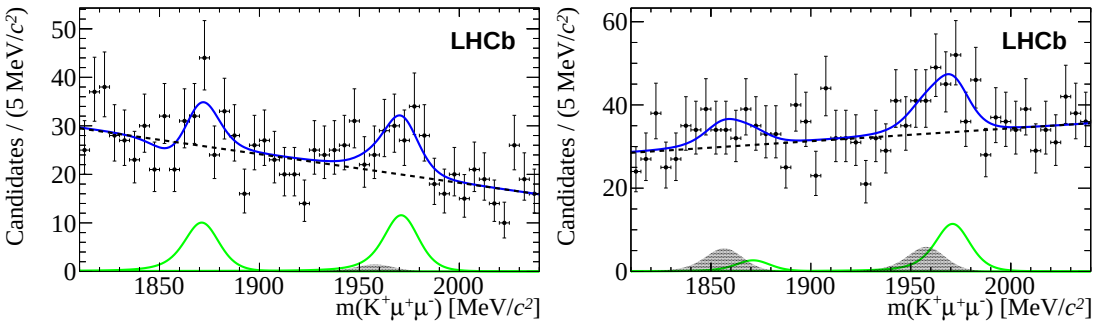

Figure 3.29: Invariant mass distributions for $D^+_{(s)} \to K^+\mu^+\mu^-$ candidates around (left) the ρ^0 resonance and (right) the ϕ resonance.

Chapter 4

The reconstruction of $D^+ \to \pi^+\pi^0$ at LHCb

A search for charge-parity asymmetry (A_{CP}) in $D^+ \to \pi^+\pi^0$ using $\pi^0 \to e^+e^-\gamma$ decays is performed using proton-proton collision data, corresponding to an integrated luminosity of 3.0 fb^{-1}, recorded by the LHCb experiment. This is the first time that this decay has been observed at a hadron collider. No evidence of CP violation is found.

4.1 Introduction

Although a challenging decay to reconstruct at the LHC due to the electrons and photon in the final state, $D^+ \to \pi^+(\pi^0 \to e^+e^-\gamma)$ decays can provide insight as a probe of direct CPV. This chapter describes a 'proof of principle' study which has not been published but its successful conclusion paves the way for a future analysis.

4.1.1 Phenomenological motivation

$D^+ \to \pi^+\pi^0$ decays can proceed via a number of Feynman diagrams, shown in Fig. 4.1. A_{CP} is expected to be even lower for this decay compared to other charm decays as it is forbidden from proceeding via the QCD loop diagram shown

in Fig. 4.2. The contribution vanishes because $\pi^+\pi^0$ is an $I = 2$ final state which cannot be reached from the initial state via the $\Delta I = \frac{1}{2}$ QCD loop operators. The decay may only proceed via the electroweak loop diagram, which is further suppressed with respect to the QCD loop diagram by a factor $\frac{\alpha_e}{\alpha_s}$. At the current level of precision, any deviation of A_{CP} from zero would thus be a sign of isospin violating NP. Note that, throughout, inclusion of conjugate processes is implied.

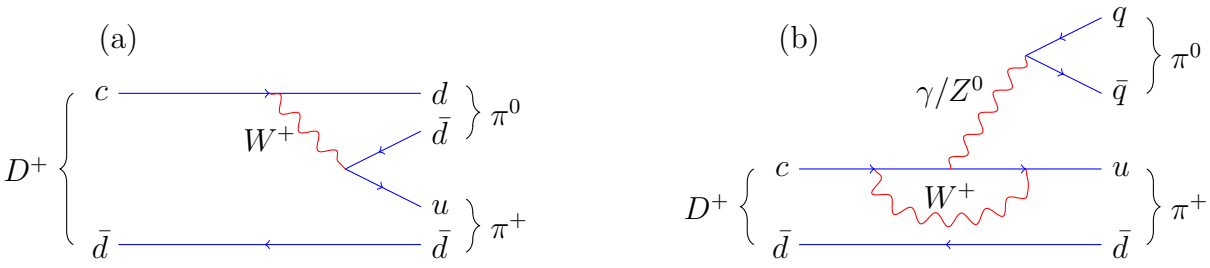

Figure 4.1: Feynman diagrams for (a) tree-level and (b) electroweak loop $D^+ \to \pi^+\pi^0$ decays (where q represents an up or down quark).

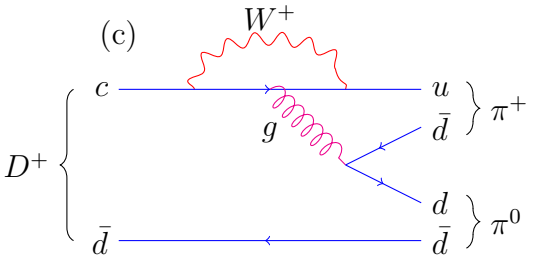

Figure 4.2: Feynman diagram showing the QCD loop $D^+ \to \pi^+\pi^0$ decay, which is forbidden due to isospin conservation.

4.1.2 Previous results

The decay $D^+ \to \pi^+\pi^0$ has only been observed at electron-positron (e^+e^-) colliders such as BaBar [78] and CLEO-c [79]. These experiments measured branching fractions of, $(1.25 \pm 0.10 \pm 0.09 \pm 0.04) \times 10^{-3}$ and $(1.18 \pm 0.03 \pm 0.05 \pm 0.03) \times 10^{-3}$, respectively. In order, the uncertainties are the statistical uncertainty, the systematic uncertainty, and the uncertainty on the branching fractions of normalisation modes. The beam corrected invariant mass distribution measured by CLEOc is

shown in Fig. 4.3. As the SM predicts, CLEOc found that the charge-parity asymmetry of the decay is consistent with zero, $A_{CP}(D^+ \to \pi^+\pi^0) = (2.9 \pm 2.9 \pm 0.3)\%$. The first uncertainty is statistical and the second systematic.

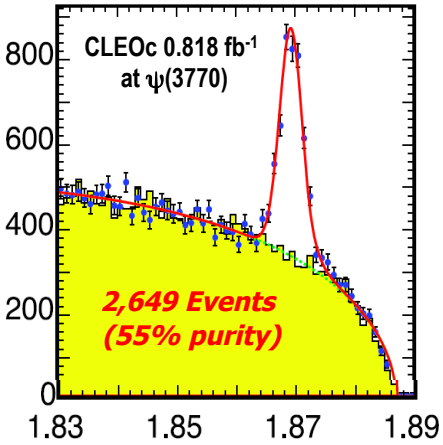

Figure 4.3: Beam-constrained mass $m_{bc} = \sqrt{E^2_{beam} - (\sum p_i)^2}$ in GeV/c^2 measured by CLEO-c. The data are shown as points (black) and the total PDF (red line) is overlaid. The combinatorial background component is also shown (shaded yellow area). Each histogram bin has a width of 0.001 GeV/c^2. To date, this is the largest sample of the decays collected. The estimated purity quoted is defined as the fraction of the number of signal events to background events within the signal region. Taken from [79].

4.1.3 π^0 reconstruction

The π^0 was the first resonant particle reconstructed by LHCb. This occurred within hours of data, using $\pi^0 \to \gamma\gamma$ decays and summing energy deposited in the ECAL. Its quick rediscovery happened because there was no need for tracking of charged particles, so no alignment was needed. The π^0 has a mean lifetime of the order 10^{-17} s [1], so decays effectively immediately. Hence, the analysis searches for a common decay vertex of the combined π^+ and π^0 decay products. By far the most prominent decay of the π^0 is to the diphoton final state, with $\mathcal{B}$ $(\pi^0 \to \gamma\gamma)$ $= (98.823 \pm 0.034)\%$ [1]. A high statistics plot of the invariant mass of diphotons is shown in Fig. 4.4. Reconstruction of this decay mode is not feasible in this analysis as it would be impossible to determine the decay vertex of the D^+ using

only the π^+, as photons, being non-ionising particles, cannot be tracked through LHCb. The decay $\pi^0 \to e^+e^-\gamma$, although substantially rarer, *c.f.* $\mathcal{B}(\pi^0 \to e^+e^-\gamma) = (1.174 \pm 0.035)\%$ [1], enables the analysis as electrons can be tracked through LHCb.

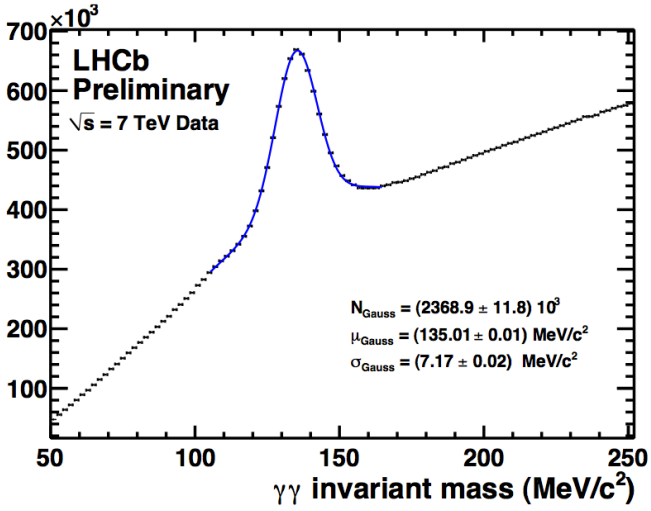

Figure 4.4: Invariant mass of diphoton pairs reconstructed in the 2010 data. The parameters of the π^0 mass fit are shown on the plot, the resolution is seen to be 7.2 MeV/c^2. Taken from [80].

4.2 Event selection

The analysis is performed using proton-proton collision data taken at two different centre of mass energies. Table 4.1 shows this. Simulation data was generated in order to guide the analysis. Details of the simulation are given in Table 4.2. The simulation can be used to measure the efficiency with which LHCb measures the decay. The results of which are shown in Table 4.3. Although approximately 2M events were generated with LHCb's acceptance, less than 100 events pass the selection of the analysis. This reduction in events is primarily a result of the trigger.

Year	$\sqrt{s}$ [TeV]	**Polarity**	$\mathcal{L}$ [fb^{-1}]
2011	7	Up	0.44
2011	7	Down	0.58
2012	8	Up	1.03
2012	8	Down	1.03

Table 4.1: Data used in the analysis.

Beam conditions	**Events per polarity**
2011	0.5 M
2012	0.5 M

Table 4.2: Simulated $D^+ \to \pi^+(\pi^0 \to e^+e^-\gamma)$ decays produced for this analysis. Both magnetic polarities are modelled.

Stage	**Efficiency (%)**
Reconstruction/Stripping	2.6×10^{-1}
Trigger cuts	5.1×10^{-3}
π^0 cut	4.1×10^{-3}

Table 4.3: Cumulative efficiencies as a percentage, measured on simulated signal generated within LHCb's acceptance. Details of the trigger and π^0 cuts are provided in Sec. 4.2.1 and Sec. 4.2.3, respectively.

4.2.1 Trigger selection

This analysis started after LHCb's data-taking period had ended. Due to this the analysis lacks a dedicated trigger line. Requiring the signal to pass specific lines within the trigger is found to be useful as it significantly reduces background levels. The efficiency of LHCb's trigger is estimated from simulation. The simulated signal is run through an emulated trigger to gauge which trigger lines it passes, see Table 4.4. The decay is required to pass the following L0, HLT1 and HLT2 criteria as trigger on signal (TOS) or trigger independently of signal (TIS):

- L0: TOS on Hadron, Electron or Photon; or Global TIS.
- HLT1: TOS on TrackAllL0 or TrackPhoton.
- HLT2: TOS on CharmHadD2HHH, CharmHadD02HH_D02KK or CharmHadD02HH_D02KPi.

Level	**Name**	**Passing** %
L0	Hadron	8
	Electron	7
	Photon	2
HLT1	TrackAllL0	56
	TrackPhoton	9
HLT2	CharmHadD2HHH	9
	CharmHadD02HH_D02KK	4
	CharmHadD02HH_D02KPi	3

Table 4.4: The subset of simulated signal candidates passing each line as TOS, given that it has passed at least one of the listed lines, as TOS, at the previous level. For example, 56% of events that trigger one of the L0 lines also trigger TrackAllL0. The L0 line rates are shown as a percentage of the number of truth-matched simulated siganl candidates which are reconstructed and pass the stripping line.

The HLT2 lines which fire are two and three body hadronic lines. The three body lines include all combinations of three charged hadrons, where each hadron can either be a $K^{\pm}$ or $\pi^{\pm}$. The final state particles which cause the trigger to fire

are the π^+ and the e^- for the two body triggers, plus the e^+ for the three body triggers. Fig. 4.5 demonstrates that the signal, when misidentified as the trigger decay, still falls within the mass window of the trigger selection. This is because although energy is removed from the system due to the omission of final state particles, such as the γ, this effect is counteracted by giving misidentified particle(s) greater mass(es). The invariant mass cuts shown in Fig. 4.5 demonstrate why a dedicated HLT2 trigger from 2015 can increase the efficiency of the trigger at selecting signal events.

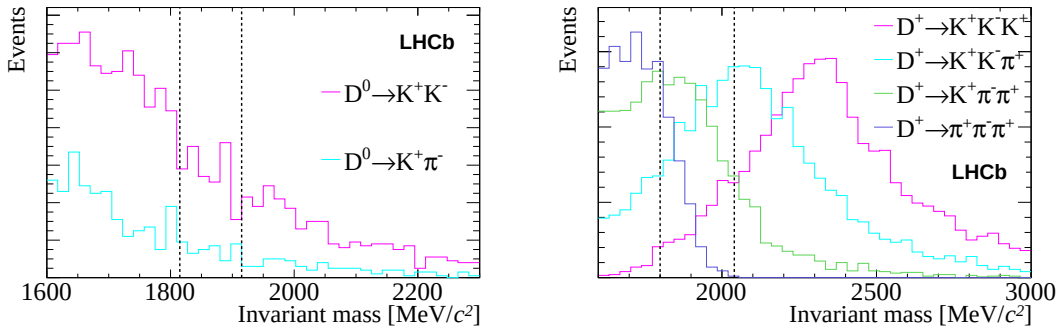

Figure 4.5: Invariant mass distributions of the misidentified π^+ and e^- (for two body triggers) and e^+ (for three body triggers) under various mass hypotheses. The window between the dotted lines shows the invariant mass cuts applied in the respective triggers.

The result of only requiring this trigger selection can be seen in Fig. 4.6. The trigger efficiencies are low at each stage. For the L0 trigger, this is expected due to the relatively low p_T of the tracks. For example, the L0ElectronTOS efficiency for B mesons decaying to electrons is an order of magnitude higher [81]. The efficiency at the HLT2 stage is low due to the lack of a dedicated trigger.

The effect of applying these trigger cuts can be seen in the plots in Figs. 4.7 and 4.8. One observes that the simulated signal is shifted to higher energies and momenta, except for the momentum and energy of the photon. The trigger applies these cuts as light quark backgrounds tend to have lower momenta. The γ is not triggered upon, so its decrease in energy and momentum is because the other

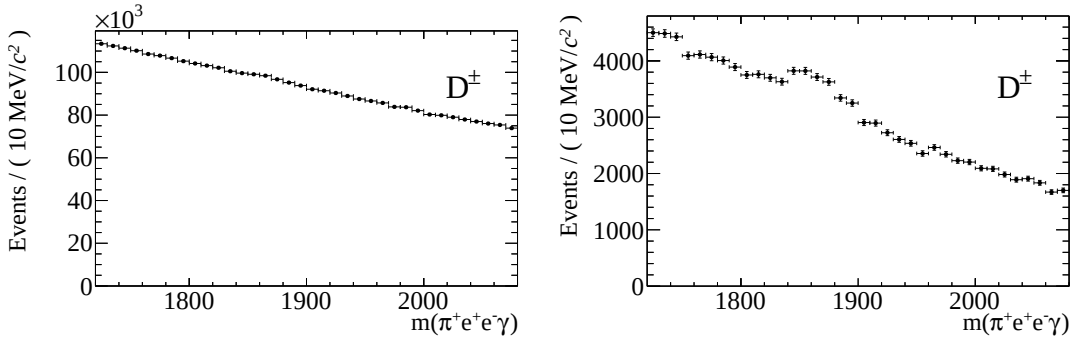

Figure 4.6: Invariant mass distributions of $m(\pi^+e^+e^-\gamma)$ candidates before the trigger requirements (left) and afterwards (right). The trigger requirements can be seen to substantially reduce the total number of events.

final state particles need a higher proportion of the total momentum to cause the trigger to fire. The flight distance χ^2 of the D^+ and the impact parameter χ^2 of the daughters is also shifted to larger values, as the trigger seeks to remove tracks coming directly from the primary vertex.

4.2.2 Stripping

The cuts applied in the stripping line are shown in Table 4.5. The cuts made in the stripping were calibrated using a sample of simulated signal decays and a sample of minimum bias data. The cuts attempt to maximise the number of signal decays that pass while keeping the overall retention rate, tested on minimum bias data, under a designated rate.

4.2.3 Final selection

A cut around the π^0 mass ($110 < m(\pi^0) < 160$ MeV/c^2) is applied to further suppress the combinatorial background, so that the shape of the signals becomes clear. Cutting on other variables is found to increase the statistical uncertainty on A_{CP} and so is not useful.

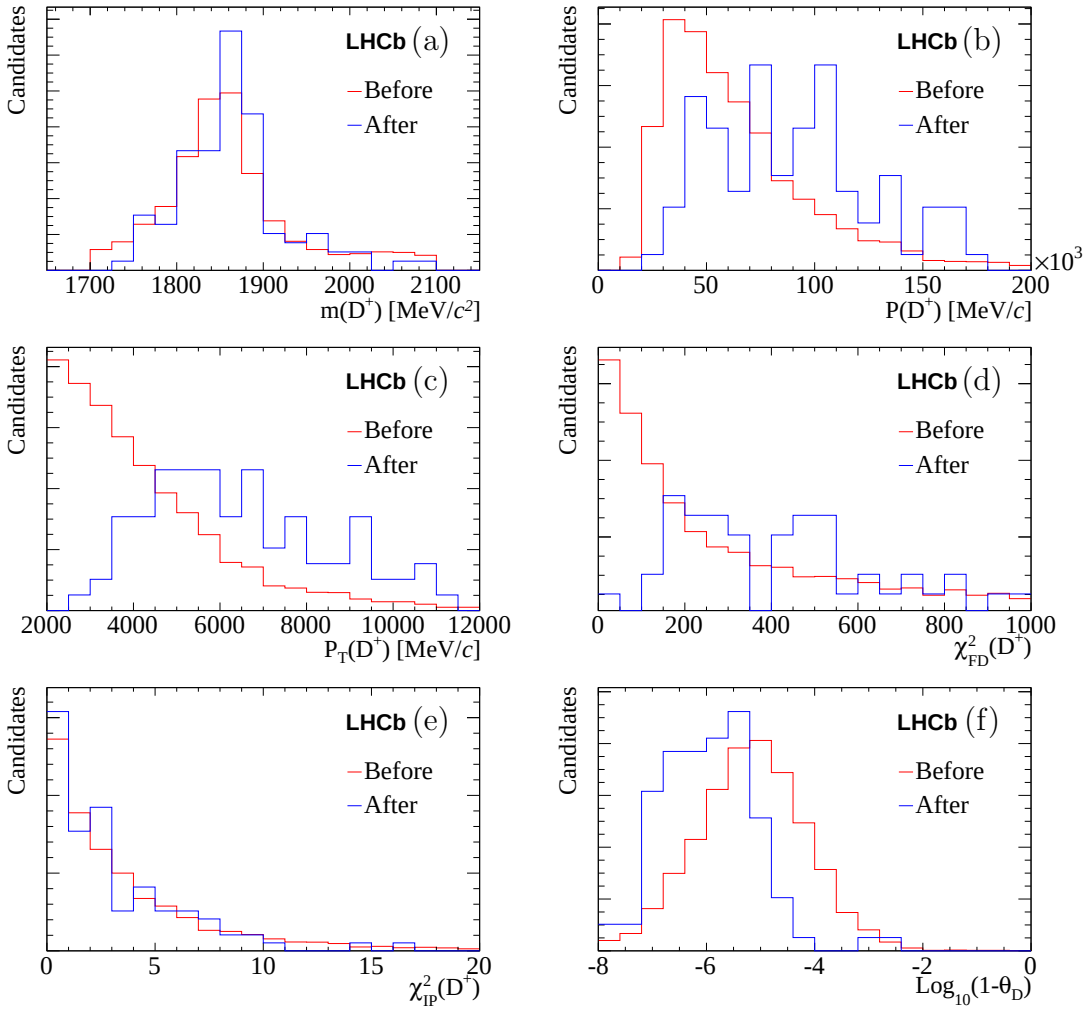

Figure 4.7: Simulated signal distributions before (red) and after (blue) applying trigger cuts. The distributions shown are (a) $m(D^+)$, (b) $p(D^+)$, (c) $p_T(D^+)$, (d) flight distance $\chi^2(D^+)$, (e) impact parameter $\chi^2(D^+)$ and (f) $\log(1-\theta_D)$. The histograms are normalised to unity.

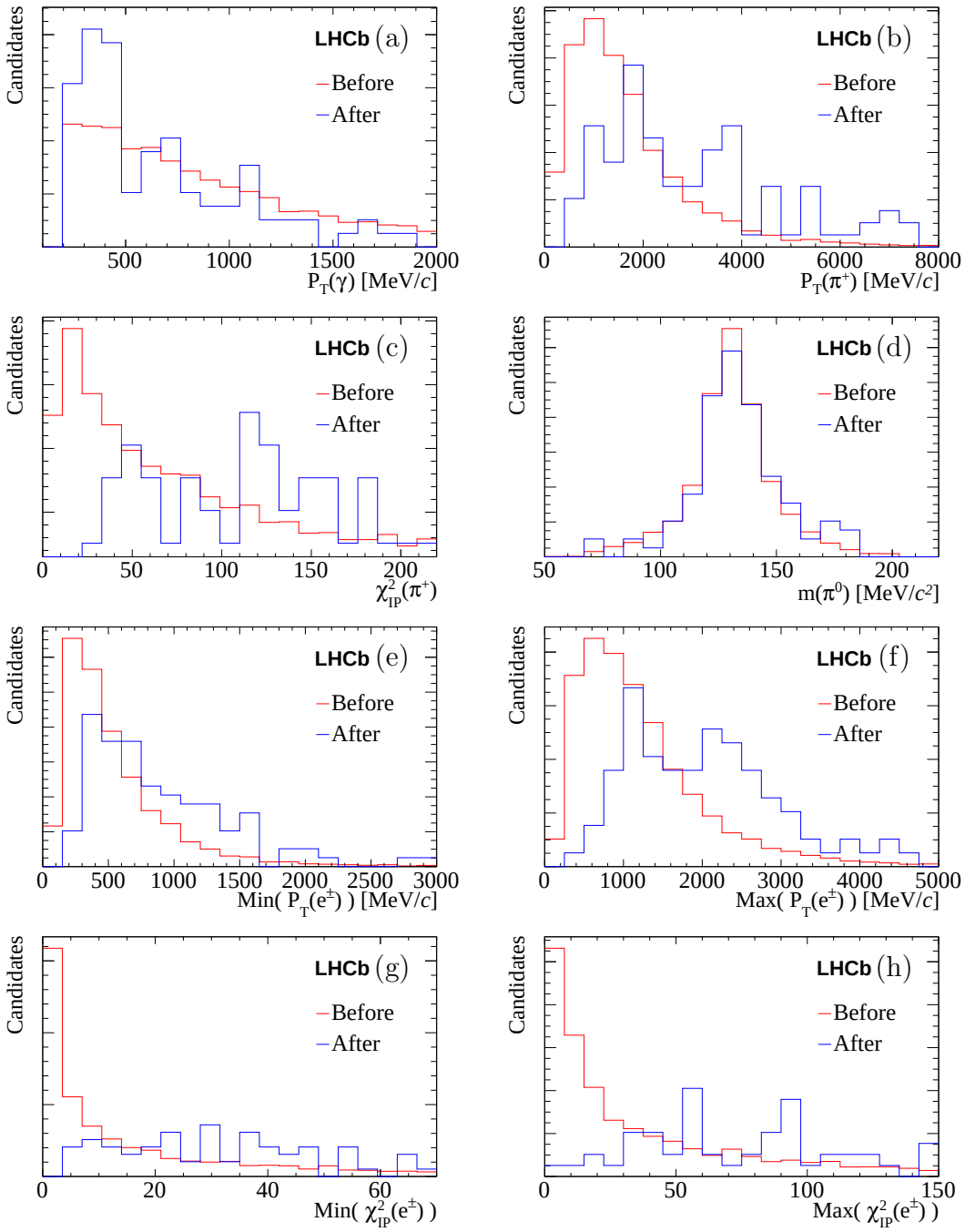

Figure 4.8: Simulated signal distributions before (red) and after (blue) applying trigger cuts. The distributions shown are (a) $p_T(\gamma)$, (b) $p_T(\pi^+)$, (c) impact parameter $\chi^2(\pi^+)$, (d) $m(\pi^0)$, (e) minimum $p_T(e^\pm)$, (f) maximum $p_T(e^\pm)$, (g) minimum impact parameter $\chi^2(e^\pm)$ and (h) maximum impact parameter $\chi^2(e^\pm)$. The histograms are normalised to unity.

Object	**Requirement**
$e^{\pm}$	CALO information = 1
	$\mathrm{DLL}_{\mathrm{e}-\pi} > -2.0$
γ	$p > 1000$ MeV
	$p_T > 200$ MeV
	Track $\chi^2_{ndf} < 5$
e^+e^-	$p_T > 200$ MeV/c
	$m < 5000$ MeV/c^2
π^0	$m > 70$ MeV/c^2
	$m < 200$ MeV/c^2
π^+	$p > 1000$ MeV/c
	$p_T > 250$ MeV/c
	Track $\chi^2_{ndf} < 5$
	IP $\chi^2_{ndf} > 4$
	$\eta > 2$
	$\eta < 5$
	RICH information = 1
	$\mathcal{P}(\mathrm{ghost}) < 0.5$
D^+	$p_T > 2000$ MeV/c
	$m > 1700$ MeV/c^2
	$m < 2100$ MeV/c^2
	SV $\chi^2 < 5$
	FD > 0.05 mm

Table 4.5: Cuts used in the stripping line, StrippingD2PiPi0.

4.3 Invariant mass distributions

The parameters describing the shapes and yields of the signal and background contributions are determined using a binned maximum likelihood fit to the invariant mass distributions.

4.3.1 π^0 candidates

The invariant mass distribution of the $e^+e^-\gamma$ candidates is shown in Fig 4.9. A gaussian and third-order polynomial are fitted to the candidates in the range $90-190\,\mathrm{MeV}/c^2$. The fitted gaussian parameters and yields are given in Table 4.6. The invariant mass is consistent within two sigma of the world average, 135.0 MeV/c^2 (with negligible uncertainty) [1]. The signal peak is wider than that shown in Fig. 4.4, possibly because of imperfect alignment in the tracking of electrons or bremsstrahlung.

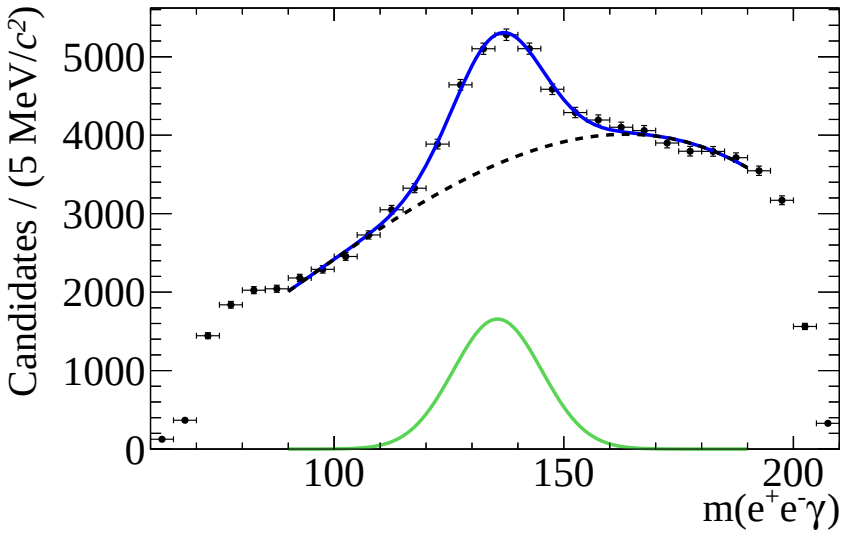

Figure 4.9: Invariant mass plots for π^0 candidates after trigger requirements. Curved edges of the distribution are due to multiple mass cuts in the stripping.

Parameter	Value
Mean	135.6 ± 0.3 MeV/c^2
Sigma	9.6 ± 0.5 MeV/c^2
π^0 yield	7981 ± 51

Table 4.6: π^0 invariant mass fit parameters.

4.3.2 $D^{\pm}$ candidates

The invariant mass of $\pi^+e^+e^-\gamma$ candidates in the range $1810 - 2040$ MeV/c^2 is fitted with PDFs describing the various components of the data. The PDF chosen to describe the signal is a crystal ball function, defined as:

$$f(x) \propto \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \text{ for } \frac{x-\mu}{\sigma} > 0, \tag{4.1}$$

$$f(x) \propto A \cdot \left(B - \frac{x-\mu}{\sigma}\right)^{-n} \text{ for } \frac{x-\mu}{\sigma} \leq 0, \tag{4.2}$$

where the first expression is a Gaussian peak of mean μ and width σ and the second is a power-law low-end tail, parameterised by,

$$A = \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(\frac{-|\alpha|^2}{2}\right), \tag{4.3}$$

$$B = \frac{n}{|\alpha|} - |\alpha|. \tag{4.4}$$

The crystal ball function is chosen as it contains a low mass tail to describe energy losses because of imperfect bremsstrahling recovery. An exponential function is used to describe the background component which may contain combinatorial or partially reconstructed backgrounds that vary smoothly across the fit range. All parameters of the fit are allowed to vary, and the shape of the signal is simultaneously fitted for D^+ and D^- candidates. The data is divided by the electromagnetic charge so that A_{Raw} can be calculated. A_{Raw} is the asymmetry between the number of reconstructed D^+ and D^- candidates, and defined as:

$$A_{Raw} = \frac{\mathcal{N}_{D^+} - \mathcal{N}_{D^-}}{\mathcal{N}_{D^+} + \mathcal{N}_{D^-}}, \tag{4.5}$$

where $\mathcal{N}_{D^+}$ and $\mathcal{N}_{D^-}$ are the measured D^+ and D^- yields, respectively. The invariant mass spectrum together with the fitted PDFs are shown in Fig. 4.10. The fitted signal parameters and yields are given in Table 4.7 and Table 4.8, respectively. The discrepancy between the mean of the observed signal distribution and that expected, 1869.62 ± 0.15 MeV/c^2 [1], is thought to be due an imperfect description of the tail.

Parameter	**Value**
Mean	1860.7 ± 1.8 MeV/c^2
Sigma	25.4 ± 2.1 MeV/c^2
Alpha	0.73 ± 0.17
Drop-off	11.4 ± 8.6
A_{Raw}	-0.0096 ± 0.0417

Table 4.7: Signal parameters for the invariant mass fit.

Decay	**Yield**
$D^+ \to \pi^+(\pi^0 \to e^+e^-\gamma)$	1939 ± 188
$D^- \to \pi^-(\pi^0 \to e^+e^-\gamma)$	1976 ± 200
Total	3915 ± 353

Table 4.8: Signal yields for invariant mass fit.

4.3.3 Contribution from $D^+ \to \pi^+(\pi^0 \to \gamma\gamma)$ decays

The expected number of $D^+ \to \pi^+(\pi^0 \to e^+e^-\gamma)$ decays is calculated using,

$$\mathcal{N}(D^+ \to \pi^+(\pi^0 \to e^+e^-\gamma)) = \mathcal{L} \cdot \sigma \cdot \mathcal{B}(D^+ \to \pi^+(\pi^0 \to e^+e^-\gamma)) \cdot \epsilon, \qquad (4.6)$$

where $\mathcal{L}$ is the total integrated luminosity from Table 4.1, σ is the cross-section for D^+ production [37] and ϵ is the overall efficiency from Table 4.3. $\mathcal{B}$ is the product of the branching fractions $\mathcal{B}(D^+ \to \pi^+\pi^0) = (1.19 \pm 0.06) \times 10^{-3}$ [1] and $\mathcal{B}(\pi^0 \to e^+e^-\gamma) = (1.174 \pm 0.035) \times 10^{-2}$. The expected yield is,

$$\mathcal{N}(D^+ \to \pi^+(\pi^0 \to e^+e^-\gamma)) = 1170 \pm 260, \qquad (4.7)$$

a factor of four smaller than the observed yield. The extra contribution to the observed yield is the result of $\pi^0 \to \gamma\gamma$ decays, where a γ converts to an electron-

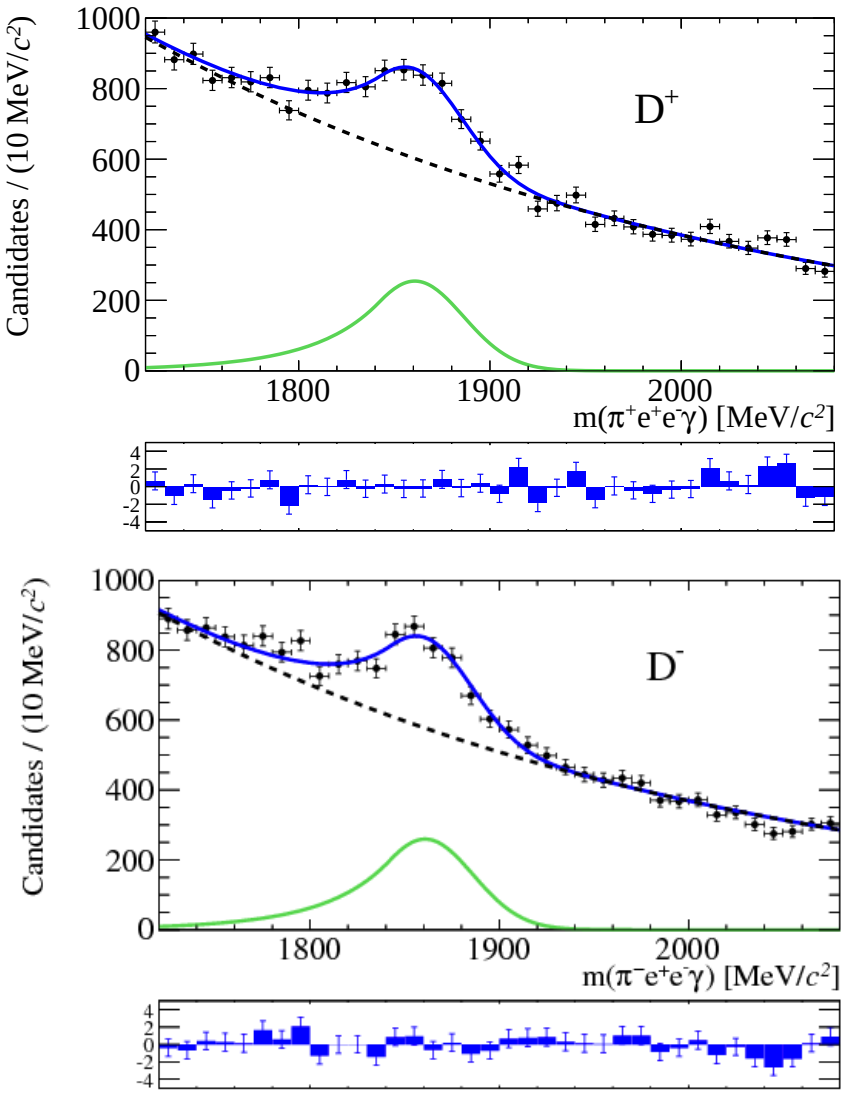

Figure 4.10: Invariant mass plots for $\pi^+e^+e^-\gamma$ (top) and $\pi^-e^+e^-\gamma$ (bottom) candidates. The data are shown as black points and the total PDF (dark blue line) is overlaid. The components of the fit are also shown: the signal (light green line) and the non-peaking background (dashed line).

positron pair before the magnet. The probability of conversion before the magnet (CBM) is $\sim 22\%$, as discussed in Sec. 2.6.2. The high momentum of the γ means the e^+e^- are highly boosted and travel in a collinear direction, thus pointing back to the decay vertex. The expected yield is calculated using,

$$\mathcal{N}(D^+ \to \pi^+(\pi^0 \to \gamma\gamma_{CBM})) = \mathcal{N}(D^+ \to \pi^+(\pi^0 \to e^+e^-\gamma)) \cdot \frac{0.22 \cdot \mathcal{B}(\pi^0 \to \gamma\gamma)}{\mathcal{B}(\pi^0 \to e^+e^-\gamma)} \cdot \mathcal{R}, \tag{4.8}$$

where the branching fraction, $\mathcal{B}(\pi^0 \to \gamma\gamma) = (98.823 \pm 0.034)\%$ and $\mathcal{R}$ is the efficiency ratio, defined by,

$$\mathcal{R} = \frac{\epsilon(D^0 \to \pi^+\pi^-(\pi^0 \to \gamma\gamma_{CBM}))}{\epsilon(D^+ \to \pi^+(\pi^0 \to e^+e^-\gamma)) \cdot \epsilon(\pi^-)}. \tag{4.9}$$

The reconstruction/stripping efficiency for $D^0 \to \pi^+\pi^-(\pi^0 \to \gamma\gamma_{CBM})$ and $D^+ \to \pi^+(\pi^0 \to e^+e^-\gamma)$ decays is 6.5×10^{-5} [82] and 2.6×10^{-3}, see Table 4.3, respectively. Assuming that the efficiency of detecting and accepting a charged pion is 70%, the expected yield is,

$$\mathcal{N}(D^+ \to \pi^+(\pi^0 \to \gamma\gamma_{CBM})) = 2590 \pm 510. \tag{4.10}$$

This is consistent with the excess of candidates observed.

4.3.4 Comparison with simulation

The distributions of kinematic and spatial variables in the data and in simulated signal can be compared. The signal is isolated in data by using the *sPlot* technique, as discussed in Sec. 3.3. The control variable used is $m(\pi^+e^+e^-\gamma)$, with candidates which fall in the area under the signal PDF given an *sWeight* closer to unity. It can be seen that the simulation and data are in good agreement.

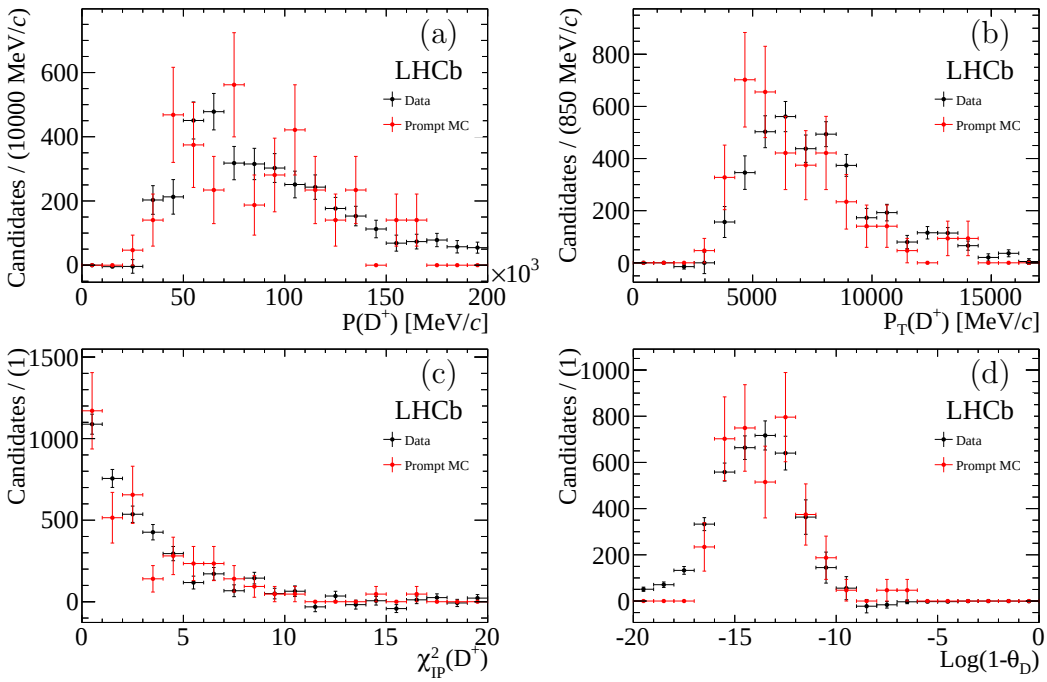

Figure 4.11: Truth-matched simulated signal (red) and background-subtracted data distributions (black) after all selection cuts. The distributions shown are (a) $p(D^+)$, (b) $p_T(D^+)$, (c) impact parameter $\chi^2(D^+)$ and (d) Log$(1-\theta_D)$.

4.4 A measurement of the charge asymmetry

To convert the raw asymmetry extracted from the fit, A_{Raw}, into an A_{CP} measurement, one must account for other sources of asymmetry.

$$A_{CP} = A_{Raw} - A^{Eff}_{Prod} - A_{Det}, \tag{4.11}$$

where A^{Eff}_{Prod} and A_{Det} are the effective production and detection asymmetries, respectively. The effective production asymmetry only applies to the events selected by this analysis, and has no physical meaning. As such, it cannot be outrightly compared to the production asymmetry of other decay modes.

Values of A_{Prod} in bins of η and p_T have been measured at a collision energy of $\sqrt{s} = 7\,\text{TeV}$ using $D^+ \to K^0_S\pi^+$ decays [83] in 48 bins. A^{Eff}_{Prod} is a function of the η and p_T of the D^+, hence the total (weighted mean) value is different for each decay mode depending upon its kinematics. The total A^{Eff}_{Prod} and its statistical

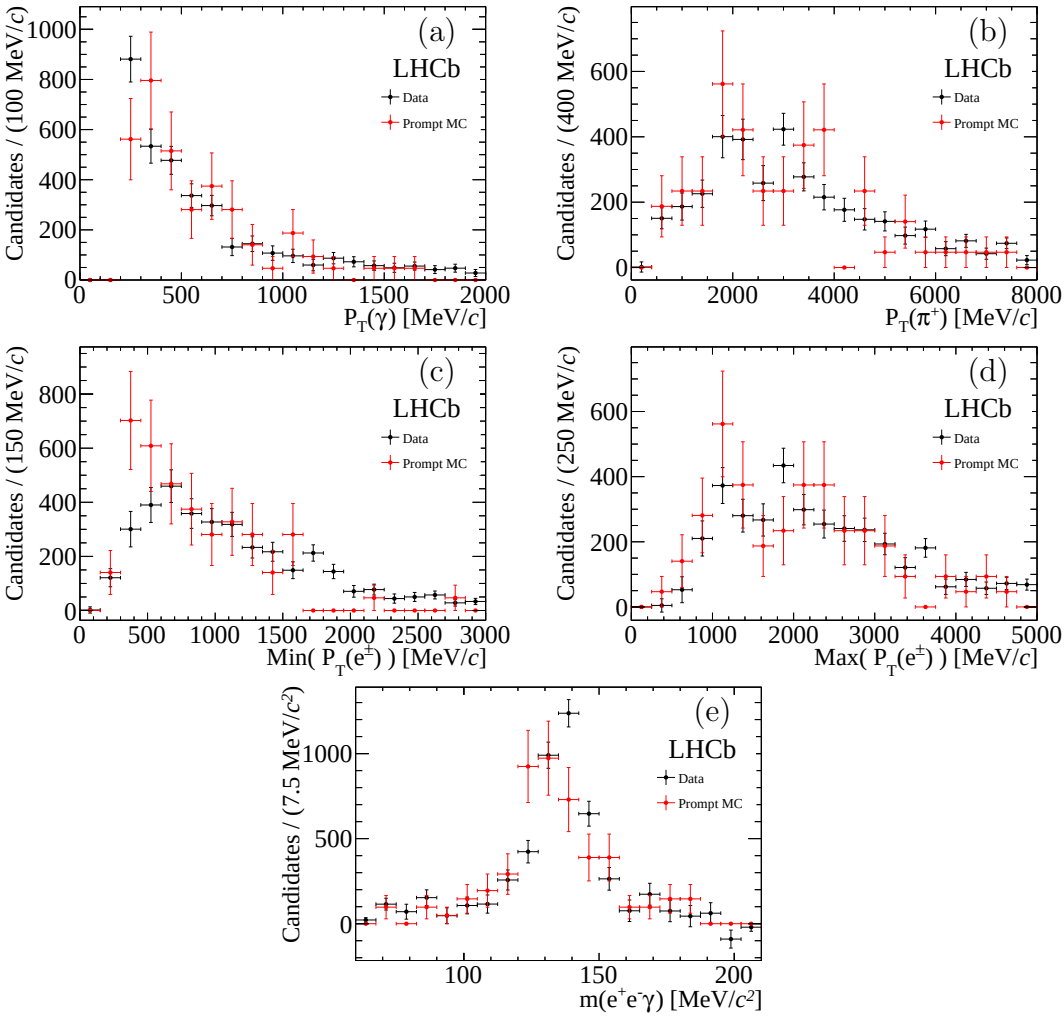

Figure 4.12: Truth-matched simulated signal (red) and background-subtracted data distributions (black) after all selection cuts. The distributions shown are (a) $p_T(\gamma)$, (b) $p_T(\pi^+)$, (c) minimum $p_T(e^\pm)$, (d) maximum $p_T(e^\pm)$ and (e) $m(e^+e^-\gamma)$. $m(e^+e^-\gamma)$ is plotted with only trigger requirements on the decay.

uncertainty for $D^+ \to \pi^+(\pi^0 \to e^+e^-\gamma)$ are found using,

$$A_{Prod}^{Eff} = \frac{\sum_{i=0}^{47} A_{Prod,i} \times n_i}{N}, \tag{4.12}$$

$$\Delta A_{Prod}^{Eff} = \sqrt{\sum_{i=0}^{47} \left(\frac{\Delta A_{Prod,i} \times n_i}{N} \right)^2}, \tag{4.13}$$

where $A_{Prod,i}$, $\Delta A_{Prod,i}$ and n_i are the production asymmetry, the statistical uncertainty on the production asymmetry and the number of signal events, respectively, in the i-th bin. n_i is found by performing a background-subtracted study in each bin using the *sPlot* technique. The signal's η-p_T distribution can be seen in Fig. 4.13. The total A_{Prod}^{Eff} is found to be $(-0.41 \pm 0.26(\text{stat}) \pm 0.18(\text{syst}))\%$. The systematic uncertainty is taken solely from the A_{Prod}^{Eff} paper [83].

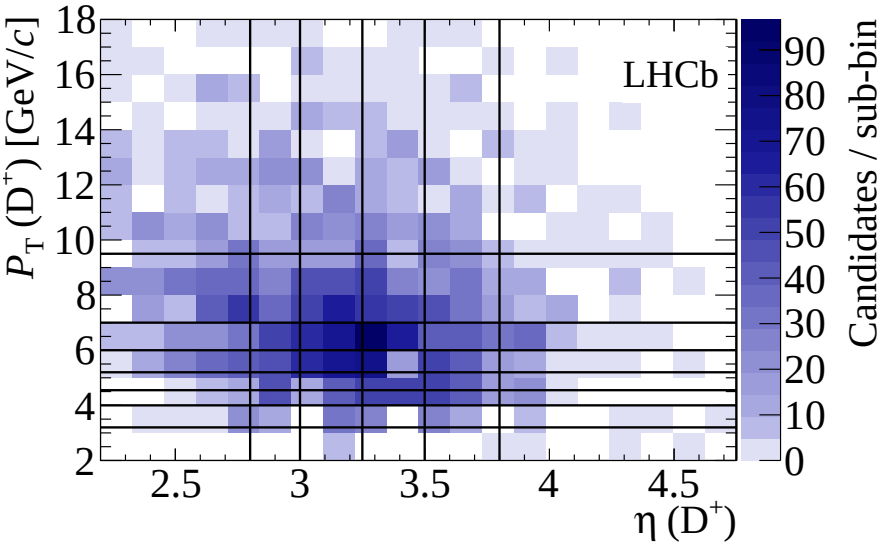

Figure 4.13: Binned η-p_T distributions of $D^+ \to \pi^+(\pi^0 \to e^+e^-\gamma)$ decays.

Detection asymmetries may affect the measurement so must be taken into account. Only the detection of the $\pi^\pm$ would have an influence as the detection asymmetry on the electrons cancel. The pion detection asymmetry, A_{Det} $(\pi^\pm)$ $= (0.00 \pm 0.25)\%$, has been measured by LHCb [84].

4.4.1 The result

The measured asymmetries, shown in Table 4.9, are used in Eq. 4.11 to obtain the inclusive *CP* asymmetry of $D^+ \to \pi^+\pi^0$ decays,

$$A_{CP}(D^+ \to \pi^+\pi^0) = (-0.59 \pm 4.19 \pm 0.34)\%.$$

Where the first uncertainty is statistical and the second is systematic. The derivation of the systematic uncertainty is discussed in Sec. 4.4.2. The measurement shows no evidence of CP violation.

Asymmetry	**Value (%)**
A_{Raw}	-0.96 ± 4.17
A^{Eff}_{Prod}	-0.41 ± 0.26
A_{Det}	0.00 ± 0.25
A_{CP}	-0.59 ± 4.19

Table 4.9: Measured $D^+ \to \pi^+\pi^0$ asymmetry values together with their statistical uncertainties. The systematic uncertainties associated to each measurement is shown in Sec. 4.4.2.

4.4.2 Dominant systematic uncertainties

A number of sources contribute to the systematic uncertainty on the A_{CP} measurement. The size of their effect is investigated and detailed here.

The choice of PDF describing the signal shape affects the number of signal events observed, and hence the value of A_{Raw}. To judge the size of this effect, other PDFs replace those given in Eq. 4.1 and Eq. 4.2. The difference in the default and amended A_{Raw} is then given as the systematic uncertainty. A Cruijff function, see Eq. 3.10, which has asymmetric tails to account for the effect of final-state radiation, is trialled as an alternative. The resulting systematic is shown in Table 4.10.

The choice of PDF describing the combinatorial background shape also affects

the value of A_{Raw}. A second order polynomial is tested. The resulting systematic is shown in Table 4.10.

The production asymmetry was only evaluated using measurements made at $\sqrt{s} = 7\,\text{TeV}$. Production asymmetry data at $\sqrt{s} = 8\,\text{TeV}$ has not been published to date, so a systematic term should be considered. This is not computed however, as the measurement's systematic uncertainty is completely dominated by its statistical uncertainty. Any future publication would have calculate this uncertainty.

Systematic uncertainty	A_{CP} (%)
Signal shape	0.04
Comb. background shape	0.13
A_{Det} $(\pi^{\pm})$	0.25
A^{Eff}_{Prod} $(D^{\pm})$	0.18
Total	0.34

Table 4.10: Systematic uncertainties on A_{CP}. The total systematic uncertainties are the sum in quadrature of the individual contributions.

4.5 Discussion

The value of A_{CP} $(D^+ \to \pi^+\pi^0)$ measured is dominated by the statistical uncertainty, with the $2011 - 2012$ dataset. This uncertainty of $\sim 4\%$ is an order of magnitude above the uncertainty on other charm CPV measurements, as previously shown in Table 1.3. Given a dedicated HLT2 trigger, a factor of ten more efficient, one can see that with a naive $\sqrt{N}$ scaling, there would be a statistical uncertainty at the one percent level, close to the sensitivities of current analyses.

Thus, a dedicated HLT2 trigger is being written so a much greater yield can be observed by LHCb from 2014 onwards.

Appendix

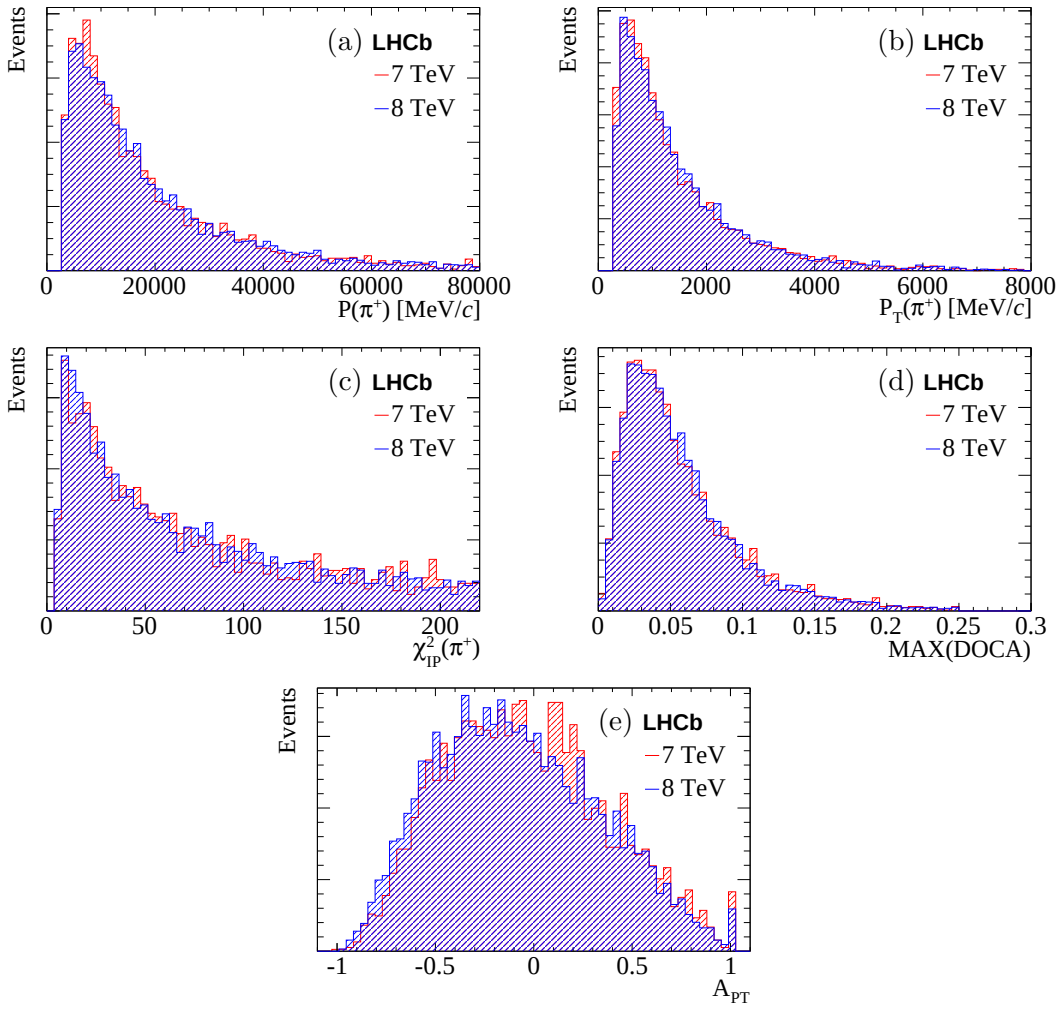

Figure 1: Distributions of simulated $D^+ \to \pi^+\mu^+\mu^-$ candidates generated at collision energies of 7 TeV (red) and 8 TeV (blue). The distributions shown are (a) $p(\pi^+)$, (b) $p_T(\pi^+)$, (c) π^+ impact parameter significance, (d) the maximum distance of closest approach between any two of the final state particles and (e) the transverse momentum asymmetry of the event.

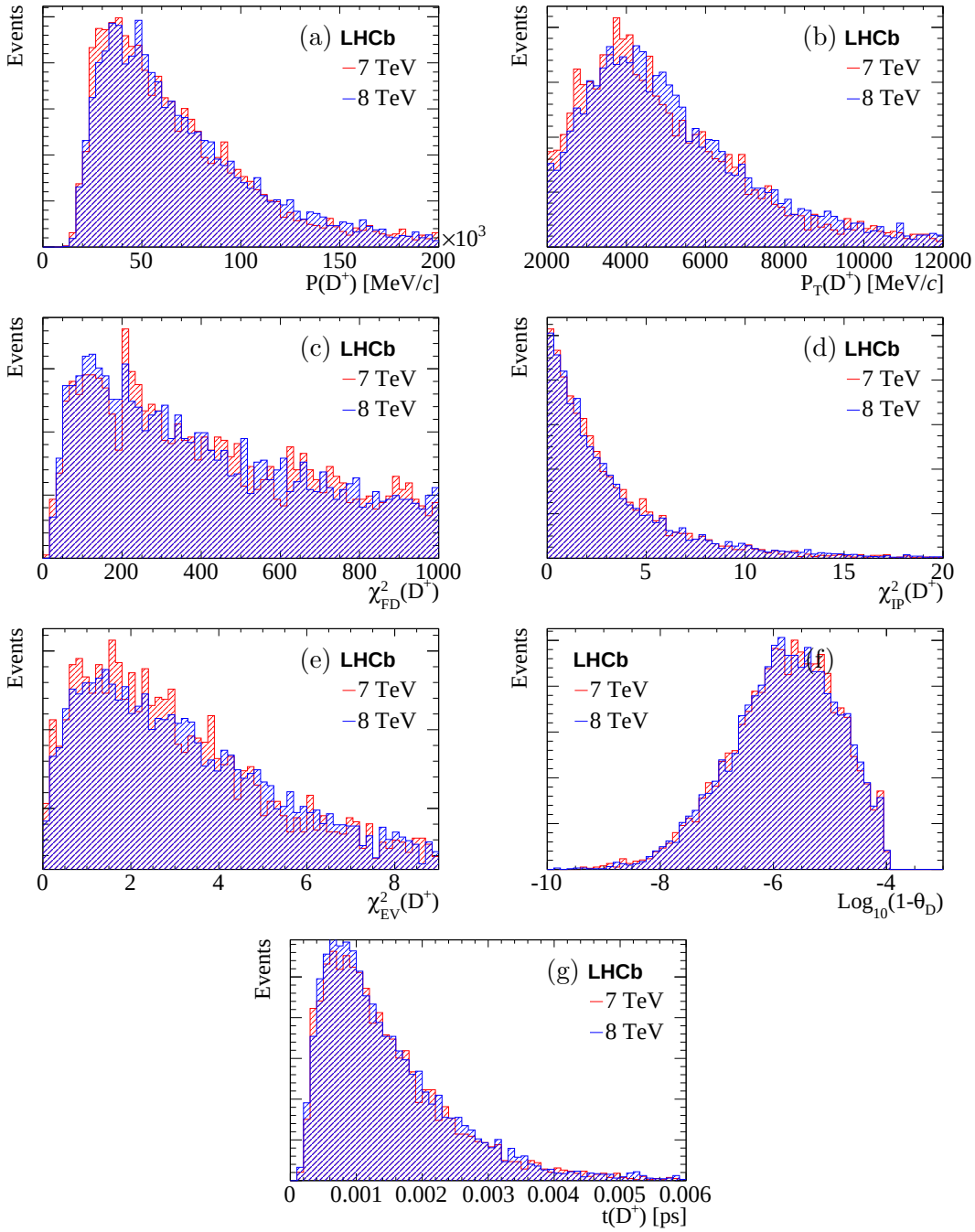

Figure 2: Distributions of simulated $D^+ \to \pi^+\mu^+\mu^-$ candidates generated at collision energies of 7 TeV (red) and 8 TeV (blue). The distributions shown are (a) $p(D^+)$, (b) $p_T(D^+)$, (c) D^+ flight distance significance, (d) D^+ impact parameter significance, (e) end vertex $\chi^2(D^+)$, (f) Log$(1-\theta_D)$ and (e) time of flight.

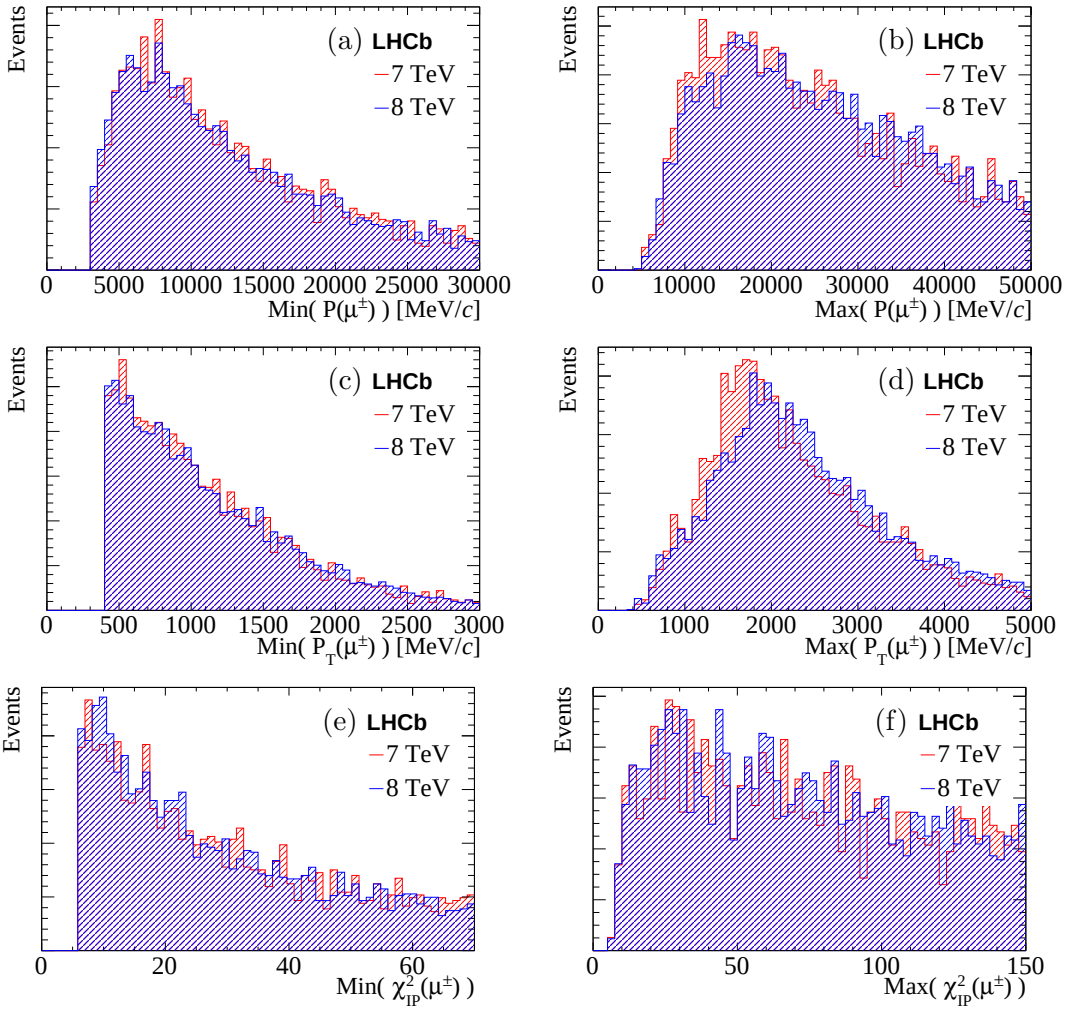

Figure 3: Distributions of simulated $D^+ \to \pi^+\mu^+\mu^-$ candidates generated at collision energies of 7 TeV (red) and 8 TeV (blue). The distributions shown are (a) minimum p ($\mu^\pm$), (b) maximum p ($\mu^\pm$), (c) minimum p_T ($\mu^\pm$), (d) maximum p_T ($\mu^\pm$), (e) minimum $\mu^\pm$ impact parameter significance and (e) maximum $\mu^\pm$ impact parameter significance.

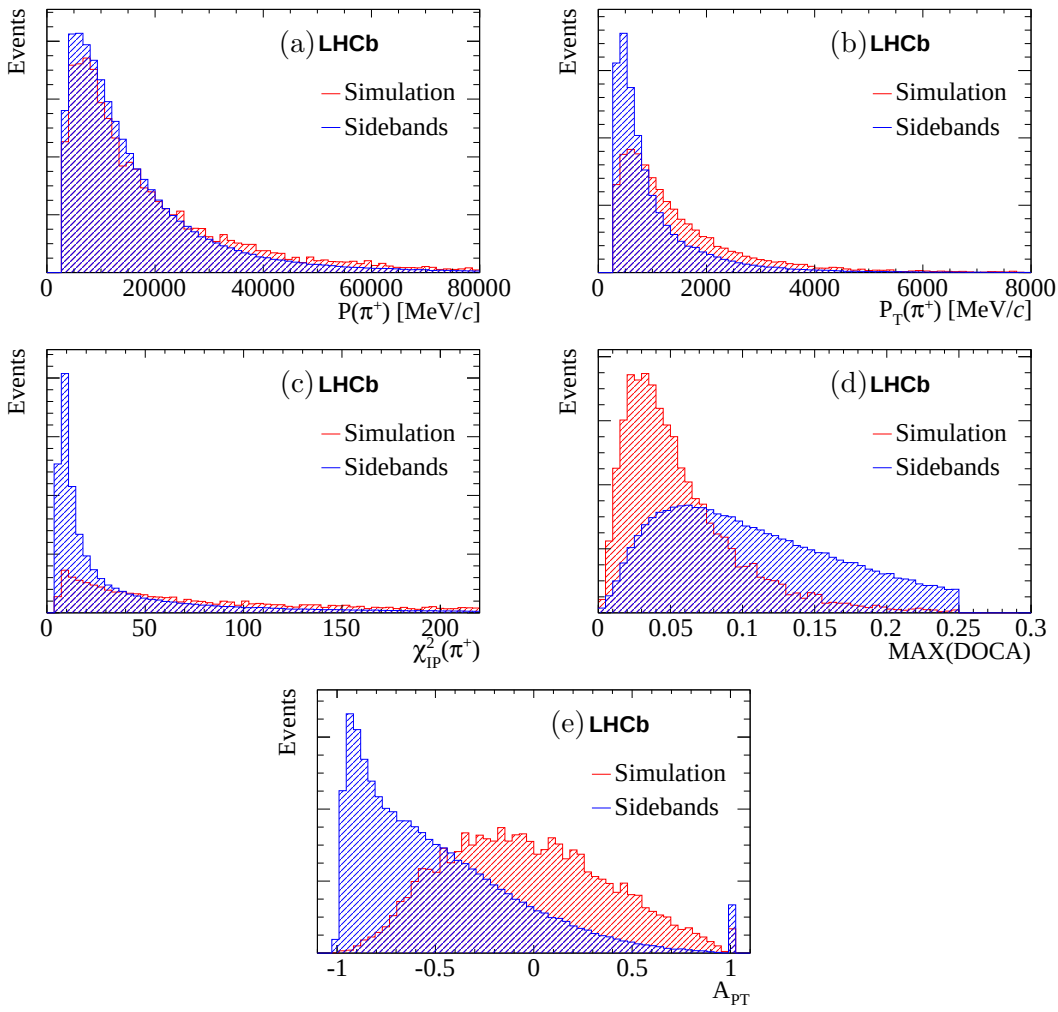

Figure 4: Distributions of simulated $D^+ \to \pi^+\mu^+\mu^-$ (red) and sideband (blue) candidates. The distributions shown are (a) p (π^+), (b) p_T (π^+), (c) π^+ impact parameter significance, (d) the maximum distance of closest approach between any two of the final state particles and (e) the transverse momentum asymmetry of the event.

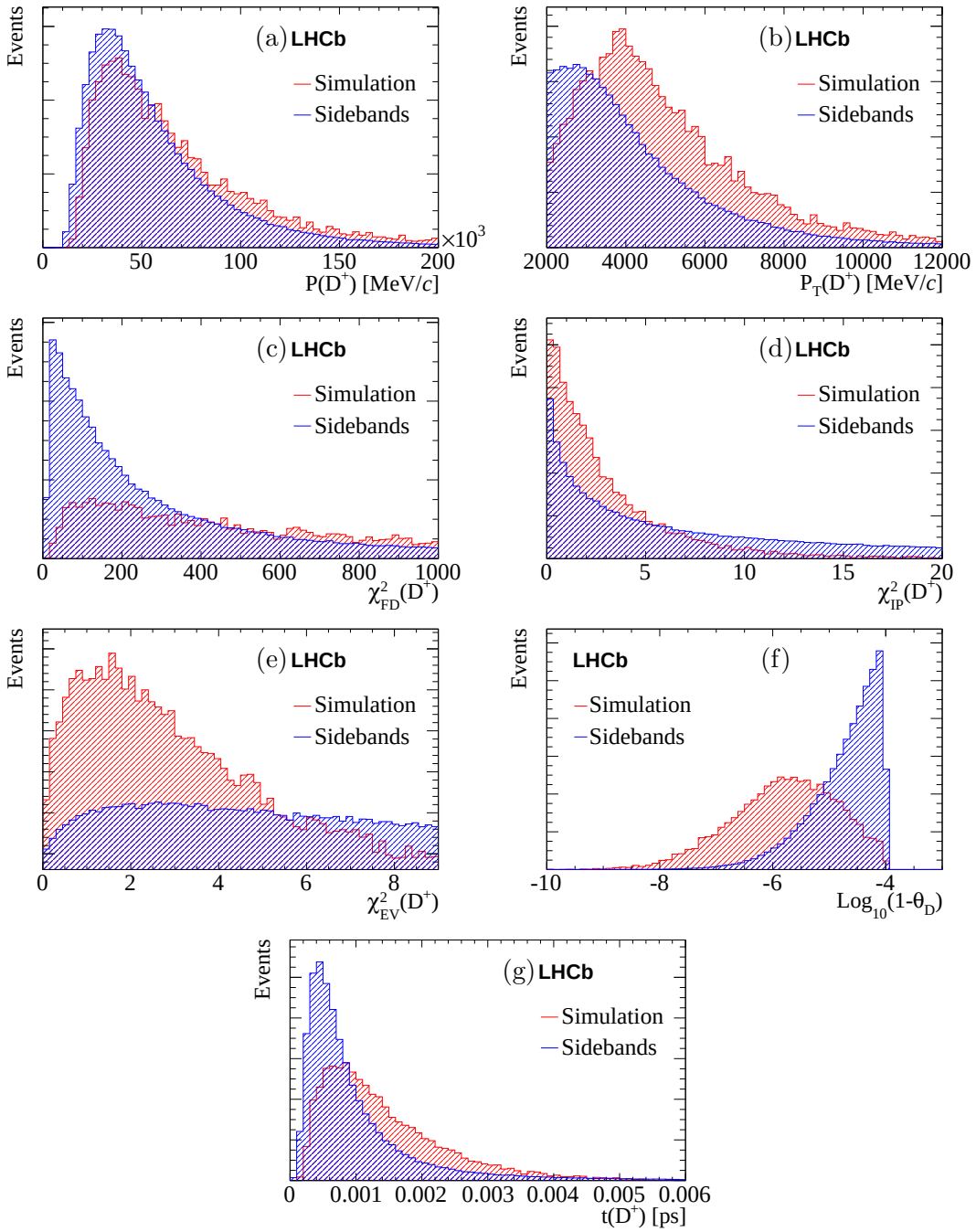

Figure 5: Distributions of simulated $D^+ \to \pi^+\mu^+\mu^-$ (red) and sideband (blue) candidates. The distributions shown are (a) p (D^+), (b) p_T (D^+), (c) D^+ flight distance significance, (d) D^+ impact parameter significance, (e) end vertex χ^2 (D^+), (f) Log(1−direction angle) and (e) time of flight.

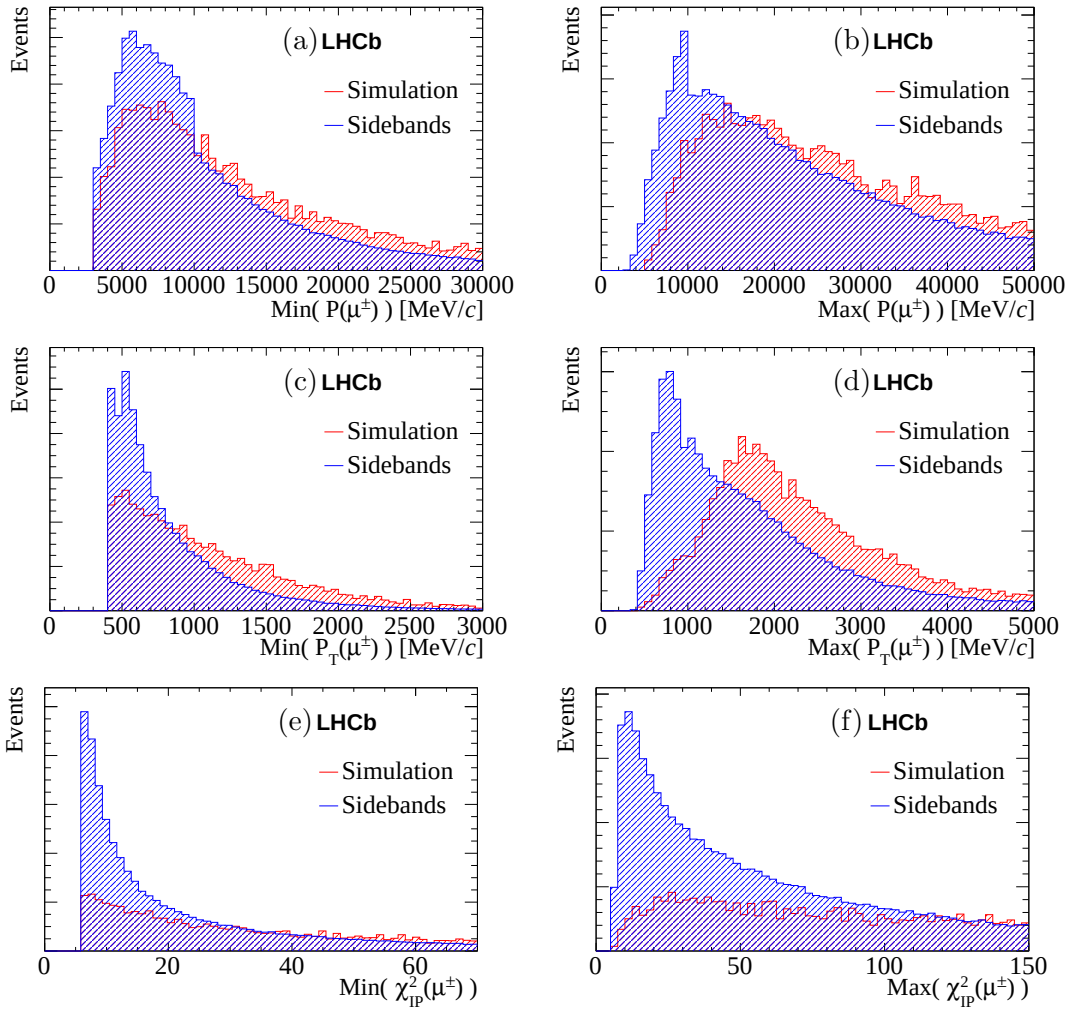

Figure 6: Distributions of $D^+ \to \pi^+\mu^+\mu^-$ (red) and sideband (blue) candidates. The distributions shown are (a) minimum p ($\mu^\pm$), (b) maximum p ($\mu^\pm$), (c) minimum p_T ($\mu^\pm$), (d) maximum p_T ($\mu^\pm$), (e) minimum $\mu^\pm$ impact parameter significance and (e) maximum $\mu^\pm$ impact parameter significance.

Candidates	Bin description	Parameter	Value	Floating?
$\pi^+\mu^+\mu^-$	low-$m(\mu^+\mu^-)$	c_0	$(1.03 \pm 0.61) \times 10^{-1}$	✓
		c_1	$(-7.2 \pm 5.8) \times 10^{-2}$	✓
		N_{comb}	70 ± 20	✓
		$N_{D_s^+ \to \pi^+(\eta \to \mu^+\mu^-\gamma)}$	434 ± 31	✓
		$\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$	$(4.0 \pm 6.0) \times 10^{-9}$	✓
		$\mathcal{B}(D_s^+ \to \pi^+\mu^+\mu^-)$	$(1.8 \pm 1.6) \times 10^{-8}$	✓
	η	c_0	$(-3.5 \pm 3.1) \times 10^{-1}$	✓
		N_{comb}	29 ± 7	✓
		N_{D^+}	48 ± 10	✓
		$N_{D_s^+}$	22 ± 5	✓
	ρ^0/ω	c_0	$(-8.9 \pm 0.9) \times 10^{-1}$	✓
		c_1	$(2.8 \pm 0.9) \times 10^{-1}$	✓
		N_{comb}	294 ± 25	✓
		N_{D^+}	166 ± 17	✓
		$N_{D_s^+}$	135 ± 14	✓
	ϕ	c_0	$(-4.9 \pm 0.7) \times 10^{-1}$	✓
		c_1	$(-9.1 \pm 8.0) \times 10^{-2}$	✓
		N_{comb}	912 ± 104	✓
		N_{D^+}	4474 ± 84	✓
		$N_{D_s^+}$	5639 ± 96	✓
	all	μ_{D^+}	1871.5 ± 0.2 MeV/c^2	✓
		$\mu_{D_s^+}$	1970.8 ± 0.1 MeV/c^2	✓
		σ_{D^+}	7.9 ± 0.2 MeV/c^2	✓
		$\sigma_{D_s^+}$	8.2 ± 0.2 MeV/c^2	✓
		$\alpha_{D^+_{(s)},L}$	$(1.2 \pm 0.2) \times 10^{-1}$	✓
		$\alpha_{D^+_{(s)},R}$	$(5.6 \pm 1.5) \times 10^{-2}$	✓
		$\varrho(\pi^+\pi^+\pi^- \to \pi^+\mu^+\mu^-)$	$(1.2 \pm 0.8) \times 10^{-2}$	✓
		β_{D^+}	2.0×10^{-6}	×
		$\beta_{D_s^+}$	1.6×10^{-5}	×

Table 1: Parameters of the lower mass invariant mass fit to $\pi^+\mu^+\mu^-$ and $\pi^+\pi^+\pi^-$ candidates. Due to the large number of parameters, those describing the contribution from $\pi^+\pi^+\pi^-$ candidates are shown in Table 2. c_1 and c_0 are coefficients of the Chebyshev series describing the combinatorial background, see Eq. 3.11. μ, σ and α terms are the parameters of the Cruijff distributions describing the signal and $\pi^+\pi^+\pi^-$ peaking background, see Eq. 3.10. N terms represent the yields of these contributions. The normalisation factors, β, are used to calculate the measured signal branching fraction using the ϕ yields and efficiency ratio. $\varrho(\pi^+\pi^+\pi^- \to \pi^+\mu^+\mu^-)$ is the misidentification rate. The shape of the $D_s^+ \to \pi^+(\eta \to \mu^+\mu^-\gamma)$ distribution is taken from a study of simulation.

Candidates	Bin description	Parameter	Value	Floating?
$\pi^+\pi^+\pi^-$	low-$m(\mu^+\mu^-)$	c_0	$(-1.2 \pm 1.0) \times 10^{-1}$	✓
		c_1	$(-9.2 \pm 8.8) \times 10^{-2}$	✓
		N_{comb}	747 ± 53	✓
		μ_{D^+}	1856.5 ± 0.4 MeV/c^2	✓
		$\mu_{D_s^+}$	1955.7 ± 0.8 MeV/c^2	✓
		$\sigma_{D^+_{(s)}}$	8.9 ± 0.4 MeV/c^2	✓
		N_{D^+}	1243 ± 49	✓
		$N_{D_s^+}$	381 ± 28	✓
	η	c_0	$(-3.6 \pm 2.0) \times 10^{-1}$	✓
		N_{comb}	139 ± 26	✓
		μ_{D^+}	1857.5 ± 0.9 MeV/c^2	✓
		$\mu_{D_s^+}$	1953.4 ± 2.3 MeV/c^2	✓
		$\sigma_{D^+_{(s)}}$	8.4 ± 1.3 MeV/c^2	✓
		N_{D^+}	215 ± 23	✓
		$N_{D_s^+}$	48 ± 10	✓
	ρ^0/ω	c_0	$(-6.4 \pm 0.7) \times 10^{-1}$	✓
		c_1	$(4.2 \pm 0.6) \times 10^{-1}$	✓
		N_{comb}	900 ± 58	✓
		μ_{D^+}	1856.4 ± 0.4 MeV/c^2	✓
		$\mu_{D_s^+}$	1956.4 ± 0.7 MeV/c^2	✓
		$\sigma_{D^+_{(s)}}$	9.3 ± 0.4 MeV/c^2	✓
		N_{D^+}	1509 ± 56	✓
		$N_{D_s^+}$	485 ± 30	✓
	ϕ	c_0	$(-2.7 \pm 0.7) \times 10^{-1}$	✓
		c_1	$(-7.7 \pm 6.4) \times 10^{-2}$	✓
		N_{comb}	1148 ± 67	✓
		μ_{D^+}	1858.6 ± 0.4 MeV/c^2	✓
		$\mu_{D_s^+}$	1958.0 ± 0.4 MeV/c^2	✓
		$\sigma_{D^+_{(s)}}$	9.2 ± 0.3 MeV/c^2	✓
		N_{D^+}	1763 ± 56	✓
		$N_{D_s^+}$	1464 ± 50	✓
	all	$\alpha_{D^+_{(s)},L}$	1.5×10^{-1}	×
		$\alpha_{D^+_{(s)},R}$	5.0×10^{-2}	×

Table 2: Parameters of the lower mass invariant mass fit to $\pi^+\mu^+\mu^-$ and $\pi^+\pi^+\pi^-$ candidates. The parameters are described in the caption of Table 1.

Candidates	Bin description	Parameter	Value	Floating?
$\pi^+\mu^+\mu^-$	high-$m(\mu^+\mu^-)$	c_0	$(1.0 \pm 0.6) \times 10^{-1}$	✓
		c_1	$(-7.2 \pm 5.8) \times 10^{-2}$	✓
		N_{comb}	997 ± 42	✓
		$\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$	$(7.8 \pm 6.1) \times 10^{-9}$	✓
		$\mathcal{B}(D_s^+ \to \pi^+\mu^+\mu^-)$	$(1.5 \pm 1.4) \times 10^{-8}$	✓
	ϕ	c_0	$(-5.3 \pm 0.7) \times 10^{-1}$	✓
		N_{comb}	1087 ± 100	✓
		N_{D^+}	5578 ± 60	✓
		$N_{D_s^+}$	6702 ± 58	✓
	all	μ_{D^+}	1871.4 ± 0.1 MeV/c^2	✓
		$\mu_{D_s^+}$	1970.8 ± 0.1 MeV/c^2	✓
		σ_{D^+}	7.9 ± 0.1 MeV/c^2	✓
		$\sigma_{D_s^+}$	8.1 ± 0.2 MeV/c^2	✓
		$\alpha_{D^+_{(s)},L}$	$(1.2 \pm 0.1) \times 10^{-1}$	✓
		$\alpha_{D^+_{(s)},R}$	$(6.7 \pm 1.2) \times 10^{-2}$	✓
		$\varrho(\pi^+\pi^+\pi^- \to \pi^+\mu^+\mu^-)$	$(1.2 \pm 0.8) \times 10^{-2}$	✓
		β_{D^+}	4.5×10^{-6}	×
		$\beta_{D_s^+}$	3.1×10^{-5}	×
$\pi^+\pi^+\pi^-$	high-$m(\mu^+\mu^-)$	c_0	$(-2.4 \pm 0.6) \times 10^{-1}$	✓
		c_1	$(-7.7 \pm 6.4) \times 10^{-2}$	✓
		N_{comb}	1184 ± 59	✓
		μ_{D^+}	1859.0 ± 0.4 MeV/c^2	✓
		$\mu_{D_s^+}$	1957.4 ± 0.3 MeV/c^2	✓
		$\sigma_{D^+_{(s)}}$	7.4 ± 0.2 MeV/c^2	✓
		N_{D^+}	1210 ± 45	✓
		$N_{D_s^+}$	1524 ± 48	✓
	ϕ	c_0	$(-2.6 \pm 0.7) \times 10^{-1}$	✓
		N_{comb}	1157 ± 65	✓
		μ_{D^+}	1858.3 ± 0.4 MeV/c^2	✓
		$\mu_{D_s^+}$	1957.9 ± 0.4 MeV/c^2	✓
		$\sigma_{D^+_{(s)}}$	9.1 ± 0.3 MeV/c^2	✓
		N_{D^+}	1752 ± 56	✓
		$N_{D_s^+}$	1443 ± 47	✓
	all	$\alpha_{D^+_{(s)},L}$	1.5×10^{-1}	×
		$\alpha_{D^+_{(s)},R}$	5.0×10^{-2}	×

Table 3: Parameters of the higher mass invariant mass fit to $\pi^+\mu^+\mu^-$ and $\pi^+\pi^+\pi^-$ candidates. The parameters are described in the caption of Table 1.

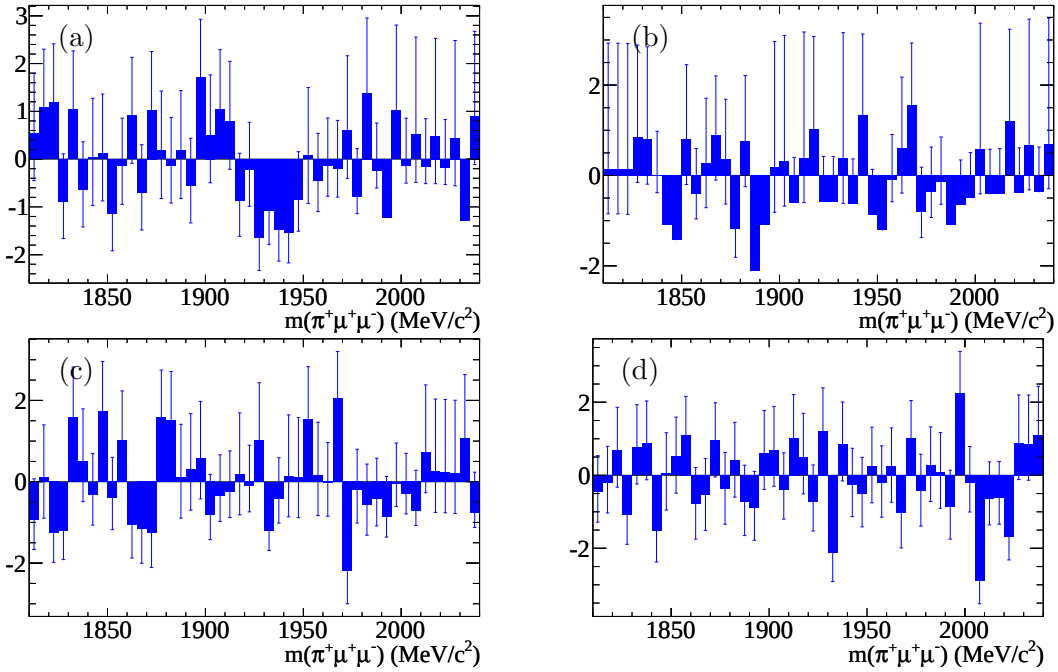

Figure 7: Residual plots of the fits to the invariant mass distributions for $D^+_{(s)} \to \pi^+\mu^+\mu^-$ candidates in the lower mass fit (excluding trigger lines with $m(\mu^+\mu^-) > 1.0$ GeV/c^2). Shown are the (a) low-$m(\mu^+\mu^-)$, (b) η, (c) ρ^0/ω and (d) ϕ bins.

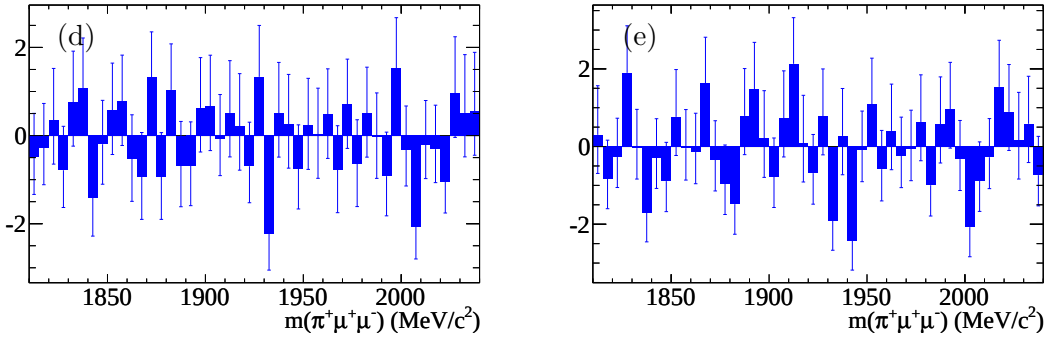

Figure 8: Residual plots of the fits to the invariant mass distributions for $D^+_{(s)} \to \pi^+\mu^+\mu^-$ candidates in the higher mass fit (including trigger lines with $m(\mu^+\mu^-) > 1.0$ GeV/c^2). Shown are the (a) ϕ and (b) high-$m(\mu^+\mu^-)$ bins.

Selection	**Low-**$m(\mu^+\mu^-)$	η	ρ^0	ϕ	**High-**$m(\mu^+\mu^-)$
Acceptance	20.10 ± 0.49	20.18 ± 1.07	19.77 ± 0.36	19.33 ± 0.28	19.82 ± 0.32
Reco/Stripping	1.310 ± 0.04	1.250 ± 0.09	1.280 ± 0.03	1.710 ± 0.03	1.800 ± 0.04
BDT	0.626 ± 0.02	0.594 ± 0.05	0.579 ± 0.02	0.713 ± 0.01	0.569 ± 0.01
Trigger	0.173 ± 0.009	0.182 ± 0.02	0.169 ± 0.007	0.218 ± 0.006	0.236 ± 0.008
High Trigger	-	-	-	0.264 ± 0.007	0.301 ± 0.01

Table 4: Efficiency, as a percentage, calculated on simulated 2011 $D^+ \to \pi^+\mu^+\mu^-$ signal.

Selection	**Low-**$m(\mu^+\mu^-)$	η	ρ^0	ϕ	**High-**$m(\mu^+\mu^-)$
Acceptance	20.42 ± 0.60	18.17 ± 1.18	19.84 ± 0.43	19.57 ± 0.34	20.09 ± 0.38
Reco/Stripping	1.210 ± 0.04	1.030 ± 0.07	1.230 ± 0.03	1.630 ± 0.03	1.680 ± 0.03
BDT	0.579 ± 0.02	0.486 ± 0.04	0.553 ± 0.01	0.661 ± 0.01	0.519 ± 0.01
Trigger	0.177 ± 0.008	0.158 ± 0.02	0.182 ± 0.006	0.224 ± 0.005	0.234 ± 0.006
High Trigger	-	-	-	0.282 ± 0.007	0.290 ± 0.007

Table 5: Efficiency, as a percentage, calculated on simulated 2012 $D^+ \to \pi^+\mu^+\mu^-$ signal.

Bibliography

[1] Particle Data Group, J. Beringer *et al.*, *Review of Particle Physics (RPP)*, Phys. Rev. **D86** (2012) 010001.

[2] E. Majorana, *Teoria simmetrica dell'elettrone e del positrone*, Nuovo Cim. **14** (1937) 171.

[3] LHCb collaboration, R. Aaij *et al.*, *Search for Majorana neutrinos in* $B^- \to \pi^+\mu^-\mu^-$ *decays*, Phys. Rev. Lett. **112** (2014) 131802, arXiv:1401.5361.

[4] LHCb collaboration, R. Aaij *et al.*, *Search for* $D^+_{(s)} \to \pi^+\mu^+\mu^-$ *and* $D^+_{(s)} \to \pi^-\mu^+\mu^+$ *decays*, Phys. Lett. **B724** (2013) 203, arXiv:1304.6365.

[5] ATLAS Collaboration. https://twiki.cern.ch/twiki/pub/AtlasPublic/CombinedSummaryPlots/AtlasSearchesSUSY_SUSY2013.pdf; accessed 7-July-2014.

[6] CMS Collaboration. https://twiki.cern.ch/twiki/pub/CMSPublic/SUSYSMSSummaryPlots8TeV/SUSY2013T2ttT6.pdf; accessed 7-July-2014.

[7] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.

[8] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. D **2** (1970) 1285.

[9] M. Bott-Bodenhausen *et al.*, *Search for decay of neutral kaons into charged lepton pairs*, Physics Letters B **24** (1967), no. 4 194 .

[10] H. Foeth *et al.*, *Search for $K_L \to \mu^+\mu^-$ and $K_L \to e^+e^-$ decays*, Physics Letters B **30** (1969), no. 4 282 .

[11] K. Lande *et al.*, *Observation of Long-Lived Neutral V Particles*, Phys. Rev. **103** (1956) 1901.

[12] K. Lande, L. M. Lederman, and W. Chinowsky, *Report on Long-Lived K^0 Mesons*, Phys. Rev. **105** (1957) 1925.

[13] J. J. Aubert *et al.*, *Experimental Observation of a Heavy Particle J*, Phys. Rev. Lett. **33** (1974) 1404.

[14] J. E. Augustin *et al.*, *Discovery of a Narrow Resonance in e^+e^- Annihilation*, Phys. Rev. Lett. **33** (1974) 1406.

[15] E. G. Cazzoli *et al.*, *Evidence for ΔS=-ΔQ Currents or Charmed-Baryon Production by Neutrinos*, Phys. Rev. Lett. **34** (1975) 1125.

[16] G. Goldhaber *et al.*, *Observation in e+ e- Annihilation of a Narrow State at 1865* MeV/c^2 *Decaying to $K\pi$ and $K\pi\pi\pi$*, Phys. Rev. Lett. **37** (1976) 255.

[17] I. Peruzzi *et al.*, *Observation of a Narrow Charged State at 1876* MeV/c^2 *Decaying to an Exotic Combination of $K\pi\pi$*, Phys. Rev. Lett. **37** (1976) 569.

[18] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[19] Z.-z. Xing, *Wolfenstein parametrization revisited*, Phys. Rev. **D51** (1995) 3958, `arXiv:hep-ph/9411340`.

[20] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the B_s^0-$\bar{B}_s^0$ oscillation frequency with the decay $B_s^0 \to D_s^-\pi^+$*, New J. Phys. **15** (2013) 053021, `arXiv:1304.4741`.

[21] LHCb Collaboration, R. Aaij *et al.*, *Observation of $D^0 - \overline{D}^0$ oscillations*, Phys. Rev. Lett. **110** (2013), no. 10 101802, `arXiv:1211.1230`.

[22] J. Brod, A. L. Kagan, and J. Zupan, *Size of direct CP violation in singly Cabibbo-suppressed D decays*, Phys. Rev. **D86** (2012) 014023, arXiv:1111.5000.

[23] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in the phase space of $B^\pm \to K^+K^-\pi^\pm$ and $B^\pm \to \pi^+\pi^-\pi^\pm$ decays*, Phys. Rev. Lett. **112** (2014) 011801, arXiv:1310.4740.

[24] LHCb collaboration, R. Aaij *et al.*, *Search for CP violation in $D^+ \to \phi\pi^+$ and $D_s^+ \to K_S^0\pi^+$ decays*, JHEP **1306** (2013) 112, arXiv:1303.4906.

[25] LHCb Collaboration, R. Aaij *et al.*, *Search for CP violation in $D^\pm \to K_S^0 K^\pm$ and $D_s^\pm \to K_S^0\pi^\pm$ decays*, arXiv:1406.2624.

[26] LHCb collaboration, *A search for time-integrated CP violation in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays*, LHCb-CONF-2013-003.

[27] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays*, arXiv:1405.2797.

[28] LHCb collaboration, R. Aaij *et al.*, *Measurement of $D^0 - \bar{D}^0$ mixing parameters and search for CP violation using $D^0 \to K^+\pi^-$ decays*, Phys. Rev. Lett. **111** (2013) 251801, arXiv:1309.6534.

[29] LHCb collaboration, R. Aaij *et al.*, *Search for the rare decay $D^0 \to \mu^+\mu^-$*, Phys. Lett. **B725** (2013) 15, arXiv:1305.5059.

[30] LHCb collaboration, R. Aaij *et al.*, *Search for the decay $D^0 \to \pi^+\pi^-\mu^+\mu^-$*, Phys. Lett. **B728** (2014) 234, arXiv:1310.2535.

[31] NA48/2 collaboration, J. Batley *et al.*, *New measurement of the $K^\pm \to \pi^\pm\mu^+\mu^-$ decay*, Phys. Lett. **B697** (2011) 107, arXiv:1011.4817.

[32] LHCb Collaboration, R. Aaij *et al.*, *First observation of the decay $B^+ \to \pi^+\mu^+\mu^-$*, JHEP **1212** (2012) 125, arXiv:1210.2645.

[33] LHCb Collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the $B^+ \to K^+\mu^+\mu^-$ decay*, JHEP **1302** (2013) 105, arXiv:1209.4284.

[34] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[35] J. J. Goodson and R. McCarthy, *Search for Supersymmetry in States with Large Missing Transverse Momentum and Three Leptons including a Z-Boson*, PhD thesis, Stony Brook U., May, 2012, Presented 17 Apr 2012.

[36] LHCb Collaboration. http://lhcb-public.web.cern.ch/lhcb-public/Images2013/IntegratedLumi_Yearly.png; accessed 7-July-2014.

[37] LHCb collaboration, R. Aaij *et al.*, *Prompt charm production in pp collisions at $\sqrt{s} = 7$ TeV*, Nucl. Phys. **B871** (2013) 1, arXiv:1302.2864.

[38] LHCb Collaboration, R. Aaij *et al.*, *Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} =$ 7 TeV in the forward region*, Phys. Lett. **B694** (2010) 209, arXiv:1009.2731.

[39] LHCb Collaboration. http://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles/07_rad_acc_scheme_right.pdf; accessed 7-July-2014.

[40] LHCb Collaboration, J. Alves, A. Augusto *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[41] G. Krocker, *CP violation in B_s^0 mixing with LHCb*, arXiv:1112.0868.

[42] A. Affolder *et al.*, *Radiation damage in the LHCb Vertex Locator*, JINST **8** (2013) P08002, arXiv:1302.5259.

[43] P. Rodríguez Pérez, *The LHCb Vertex Locator performance and Vertex Locator upgrade*, Journal of Instrumentation **7** (2012) C2008, arXiv:1209.4845.

[44] LHCb Collaboration. http://lhcb-public.web.cern.ch/lhcb-public/en/detector/Trackers2-en.html; accessed 7-July-2014.

[45] LHCb Outer Tracker group, R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) 01002, arXiv:1311.3893.

[46] R. Froehwirth, *Application of kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987), no. 23 444 .

[47] LHCb Collaboration, J. Van Tilburg, *Matching VELO tracks with seeding tracks*, LHCb-2001-103 (2001).

[48] LHCb Collaboration, R. Aaij *et al.*, *Measurement of b-hadron branching fractions for two-body decays into charmless charged hadrons*, JHEP **1210** (2012) 037, arXiv:1206.2794.

[49] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, European Physical Journal C **73** (2013) 2431, arXiv:1211.6759.

[50] A. M. Ruland and the BaBar calorimeter group, *Performance and operation of the BaBar calorimeter*, Journal of Physics: Conference Series **160** (2009), no. 1 012004.

[51] CMS Collaboration. http://www.slac.stanford.edu/econf/C060717/papers/L012.PDF; accessed 7-July-2014.

[52] LHCb Collaboration, O. Deschamps *et al.*, *Photon and neutral pion reconstruction*, LHCB-2003-091 (2003).

[53] F. Archilli *et al.*, *Performance of the Muon Identification at LHCb*, Journal of Instrumentation **8** (2013) 20P, arXiv:1306.0249.

[54] E. Aslanides *et al.*, *The Level-0 muon trigger for the LHCb experiment*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **579** (2007), no. 3 989 .

[55] LHCb Collaboration. http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html; accessed 7-July-2014.

[56] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[57] S. Fajfer, N. Kosnik, and S. Prelovsek, *Updated constraints on new physics in rare charm decays*, Phys. Rev. **D76** (2007) 074010, arXiv:0706.1133.

[58] G. Burdman, E. Golowich, J. L. Hewett, and S. Pakvasa, *Rare charm decays in the standard model and beyond*, Phys. Rev. **D66** (2002) 014009, arXiv:hep-ph/0112235.

[59] S. Fajfer, S. Prelovsek, and P. Singer, *Rare charm meson decays $D \to Pl^+l^-$ and $c \to ul^+l^-$ in SM and MSSM*, Phys. Rev. **D64** (2001) 114009, arXiv:hep-ph/0106333.

[60] A. Paul, I. I. Bigi, and S. Recksiegel, *On $D \to X_u l^+ l^-$ within the Standard Model and Frameworks like the Littlest Higgs Model with T Parity*, Phys. Rev. **D83** (2011) 114006, arXiv:1101.6053.

[61] D0 collaboration, V. Abazov *et al.*, *Search for flavor-changing-neutral-current D meson decays*, Phys. Rev. Lett. **100** (2008) 101801, arXiv:0708.2094.

[62] FOCUS Collaboration, J. Link *et al.*, *Search for rare and forbidden three body dimuon decays of the charmed mesons D^+ and D_s^+*, Phys. Lett. **B572** (2003) 21, arXiv:hep-ex/0306049.

[63] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.

[64] D. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[65] J. H. Friedman, *Greedy function approximation: A gradient boosting machine*, Annals of Statistics **29** (2000) 1189.

[66] J. H. Friedman, *Stochastic gradient boosting*, Comput. Stat. Data Anal. **38** (2002) 367.

[67] A. Hoecker *et al.*, *TMVA - Toolkit for Multivariate Data Analysis*, ArXiv Physics e-prints (2007) arXiv:physics/0703039.

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

[69] ARGUS collaboration, H. Albrecht *et al.*, *Search for hadronic $b \to u$ decays*, Physics Letters B **241** (1990), no. 2 278 .

[70] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv:physics/0306116.

[71] M. Fischler and D. Sachs, *An Object oriented minimization package for HEP*, eConf **C0303241** (2003) MOLT004, arXiv:hep-ph/0306054.

[72] BaBar Collaboration, P. del Amo Sanchez *et al.*, *Study of $B \to X\gamma$ decays and determination of $|V_{td}/V_{ts}|$*, Phys. Rev. **D82** (2010) 051101, arXiv:1005.4087.

[73] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, The Annals of Mathematical Statistics **9** (1938) 60.

[74] E. Greening, O. Kochebina, M. John, and B. Viaud, *Searches for $D^{\pm}_{(s)} \to h^{\pm}\mu^{+}\mu^{-}$ and $D^{\pm}_{(s)} \to h^{\pm}\mu^{\pm}\mu^{\pm}$*, LHCb-ANA-2012-025. Linked to LHCb-PAPER-2012-051.

[75] T. Junk, *Confidence level computation for combining searches with small statistics*, Nucl. Instrum. Meth. **A434** (1999) 435, arXiv:hep-ex/9902006.

[76] A. L. Read, *Presentation of search results: The CL(s) technique*, J. Phys. **G28** (2002) 2693.

[77] BaBar Collaboration, J. Lees *et al.*, *Searches for Rare or Forbidden Semileptonic Charm Decays*, Phys. Rev. **D84** (2011) 072006, arXiv:1107.4465.

[78] BaBar Collaboration, B. Aubert *et al.*, *Measurement of the $D^{+} \to \pi^{+}\pi^{0}$ and $D^{+} \to K^{+}\pi^{0}$ branching fractions*, Phys. Rev. **D74** (2006) 011107, arXiv:hep-ex/0605044.

[79] CLEO Collaboration, H. Mendez *et al.*, *Measurements of D Meson Decays to Two Pseudoscalar Mesons*, Phys. Rev. **D81** (2010) 052013, arXiv:0906.3198.

[80] F. Soomro, B. Cameron, and A. Golutvin, *Radiative decays of B mesons at LHCb*, PhD thesis, Imperial Coll., London, London, 2011, Presented 03 Jun 2011.

[81] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^0 \to K^{*0}e^+e^-$ branching fraction at low dilepton mass*, JHEP **1305** (2013) 159, arXiv:1304.3035.

[82] J. Brodzicka and M. Gersabeck, *Private communication*, 9-July-2014.

[83] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the $D^\pm$ production asymmetry in 7 TeV pp collisions*, Phys. Lett. **B718** (2013) 902, arXiv:1210.4112.

[84] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the $D_s^+ - D_s^-$ production asymmetry in 7 TeV pp collisions*, Phys. Lett. **B713** (2012) 186, arXiv:1205.0897.